IJACSA
WHERE WISDOM SHARES
International Journal of Advanced Computer Science and Applications

Volume 4 Issue 7

July 2013

ISSN 2156-5570(Online)
ISSN 2158-107X(Print)

SAI

www.ijacsa.thesai.org

INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS

THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org

Editorial Preface

From the Desk of Managing Editor…

It is our pleasure to present to you the July 2013 Issue of International Journal of Advanced Computer Science and Applications.

Today, it is incredible to consider that in 1969 men landed on the moon using a computer with a 32-kilobyte memory that was only programmable by the use of punch cards. In 1973, Astronaut Alan Shepherd participated in the first computer "hack" while orbiting the moon in his landing vehicle, as two programmers back on Earth attempted to "hack" into the duplicate computer, to find a way for Shepherd to convince his computer that a catastrophe requiring a mission abort was not happening; the successful hack took 45 minutes to accomplish, and Shepherd went on to hit his golf ball on the moon. Today, the average computer sitting on the desk of a suburban home office has more computing power than the entire U.S. space program that put humans on another world!!

Computer science has affected the human condition in many radical ways. Throughout its history, its developers have striven to make calculation and computation easier, as well as to offer new means by which the other sciences can be advanced. Modern massively-paralleled super-computers help scientists with previously unfeasible problems such as fluid dynamics, complex function convergence, finite element analysis and real-time weather dynamics.

At IJACSA we believe in spreading the subject knowledge with effectiveness in all classes of audience. Nevertheless, the promise of increased engagement requires that we consider how this might be accomplished, delivering up-to-date and authoritative coverage of advanced computer science and applications.

Throughout our archives, new ideas and technologies have been welcomed, carefully critiqued, and discarded or accepted by qualified reviewers and associate editors. Our efforts to improve the quality of the articles published and expand their reach to the interested audience will continue, and these efforts will require critical minds and careful consideration to assess the quality, relevance, and readability of individual articles.

To summarise, the journal has offered its readership thought provoking theoretical, philosophical, and empirical ideas from some of the finest minds worldwide. We thank all our readers for their continued support and goodwill for IJACSA. We will keep you posted on updates about the new programmes launched in collaboration.

Lastly, we would like to express our gratitude to all authors, whose research results have been published in our journal, as well as our referees for their in-depth evaluations.

We hope that materials contained in this volume will satisfy your expectations and entice you to submit your own contributions in upcoming issues of IJACSA

Thank you for Sharing Wisdom!

Managing Editor
IJACSA
Volume 4 Issue 7 July 2013
ISSN 2156-5570 (Online)
ISSN 2158-107X (Print)

Editorial Board

Reviewer Board Members

CONTENTS

Relation Inclusive Search for Hindi Documents

Pooja Arora
Research Scholar, Banasthali Vidyapith
Assistant Professor, MCA Department AKGEC
Ghaziabad, India

Om Vikas
Former Director, ABV, IITM, Gwalior
Senior Member, IEEE
India

***Abstract*—Information retrieval (IR) techniques become a challenge to researchers due to huge growth of digital and information retrieval. As a wide variety of Hindi Data and Literature is now available on web, we have developed information retrieval system for Hindi documents. This paper presents a new searching technique that has promising results in terms of F-measure. Historically, there have been two major approaches to IR - keyword based search and concept based search. We have introduced new relation inclusive search which performs searching of documents using case role relation, spatial relation and temporal relation of query terms and gives results better than previously used approaches. In this method we have used new indexing technique which stores information about relation between terms along with its position. We have compared four types of searching: Keyword Based search without Relation Inclusive, Keyword Based search with Relation Inclusive, Concept Based search without Relation Inclusive and Concept Based search with Relation Inclusive. Our proposed searching method gave significant improvement in terms of F-measure. For experiments we have used Hindi document corpus, Gyannidhi from C-DAC. This technique effectively improves search performance for documents in English as well.**

Keywords—Relation inclusive search; RSearch; spatial & temporal prepositions and postpositions; Hindi document retrieval; case roles.

I. INTRODUCTION

World Wide Web is used to share huge repository of texts in many languages as everybody can post content in any language. Search across multiple languages is desirable with the increase of many languages on the web. In order to enable a wider proportion of population to benefit from Information technology, it is desirable that human-machine interface permits one's native language of communication. In the context of a multi-lingual country like India, this can be of immense value. The users will be greatly helped if they are able to retrieve information in a language in which they are comfortable. In the context of Indian languages, Hindi language has been given much emphasis leading to the development of significant number of Hindi documents. According to Ethnologue statistics, among the list of the top 10 languages, Hindi is coming forth with 260 million first language speakers and English is coming third with 335 million first language speakers [14]. Hindi language is spoken by 41% of population of India whereas about 5% of population understands English as their second language [10]. A wide variety of Hindi Data and Literature is now available on web [6]. The number of users who want the information in Hindi language is increasing. We have developed an Information Retrieval System (IRS) for Hindi documents. Various search engines are available on the internet as independent search engine sites in Hindi like Google [11], Raftaar [12] and Hinkhoj [13]. The retrieval accuracy of these search engines does not satisfy users need. Increasing the appropriateness of the results returned by these search engines is critical to dealing with the huge repository of the data.

Historically, there have been two major approaches to IR – **keyword based search** and **concept based search**. In keyword based search, search engines use words or multi-words phrases that occur in documents and queries as atomic elements and the content of the documents are described by list of keywords. The search procedure, used by these search engines, is principally based on the exact matching of document and query terms and does not take into consideration the various meanings or possible concepts that a word represents [4], [5]. If the user chooses a valid synonymous word that is not in any document then it would fail the search. In general, this approach has many problems as the user may not get the most relevant and useful content related to the query. The first problem is low precision, which is due to the irrelevance of many of the search results. This results in a difficulty finding the relevant information. The second problem is low recall, which is due to the inability to index all the information available on the Web. This results in a difficulty finding the unindexed information that is relevant. Also the level of deepness of analysis of the language is very low, so the relevant information is not retrieved by the search engines.

On the other hand, concept based search fetch document based on their meaning rather than the presence of the keywords. Here, the meaning of words is analysed and not only their syntactic representations. This type of searching uses query expansion techniques where the initial query is appended with related, contextual, or synonymous terms so as to make the new query more complete to define the required concept [6], [7], [8]. But, this approach also uses the words that occur in query and in documents as atomic elements and no relation between them is used for retrieval. However, concept-based approaches allow reaching a higher precision than keyword based approaches.

There is a need to consider relation between terms present in the documents. Documents could contain three types of relations: **case role relation, spatial relation or temporal relation**. For example, "***टेबल के ऊपर किताब***" (*table ke upar kitaab*) => "***book on table***" contains spatial relation i.e. position of ***किताब*** w.r.t ***टेबल*** given by ***के ऊपर*** (postposition) and position of ***book*** w.r.t. ***table*** given by ***on*** (preposition); ***"श्याम ने राम को किताब दी"*** (*Shyam ne Ram ko kitaab di*) contains case role (karaka) relation. Here, case role of ***श्याम*** (*Shyam*) is **कर्ता** (nominative case) and ***राम*** (Ram) is **कर्म** (objective case); "***राम***

से पहले श्याम घर जाएगा" (*Ram se pehle Shyam ghar jayega*) contains temporal relation given by "***पहले***". If the documents are indexed without considering relational information and remove prepositions/postpositions as stopwords before indexing of the documents to reduce index file size, then the relation existing between terms is lost. Taking this point into account, we propose **relation inclusive search** as a new and promising way of improving search on the IRS. We call it ***RSearch*** (relation inclusive search). The main idea is to keep the same infrastructure which has made previous methods so successful, thus improving the system performance. Informally, ***RSearch*** do not use terms present in query and in documents as atomic elements, but the semantics of relationships between query terms is considered for improving search. While indexing it stores all the relational information existing between terms present in the documents in the index file and use that information for retrieving relevant documents.

To reduce the size of index file we have removed stop words before indexing the documents. We have categorized stopwords in two categories – Relational (**नीचे, ऊपर, पर, आगे, अंदर, ने, को, है, था** etc) & Non-relational (**ही, तब, यह** etc). These stopwords have different impact on the information retrieval process. Relational stopwords indicate semantic relevance that is necessary for efficient information retrieval. Removing relational stopwords from the document would result in loss of such relevant semantic information resulting in decrease of relevance efficiency of the system. While removing non-relational stopwords would reduce the document length resulting into faster search. So, we remove only non-relational stopwords.

II. Relation Inclusive Search

There are multiple objects in a query which are dependent on each other. Relations between objects are given by prepositions (in English) and postpositions (in Hindi) depending upon the language of the document. We have proposed a new relation inclusive searching technique ***RSearch*** which stores relation between terms present in the document in the index file and also considers that relation for retrieving relevant documents from the corpus. It is based on new indexing scheme. It fetches documents based on **Case Role** (**Karaka) relation**, **Spatial relation** and **Temporal relation** existing between query terms. All these types of relations are discussed in the following subsections in detail. Before that, let us consider the document collection shown in figure 1. In figure 2, we showed examples of ten queries, which are submitted to this document collection. We will use the same document collection and sample queries in the following subsections.

III. Types Of Relations

There are three types of relations we have considered in relation inclusive search.

A. Case Role (Karaka) relation

Hindi language has eight types of case roles which are shown below with related **परसर्ग**/suffixes**.** (see table I) [9]

D1: दो दिन पहले बारिश हुई थी। **do din pehle barish huee thee**
D2: किताब टेबल पर रखी है। **kitaab table par rakhi hai**
D3: हिंदी की किताब टेबल पर रखी थी। **Hindi kee kitaab table par rakhi thee**
D4: एक कुत्ता टेबल के नीचे से कूदकर भाग गया। **ek kutta table ke niche se koodkar bhaag gaya**
D5: श्याम ने एक जानवर को मारा। **Shyam ne ek janwar ko mara**
D6: श्याम के बाद राम घरजाएगा। **Shyam ke baad Ram ghar jayega**

Fig.1. A document collection

Query #	Query	Type of relation present
1	दो दिन के बाद बारिश होगी do din ke baad barish hogi	Temporal relation
2	वृक्ष के पास जानवर vriksh ke paas janwar	Spatial relation
3	हिन्दी की पुस्तक टेबल के नीचे है Hindi kee pustak table ke niche hai	Spatial relation
4	टेबल पर किताब रखी है table par kitaab rakhi hai	Spatial relation
5	राम निवास के पास खड़ा है Ram niwaas ke paas khara hai	Spatial relation
6	जानवर ने श्याम को मारा janwar ne Shyam ko mara	Case role relation
7	औरत ने चाकू से फल काटा aurat ne chaku se phal kata	Case role relation
8	डिब्बे के अंदर पुस्तक dibbe ke ander pustak	Spatial relation
9	रूम में टी. वी. रखा है room mein T.V. rakha hai	Spatial relation
10	राम से पहले श्याम घर जाएगा Ram se pehle Shyam ghar jayega	Temporal relation

Fig.2. Sample queries

Consider the following sentence:

Sentence 1: **जानवर ने श्याम को मारा।**
janwar ne Shyam ko mara
Animal killed Shyam.

Here, case role of **जानवर** is **कर्ता** (nominative case) and **श्याम** is **कर्म** (objective case).

As Hindi is free order language; order of words contains only secondary information such as emphasis etc. Primary information relating to 'gross' meaning (e.g., one that includes semantic relationships) is contained elsewhere [1], [2], [3].

Therefore, same sentence could be written as:

Sentence 2: **श्याम को जानवर ने मारा।**

An answer for query Q6 in figure 2 using keyword based searching would be D5 as this document contains all the query terms. But this is not the correct answer as case roles of **जानवर** and **श्याम** do not match. In D5, the case role of **श्याम** is **कर्ता** (nominative case) and **जानवर** is **कर्म** (objective case) but reverse is the case in Q6.

So, to improve the information retrieval effectiveness in terms of precision and recall, we have stored the case role relation between terms in the index file.

1) Indexing a Case Role relation

While indexing the documents, the **case role** (giving the karaka relation) of the word is also added in its posting list along with its position in a particular document. For example,

TABLE I. CASE ROLES WITH RELATED SUFFIXES

S.No.	कारक / Case Role	परसर्ग (In Hindi)	Suffixes (In English)
(i)	कर्ता (karta)/ nominative case (K1)	ने, ne	by
(ii)	कर्म (karma)/ objective case (K2)	को, ko	to
(iii)	करण (karana)/ instrumental case (K3)	से (se), के साथ (ke sath), के द्वारा (ke dvara)	with
(iv)	संप्रदान (sampradaan)/ dative case (K4)	के लिए (ke liye), को (ko)	for, to
(v)	अपादान (apaadaan)/ ablative case (K5)	से (se)	taking from (separation / detachment)
(vi)	सम्बन्ध (sambandh)/ relative case (K6)	का (ka), के (ke), की (ki), रा (ra), रे (re), री (ri)	belong to (something / somebody)
(vii)	अधिकरण (adhikarana)/ locative case (K7)	में (mein), पर (par)	in, on
(viii)	सम्बोधन (sambodhan)/ vocative case (K8)	हे! (hey!), अरे! (are!)	oh!

जानवर is present at 4th position in document D5 with case role as **कर्म (K2)** is represented as D5:[[4,K2]] and **श्याम** is present at 1st position in document D5 with case role as **कर्ता (K1)** is represented as D5:[[1,K1]]**.** (see figure 3)

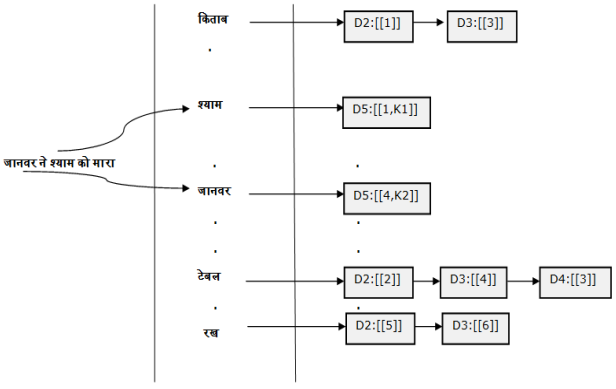


Fig.3. Inverted Index file in relation inclusive search

2) Steps for matching

To retrieve the documents for query Q6, matching involves the following steps:

*1) Retrieve the posting lists of **जानवर** and **श्याम.***

2) Get the set of documents from the posting lists containing both the terms and also their case roles in those documents.

3) Find out the case roles of query terms.

4) For each document check if the case roles of both the terms are matching with their case roles present in the query.

In this example, document D5 contains both the terms. In D5, case role of **जानवर** is **कर्म (K2)** and that of **श्याम** is **कर्ता (K1).** But the **जानवर** is **कर्ता (K1)** and **श्याम** is **कर्म (K2)** in Q6. As the case roles of **जानवर** and **श्याम** are not matching, an answer for query *Q*6, computed by relation based search is the empty set which is the correct answer.

B. Spatial relation

Spatial relation shows position of an object w.r.t another object.

For example,

Document D2 gives position of **किताब** w.r.t **टेबल** (see figure 1).

1) Indexing a Spatial relation

All the relational stopwords giving spatial relation are indexed in separate inverted file. For example, in posting list of **पर** - D2:[[2,1]] means in document D2 **पर** gives the relation between words present at positions 2 (**टेबल**) and 1 (**किताब**) respectively (see figure 4).

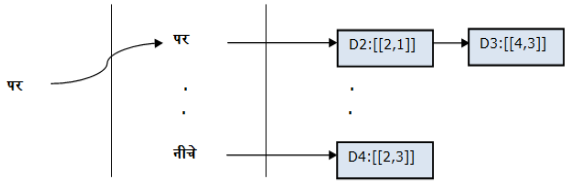


Fig.4. Inverted Index of Relational Stopwords

2) Representation of Spatial relation

We are extracting relations in the form of triples:

$Word_1$=>RELATION<=$Word_2$ and they are used as indexing terms.

For example:

In documents D2 and D3, relation **टेबल =>पर<= किताब** exists, that means **किताब** is lying on the **टेबल.**

In document D4**, टेबल =>नीचे<= कुत्ता** exists, that means **कुत्ता** is sitting below the **टेबल.**

3) Spatial relations extracted

Currently, we have extracted the following relations:

1) *नीचे* *(below)*
2) *ऊपर/पर* *(on)*
3) *पास* *(near)*
4) *दूर* *(far)*
5) *आगे* *(front)*
6) *पीछे* *(back)*
7) *अंदर* *(inside)*
8) *बाहर* *(outside)*

4) Properties of Spatial relation

Following are the properties of spatial relation:

1. If relation X=>REL1<=Y exists, then relation Y=>REL2<= X also exists such that REL1 is antonym (opposite relation) of REL2.

For example, in document D2 if relation **टेबल=>पर<=किताब** exists, then relation **किताब=>नीचे<=टेबल** also exists as relations **पर** and **नीचे** are opposite relations.
Various opposite relation are:

a) ऊपर/पर opposite of नीचे
b) पास opposite of दूर
c) आगे opposite of पीछे
d) अंदर opposite of बाहर

2. If relations X=>REL<=Y and Y=>REL<=Z exists then relation X=>REL<=Z also exists.

For example, if relations **टेबल=>पर<=किताब** and **किताब=>पर<=कलम** exists, then relation **टेबल=>पर<=कलम** also exists.

3. If relation X=>REL<=Y exists, then relation Y=>REL<=X also exists for relations in set REL={पास, दूर}

For example, if relation **राम=>पास<=रहीम** exists, then relation **रहीम=>पास<=राम** also exists.

5) Steps for matching

For example,

Consider the query Q4 in figure 2. Following steps are involved for searching of relevant documents:

Steps:

*1) Perform stemming on query terms. Query Q4 reduced to: **टेबल पर किताब रख** (see figure 5).*

2) Retrieve posting lists of query terms.

*3) Using posting lists, we get the document set containing query terms **टेबल, पर, किताब** and **रख** {D2, D3}. We left D4 as it does not contain terms **किताब, पर** and **रख**.*

*4) Match query relation with the relation present in the resultant documents. Relation in query is **टेबल =>पर<= किताब**.*

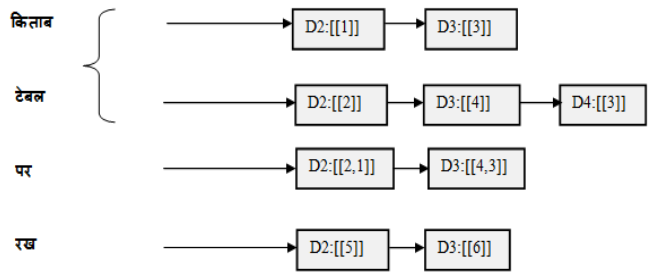


FIG.5. Posting lists of query terms

Now the relation retrieved from posting lists is checked. Posting list of **पर** contains D2:[[2,1]], D3:[[4,3]]. That means D2 contains relation **पर** between words at position 2 and 1 which are **टेबल** and **किताब** respectively giving relation **टेबल=>पर<=किताब**. Similarly, D3 contains relation **पर** between words at position 4 and 3 which are **टेबल** and **किताब** respectively giving relation **टेबल =>पर<= किताब**. As the relation is matching in both the documents, {D2, D3} resultant document set is retrieved.

C. Temporal relation

All the temporal related words (**पहले, बाद मे**) come under relational stop words.

For example,
दो दिन पहले बारिश हुई थी।
do din pehle barish huee thee

Here, "**दो दिन पहले**" gives the temporal relation. If we remove it, the temporal relation would be lost and the system does not retrieve relevant document. In order to improve the efficiency of the system we don't discard temporal related words (**पहले, बाद मे**) as stop words.

1) Indexing a Temporal relation

All the relational stopwords giving temporal relation are indexed in separate inverted file. The posting lists of temporal related words contain temporal information. We have considered two types of temporal relation- **Number_of_days** & **Order_of_entities**.

1. Number_of_days: If the document/query contains words like {**दिन**, **दिवस**} along with words {**पहले, बाद**}, then the temporal relation is "**Number_of_days**". In that case we store number of days in the posting list of temporal related words along with their positions.

For example,

In document D1, temporal related word **पहले** is present at 3rd position and number of days are 2. Same information is stored in the posting list of **पहल** (stemmed word) (see figure 6).

2. Order_of_entities: If the document/query contains words like **से, के** along with words {**पहले, बाद**}, then the temporal relation is "**Order_of_entities**". In that case we store order of entities in the posting list of temporal related words along with their positions. However, if the query contains temporal related word **पहले**, then we check for the posting list of **बाद** also and vice versa as **पहले** and **बाद** are opposite words.

For example,

In document D6, temporal related word **बाद** is present at 3rd position giving the order of entities between **श्याम** & **राम** which are stored at 1st and 4th position respectively. Same information is stored in the posting list of **बाद**. Relation between **श्याम** & **राम** is **श्याम => बाद <= राम** (see figure 6).

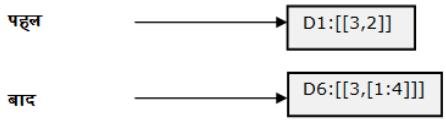


Fig.6. Indexing a temporal relation

2) Steps for matching

First we identify if there are temporal related words present in the query. If query contains temporal words then retrieve the posting list of those words from their index file.

For example,

Consider the query Q10 in figure 2. Following steps are involved for searching of relevant documents:

1) Stem the query terms and remove non-relational stop words. Then query Q10 is reduced to:

राम से पहल श्याम घर जाए

2) Retrieve posting lists of query terms excluding relational stop words (see figure 7).

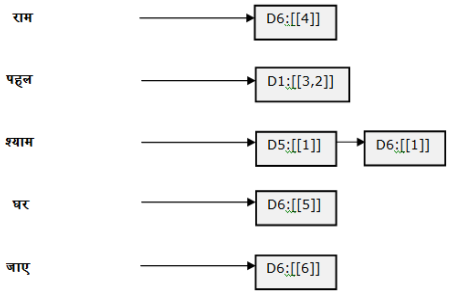
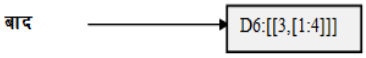


Fig.7. Posting lists of query terms

*3) Retrieve the posting list of term opposite of **पहले** i.e. **बाद.***



*4) Using posting lists, we get the document set containing query terms **राम श्याम घर जाए** {D6}. From the posting list of **बाद** relation extracted is **श्याम**=> **बाद<=राम** which is opposite of **राम =>पहल <= श्याम.** So, relation present in D6 is matching with relation present in query.*

5) Resultant document set {D6} is retrieved which is the correct answer.

IV. Evaluation Measures

A. **Recall:** It the fraction of the documents that are relevant to the query that are successfully retrieved. [4]

$$Recall = \frac{Number\_Relevant\_Retrieved}{Number\_Relevant}$$

B. **Precision:** It is the fraction of the documents retrieved that are relevant to the user's information need.

$$Precision = \frac{Number\_Relevant\_Retrieved}{Number\_Total\_Retrieved}$$

C. **F-measure:** It is the weighted harmonic mean of precision and recall.

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

V. Experimental Evaluation

We have implemented Information Retrieval System in Java. For experiments, we have used Hindi test collection of about 100 documents extracted from gyannidhi corpus from CDAC. Details of the Hindi corpus are mentioned in table II. Queries on which estimation is done are shown in figure 2. Table III shows normalization of the sample queries of figure 2 for different types of retrieval. Query normalization means removal of stop words and stemming of the remaining words.

We have compared four types of searching: Keyword Based search without Relation Inclusive, Keyword Based search with Relation Inclusive, Concept Based search without Relation Inclusive and Concept Based search with Relation Inclusive.

Keyword based retrieval follow standard normalization in which all the stop words are removed and words are stemmed. In concept based search, query is expanded using synonyms of the words while in relation inclusive search case role relation, spatial relation and temporal relation among words are considered in addition.

A. Estimation of F-measure

Sample queries shown in figure 2 are run on the four environments as mentioned above to get the results. Table IV gives the estimation of F-measure for the four types of searching when they run on same set of queries. From the estimation, it can be seen that F-measure for relation inclusive searching is much higher as compared to that of without relation inclusive. In some cases where synonyms of words are not being used, F-measure is coming out to be same for keyword based searching and concept based searching. For example, for queries Q4, Q6 and Q10 F-measures are 0.57 and 0.80, 0.75 and 0.86, 0.67 and 0.77 respectively for keyword based searching and concept based searching.

TABLE II. DETAILS OF HINDI DOCUMENT CORPUS

	Hindi Document Corpus
No. of Documents	100
No. of index terms	25374
No. of queries	20
Average No. of terms/doc	336

TABLE III. QUERY NORMALIZATION

Query #	Query	Query Normalization			
		Keyword based -Relation Inclusive	Keyword based +Relation Inclusive	Concept based -Relation Inclusive	Concept based +Relation Inclusive
1	दो दिन के बाद बारिश होगी It will be raining after two days	दिन बारिश rain day	दो दिन बाद बारिश rain after two day	दिन बरसात rain day	दो दिन बाद बरसात rain after two day
2	वृक्ष के पास जानवर animal near tree	वृक्ष जानवर animal tree	वृक्ष पास जानवर animal near tree	पेड़ जानवर animal tree	पेड़ पास जानवर animal near tree
3	हिन्दी की पुस्तक टेबल के नीचे है Hindi book is lying below table	हिन्दी पुस्तक टेबल Hindi book table	हिन्दी पुस्तक टेबल नीच Hindi book below table	हिन्दी किताब टेबल Hindi book table	हिन्दी किताब टेबल नीच Hindi book below table
4	टेबल पर किताब रखी है book is lying on table	टेबल किताब book table	टेबल पर किताब book on table	टेबल किताब book table	टेबल पर किताब book on table
5	राम निवास के पास खड़ा है Ram is standing near house	राम निवास खड़ा Ram stand house	राम निवास पास खड़ा Ram stand near house	राम घर खड़ा Ram stand home	राम घर पास खड़ा Ram stand near home
6	जानवर ने श्याम को मारा animal killed Shyam	जानवर श्याम मारा animal kill Shyam	जानवर ने श्याम को मारा animal kill Shyam	जानवर श्याम मारा animal kill Shyam	जानवर ने श्याम को मारा animal kill Shyam
7	औरत ने चाकू से फल काटा lady cut fruit with knife	औरत चाकू फल काटा lady cut fruit knife	औरत ने चाकू से फल काटा lady cut fruit knife	महिला चाकू फल काटा lady cut fruit knife	महिला ने चाकू से फल काटा lady cut fruit knife
8	डिब्बे के अंदर पुस्तक book inside box	डिब्ब पुस्तक book box	डिब्ब अंदर पुस्तक book inside box	डिब्ब किताब book box	डिब्ब अंदर किताब book inside box
9	रूम में टी. वी. रखा है T.V. inside room	रूम टी. वी. T.V. room	रूम में टी. वी. T.V. inside room	कमरा टी. वी. T.V. room	कमरा में टी. वी. T.V. inside room
10	राम से पहले श्याम घर जाएगा Shyam will go home before Ram	राम श्याम घर जाए Shyam go home Ram	राम पहल श्याम घर जाए Shyam go home before Ram	राम श्याम घर जाए Shyam go home Ram	राम पहल श्याम घर जाए Shyam go home before Ram

TABLE IV. ESTIMATION OF F-MEASURE

Query #	Keyword based -Relation Inclusive	Keyword based +Relation Inclusive	Concept based -Relation Inclusive	Concept based +Relation Inclusive
1	0.50	0.71	0.67	0.75
2	0.40	0.67	0.59	0.77
3	0.55	0.67	0.63	0.77
4	0.57	0.80	0.57	0.80
5	0.55	0.62	0.57	0.73
6	0.75	0.86	0.75	0.86
7	0.43	0.62	0.57	0.67
8	0.40	0.67	0.60	0.75
9	0.46	0.67	0.62	0.80
10	0.67	0.77	0.67	0.77

Figure 8 gives the graphical representation of the table IV. A graph is plotted between F-measure and the search items to give a comparison of the four search environments. This experiment indicates that system performance is increased in terms of F-measure using relation inclusive searching both in keyword based searching and concept based searching. The performance gains came from query classes which had relations between query terms that could be either case role, spatial or temporal. The experiments showed the benefit of relation based searching by improving precision and recall values. The relation inclusive searching method will give better search performance among documents in English and other similar languages where case roles are represented by spatial/temporal prepositions.

VI. CONCLUSION

In this paper, we have introduced new relation based technique which performs searching of documents using relational, spatial and temporal relations of query terms and gives results better than previously used keyword based and concept based approaches. New indexing technique is used in our method which stores information about case role, spatial and temporal relation along with term position. We have compared four types of searching: Keyword Based search without Relation Inclusive, Keyword Based search with Relation Inclusive, Concept Based search without Relation Inclusive and Concept Based search with Relation Inclusive. For experiments, we have used Hindi test collection extracted from gyannidhi corpus from CDAC. Our proposed searching method gave significant improvement in terms of F-measure. The performance gains came from query classes which had relations between query terms that could be either case role, spatial or temporal. The above relation inclusive searching method will give better search performance among documents in English and other similar languages where case roles are represented by spatial/temporal prepositions.

References

[1] A. Bharati, R. Sangal, "Parsing free word order languages in the Paninian framework," in Proceedings of the 31st annual meeting on Association for Computational Linguistics, Association for Computational Linguistics Stroudsburg, PA, USA, doi>10.3115/981574.981589, pp: 105-111, 1993.

[2] A. Bharati, R. Sangal, "A karaka based approach to parsing of Indian languages," in Proceedings of the 13th conference on Computational linguistics - Volume 3, Association for Computational Linguistics Stroudsburg, PA, USA, ISBN:952-90-2028-7 doi>10.3115/991146.991151, pp: 25-29, 1990.

[3] G. Cardona,"Panini: His Work and Its Tradition (Vol. 1: Background and Introduction)," Motilal Banarsidas, Delhi, 1988.

[4] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze. Introduction to Information Retrieval. Cambridge University Press, 2008.

[5] G. Salton and M. McGill.b Introduction to Modern Information Retrieval. McGraw-Hill, 1983.

[6] Sujoy das, Anurag Seetha, M. Kumar and J.L Rana. "Post Translation Query Expansion using Hindi Word-Net for English-Hindi CLIR System," FIRE 2010, Gandhinagar, Gujrat, India, February 2010.

[7] Yonggang Qiu and H.P. Frei, "Concept Based Query Expansion," in Proceedings of ACM-SIGIR'93-6/93 /Pittsburgkr, PA, USA, pp: 160-169, 1993.

[8] Zhiguo Gong, Chan Wa Cheang, and Leong Hou U, "Web Query Expansion by WordNet," In Proceedings of the DEXA 2005, LNCS 3588 166 – 175.

[9] http://chitrapurmath.net/sanskrit/supp004.pdf

[10] http://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers_in_India

[11] www.google.co.in

[12] www.raftaar.in

[13] www.hinkhoj.com

[14] "Summary by language size", http://www.ethnologue.com/statistics/size

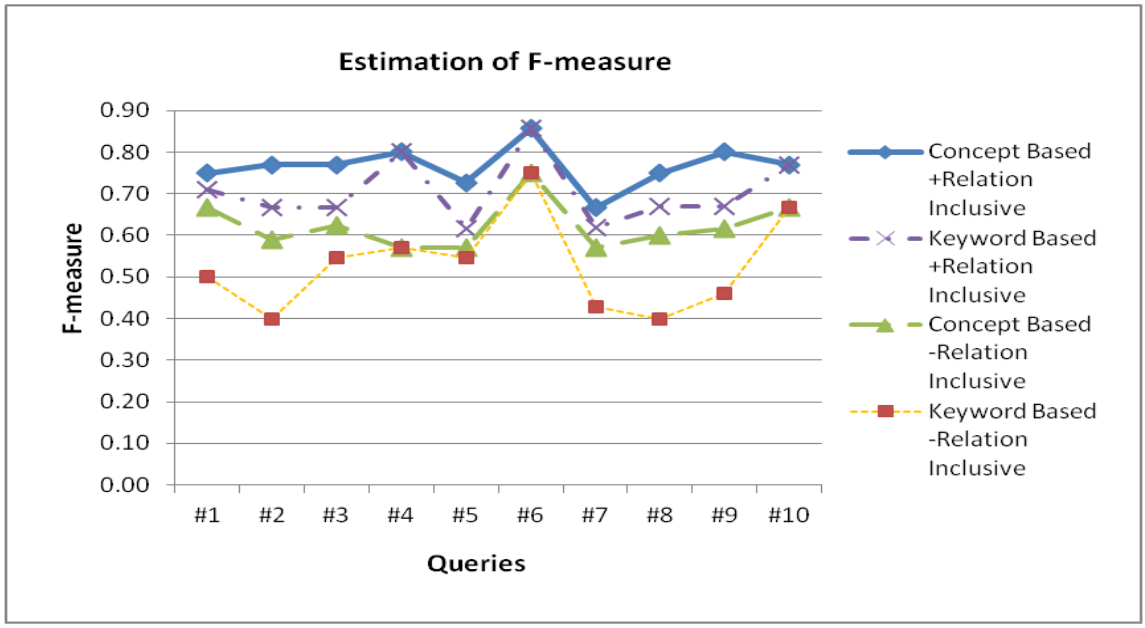


Fig.8. Comparison of F-measure for four types of searching

Applying Genetic Algorithms to Test JUH DBs Exceptions

Mohammad Alshraideh
The University of Jordan, King Abdullah
II School for Information Technology,
Department of Computer Science, Amman 11942 Jordan,

EzdeharJawabreh
Palestine Polytechnic University,
Applied Science Department

Heba M. AL Harahsheh

Basel A. Mahafzah
The University of Jordan, King Abdullah
II School for Information Technology,
Department of Computer Science, Amman 11942 Jordan,

***Abstract*— Database represents an essential part of software applications. Many organizations use database as a repository for large amount of current and historical information. With this context testing database applications is a key issue that deserves attention. SQL Exception handling mechanism can increase the reliability of the system and improve the robustness of the software. But the exception handling code that is used to respond to exceptional conditions tends to be the source of the systems failure. It is difficult to test the exception handling by traditional methods. This paper presents a new technique that combines mutation testing and global optimization based search algorithm to test exceptions code in Jordan University Hospital (JUH) database application. Thus, using mutation testing to speed the raising of exception and global optimization technique in order to automatically generate test cases, we used fitness function depends on range of data related to each query. We try to achieve the coverage of three types of PL/SQL exceptions, which are No_Data_Found (NDF), Too_Many_Rows (TMR) and Others exceptions. The results show that TMR exception is not always covered this due to existence of primary key in the query, also uncovered status appear in nested exceptions.**

Keywords—Database Application; Exception Handling;Mutation Testing; Genetic Algorithms; Select Statement.

I. INTRODUCTION

Dynamic test data generation methods use information from the execution of the program under test. A simple example of a dynamic method is random test data generation [21]. In this method, candidate test data is generated randomly by sampling values from the input domain. Each candidate test case is then executed and only those that cover a required program element are retained.

The problem with this approach is clear with complex programs or complex adequacy criteria, where an adequate test input may have to satisfy very specific requirements. In such a case, the number of adequate inputs may be very small compared to the total of inputs, so probability of selecting an adequate input by chance can be low. As an example, consider the problem of generating an input to execute the target branch A of the Flag program shown in Fig.1. The target branch A is executed when a = 1. The problem is to find input values, x and y such that a is set to 1. Note that f may be a complex and poorly understood function of x and y. Depending on the size of the domains of x and y and on the behaviour of f, it is possible that there is only a very small probability that a randomly generated input will set the variable a to 1 and thus execute the target branch A in Fig.1 In general, random test data generation is generally considered to be ineffective at covering all branches in realistic programs [22].

```
void Flag(int x, int y)
{
        a = f(x, y);
        if (a == 1)
                flag = true;  // target A
                        }
        ...
        if (flag) {
                // target B
                }
}
```

Fig.1. An example of a program (called Flag) with a "flag variable" problem.

Search-based software testing is a dynamic method of test data generation in which search methods or optimization techniques are used to generate tests and have been successfully applied in structural testing [[24] [25[23] [26][27][28][29]]. As with all search methods, search based software testing relies crucially on an evaluation or cost function to compare candidate test cases.

Database application is an important class of software that requires intensive testing. Usually database application is defined as a program that communicates with data stored in the database. Typically, this communication is done by using Structured Query Language (SQL). SQL includes Data Definition Language (DDL) that creates database schema or integrity constraints and Data Manipulation Language (DML) that retrieves or modifies records stored in the database, such

as SELECT, INSERT, DELETE, MERGE and UPDATE statements.

An important part of any database application that requires testing is the exceptions handling part. This part of code is responsible for recovering system from unusual events (exceptions) that may occur while communicating with the database. It also helps in designing more modular programs since it can be placed separately. Therefore, it is important to test this code effectively. The test period is the most important period in the software life and its cost is the biggest. In the study of Sinha and Harrold [19], they observed that 23.3% and 24.5% of the classes contained try and throw statements. If the efficiency of testing exception handling is improved, then it will improve the efficiency of the testing of the whole system.

Although the purpose of the exception handling code is to improve the robustness of the software, people noticed that the exception handling code contains more errors than the other parts of the software. For example, in a case-study by Toy [2], more than 50% of the operational failures of a telephone switching system were due to faults in exception handling and recovery algorithms. Another example, the Ariane 5 launch vehicle was lost due to an un-handling exception destroying $400 million of scientific payload [1].

Generally, exceptions' handling testing require methods that are different from the usual ones. For example, to test exceptions handling code in conventional program, a research by Tracey et al. [1], used a technique that is based on global optimization algorithm to automatically generate test cases for the purpose of testing the handling of runtime exceptions in safety critical systems. Other method by Zhang [3], presented a mutation technique to accelerate the raising of exceptions in order to save time spent for exception to occur.

To illustrate the problem, consider the problem of generating test data to execute the target no_data_found exception in Fig.2. The PL/SQL block in Fig. 2 selects information about the salary and commission of employee whose number is 102055. If no information is found it raises the No_Data_Found exception and the two variables assigned to zero. Note that No_Data_Found is a system defined exception.

```
Begin
        select salary, commission into v_salary,
v_commssion
        from employee where  employee_number = 102055;
Exception
                whenNo_Data_Found then
                v_salary = 0;
v_commission =0;
End;
```

Fig.2. An example of a program exception.

According to our knowledge, no research till now has focused on testing exceptions handling in database application. For this purpose, this research presents an approach that covers exceptions code in PL/SQL Oracle 10g database for JUH application. Our approach takes benefits from ideas used in conventional exceptions handling testing. Also, our research work rephrases the idea of mutation testing used by Zhang [3] and global optimization searching algorithm used by Tracey [1], and integrates them in testing exceptions that results from querying the database; i.e. from SELECT statements. In our research, three types of exceptions where studied: No_Data_Found (NDF), Too_Many_Rows (TMR) and others.

The rest of the paper is organized as follows: Section 2 discusses the background and related work. Then, in Section 3, the implementation of our system is described. Section 4 shows the conducted experiments. The obtained results are discussed and analyzed in Sections 5. Finally, Section 6 draws the conclusions and future research.

II. Background And Related Work

A common definition of exception is the union of "error," "exceptional case," "rare situation," and "unusual event" [17]. The entity that is raising an exception stops and waits for the completion of the exception processing. Exceptions are usually divided into two types: predefined and user-defined exceptions [18]. The predefined exceptions are declared implicitly and are also raised implicitly when the language rules are violated at run-time and in response to hardware errors. The user-defined exceptions are defined and detected at the application level; they can be raised explicitly in the application via the raise statements. Exception handling is the immediate response and consequent action taken to handle the exceptions. An exception handler is the code attached to an entity for one or several exceptions and is executed when any of these exceptions occur within the entity.

Test-data alone cannot test the raising of exceptions in response to hardware errors. In this paper, we focus on testing exception that violates run-time SQL rules (predefined). The input domain of most of programs, D, is likely to be very large, but the input domain which can causes an exception is likely to be small. It is very difficult to find the test-data that can raise an exception in large input domain. In this research, mutation is used only to give more paths to generate test cases.

A. Mutation Testing Overview

Mutation testing is a white-box fault-based testing technique originally proposed by DeMillo et al. [4]. The primary goal of mutation testing is to assist in developing adequate test suite, or it can be used to determine the effectiveness of a given test case by measuring its ability to detect faults. It operates by generating many versions (mutants) of the original program each has a fault that is injected by changing a syntactic operator (mutation operator) in the main program. Given the set of test cases to determine their effectiveness; they are initially executed against the original program to get the expected output [5], then they are executed against the mutants in the hope that they will give an output that is different from the original program's output. If this happen the mutant is said to be killed and the test case is an effective one, otherwise the mutant is a live and we need to generate more test cases to kill it. Some mutants will always give the same output as original program's output, which are called equivalent mutants.

One measure for this mutation process is mutation score which is the proportion of killed mutant over all mutants except equivalent ones. Mutation score supplies the tester with feedback about how the testing was completed and the adequacy of the test cases [5].

B. Mutation Testing for Database Applications

Many researches in testing database used mutation testing as a method to evaluate the effectiveness of the automatic test case generation techniques. In the literature Tsai et al. [6], introduced mutants by changing a set of mathematical operators in the code to determine the fault detection ability of the test cases. In testing database transitions with AGENDA, Chays and Deng injected faults in queries as a final step to demonstrate the adequacy of their proposed approach [7].

The most related works to this research are the researches which concerned with generating mutation operators specifically related for SQL queries. Chan et al. [8] proposed a mutation technique based on the conceptual data model. Their proposed technique employed the constraints that are in the Enhanced Entity Relation (EER) diagram to get SQL semantic based mutation operators. Tuya et al. [9] proposed a large set of SQL mutation operators designed for SELECT statement with the purpose of determining the adequacy of test suite and as a mean for injecting faults in order to compare different database testing techniques. The operators covered wide range of SQL particularities. Originally their mutants have four types: mutants for main SQL clauses (SELECT, JOIN, and subquery predicates), mutants for operators in the conditions or expressions, mutants related for NULL values and finally mutants for replacement of identifiers, as shown in Table 1. These mutants are further classified into types and subtypes, for more details see [9] [10]. Tuya and his colleagues tested the proposed mutants against a set of queries drawn from the NIST SQL conformance. In addition they tried to improve the feasibility of their mutants by running different experiments that aim to reduce the number of mutants (selective mutation) or to reduce the number of test cases (by reordering mutants). In this research, Tuya's mutants were studied and reused to test exceptions that raised from SELECT statements.

TABLE I. MUTATION OPERATORS USED IN THE EXPERIMENTS.

Category	Types
SC: SQL Clause Mutation Operators	SEL: SELECT clause.
	JOI: JOIN clause.
	SUB: Subquery predicates
	GRU: Group by clause.
	AGR: Aggregate functions
	UNI: Union, Union All.
	ORD: Order by clause.
OR: Operator Replacement Mutation Operators	ROR: Relational Operator Replacement
	LCR: Logical Connector Operator
	UOI: Unary Operator Insertion
	ABS: Absolute Value Insertion
	AOR: Arithmetic Operator Replacement
	BTW: Between predicate
	LKE : Like predicate
NL – NULL Mutation Operators	NLF: Null check predicates
	NLS: Null in select list
	NLI/NLO: Nulls in the input data
IR: Identifier Replacement Mutation Operators	IRC: Column replacement
	IRT: Constant replacement
	IRP: Parameter replacement
	IRD: Hidden column replacement

C. Genetic Algorithm

Genetic Algorithm (GA) was developed initially by Holland in the 1960s and 1970s [11]. It is a heuristics global optimization technique that attempts to find a good approximation to the optimal solution in a search problem. GA simulates the evolutionary process through the implementation of selection, recombination and mutation processes.

In literature GAs were used to automate the process of test data generation by searching the domains of the applications for suitable values that meet some testing criteria. Tracey et al. [1] used a GA to automatically generate test cases for testing exception handling code in safety critical systems. The technique is based on a global optimization technique; a GA that searches for the closest test data that will cause a run time violation to occur. Khor and Grogono [12] introduced genet, which is an Automated Test Data Generator (ATG) to generate test data for branch coverage. Their technique did not use program graphs because it was programming language independent and it used a GA to search for test data in the variables domains.

Masud et al. [13] introduced a strategy for mutation testing using GA, in which they instrument mutants, divide the program into small unit and then try to kill each mutant unit using genetic algorithm with special fitness function. Domínguez-Jiménez J. et al. [14] proposed a framework for mutant genetic generation for WS-BPEL (an XML language for web services). The approach defines a set of mutation operators for WS-BPEL. It automatically generates mutants using a GA that reduces the computation cost of executing by selecting a set of possible mutants. Bottaci [20] proposed fitness function which is defined in a way that a test case is able to kill a mutant if it satisfies the same three conditions used by Offutt in CBT [5], namely, the reachability, the necessary and the sufficiency conditions.

D. Genetic algorithm fitness function

The fitness functions in GA used for testing coverage three types of exceptions are as follows: Maximum and minimum values used as fitness to test coverage of NDF exception, if the value that generated by GA for column in the query or mutant under testing greater than the maximum value of this column in the database or less than the minimum value then NDF exception converge in this query. So, the fitness value is equal the min [distance between test case value and maximum column value, distance between test case value and minimum column value] +1.

Fitness for TMR is COUNT for the value that generated to column for query, if the value that generated appears two times or greater and that depending of structure of query then TMR exception coverage. So, the fitness is equal the distance between test case value and closed column value has count greater than 1. Table 2 shows an example of Testtable to illustrate how to calculate the fitness function to cover exception types. Three columns in Testtable (C1, C2 and C3) with number data types, these columns values have different values.

TABLE II. TESTTABLE VALUE FOR THREE COLUMNS C1, C2, AND C3.

C1	C2	C3
1	4	90
-20	-890	54
1	56	843
2601	786	-30
-5200	2000	1800
-20	199	-234
400	213	-90

```
Begin
…
        Select C1 into localvarible from Testtable
        Where
                C1 =Testcase;
…
Exception
        When no_data_found then
                //Target1 Executed
        When too_many_rows then
                //Target2 Executed
        When Others
                //Target3 Executed
End;
```

Fig.3. Example of exception types.

Fig. 3 illustrated a query with 3 exceptions. To cover Target1 in Fig. 3, this means no_data_found exception need to be raised, if testcase = 700 for example, then the fitness value = min((2601 -700), (700 – (-5200))) +1, = min(1901, 5900) +1= 1902, So the cost value is 1902 to execute Target1. If we need to execute Target2 in Fig. 3 with the same Testcase (700), then the fitness value will be calculated as follows: Fitness value = the difference between 700 and the closed column value which has frequency more than 1. Which is equal to abs(700 -1) =699. Table 3 shows test cases examples and NDF and TMR fitness values.

TABLE III. EXAMPLES ILLUSTRATED FITNESS FUNCTION FOR THE LISTED VALUES.

Test case value	NDF Fitness value	TMR Fitness value
700	1902	699
100	2501	99
1	2601	0
3000	0	2999
-2000	3001	2980

E. Jordan University Hospital Computer System

The core of Jordan University Hospital (JUH) information system is bought in 1994, and then the JUH IT team developed the Hospital Information System (HIS) using Oracle forms, and upgrades it to Oracle 10g. HIS developed to provide best medical services for patients and physicians. Delivering these services require hospitals to review the way they manage their business processes and supply more efficient features to physicians, patients, and hospitals officials as well as other decision makers. In order to provide such services, the health facility must focus on developing a solution to connect all its resources and makes it available to all who needs utilizing it using latest technology. This kind of solution will enhance the performance and optimize the efficiency and will reduce the cost of ownership.

IT department in JUH creates a solution suite that transforms the hospital to a community allowing the access to all resources and data as needed. HIS is a comprehensive solution developed specifically for health facilities in the region. It is flexible, comprehensive, multilingual, integrated and secured solution that supports clinical, financial, administration and higher management needs.

In general, a hospital management system can be sub-categorized into the following groups (Fig. 4):

- Medical Information System (Administrative and Clinical).
- Enterprise Resource Planning (ERP) (Material, Financial and Human Resources).
- Support System.

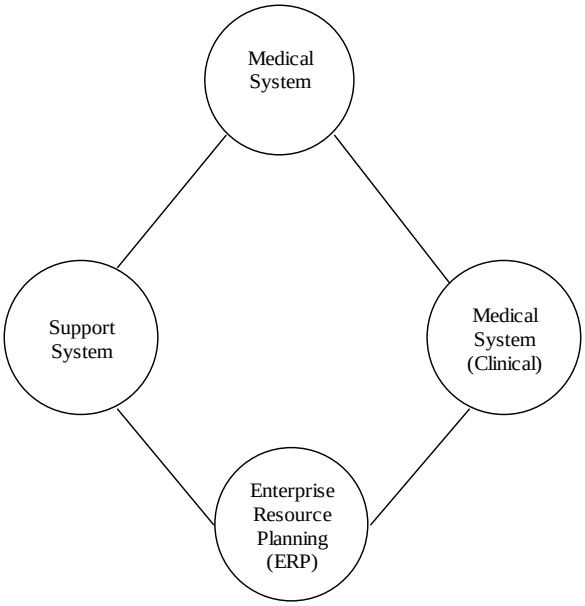


Fig.4. Hospital management system sub-categorizes.

Medical systems are developed to deliver all needed services to the hospital community (physicians, patients and administration). The systems manage all patients' data and information during their treatment episode in a professional and efficient manner. Medical systems strategically support a full range of hospital functions. It contains a repository of all patients' clinical, billing and demographic data, reducing paper work, manual effort and errors. Furthermore; it allows for better staff utilization allowing for more time to focus on planning and goals achievements. This enables the hospital to provide better quality and more efficient services, needed by patients and physicians. Medical systems are integrated with financial, administration, human resources, and material management systems. It contains vast collection of data including patient data, treatment data, hospital visit data, patient transactions data, hospital data, and statistical information.

HIS medical systems provide many key functions including:

- Medical administrative including:
  - Patient master index
  - Admission, discharge and transfer
  - Scheduling and appointments
  - Medical records
  - Medical reports
  - Medical statistics
  - Catering
  - Order entry and results communication
- Medical clinical including:
  - Out-patient clinics
  - Accidents and emergency
  - Operation theater
  - Maternity
  - Doctors desktop
  - Nurse station
  - Laboratory
  - Radiology
  - Pharmacy
- Patient accounting including:
  - Pricing and package deals
  - Patient billing
  - Insurance contract management
  - Claims management.

In order to test our system in this research we will select different procedures and functions, which will be described later in this paper.

III. Proposed System Overview

Fig. 5 depicts the basic structure of our system. Following is a brief description of the system's components:

- **SQL Mutation Tool:** which is a web service tool designed by Tuya et al. [9]. This tool accepts as inputs the SQL query and DB schema file. It will generate a set of mutants for the given query according to predefined SQL mutation operators as suggested by Tuya, as seen in Table 1. The DB schema and queries in our system were related to Human Resources (HR) sample database in Oracle 10g.

- **Mutant's Pool:** the required numbers of queries that satisfy a set of conditions are saved in the mutant pool waiting to be processed by the next stage.

Extracting Parameters: the SQL mutation tool supports the existence of parameters in the SQL query. For example, in the following query,there are two parameters *x* and *y* of type integer, one for location_id and the other for department _id. The pair of question marks symbol denotes the existence of parameter as it supported in the SQL Mutation Tool. The letter just before the last question mark indicates the type of parameter, where *I*stands for integers, *d* for decimal, *c* for characters and strings and *u* for date data type. This stage will extract parameters information and packed it in a record called TC structure, in the following query it will be (integr, integr).

SELECT manager_id from departments where location_id=?xi? ordepartment_id=?yi?

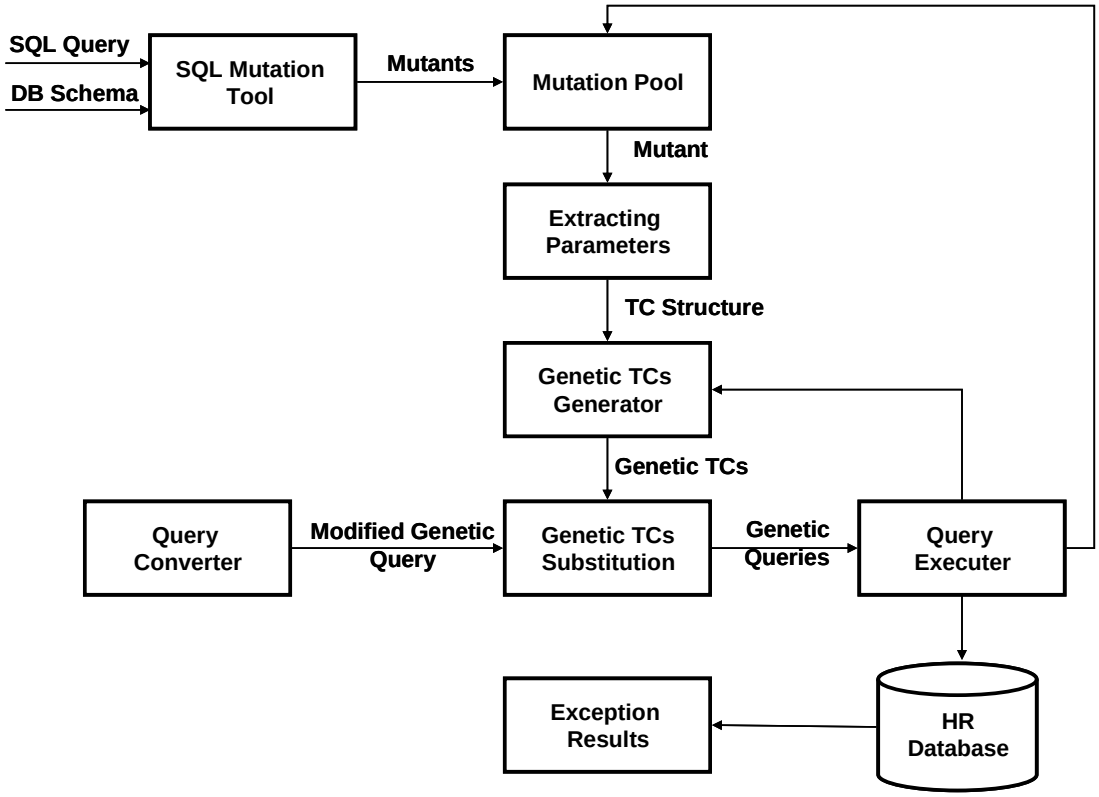


Fig.5. Basic structure of our system.

- **Genetic TCs Generator:** for each mutant TC structure, a GA will generate random population of the appropriate type within some specified range. The initial population consists of ten individuals (test cases). The population in the successive generations is resulted from crossover and mutation processes. The GA used here has no fitness function, so all the initial random ten individuals are used to participate as parents. To generate the offspring, a one point cross-over operation is performed between each pair of parents. The resulted offspring is then randomly mutated and used as the new population for the next generation. After performing many experiments to find the suitable TC, we found that running the GA for fifty generations will achieve the desired coverage of the intended exceptions. This TC called genetic test case.
- **Query Converter:** the mutants generated by SQL mutation tool are just SELECT statements and in order to raise any of NDF or TMR exceptions, a modifier is used in this system to convert SELECT statement into SELECT INTO statements. This is done by performing the necessary steps of defining the required number of variables according to the SELECT list using the schema file to extract variables types.
- **Genetic TCs Substitution:** for the mutant who is currently processed and modified, the parameters in the mutant will be substituted by the corresponding ones in each of the genetic TCs. The resulted query for a given mutant is called modified genetic query.
- **Query Executer:** the executer uses the modified genetic queries as input. Its main operation is to replace them one by one in PL/SQL block that contains exceptions handling code for the three exceptions: NDF, TMR and Others as shown in Fig. 6.

```
Declare
        Variable Definition List from preprocessor
Begin
        The modified genetic query
Exception
        whenno_data_found then
```

```
                …
:NDFFlag:=1;
        whentoo_many_rows then
         …
        :TMRFlag:=1;
        when others then
         …
        :OthersFlag:=1;
End;
```

Fig.6. Exception flag setting

The overall process of our system can be summarized in Fig. 7. The algorithm in Fig. 7 will iterate until the population has evolved to form a solution to the problem (i.e., solutions that achieved coverage of three intended exceptions), or until a maximum number of iterations have taken place (suggesting that a solution is not going to be found given the resources available).

```
For each SQL Query generate all possible mutations and save the required ones
in pool P.
For each mutant m in P and stop criteria not met do {
        Extract TC structure from m.
        Convert SQL SELECT in m to SELECT … INTO … command.
Simple Genetic Algorithm ( )
  {
        Initialize TC population;
        Evaluate population;
        While termination criterion not reached
        {
                Select solutions for next population;
                For each solution s in the selected population
                {
                        Substitute s in mutant m;
                        Execute mutant m and watch the result if any
        exception appear.
                        Exit if the flags of three exceptions are set.
                }
        Find next population by performing crossover and mutation;
        Evaluate population;
        }
  }
}
```

Fig.7. Algorithm used to find test case to coverage SQL Exceptions

IV. EXPERIMENTAL ENVIRONMENT

A. Test Objects

This section describes the test objects and the input domain sizes used. The following are source code for the test objects:

- ***OutPricing:*** Determines the pricing of treatments at outpatient clinics. Depending on his insurance the patient, this function calculates the amount of money the patient has to pay (depending on the type of insurance, the patient pays different ratios for his treatment) and the amount of money the insurance has to pay for the patient's treatment as shown in Fig. 8.
- ***InPricing:***This procedure calculates the invoice value of the patient inside the hospital based on the type of patient insurance and the type of medical procedure offered to patients (accommodation, scouting, doctors' fees, operations, laboratory, radiology, medicine, etc.). Also, this program calculates the percentage paid by the patient and the percentage paid by the insurance company, if any. Moreover, this function bills the patient with the amount of money he has to pay and bills the insurance company with the amount of the money it has to pay.
- ***JU-Med-fees-deduction:*** This package used for Jordan University staff, where there is an allocated account number for each staff in the system of JUH. It calculates bill value based on Jordan University insurance. Then deported the total amount of bill after deduct the hand-collect from the patients into tables to be used in Jordan University financial department later on.
- ***at-info-ibr:*** This function calculates the invoice value for private patients (in patients and out patients), then bills the patients with the amount of money he or she has to pay.
- ***Lab-interface:*** The main goal of this function is to transfer the results from medical machines (lab devices) to HIS system automatically (without user interaction). So, the function receives the message from medical devices then converts it to be entered to HIS system.
- ***salup_new_calc_all:*** This procedure calculates staff incentives as follows: it selects the category that owns the nursing, administrative, officer or a medical technician, by the department and qualifications. Then it determines the share of the incentives that the employee is entitled, as his career (Branch Chief, Chief, Division of, etc.). It discounts days leave without pay from the employee share incentives as shown in Fig. 9.

```
[1]  select prc_division,prc_category into p_div,p_cat from store_pricing_groups where
[2]  div_id=p_divn  and  grp_id=p_grp
[3]  …
[4]  select decode(p_insur_type, '1', prc_limit_out_e, '2', prc_limit_out_f)
[5]      into p_max_cov
[6]      from prc_limits
[7]      where prc_group    = p_group_id
[8]        and prc_division = p_div;
[9]  p_max_cov := nvl(p_max_cov, 99999);
[10]  Exception
[11]    when no_data_found then
[12] p_error_no := 1; raise exit_proc;
[13]    when others then
[14] p_error_no := 11; raise exit_proc;
[15] ….
[16] when exit_proc then
[17]    if p_error_no = 1 then
[18] raise_application_error( -20001, ‘Coverage limitsdo not exist.');
[19] elsifp_error_no = 2 then
[20] raise_application_error( -20003, ‘Ratepricingdoes not existforthis materials!!');
[21] elsifp_error_no = 3 then
[22] raise_application_error( -20004, 'Materialis not definedin the tableprice!!');
[23] elsifp_error_no = 5 then
[24] raise_application_error( -20006, ‘Pricing datais incompleteforthis patient!!');
[25] elsifp_error_no = 11 then
[26] raise_application_error( -20001, '11');
[27] elsifp_error_no = 21 then
[28] raise_application_error( -20003, '21');
[29] elsifp_error_no = 31 then
[30] raise_application_error( -20004,  '31');
[31]    end if;
[32]  when others then
[33] raise_application_error( -20005,   sqlerrm);
```

Fig.8. An example of a program (exception) **OutPricing**.

```
[1]  …
[2]  SELECT chng_date, old_emp_admin, old_emp_department, old_emp_job
[3]  into v_chng_date, v_old_emp_admin, v_old_emp_department, v_old_emp_job
[4]  FROM EMP_JOB_CHANGES e
[5]   WHERE emp_id = salup_rec.emp_id
[6]     and CHNG_DATE BETWEEN TO_DATE('02-'||TO_CHAR(P_MONTH-5, '00')||'-'||
[7]  TO_CHAR(P_YEAR, '0000'), 'DD-MM-YYYY')
```

```
[8]          and salup_rec.emp_join_date<chng_date
[9]      Exception
[10]     When no_data_found then
[11]         Begin
[12]           select days into v_leave_days
[13]           from salup_new_leaves_200806
[14]         where emp_id = salup_rec.emp_id
[15]                and month = p_month
[16]                and year = p_year;
[17]                and salup_rec.emp_status<> 2;
[18]          exception
[19]              when no_data_found then
[20]     v_leave_days:= 0;
[21]          end;
[22]     When two_many_rows then
[23]     …
[24]     End;
```

Fig.9. Nested exceptions example (fragment of JU-med-fees-deduction program).

B. Hardware and Software Environments

In this section, the specifications of the experimental environment utilized by this work are presented. These specifications include both the hardware and software modules used in implementing the simulator. More specifically, the hardware specifications that are used in the experiments include a Dual-Core Intel Processor (CPU 2.66 GHz), 2 MB L2 Cache per CPU, and 1 GB RAM. Moreover, the software specifications that are used in the experiments include windows XP.

Moreover, in order to assess the reliability of the cost functions introduced in the previous section, an empirical investigation was done. A number of test procedures and functions were assembled from JUH and an attempt was made to generate inputs to achieve branch coverage. These programs are described in Table 4. The size of each program is given as Lines of Code (LOC), number of select statements, and number of exceptions (three types of exceptions), and the last column represents the total number of mutations for each program generated by the tool in [10]. The range of number of mutations for each Query is from 24 to 79 mutations.

The search was directed to generate data for one exception at a time. The order in which the exceptions of the program were targeted was arbitrary, except that no nested exception was targeted before the containing exception as shown in Fig. 9. This is not; in general, a good strategy since it will become stuck at an infeasible exception.

TABLE IV. THE FUNCTIONS USED FOR EMPIRICAL INVESTIGATION.

Program Name	Lines of Code	Number of Select Statements	Number of Exceptions	Total number of mutations
OutPricing	295	14	6	12138
InPricing	362	22	33	11222
JU-Med-fees-deduction	307	12	18	7368
Pat-info-ibr	259	9	9	11655
Lab-interface	1389	45	69	38892
Salup_new_calc_all	707	17	24	25453

A steady-state style genetic algorithm, similar to Genitor [24], was used in this work. The cost function values computed for each candidate input were used to rank candidates within the population in which no duplicate genotypes are allowed.

A probabilistic selection function selected parent candidates from the population with a probability based on their rank, where the highest ranking having the highest probability. More specifically, for a population of size n, the probability of selection (P_s) is shown in Equation 1.

$$P_s = \frac{2(n - rank + 1)}{n(n-1)} \qquad (1)$$

In this work, a fixed population size of 100 was used. This parameter was not "tuned" to suit any particular program under test. In a steady state update style of genetic algorithms (as used in this work); new individuals that are sufficiently fit are inserted in the population as soon as they are created.

The criterion to stop the search was set up to terminate the search after 50 executions of the program under test, when only if full coverage was not achieved. Individuals were recombined using binary and real-valued (one-point and uniform) recombination, and mutated using real-valued mutation. Real-valued mutation was performed using "Gaussian distribution" and "number creep". These queries (Select statements in program described in Table 4) contain 109 different SELECT statements related to JUH database with 159 exceptions as a whole. The list of the queries includes conditions of types: WHERE, HAVING, and ON. In the case of WHERE condition, different SQL clauses were implemented as: [not] BETWEEN, [not] IN, [not] LIKE, IS [not] NULL, Logical connector AND and OR and the using of expressions with relational operators (=, >, <, <=, >=, <>).

Each query along with the schema file was executed. The generated mutants for each one are saved, and then they are processed; each one in separate; by passing through the stages explained in our system in Section 3. In our experiments we executed all the resulted genetic queries for each mutant in order to assess the area of exceptions coverage as described in the algorithm in Section 3.

Table 5 shows all possible mutants for the first query 1 (line 1) in Fig. 8, where id: the identification numbers of each mutant. The same ids that are generated from sql mutation tool [10] are used. Mutant subtype: refers to type of mutants when it is applied to particular sql clause (for more information see table 1).

TABLE V. ALL MUTATIONS USED FOR (SELECT PRC_DIVISION, PRC_CATEGORY INTO P_DIV, P_CAT FROM STORE_PRICING_GROUPS WHERE DIV_ID=P_DIVN AND GRP_ID=P_GRP) QUERY.

ID	Cat	Type	Subtype	Mutated SQL
1	SC	SEL	SLCT	SELECT DISTINCT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
2	NL	NLS	NLSS	SELECT COALESCE(prc_division ,'9999') AS prc_division , prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
3	NL	NLS	NLSS	SELECT prc_division , COALESCE(prc_category ,'9999') AS prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
4	IR	IRC	IRCCS	SELECT STORE_PRICING_GROUPS.PRC_CATEGORY ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
5	IR	IRC	IRCCS	SELECT STORE_PRICING_GROUPS.DIV_ID ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
6	IR	IRC	IRCCS	SELECT STORE_PRICING_GROUPS.GRP_ID ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
7	IR	IRC	IRCPS	SELECT ?xc? , prc_category FROM store_pricing_groups WHERE div_id= ?xc? AND grp_id= ?xc?
8	IR	IRD	IRDDS	SELECT STORE_PRICING_GROUPS.INV_ITEM ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
9	IR	IRC	IRCCS	SELECT prc_division , STORE_PRICING_GROUPS.PRC_DIVISION FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
10	IR	IRC	IRCCS	SELECT prc_division , STORE_PRICING_GROUPS.DIV_ID FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
11	IR	IRC	IRCCS	SELECT prc_division , STORE_PRICING_GROUPS.GRP_ID FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
12	IR	IRC	IRCPS	SELECT prc_division , ?xc? FROM store_pricing_groups WHERE div_id= ?xc? AND grp_id= ?xc?
13	IR	IRD	IRDDS	SELECT prc_division , STORE_PRICING_GROUPS.INV_ITEM FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id= ?xc?
14	NL	NLI	NLIW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE (STORE_PRICING_GROUPS.DIV_ID IS NULL OR div_id = ?xc?) AND grp_id= ?xc?
15	NL	NLO	NLIW1	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE (STORE_PRICING_GROUPS.DIV_ID IS NULL OR NOT div_id = ?xc?) AND grp_id= ?xc?
16	NL	NLO	NLIW2	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE (STORE_PRICING_GROUPS.DIV_ID IS NULL) AND grp_id = ?xc?
17	NL	NLO	NLIW3	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE (STORE_PRICING_GROUPS.DIV_ID IS NOT NULL) AND grp_id = ?xc?
18	NL	NLI	NLIW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND (STORE_PRICING_GROUPS.GRP_ID IS NULL OR grp_id= ?xc?)
19	NL	NLO	NLIW1	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND (STORE_PRICING_GROUPS.GRP_ID IS NULL OR NOT grp_id= ?xc?)
20	NL	NLO	NLIW2	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND (STORE_PRICING_GROUPS.GRP_ID IS NULL)
21	NL	NLO	NLIW3	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND (STORE_PRICING_GROUPS.GRP_ID IS NOT NULL)
22	IR	IRC	IRCCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE STORE_PRICING_GROUPS.PRC_DIVISION = ?xc? AND grp_id= ?xc?
23	IR	IRC	IRCCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE STORE_PRICING_GROUPS.PRC_CATEGORY = ?xc? AND grp_id= ?xc?
24	IR	IRC	IRCCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE STORE_PRICING_GROUPS.GRP_ID = ?xc? AND grp_id= ?xc?

25	IR	IRD	IRDDW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE STORE_PRICING_GROUPS.INV_ITEM = ?xc? AND grp_id= ?xc?
26	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id<> ?xc? AND grp_id= ?xc?
27	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id> ?xc? AND grp_id= ?xc?
28	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id< ?xc? AND grp_id= ?xc?
29	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id>= ?xc? AND grp_id= ?xc?
30	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id<= ?xc? AND grp_id= ?xc?
31	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE (1=1) AND grp_id = ?xc?
32	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE (1=0) AND grp_id = ?xc?
33	IR	IRP	IRPCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = STORE_PRICING_GROUPS.PRC_DIVISION AND grp_id = ?xc?
34	IR	IRP	IRPCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = STORE_PRICING_GROUPS.PRC_CATEGORY AND grp_id = ?xc?
35	IR	IRP	IRPCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = STORE_PRICING_GROUPS.GRP_ID AND grp_id = ?xc?
36	OR	LCR	LCRW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? OR grp_id= ?xc?
37	OR	LCR	LCRW	SELECT prc_division , prc_category FROM store_pricing_groups WHERE (1=1)
38	OR	LCR	LCRW	SELECT prc_division , prc_category FROM store_pricing_groups WHERE (1=0)
39	OR	LCR	LCRW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc?
40	OR	LCR	LCRW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE grp_id = ?xc?
41	IR	IRC	IRCCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND STORE_PRICING_GROUPS.PRC_DIVISION = ?xc?
42	IR	IRC	IRCCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND STORE_PRICING_GROUPS.PRC_CATEGORY = ?xc?
43	IR	IRC	IRCCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND STORE_PRICING_GROUPS.DIV_ID = ?xc?
44	IR	IRD	IRDDW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND STORE_PRICING_GROUPS.INV_ITEM = ?xc?
45	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id<> ?xc?
46	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id> ?xc?
47	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id< ?xc?
48	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id>= ?xc?
49	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id<= ?xc?
50	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND (1=1)
51	OR	ROR	RORW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND (1=0)
52	IR	IRP	IRPCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id = STORE_PRICING_GROUPS.PRC_DIVISION
53	IR	IRP	IRPCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id = STORE_PRICING_GROUPS.PRC_CATEGORY
54	IR	IRP	IRPCW	SELECT prc_division ,prc_category FROM store_pricing_groups WHERE div_id = ?xc? AND grp_id = STORE_PRICING_GROUPS.DIV_ID

V. RESULTS AND EVALUATION

GA search strategy was investigated empirically by generating test data for the functions and procedures shown in Table 4. What is not clear, however, is how long such a search would take. In addition, the empirical investigation will provide information about the GA search.

The results in Table 6 show the average number of executions required to find exception coverage test data. Table 6 shows the total number of exception ineach program, the number of exceptions successfully converge by GA, the total mutations used to coverage these exceptions, and the number of executions required to find test data. From this table, we notice that not all exceptions are covered by GA, (**OutPricing, JU-med-fees-deduction, and Salup_new_calc_all)**,the type of exceptions that are not cover is like line 17 in Fig. 8, which represent branch condition in the exception, and line 19 in Fig. 9 which represents nested exceptions.

Table 7 shows the types of exceptions that are not covered by GA from Table 6. From Table 7 we notice that most type of mutations that are not covered is TMR mutation type. The reason that, the uncovered of TMR (the invisibly in TMR) is due to: the existence of primary keys in the query, or columns with unique values where it is impossible to get TMR in the result.

Table 8 shows the number of program executions to find test data for exception after excluding the exceptions that are not covered in Table 6. The number of program executions ranged from 10475 used 213 mutations in OutPricing program down to 1493 used only 49 mutations in InPricingprogram

TABLE VI. EXCEPTIONS COVERAGE RESULTS.

Program Name	Total Number of Exceptions	Number of Exceptions Successfully Coverage	Total Number of Mutations Used to Coverage Exceptions	Total Number of GA Generations
OutPricing	6	5	290	14325
InPricing	33	33	49	1493
JU-Med-fees-deduction	18	14	368	18134
Pat-info-ibr	9	9	123	3492
Lab-interface	69	69	181	7946
Salup_new_calc_all	24	21	545	16914

TABLE VII. TYPES OF NOT COVERAGE EXCEPTIONS.

Program Name	Total Number of Exceptions	Number of Exceptions Successfully Coverage	Types of Not Coverage Exceptions		
			NDF	TMR	Others
OutPricing	6	5	-	1	-
InPricing	33	33	-	-	-
JU-med-fees-deduction	18	14	1	3	-
Pat-info-ibr	9	9	-	-	-
Lab-interface	69	69	-	-	-
Salup_new_calc_all	24	21	-	3	-

TABLE VIII. EXCEPTIONS COVERAGE AFTER EXCLUDING NOT COVERAGE EXCEPTION IN TABLE VI.

Program Name	Number of Exceptions	Total number of mutations used to coverage Exceptions	Total number of GA generations
OutPricing	5	213	10475
InPricing	33	49	1493
JU-med-fees-deduction	14	164	7934
Pat-info-ibr	9	123	3492
Lab-interface	69	181	7946
Salup_new_calc_all	21	159	7614

Although the discussion and conclusions of our results were related to JUH database, it is applicable to any other database, since the presence of database constraints such as primary keys and the effects of mutation operators on the conditions will be the same regardless the contents of the database.

VI. Conclusions And Future Works

In this paper, we have designed a system to generate automatically TCs that cover three known exceptions (which are **No_Data_Found (NDF), Too_Many_Rows (TMR)** and **Others** exceptions) in PL/SQL Oracle database. This system combines the mutation testing in order to speed the raising of exceptions, with a genetic algorithm that will automatically generate TCs. Experiments have been done to evaluate the system on JUH database application. The obtained results were subject to analytical studies. The studies illustrate that not all exceptions are covered.

It is concluded that this system achieved the desired coverage of the intended exceptions. The TMR was the most difficult one to cover since it has a lot of reasons that make it an invisible path, such as: the existence of primary keys or unique values, the nature of the query and existence of different categories of mutants.

It is still believed that the interpretation of the obtained results needs more improvement in complicated exceptions or branch inside exceptions. The future works are based on extending this research work to solve these problems. However, ongoing researches have been established to improve the system in different areas, such as including other types of Oracle exceptions to cover.

References

[1] N Tracey., J. Clark, K. Mander , and J. McDermid,"Automated Test-Data Generation for Exception Conditions", Software-Practice and Experience, 30(1), 2000, pp.61-79.

[2] W. N.Toy, "Fault-tolerant design of local ESS processors, The Theory and Practice of Reliable System Design", D.P. Siewiorek and R.S. Swarz, eds., pp. 461–496, Bedford, Mass.: Digital Press, 1981.

[3] S. J Jiang.,Y. P Zhang, D. S Yan, and Y. P. Jiang,." An Approach to Automatic Testing Exception Handling", ACM SIGPLAN Notices, Vol. 40(8), pp. 34- 39, 2005.

[4] A. DeMillo Richard, D. S. Guindi, K. N. King McCracken , and A. Offutt Jefferson , "An Extended Overview of the Mothra Software Testing Environment", In Proceedings of the second Workshop on Software Testing, Verification, and Analysis, Los Alamitos, pp.142-151, 1988.

[5] A. DeMillo Richard and A. Jefferson Offutt, "Constraint-Based Automatic Test Data Generation", IEEE Transactions on Software Engineering, Vol. 17 (9), pp. 900-910, 1991.

[6] W.T. Tsai, D. Volovik, T. F. Keefe ," Automated test case generation for programs specified by relational algebra queries", IEEE Transactions on Software Engineering, 16(3), pp. 316-324, 1991.

[7] Chays David and Y. Deng , "Testing Database Transactions with AGENDA", In Proceedings of the 27th International Conference on Software Engineering (ICSE), IEEE Computer Society, pp.70- 87,2005.

[8] W. K Chan., S. C Cheung., and T. H Tse, "Fault-Based Testing of Database Application Programs with Conceptual Data Model", In Proceedings of the 5th International Conference on Quality Software (QSIC 2005), IEEE Computer Society Press, Los Alamitos, California, pp. 187-196, 2005.

[9] Tuya Javier, M. J. Suarez-Cabal and de la Riva Claudio, " Mutating database queries, Information and Software Technology" , vol. 49 (4), pp. 398 -417,2007.

[10] Tuya Javier, M. J. Suarez-Cabal, and de la Riva Claudio , "Sqlmutation: A Tool To Generate Mutants Of SQL Database Queries", In Proceedings of Second Workshop on Mutation Analysis (Mutation 2006 - ISSRE Workshops 2006), IEEE Computer Society Washington, DC, USA, p. 1, 2006.

[11] J. H. Holland ,"Adaptation in Natural and Artificial Systems", University of Michigan Press.1975.

[12] Khor S. and Grogono P. (2004). Using a Genetic Algorithm and Formal Concept Analysis to Generate Branch Coverage Test Data Automatically, In the Proceedings of the 19th International Conference onAutomated Software Engineering, pp. 346-349.

[13] M. Masud, , A. M NayakZaman,., and N. Bansal, "Strategy for mutation testing using genetic algorithms", IEEE Electrical and Computer Engineering, vol. 13(2), pp. 1049 – 1052,2005.

[14] J. Dom´ınguez-Jim´enez ,A. Estero-Botaro, and I. Medina-Bulo,"A Framework for Mutant Genetic Generation for WS-BPEL", Springer-Verlag Berlin Heidelberg, Vol. 5404, pp. 229–240, 2009.

[15] Oracle. Oracle® Database Sample Schemas, http://download.oracle.com, [Online October, 2012].

[16] HYPERLINK "http://in2test.lsi.uniovi.es/sqlmutation"

i. http://in2test.lsi.uniovi.es/sqlmutation, [Online June, 2013].

[17] J. Lang and D. B. Stewart,"A study of the applicability of existing exception-handling techniques to component-based real-time software technology", ACM Transactions on Programming Languages and Systems, 20(2), pp. 274-301,1998.

[18] J. B. Goodenough,"Exception handling: issues and a proposed notation", Communications of the ACM, v. 18 (12), pp. 683-696, 1975.

[19] S. Sinha andM. J. Harrold, "Analysis and testing of programs with exception-handling constructs", IEEE Transactions on Software Engineering, 26(9), pp. 849-871,2000.

[20] L. Bottaci , "A genetic algorithm fitness function for mutation testing", In Proceedings of SEMINAL: Software Engineering using Metaheuristic Innovative Algorithms, Workshop 8, ICSE 2001, 23rd International Conference on Software Engineering, Toronto, Canada, pp. 3-7, 2001.

[21] J. E. Duran ,S. C. Ntafos, "An evaluation of random testing". IEEE Transactions on Software Engineering 10(4):438 – 444,1984.

[22] P. D. Coward, " Symbolic execution and testing". Information and Software Technique 33(1):53-64, 1991.

[23] R. Feldt, R. Torkar ,T. Gorschek, W. Afzal, " Searching for cognitively diverse tests: Towards universal test diversity metrics". In: IEEE International Conference on Software Testing Verification and Validation Workshop, ICSTW '08, Computer Society Washington, DC, USA, pp 178-186, 2008.

[24] M. Alshraideh M, and L. Bottaci , " Automatic software test data generation for string data using heuristic search with domain specific search operators". Software Testing, Verification and Reliability 16(3):175-203, 2006.

[25] M. Alshraideh , B. A. Mahafzah and S. Al-Sharaeh, "A Multiple-Population Genetic Algorithm for Branch Coverage Test Data Generation", Software Quality Control, Vol. 19, No. 3, pp. 489-513, 2011.

[26] M. Alshraideh,L.BottaciandB A.Mahafzah, "Using Program Data-State Scarcity to Guide Automatic Test Data Generation", Software Quality Control, Vol. 18, No. 1, pp. 109-144, 2011.

[27] M. Alshraideh ,B. A. Mahafzah., H. S. EyalSalman ,I. Salah , "Using Genetic Algorithm as Test Data Generator for Stored PL/SQL Program Units", Journal of Software Engineering and Applications, Vol. 6, No. 2, pp. 65-73, 2013.

[28] P. McMinn ,"Search-based Software Test Data Generation: a Survey: Research Articles. Software Testing", Verification & Reliability, Volume 14, Number 2, Pages 105-156, 2004.

[29] Korel B. (1990), Automated Software Test Data Generation. IEEE Transactions on Software Engineering, Vol. 16(8), pp. 870-879.

Visual Exploration of Complex Network Data Using Affective Brain-Computer Interface

Sergey V. Kovalchuk, Denis M. Terekhov
E-Science Research Institute,
National Research University of IT, Mechanics and Optics St. Petersburg, Russia

Aleksey A. Bezgodov, Alexander V. Boukhanovsky
E-Science Research Institute,
National Research University of IT, Mechanics and Optics St. Petersburg, Russia

***Abstract*—This paper describes the current state of the work aimed towards an affective application of BCI to the task of complex data visual exploration. The developed technological approach exploits the idea of supporting tacit and complex domain-specific knowledge acquisition during the examination of visual images built using large input data sets. The presented experimental research on the complex network data exploration process shows the capabilities of the presented approach through the analysis of a user's affective state estimation.**

Keywords—affective brain-computer interface; complex network data; visualisation; virtual reality; domain-specific knowledge, human-computer interaction.

I. Introduction

Contemporary scientific tasks deal with a huge amount of complex data. These data sets present results of observations, analysis and simulation. The approaches to the handling of large and complex data sets are important issues of e-science [1]. Usually, the main attention of an e-science project is focussed on the development of software tools for the processing of this data (simulation, data mining, etc.). Nevertheless, the final processed results are often quite complex and require the development of special user-centric data-analysis tools. The common approach to the development of such tools explores the idea of data visualisation within a 2D or 3D space. There are a lot of techniques of complex data presentation with a long history of development (see, for example, the brilliant works of Edward R. Tufte [2]). A strong push for the development of visual analysis technologies was given by the appearance of virtual reality, which enables advanced an interactive visual exploration of complex spatiotemporal datasets. However, the complexity of visualised data still makes cognitive analysis (i.e. coming to a particular conclusion(s) during analysis) a difficult task to perform.

Usually, cognitive analysis is focussed on several "objects of interest" to be analysed with greater accuracy (e.g. unusual structure or behaviour). Nevertheless, in the case of data complexity there occur several important issues with regard to the visual exploration of complex data structures. Firstly, it is often rather difficult to automatically identify the importance of particular elements of complex visual scenes. As a rule, experts in a particular problem domain usually fulfill this task with ease, but non-expert users often face this problem during visual data analysis. Secondly, even for experts it is sometimes quite difficult to focus consciously on every important object within the scene. Frequently, experts can perform analysis without explicit attention to all the aspects of the analysed scene by using some tacit knowledge and their own experience. In this case it is very difficult to identify all the aspects of the decision even by interviewing the expert. Both issues aggravate in cases of multiple "points of interest" to be discovered and need to be taken into account in order to make a correct conclusion. Moreover, considering these issues from the point of view of knowledge acquisition, both of them deal with the loss of some parts of knowledge. In the first case, there are pieces of knowledge which can be discovered by analysing visualised data. In the second case, a part of expert tacit knowledge is lost. To overcome these issues, a visualisation system should have information about the objects within the visual scene which are most important for the user within a particular task.

Nowadays, brain-computer interface (BCI) appears to be a technology which allows for the augmenting of virtual reality with brand new input from the user [3]. Often this technology is used for navigation or similar purposes. Though considering the presented issues, one of the interesting abilities of BCI is the estimation of the affective state of the user that include, for example, engagement and excitement levels, or even high-level characteristics like arousal and valence [4], or workload and vigilance [5]. This allows the use of BCI as one of the tools available within an affective computing approach [6]. In this article the idea of using the affective state of experts, estimated by brain-computer interface for support of complex scientific data exploration, is discussed. We believe that such implementation can significantly enrich the virtual reality technology by enabling tacit knowledge acquisition and using them to make the process of data analysis more effective.

II. Visual Exploration of Scientific Data

In considering scientific data visualisation, several abstraction levels of data can be defined.

1) Raw data. These data present the initial values to be visualised. In particular, analysis of this data is the main goal of visual data exploration.

2) Semantic objects. Initial data can be arranged according to the structure of the investigated system, its elements and properties. The structure can be defined using terms specific to a particular problem domain.

3) Visual objects. Each object can have several visual representations. These representations may present images of a related natural object or visualise some characteristics of the

data. The data can also influence the behavior of the visualised object.

4) Visual scene. A complete dynamic visual scene which shows all the images related to the scene, analysed objects and presented data.

While constructing a visual representation of the data it is required to pass all the levels from 1 to 4. This can be done automatically (by software) or semi-automatically (e.g. by implicit identification of semantic objects during visualisation of software development). Nevertheless, the result which users can see and explore is a complete visual scene. It is interesting that, while analysing the scene, the user passes almost the same levels but in reverse. Firstly he/she perceives the scene (level 4) using visual stimuli. Then the scene is decomposed into the particular visual objects (level 3). They are classified and analysed according to the experience of the user. The experience of the user includes knowledge of a particular problem domain which allows them to recognise the correct semantic meaning of each object (level 2). Finally, the analysis of the visual images in accordance with their semantic meaning gives the user an insight into the initial data (level 1). And here, having this insight, the user understands the meaning of the data that was initially presented. That leads him/her to a particular conclusion, or way of reasoning or decision making. The idea of the presented work is to analyse the affective state of the user and to link it with the process of data exploration which is performed within these four levels of abstraction.

There are several scenarios that benefit from the implementation of the presented approach. Here are two examples which show a possible use:

1) Knowledge acquisition. Retrieving knowledge often becomes a difficult task as the knowledge may be very complex and contain tacit parts. For example, some experts make a decision using their experience without explicit attention or consideration of some statements that are considered as obvious. Analysing of domain-specific data could involve a lot of expert knowledge of different kinds. So it could be possible to track the interest of the expert and turn their minds to particular features of the visual scene during the following interview.

2) Experts' and specialists' training. By analysing the affective state of the expert during the exploration of training data sets, it might be possible to identify features which require attention. These features can be explicitly marked for educational purposes. Studying experts, decision makers, or even students could see which particular features are usually important within domain-specific data.

A. Virtual Reality Semantic Structure

Today, virtual reality technology can be considered as a powerful tool for interactive data visualisation. We try to exploit the idea of complex scientific data presentation within virtual reality (interactive 3D scenes), but the approach can be mapped with ease onto any 2D or 3D data representation.

Within the presented approach the core structures are semantically described as domain-specific objects which organise incoming data. Frequently within e-science tasks [1] the data represents the result of a computer simulation. In this case, the process of data exploration is tightly coupled with the process of simulation. The proposed approach uses VSO (Virtual Simulation Objects) concept and technology [7] which allows the organisation of the simulation and data analysis process around the domain-specific semantic structure of the investigated system. Within the VSO concept, the system is described as a set of objects which interact with each other. Each of the VSOs is related to some real-world object, which forms a real system. So the set of VSOs can be considered as an image of the system to be investigated using simulation. This concept claims the following requirements:

1) System's description. This should be considered as a structural model of the investigated system, containing the objects which are interacting with each other and with the environment. The environment can be described as a separate VSO or as a set of parameters.

2) VSO can be considered as a distributable set of knowledge which can be integrated within a processing system to make it support the simulation of particular objects. This allows us to make the system's description interpretable in two ways: a) the system should have sense within a problem domain (for the user); b) the description should be machine-interpretable (for the simulation process and data analysis).

3) Typical e-science tasks consist of three stages: modelling, simulation and result analysis. VSO concepts should present continuous technological and informational support for all three stages of this process.

The structure of the virtual simulation object can be considered as a tuple:

$$VSO = \langle B, V, Q, M, E \rangle, \quad (1)$$

where B is a set of available bases (spatial, temporal or group [8]); V is a set of values which can be defined on the bases from B; Q is a set of quality metrics for values from V; M is a set of models, which operates with values from V; E is a set of interconnections between models. The system's description includes a set of VSOs.

A set of visual objects (images) can be associated with the VSOs set:

$$S(VSO) = \{ s_O^{3D} \} \cup S_V^{3D} \cup S_C^{3D}, \quad (2)$$

Where s_O^{3D} – basic visual image of the VSO; S_V^{3D} – set of visual images for values; S_C^{3D} – set of visual cognitive structure that extend basic VSO representation. The position of visual images can be described as the following function:

$$f_{pos}(VSO) : \bigcup B \to \mathbb{R}^4. \quad (3)$$

This maps bases of the VSO onto 3D space and 1D time. Semantically, the system is described as a set of interactive objects $VSO_{SC} = \{ VSO_i, i = 1 \div N \}$. So the visual scene can be

described as a set $S_{SC} = \bigcup_i S(VSO_i)$. For the purpose of scene analysis the following function can be defined:

$$S^{-1} : S_{SC} \to VSO_{SC}. \quad (4)$$

This allows us to identify the semantic object for any visual object.

During the process of visual exploration the user is characterised by the position and gaze direction $p(t) \in \mathbb{R}^3 \times \mathbb{R}^3$ within the 3D scene. But, within a consideration of the task, a much more interesting problem is to identify the attention position (focus) $a_G(t) : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$. Moreover, using this "spatial" attention focus, visual objective focus can be defined as $a_O(t) : \mathbb{R}^3 \to S_{SC}$, which identifies the visual object (image) which the user is looking at. Finally, function (4) gives semantic focus, which identifies the semantic object and all the related data:

$$\begin{aligned} &a_{SO}(t) = S^{-1}(a_O(t)) : \mathbb{R}^3 \to VSO_{SC}, \\ &a_{SV}(t) : \mathbb{R}^3 \to a_{SO}(t).V \times a_{SO}(t).B, \text{ if } a_O(t) \in S_V^{3D}, \\ &a_{SC}(t) : \mathbb{R}^3 \to \left(a_{SO}(t).V \times a_{SO}(t).B\right)^*, \text{ if } a_O(t) \in S_C^{3D}, \end{aligned} \quad (5)$$

Where $a_{SO}(t)$ represents attention to the object; $a_{SV}(t)$ – attention to the values within the bases of the object; $a_{SC}(t)$ – attention to the visual cognitive structures.

The presented formal description allows mapping of abstraction levels 2-4 with an arrangement of user's attention during data exploration. Thus, it allows us to trace the user's exploration process through space, and visual and semantic objects. This trace forms the initial data set for an affective state analysis which is performed using BCI.

One important issue here is that in many cases it is quite difficult to estimate a single position of attention (especially concerning 3D space). There are several reasons which make this task more complex. Firstly, it is difficult to estimate the correct gaze direction precisely as there are errors in measurements. Secondly, even if the gaze direction could give a single point on the screen, the viewport of the user is somehow wider. Finally, if there are a lot of small objects on the screen, it is difficult to guess which one the user is looking at. This can be more troublesome if the 3D space is considered as the distance to the object to be taken into account. To overcome these issues several techniques can be used:

1) Gaze weight estimation. During the process of exploration the gaze direction produces a weight function $w(d)$ which depends on the distance from the gaze line (within a 2D or 3D space). Thus, the function takes into account the spatial fuzziness of the attention. It has a maximum at position $d = 0$ and declines while d is increased. The form of the function can be different. Within our experimental research we used the following weight functions:

$$\begin{aligned} &w_1(d) = e^{-\frac{d^2}{d_0}}, \\ &w_2(d) = \begin{cases} 1 - \dfrac{|d|}{d_0}, \text{ if } |d| < d_0, \\ 0, \text{ otherwise,} \end{cases} \end{aligned} \quad (6)$$

where d_0 allows us to tune "wideness" of attention.

1) 3D analysis. The process of interactive exploration within virtual reality gives us additional capabilities for gaze position estimation. For example, if the viewport of the 3D scene is spinning around some point and the user is looking at one object from different directions, it gives a better precision of attention position estimation as there exists a point of intersection of gaze lines. If the affective track is recorded and then aggregated using the weight function, the position or object at the intersection place will gather more scores.

B. Affective State Estimation and Mapping

Today, brain-computer interface allows us to estimate different parameters of a user's affective state [4, 5]. The popular approach is to use electroencephalography (EEG) devices for this purpose as they are easy to use and provide quite good opportunities for building BCI. The devices which implement this technology become lighter and easier to use (see e.g. Axio project [9]).

Within the proposed approach we try to join affective state estimation with the process of visual data exploration. During the exploration, a set of characteristics $\xi(t) : A \to \Xi$ can be tracked. Here A is a set of available attention aims (possible focus), Ξ is a set of BCI-based affective characteristics (engagement, excitement, etc.). It is possible to associate characteristics with different attention aims: spatial attention – $\xi_G(t) : \mathbb{R}^3 \to \Xi$; visual objective attention – $\xi_O(t) : S_{SC} \to \Xi$; and semantic attention which, using (5), can be defined as follows:

$$\begin{aligned} &\xi_{SO}(t) : VSO_{SC} \to \Xi, \\ &\xi_{SV}(t) : a_{SO}(t).V \times a_{SO}(t).B \to \Xi, \\ &\xi_{SC}(t) : \left(a_{SO}(t).V \times a_{SO}(t).B\right)^* \to \Xi. \end{aligned} \quad (7)$$

The function $\xi(t)$ can be used to construct 3D spatial maps of affective state and estimation of attitude towards different visual and semantic objects. As the goal of affective state processing within the considered task is to analyse the user's attitude towards the data structures, the estimations (7) are usually more important.

One of the important things about the affective state processing is aggregation of the tracked states Ξ to estimate affective characteristics of visual and semantic objects in accordance with attention track gathered as it was described earlier. Within the proposed approach, two types of analytical aggregation procedures are used.

Fig.1. Experimental facilities: a) Emotiv EPOC BCI; b) 3D-Wall and Microsoft Kinect; c) DIY eye-tracking system

1) Weighted summation. This used equations (6) for every affective parameter. This allows us to score almost every object within the scene. So, the resulting score can be defined as:

$$K(s) = \sum_t w\left(\left|a_G(t) - s\right|\right) \cdot \xi(t) \,, \qquad (8)$$

Where s is an object for score estimation; $\left|a_G(t) - s\right|$ is a distance from gaze (as a 2D point or 3D line) to the estimated object which is used in weight function (6); $\xi(t)$ is a characteristics vector (7) tracked by BCI.

2) Statistical analysis of affective parameters registered within a limited distance from the object: $\left|a_G(t) - s\right| < d_0$. This analysis could show affective characteristics induced by the object with more details.

III. Experimental Research

A. Hardware facilities

For the purpose of experimental research of the proposed approach, several pieces of specific hardware were used. The core hardware within the experiments was the Emotiv EPOC BCI [10] which uses EEG technology (see fig. 1a). Among other available features, this device allows us to estimate the affective state of the user including the values of the following characteristics: excitement, engagement, meditation and frustration. To perform the experiments two test beds were used:

1) A virtual reality environment was built using a 3D-wall (fig. 1b) which allows us to observe the virtual environment using polarised glasses. To interact with the environment, the Kinect [11] interface was used to support movement within the 3D space. The wide screen of the 3D-wall allows us to use gyroscope sensors of the epoc to track the head position as a basic gaze-tracking source.

2) A PC-based environment was built using a general-purpose PC. The 3D environment mouse was used for interaction. To track the gaze of the user, an eye tracking system was built using a Genius iSlim 321r camera (fig. 1c), taking as an example the experience described in [12]. To track the gaze direction, the ITU Gaze Tracker [13] was used.

B. Software implementation

The presented approach was implemented within the application for a visual analysis of complex networks (such as social networks, organisational structures, etc.). The software was developed using the Microsoft .NET Framework and uses the capabilities of the CLAVIRE cloud computing environment [14] developed at the e-Science Research Institute. The environment allows us to perform the complex simulation process in an automatic way and can present the resulting data simultaneously to the simulation by the use of Interactive Workflow (IWF) technology. This approach allows the joining of external data services or clients dynamically. Within the considered task external client is presented by the developed visual data exploration software.

The developed visualisation software presents the network in the 3D space. The network can be painted (both edges and nodes) and laid out according to predefined rules. The user is allowed to move "through" the network space and to zoom it in and out. Also, the user is allowed to switch the layout and colouring scheme, and to obtain further information on selected nodes.

C. Test case

For the purpose of the experiments, a network of sexual contacts used for HIV simulation [15] was visualised using developed software. The network for that case contains individuals (nodes) and their contacts (edges). During the experimentation process the user explores presented data within the 3D space by moving and zooming in and out of the presented visual structure. Fig. 2a shows a sample screenshot of the visualised complex network. Here, the network has a connection-based layout (nodes with more connections are displayed on the top). While the user was exploring the visual scene, the user's gaze was tracked. Fig. 2b shows the gaze track that was recorded while exploring the static view of the visual scene presented in fig. 2a. Then, gaze track was transformed into a viewing heat map (see fig. 2c for the presented gaze track).

This technique is often used within usability researches. It is supposed to show the places on the screen which attract most of the user's attention. The gaze track and heat map can be also constructed for a 3D space. Within the research they are implicitly constructed during the score computation.

Simultaneously, to gaze tracking the values given by BCI are also tracked. Within the test case, engagement value is mostly used as it seems to be semantically closer to the interest level. Averaging these values can give an additional affective heat map (see fig. 2d). The affective heat map resembles the view heat map in shape, but it has differences in the heat value (as they are more than just a sum).

More interesting results can be obtained while analysing the affective state of the user within the relationship to the particular objects. In this case network nodes are considered as such objects. Figure 2e shows the distribution of affective state

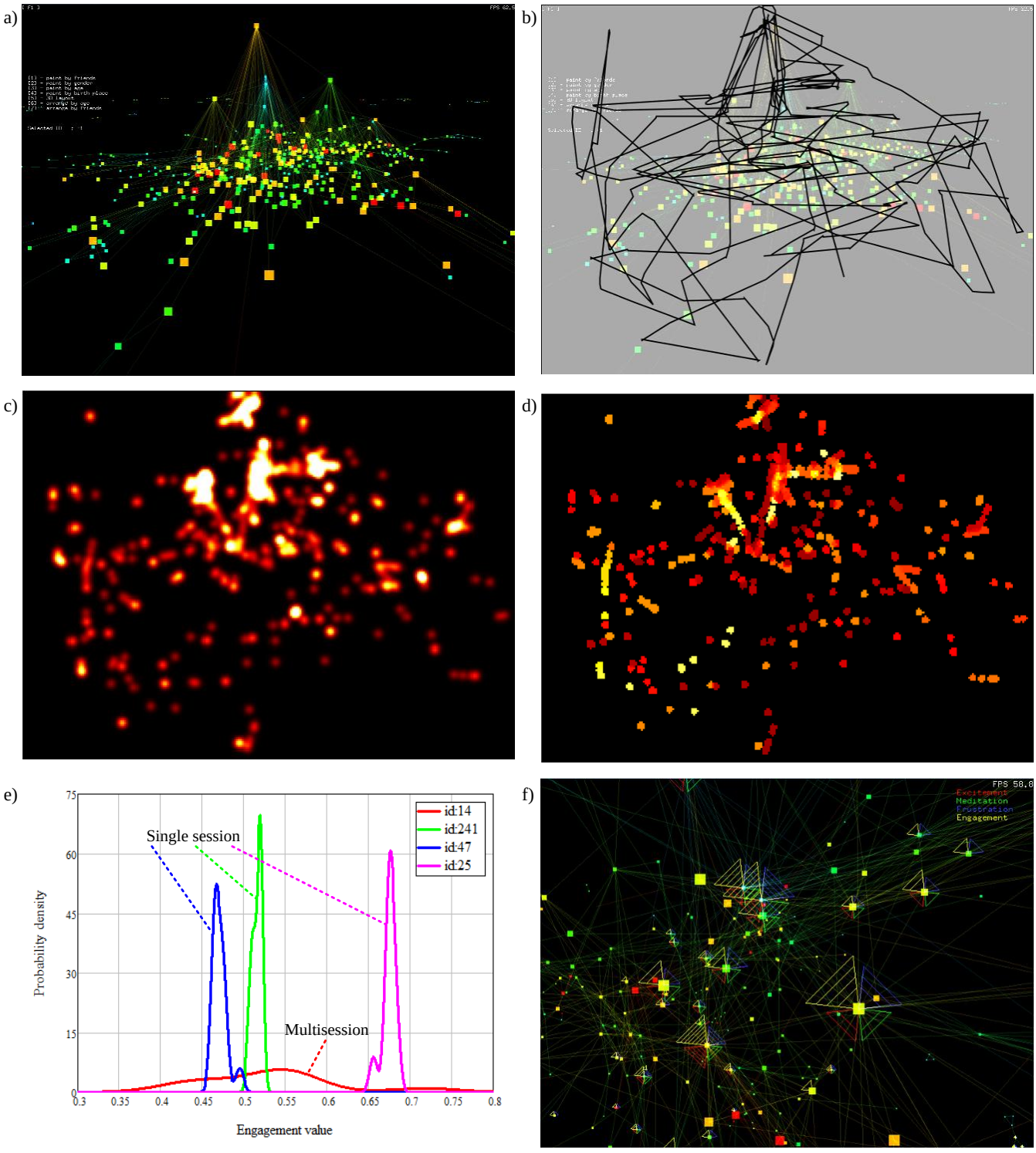


Fig.2. Experimental data: a) initial data visualization; b) view track; c) view heat map; d) affective heat map (engagement); e) distribution of measured value for top-viewed nodes (engagement); g) node scores visualization

values (engagement) recorded in relationship to particular nodes (mentioned by id). The experiment shows that: a) there are "sessions" within the exploration process, because the user can return to the already viewed node showing a different value of affective characteristics (fig. 2e shows 3 nodes with a single view session, which have very narrow distributions, and one node with a multiple session – it has quite a smooth distribution); b) the distribution of BCI-tracked values within a session can be distinguished from one another. Results of affective state analysis allow scoring of viewed objects (nodes): fig. 2f shows a visualisation of the same network with the scores for different affective characteristics (excitement, engagement, meditation and frustration). The value of characteristics is shown by the number of triangles of different colour orientated in four directions.

IV. Discussion

Contemporarily, affective BCI is quite a new and developing technology which gives us the ability to interact with the user on a completely new level. While initially BCI was mainly used with medical issues or to support disabled people, today there are tasks where it supports, optimises, and even enhances the general user's ability [16]. Moreover, this technology forces the appearance of completely new areas of research like neuromarketing [17] or augmented cognition [18].

On the other hand, as the amount of information available for analysis today is increasing more rapidly, new forms of data analysis are required. One of the approaches towards overcoming the problem is the visualisation of complex data using specific artistic techniques (see [2] and examples at [19]). However, the analysis process still remains quite complicated, as within domain-specific tasks the expert's knowledge or specific data-mining solutions are required to recognise some important features.

The idea that lies behind the developing approach is to apply affective BCI technology to support the data analysis process (e.g. to support the expert knowledge acquisition process). As there is a lot of tacit or complicated knowledge in the wide area of problem domains, the idea promises to be fertile and it can be used immediately: during the exploration or by analysing recorded tracks; within experts' interviews or within the study process; with data-based aggregation or for analysing the user state and skills.

Yet, the analysis of the idea uncovers a set of questions. Which part of the user's impression is caused by the data analysis, and which part is caused solely by the visual expressiveness of the image? What influences do the domain skills and experience of the user have? How should the user's current state and mood be taken into account? We are trying to answer these and other questions within our current research.

V. Conclusions and Future Work

The work presented in this paper is devoted to the development of technology which augments the visual exploration of complex scientific data with affective BCI. The scope of the research lies between a set of developing areas such as BCI, affective computing, augmented cognition, human-computer interaction, complex data visualization, virtual reality etc. Within the solution the significant role is played by BCI technology. Today the area of applied BCI is presented by many solutions (see, for example, works [3-5, 16-18]). Still the issue of affective complex data exploration using this technology is uncovered.

The presented approach is aimed at the implementation of a new of presenting complex data by leading the user to the most significant objects of visualised data. The developed technological solution and experimental research performed using the complex network visualisation application show the ability to use affective BCI within this task.

The work is still in progress. Upcoming future plans include a deeper analysis of time-based and session-based measuring of affective user state, development of general technology for virtual reality augmentation on a basis of performed experimental technology, and further experimental research for a more complete understanding of the affective data exploration process (including the questions mentioned earlier).

Acknowledgement

This work was supported by the project granted from the Ministry of Education and Science of the Russian Federation under agreement 14.B37.21.1870 and the project awarded with the Presidential Scholarship number SP-2198.2012.5.

References

[1] T. Hey, S. Tansley, K. Tolle, "The Fourth Paradigm. Data-Intensive Scientific", Discovery, Microsoft, 2009, p. 252.

[2] E.R. Tufte. "The Visual Display of Quantitative Information", 2nd edition, 2001, p. 200.

[3] A. Lecuyer, F. Lotte, R.B. Reilly, R. Leeb, M. Hirose, M. Slater, "Brain-Computer Interfaces, Virtual Reality, and Videogames", IEEE Computer, Vol. 41, Issue 10, 2008, pp. 66-72.

[4] A. Yazdani, J.S. Lee, J.M. Vesin, T. Ebrahimi, "Affect recognition based on physiological changes during the watching of music videos", ACM Transactions on Interactive Intelligent Systems, 2(1), 2012, pp. 1-26.

[5] J.B.F. van Erp, H.(J.A.) Veltman, M. Grootjen, "Brain-Based Indices for User System Symbiosis", Brain-Computer Interfaces. Applying our Minds to Human-Computer Interaction, Human-Computer Interaction Series, 2010, pp. 201-219.

[6] R.W. Picard, "Affective computing", MIT Media Laboratory, Perceptual Computing, 1995. [http://affect.media.mit.edu/pdfs/95.picard.pdf]

[7] S.V. Kovalchuk, P.A. Smirnov, S.S. Kosukhin, A.V. Boukhanovsky, "Virtual Simulation Objects concept as a framework for system-level simulation", IEEE 8th International Conf. on E-Science, 2012, pp. 1-8.

[8] G. Klir, "Architecture of Systems Problem Solving", Plenum Press, New York, 1985.

[9] Axio, revolutionary focus [http://www.axioinc.com/]

[10] Emotiv. EEG System. Electroencephalography [http://www.emotiv.com/]

[11] Kinect for Windows. Voice, Movement & Gesture Recognition Technology [http://www.microsoft.com/en-us/kinectforwindows/]

[12] R. Mantiuk, M. Kowalik, A. Nowosielski, B. Bazyluk, "Do-It-Yourself Eye Tracker: Low-Cost Pupil-Based Eye Tracker for Computer Graphics Applications", Advances in Multimedia Modeling. Lecture Notes in Computer Science, Vol. 7131, 2012, pp. 115-125.

[13] J. San Agustin, H. Skovsgaard, E. Mollenbach, M. Barret, M. Tall, D.W. Hansen, J.P. Hansen, "Evaluation of a low-cost open-source gaze tracker", in Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications (Austin, Texas, March 22 - 24, 2010), 2010, pp. 77-80.

[14] K.V. Knyazkov, S.V. Kovalchuk, T.N. Tchurov, S.V. Maryin, A.V. Boukhanovsky, "CLAVIRE: e-Science infrastructure for data-driven computing", Journal of Computational Science, Vol. 3, Issue 6, 2012, pp. 504–510.

[15] P.M.A. Sloot, S.V. Ivanov, A.V. Boukhanovsky, D. van de Vijver, C. Boucher, "Stochastic simulation of HIV population dynamics through complex network modelling", International Journal of Computer Mathematics, Vol. 85, Issue 8, 2008, pp. 1175-1187.

[16] B. Blankertz, M. Tangermann, K.R. Muller, "BCI Applications for the General Population", Brain-Computer Interfaces - Principles and Practice, 2012, pp. 363-372.

[17] D. Ariely, G.S. Berns, "Neuromarketing: the hope or the hype of neuromarketing in business", Nature Reviews Neuroscience, 11(4), 2010, pp. 284-292.

[18] A.A. Kruse, D.D. Schmorrow, "Session overview: Foundations of augmented cognition", Foundations of Augmented Cognition, 2005, pp. 441–445.

[19] Visual Complexity [http://www.visualcomplexity.com/]

LED to LED communication with WDM concept for flash light of Mobile phones

Visible Light Communication

Devendra J Varanva
PG Student: ECE Department,
School of Engineering, RK University,
Rajkot, India

Kantipudi MVV Prasad
Assi. Prof.: ECE Department,
School of Engineering, RK University,
Rajkot, India

***Abstract*— after observing recent developments in Free Space Optical Communication especially Visible Light Communication, It is clear that LED is main component as a source. LED being solid state device makes endless list of possibilities. But here we will get through its ability to sense light as well, and use of Wavelength Division Multiplexing (WDM) in mobile flash is also suggested, this opens door to many applications.**

Keywords—Free Space Light Communication; Visible Light Communicatio; Solid State Device; LED as Sensor; LED-LED Communication; WDM; VLC with Mobile handset/Smartphone.

I. INTRODUCTION

Free Space Optical Communication is emerging technology. It works in similar way like optical fiber network works but here air/free space shall be used as transmission medium. Today in different systems Laser, Infrared, and Ultraviolet Light used for communication over the air or in free space. This is actually needed at some places where wired connection is not easier or it is temporary connection. Like some years back use of IR communication in mobile handsets to share files was very popular until Bluetooth came. It has low speed but it was enough if we talk about that time. Best part of it was wireless connection which is better than connecting two phones with wire for such a short time of use. Being wireless has been dream of every human being. But here we are talking about free space communication using visible LEDs; popularly known as Visible Light Communication (Visible light is a part of electromagnetic spectrum which is perceivable by humans). People are talking much about it and they should, as this technology has very high potential to dominate communication market which is dominated by Radio Frequencies now.

Visible Light Communication is not a new technology. Since ancient times humans are using it in simple form. For example, in old times to give war signals, they used reflection of sunlight using brushed iron piece or smoke in day time and at night they were using fire to give signals. That was the best and very effective way to inform others by giving signal like this for that time as technology was not developed enough.

Today LED is very power efficient unlike fluorescent lights, tungsten bulb, etc. It is solid state device which can switch at very high speed that human eyes can't perceive. This property is matter of interest. We can send data with OOK (On-Off Keying) [1], OFDM [2], etc modulation schemes. To send data via Visible light communication we should switch LED On and Off. When LED switched On, receiver will detect as "1" and when it switched Off, receiver should detect as "0".It is not as simple as sending and detecting, you need to use amplifier/filter at receiver end. Best part of Visible Light Communication is you can get advantage of high speed data communication at the same time you can use it for lighting purpose. By using this technology we can save lot of energy wasted in using Wi-Fi, as radio frequency can be used for only data communication purpose not lighting.

As we all know LED can be used for sensing purpose too. Publish by Forrest W. Mims [3] [4] in late 1970's but we have forgotten that. LED is nothing but PN diode. Working of Photodiode/sensor is quite similar. In our experiments we decided to use sensing capability of LEDs, Especially clear lens LED used in our experiments which gives best results.

II. POTENTIAL ADVANTAGES OF VLC

Visible Light Communication has very unique advantages (especially over RF); some of them are explained here.

A. Two functions : Lighting & Data

We use visible color LEDs for lighting purpose (as it supposed to be) as well as for Data Communication .Combining illumination and Communication is dual advantage for us unlike Radio Frequencies.

B. Security

If you are under the light/receiving the light means only you are receiving the data. e.g., one can use their light as hotspot to transmit data and you can receive data only where light reaches. So someone outside the room can't receive it if light is not going outside unlike Radio Frequencies where you cannot do same.

C. It can work in hazardus zones

We know that Radio Frequencies are not allowed in hazardous zone like nuclear plants, petroleum refineries, Hospitals, etc. But light used everywhere so we can install VLC hotspot for communication purpose, or to send current location data for automatic robot or other kind of machine and also can be used to send warning signal.

D. Parallel Communication is also possible

Researchers have already shown wavelength division multiplexing (WDM) capabilities for VLC. We can send

different data on different line at the same time using different wavelength of the light [5].

III. CHALLANGES

Advantages come at a cost in VLC, as it has some challenges to be resolved. (1)VLC is Line of Sight communication require you to align transmitter and receiver in straight line of sight. Without it VLC won't work. (2) Communication at long distances can't be possible due to dust or foggy environment, or in rain. In simple terms we can say if you can see the light you can receive the data. Despite many challenges, VLC is still very effective technology for future.

IV. LED AS A SENSOR

Many researchers are working on this technology. It has been since Forrest M. Mims have shown that LED can act like a sensor to. This is the fact that clear lens LED can sense light with lower wavelengths than it emits. In short LED's sensitivity region is slightly wider then its spectral emission profile. So it can be used as a low cost option for some simple applications in place of LDR(Light Dependent Resistor) or Photodiodes(PD can detect wide spectrum of light).as LEDs are widely available at low cost. The fact is LEDs are intended to work for lighting purpose but we are using it in VLC for data communication too. And right now we are talking about using it as a photo sensor. But matter of interest is its capabilities to be used as a low speed photodiode which can have speed not more than some KBs. here we want to mention that LEDs can be used for bi-directional communication too using microcontroller by changing its polarity with two pins (both pins of LED connected to two I/O pins) [6] in which LED will show us property like capacitor, we first charge it and after changing its polarity we will calculate time to discharge. Here when you through light on it, will make it to discharge faster. We can calculate intensity of light by this way [7].

As shown in Table I. different color clear lens LED showed us different properties. We took Voltage change in LED when light is thrown to it. As we noted earlier LED can sense only wavelength emitted by it or from slightly near spectrum. Different Color LEDs have different characteristics as we can see in the table. Where green LED (Tx) to Orange LED (Rx) showed us best performance, also Blue LED (Tx) to Green LED (Rx).Specially for bi-directional communication (same LED for transmitter and receiver), Red LED might be good choice. As it can receive same light better emitted by its self.

Here the biggest advantage over Photodiode is, we don't need glass of other king of filter for filtering particular color of light, LED will do by itself. We have connected Anode to voltmeter and cathode to ground to measure voltage (displayed in Table. I). Resistor was used to control current of LED. It should be noted that these are the values for sensing purpose only we have to look at minor photocurrent generated by it and the frequency response of LED as sensor too. Different Color LEDs have different characteristics.

TABLE I. DIFFERENT COLOR LEDS AS PHOTO SENSOR

LED	**Detector**					
	COLOR	***W***	***R***	***G***	***O***	***B***
Emitter	**WHITE**	11	430	405	420	30
	RED	0	750	0	30	0
	GREEN	0	450	105	1380	0
	ORANGE	0	140	0	50	0
	BLUE	54	30	1250	165	270

a. all values are in mill-volts(all are average/approximate values)

b. performed in normal room light condition

c. distance between Tx and Rx was 7.5cm

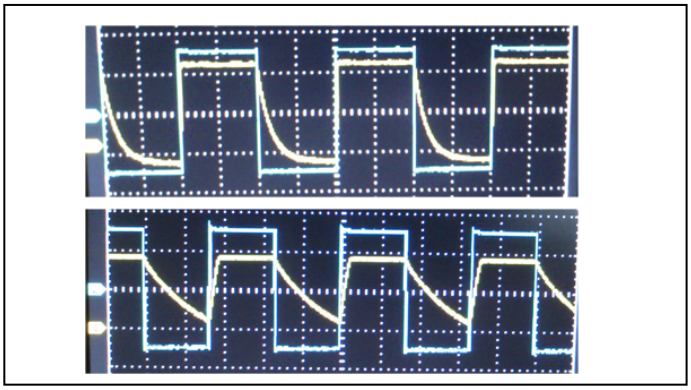


Fig.1. Waveforms of LED as Sensor

In Fig.1 two waveforms received by LED is shown, which are 500Hz and 3 KHz respectively. We have noted that quality of waveform degraded at higher frequency. We can go further this frequency too but we have to use filter and amplifier as far as communication is concern. The waveforms are for indication purpose only, because it depends on length between transmitter and receiver, external lighting condition, and noise. We can't ignore the fact that LED is not coherent light source like LASER is. So long distance can't be achieved. But we can use reflector at transmitter side to focus light and magnifying optics at receiver side to concentrate light [8].

V. VLC FOR MOBILE HANDSETS

Today all mobile phones uses flash for camera or as emergency light for night. There is not any additional use of it. In LED white light is produced by blue emission which excite yellow phosphor layer on top of the LED. So we can say white light is generated by combining Blue and Yellow radiation.

In suggested scheme we can use Mobile Handset flash to transmit data. Researchers have already achieved higher data rate with White LEDs[9] .But as we know that white light can be generated by Red, Green, Blue lights too, which is what we are interested in. In proposed method we suggest to use this 3 color light to produce white light for future mobile applications. In mean while someone can use these LEDs as sensors to receive data too when flash is not being used. As shown in Fig. 2 we have used RGB led which was producing White Light (this is the fact everybody know already, but included for better understanding of readers).

Fig.2. R-G-B LEDs to produce white light.

We can use this technique in mobile handsets in which LEDs are to be connected to microcontroller of the mobile to receive data as well. As most of the mobile handsets don't use flash all the time we can use this for low level data communication like to exchange text or midi tones. But it is also the fact that Blue LED is not good as far as sensing is concerned. And Green can't receive light from Red and Orange LEDs. But Red LED is better receiver. So we can use Red LED to receive data. And other LED to produce white light when data is not being received.

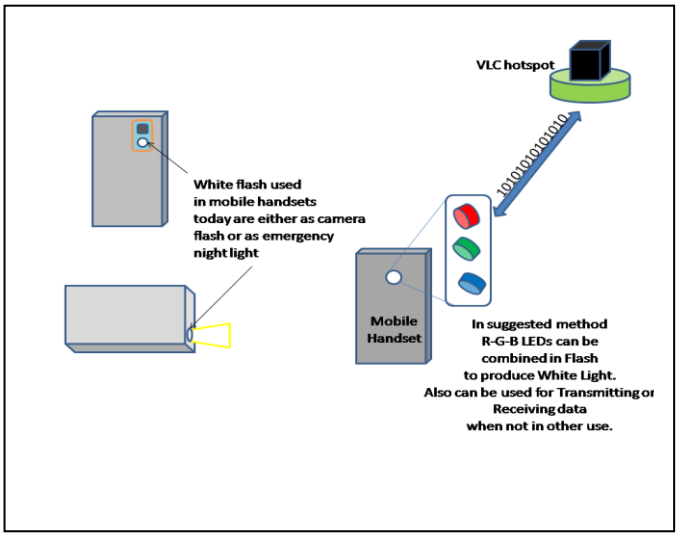


Fig.3. Suggested model for flash in Mobile Handsets.

We can see in Fig.3 where we have shown such model, which can be implemented with this technology. Performance of LED can be improved with filter and amplifiers. But there is also a limit. More and more manufacturer should do research and development to produce such hybrid LEDs which can be used for bi-directional communication. We have used 5mm clear lens LED but one can also try different LED to get improved result along with good photo diode amplifier.

We should also note that all the home appliances we use today have LED as power or other function indicator. So we can implement this technology there as well. Applications are endless only limited by the imagination of someone.

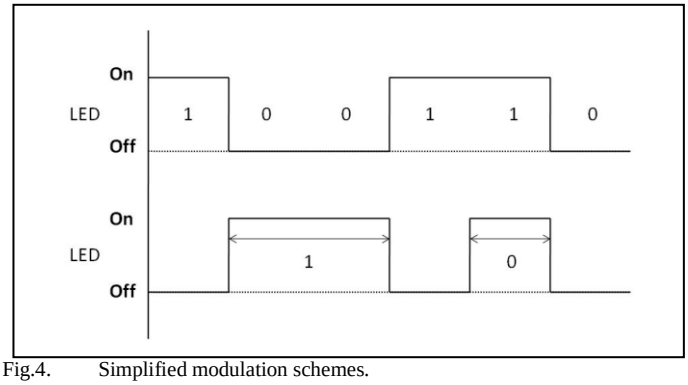


Fig.4. Simplified modulation schemes.

Modulation scheme is also important in this regard. We can do parallel communication here so appropriate scheme should be selected according to our application. OOK, PWM, OFDM, etc are the popular among researchers for visible light communication [10]. We can even build our low level protocol to develop applications. This is what we are doing our research on.

VI. Results

As we noted earlier that some LED can "sense" only some LED lights, in our experiments we had two set up where in first set-up two microcontrollers used as Transmitter and Receiver. From transmitter we have 4 switches which are supposed to control 4 equipment (LED in our case).the transmission medium was Red LED as transmitter and another Red LED as receiver. At the same time in second set up we have used Blue LED as transmitter connected with function generator with certain frequency and Green LED as receiver. Distance between both transmitter and receiver was ~3cm.both Tx and Rx worked individually without any loss even placed side by side in such a way that both transmitters have 1cm distance only and same was done on receiver side. We should note here that when we decreased distance so lower to place them very near at that time Red LED was unable to sense switching command send by the transmitter Red LED due to the interference of Blue and Green LEDs. But with approx 3cm distance we were able to switch on LED/devices at receiver side (with Red LED to Red LED) at the same time we were able to receive waveforms at second receiver (with Blue LED to Green LED).we selected Red LED for first configuration intentionally as we wanted to try same color for transmission and reception purpose.

All experiments are in primary stage as one has to make this technology more effective and resolve noise problem. Here choosing modulation scheme is also very important. There might be many use of this technology but in current time we

use camera flash not even for data transmission purpose. We encourage more and more research to be done for this.

Here LED was working with such low frequency so we can do communication with two methods shown in Fig .4. Where first is a scheme in which we can set timer for both side and can send data with changing the state of LED on or off. Popularly On-Off keying. Synchronization is most important unless we will be compromising with data rate. In second scheme we don't have to set timer as we defining certain time for both 1 and 0.therefore if LED will successfully sense light for defined time it will be easy for us to consider it as 1 or 0.popularly known as Morse code where dot-dash used in similar way. This is simple scheme which can be used for development purpose.

VII. Future Work

Data rates are matter of concern here so work should be done to achieve better data rates by using different LEDs, better /high end hardware like FPGAs, or appropriate modulation scheme.

In our experiments we were limited by some cm distance because we didn't use reflector and magnifying optics and not even amplifier/filter. One can achieve better light gathering by this like astronomy telescope do.

VIII. Conclusion

In this paper we have shown how VLC is useful over Radio Frequencies at some places. We have also shown that different color LED can sense different lights. According to our application we can select different color LEDs. In suggested model we have shown idea of improved flash with WDM which can be used as data transceiver. We are also continuing research to implement this technology successfully and finding other ways to improve results. Other applications can be done also using this technology, and list of it will never end.

References

[1] H. Le Minh , D. O. ;Brien , G. Faulkner , L. B. Zeng , K. Lee , D. Jung , Y. Oh and E. T. Won "100-Mb/s NRZ visible light communications using a post equalized white LED", IEEE Photon. Technol. Lett., vol. 21, no. 15, pp.1063 -1065 2009.

[2] Elgala, H., Mesleh, R., Haas, H., Pricope, B.," OFDM Visible Light Wireless Communication Based on White LEDs", IEEE 65th Vehicular Technology Conference, 2007. VTC2007-Spring., 22-25 April 2007, 2185 - 2189.

[3] Forrest M. Mims III," Siliconnections: Coming of Age in the Electronic Era" , McGraw-Hill, New York, NY, 1986.

[4] Forrest M. Mims III, "LED Circuits and Projects", Howard W. Sams and Co., Inc., New York, NY, pp. 60-61, 76-77, 122-12.

[5] Khan, T.A., Tahir, M., Usman, A., "Visible light communication using wavelength division multiplexing for smart spaces", Consumer Communications and Networking Conference (CCNC), 2012 IEEE,14-17 Jan,230 – 234.

[6] P. Dietz, W. Yerazunis, and D. Leigh, "Very low-cost sensing and communication using bidirectional leds," in UbiComp 2003: Ubiquitous Computing. Springer, 2003, pp. 175–191.

[7] S. Schmid, G. Corbellini, S. Mangold, T. R. Gross, "An LED-to-LED Visible Light Communication System with Software-Based Synchronization", GLOBECOM Workshop (OWC), 3-7 Dec, 2012, Anaheim, CA, USA.

[8] DJ Varanva, KMVV Prasad, "Various Aspects of Visible Light Communication and its applications", e-DCSECT-2013, International Journal of Electronics and Communication Technology (IJECT) Volume 4, Spl – 2 / Jan - March 2013,105-107.

[9] A. M. Khalid, G. Cossu, R. Corsini, P. Choudhury, E. Ciaramella," 1-Gb/s Transmission Over a Phosphorescent White LED by Using Rate Adaptive Discrete Multitone Modulation", Volume 4,Number 5,October 2012,1465-1473.

[10] O'Brien, D. C. Zeng, L., Le-Minh, H., Faulkner, G., Walewski, J.W., Randel S.,"Visible Light Communications: challenges and possibilities", Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium, 15-18 Sept. 2008, 1-5.

Rule Based System for Recognizing Emotions Using Multimodal Approach

Preeti Khanna
Information System
SBM, SVKM's NMIMS
Mumbai, India

Sasikumar, M.
Director, R & D
Center for Development of Advance Computing, CDAC
Mumbai, India

***Abstract*—Emotion is assuming increasing importance in human computer interaction (HCI), in general, with the growing feeling that emotion is central to human communication and intelligence. Users expect not just functionality as a factor of usability, but experiences, matched to their expectations, emotional states, and interaction goals. Endowing computers with this kind of intelligence for HCI is a complex task. It becomes more complex with the fact that the interaction of humans with their environment (including other humans) is naturally multimodal. In reality, one uses a combination of modalities and they are not treated independently. In an attempt to render HCI more similar to human-human communication and enhance its naturalness, research on multiple modalities of human expressions has seen ongoing progress in the past few years. As compared to unimodal approaches, various problems arise in case of multimodal emotion recognition especially concerning fusion architecture of multimodal information. In this paper we will be proposing a rule based hybrid approach to combine the information from various sources for recognizing the target emotions. The results presented in this paper shows that it is feasible to recognize human affective states with a reasonable accuracy by combining the modalities together using rule based system.**

Keywords—Human Computer Interaction (HCI); Multimodal emotion recognition; Rule based system; Emotional state; Modalities.

I. INTRODUCTION

With the increasing role of computer system in society, HCI has become an integral part of our daily lives. In today's scenario, computers are not only used to perform tasks, but also to learn, communicate, explain, argue, debate, observe and also design. The major concern of HCI now is the need to improve the interactions between humans and computers through justifications and explanations. Thus we observe a significant growth of new forms of 'natural' and 'indirect' interfacing. HCI is experimenting with alternate input mechanisms and multimodal input mechanisms through speech, gesture, posture and facial expression. These help in substituting the largely impersonal devices such as a keyboard and a mouse, for a non-tech savvy.

One of the significant ingredients which could enhance the interaction between human and computer is emotions. Emotions play a vital role in the communication among human beings. However so far, emotions have not played a substantial role in HCI. Incorporating the emotions in HCI is a challenging task. Research studies have been undertaken to investigate and develop various approaches and technology to incorporate emotions in HCI. Some of the recent trends with this respect focus on how a computer can automatically detect the emotional state of a user and then adapt its behaviour accordingly. There is increasing research interest and various applications along these lines.

Some of the prominent areas include e-commerce, help desks, customer support, e-learning, etc. For example, an emotion-aware interface can enhance the sensitivity of an automatic tutor which can adjust the content of the tutorial and speed and style at which it is delivered. As helper/assistant robots (e.g. AIBO) are becoming common toys, often aimed at helping elderly persons in their day to day tasks, the ability to relate to their emotions becomes something of paramount importance. Computer games may adapt playing conditions to the emotional level of the players. Surveillance is another application domain in which the reading of emotions may lead to better performance in predicting the future actions of subjects. In this way, the emotion driven technology can enhance the existing systems for the identification and prevention of terrorist attacks in public places. Certainly not all computes need to pay attention to emotions, or have emotional abilities. Some machines are useful as rigid tools, and it is fine to keep them that way.

The issue of enhancing HCI with emotion raises a number of questions. What are the sources of information that a machine can use to decode the emotional state of the user? What kind of information (emotional cues) are available from these sources? How does one use these sources to estimate the emotional state? What are the emotional states of interest for us from the perspective of enhancing HCI? How to combine multiple modalities? Does the performance of the multimodal emotion recognition for a specific set of target emotions depend on type of fusion model? In what ways can a machine use the knowledge of the user's emotional state to modify its behaviour? There is a plethora of existing work that bears on one or more of these questions.

The paper begins by defining problem domain regarding multimodal emotion recognition. Section II discusses the complete framework of our rule based system. This approach is based on certainty factor i.e. the MYCIN approach. Section III talks about the overall framework of emotion recognition independent of any modalities. Then we explained our approach of rule based multimodal emotion recognition using the case scenarios of 'facial expressions' in section IV. We did few experiments on multimodal data and tested on our rule

based system which is explicitly mentioned in section V. We conclude the paper by summarizing the results and consider some challenges facing the researchers in this area.

The Problem Domain for Multimodal Emotion Recognition

Humans recognize emotion, fusing information from multiple sources: speech signals, facial expressions, gesture, bio-signals and others. Inadequacies of unimodal recognition systems drive the need to go for multimodal recognition. In literature, some attempts like [1], [2], [3] and [4] have considered the integration of information from facial expressions and speech. This paper explores how to combine the information from various sources (e.g. facial expression [24], speech [21, 22, 23] and others [20]) to achieve better recognition of emotional state using rule based approach.

There are two broad approaches to design of a multimodal recognition engine: feature based and decision based. Feature level fusion involves simply merging the features of each modality into a single feature vector. In this method of fusing, all the features are mixed together irrespective of their nature and type. For example, feature can be position of some feature points on the face or the prosodic features of a speech signal. Feature sets can be quite large as we will see later. This high cardinality can result in soaring computational cost for this fusion approach [5]. Decision level fusion is based on the fusion of decisions from each modality where the input coming from each modality is processed independently and these unimodal recognition results are combined at the end [6]. This fusion has advantage of avoiding synchronization issues over the feature level fusion. Decision level fusion ignores possible relationships between features coming from different modalities. Several works [7], [8] and [9] have discussed multimodal fusion; in particular [10] discusses many issues and techniques of multimodal fusion.

Finding an optimal fusion type for a particular combination of modalities is not straightforward. A good initial guess can be based on the knowledge of the interaction and synchronization of those modes in a natural environment. Hybrid fusion attempts to combine the benefits of both feature level and decision level fusion methods, may be a good choice for some multimodal fusion problems. However, based on existing knowledge and methods, how to model multimodal fusion for target set of emotions is still an open problem. We propose a hybrid approach for multimodal fusion. This model is based on modeling each modality through a set of rules. In this process of formulation of rules, feature analysis plays a very important role. These rule sets were tested and listed later in our running example of facial expression.

II. Rule Based System: Base for Our Hybrid Model

A rule based system consists of *if-then rules*, a bunch of *facts*, and an *interpreter* controlling the application of the rules. A simple *if-then* rule has the form 'if *x* is *A*, then *y* is *B*'. The if-part of the rule, '*x* is *A*', is called the *antecedent* or *premise*, while the then-part of the rule, '*y* is *B*', is called the *consequent* or *conclusion*. When the premise is known to hold in a scenario, the conclusion can be drawn. This is the normal interpretation of a rule. One of the major strength of rule based representation is its ability to represent various uncertainties. Uncertainty is inherently part of most human decision making. This uncertainty could arise from various sources like incomplete data or domain knowledge used being unreliable. So the if – then rules is often represented like 'If A, B, C ----> then D, with certainty X', where X represents the degree of belief or confidence in the rule [11].

A. Approaches for Handling Uncertainty

To handle these uncertainties, there are two broad approaches - those representing uncertainty using numerical quantities and those using symbolic methods. In numerical approaches, one models the uncertainty by numbers and provides some algebraic formulae to propagate these uncertainty values to the conclusions. These approaches are useful for handling the issues related to "unreliable or inaccurate knowledge". For example, Bayesian reasoning [12], Evidence theory [13] and Fuzzy set approaches [14] are numerical models. On the other hand, symbolic characterization of uncertainty is mostly aimed at handling incomplete information, e.g., Assumption Based Reasoning [15], Default Reasoning [16] and Non-monotonic Logic [17]. For example, if there is not enough information available, the system makes assumptions that can be corrected later, when more information is received.

In our domain, the basic problem is that there are hardly any features or feature combinations which can infer any emotion to complete certainty. Therefore, we concentrate on numerical approaches for handling the uncertainty. We have adopted the '*Confirmation Theory'* as used in MYCIN approach [12] to deal with uncertainty in our domain. This approach works well with rule based representation of domain knowledge.

B. Reasoning with Certainty Factors (CF): The MYCIN Approach

Shortliffe and Buchanan [12] developed the Certainty Factor (CF) model in the mid-1970s for MYCIN, an expert system for the diagnosis and treatment of infections of the blood. Since then, the CF model has been widely adopted for uncertainty management in many rule based systems. Each rule is assigned CF by domain experts. This is meant to represent the uncertainty of the rule. Higher CF indicates that the conclusion can be asserted with higher confidence when the conditions are true. Similarly every fact in the domain is also given CF indications how confident one is in that. Reference [12] intended a CF to represent the change in belief in a hypothesis given some evidence. In particular, a CF between 0 and 1 means that the person's belief in *h* given *e* increases, whereas a CF between -1 and 0 means that the belief decreases. A value of +1.0 indicates absolute belief and -1.0 indicates absolute disbelief. The method generally used to propagate the measure of uncertainty in the antecedents and the uncertainty attached to the rule to the conclusions being derived is briefly explained below. This propagation is done in two steps [11].

- The different antecedents in the rule, in general, have different values of uncertainty attached to them. Some formula is required to combine these measures and provide a consolidated uncertainty number. This option

considers the strength of the weakest link in a chain as the strength of the chain. This is defined as:

CFantecedents = {minimum of CFs of all antecedents} (1)

- Then this measure (uncertainty for the set of antecedents) is combined with the measure of uncertainty attached to the rule to give a measure of uncertainty for the conclusion of the rule.

*CF of the conclusion from rule = {CF associated with rule R1} * {CFantecedents}, provided CFantecedents >= threshold* (2)

It can be seen that the CF obtained for a conclusion from a particular rule will always be less than or equal to the CF of the rule. This is consistent with the interpretation of the CF used by MYCIN, that is, the CF of a rule is the CF to be associated with the conclusion if all the antecedents are known to be true with full certainty. In a typical rule based system, there may be more than one rule in the rule base that is applicable for deriving a specific conclusion. Some of them will not contribute any belief to the conclusion, because CF of antecedents has a CF less than the threshold. The contributions from all the other rules for the same conclusion have to be combined. For MYCIN model, initially CF of a conclusion is taken to be 0.0 (that is, there is no evidence in favour or against) and then as different rules for the conclusion fires, the CF gets updated. MYCIN uses a method that incrementally updates the CF of the conclusion as more evidence for and against is obtained. Let CFold be the CF of the conclusion so far, say, after rules R1, R2,...Rm have been fired. Let CF*in* be the CF obtained from firing of another rule Rn. The new CF of the conclusion (from rules R1, R2……...Rm and Rn), CFnew, is obtained using the formulae given below.

*CFnew = CFold + CFin * (1 - CFold)*
when (CFold, CFin >0) (3)

*CFnew = CFold + CFin * (1 + CFold)*
when (CFold, CFin < 0) (4)

CFnew = (CFold + CFin) / (1 – min (|CFold|, |CFin|))
otherwise (5)

We adopt this calculus in our model and explained later with a running example in section IV. Before that we first discuss the overall framework of emotion recognition system.

III. Framework For Emotion Recognition

The overall conceptual framework for emotion recognition includes pre processing, feature extraction, feature analysis and selection of the features, formulation of rules and measuring performance to classify the target emotional states. We will explain each of these in brief as below. We will use facial expressions as the running example to illustrate these stages, etc. The framework remains same across all modalities [20, 21, 22].

A. Pre Processing and Feature Extraction

The first step is pre processing. The objective of this step is to make the input data in a standard format and suitable for extracting the desired features. Usual preprocessing steps include size normalization of the frontal image, noise removal from speech signal etc.

The next step is feature extraction. We need to identify useful features from each of these input sources (pre-processed input data – image, audio and others). For example, location of feature points such as eyes, eye corners, eyebrows, eyebrow corner, mouth corners, upper and lower lip, nose and nostrils, etc. are important for facial expression analysis. The work in this step involves identifying relevant features and formulating algorithms to extract these features from their respective input data.

C. Feature Analysis and Selection

Once the basic feature set is ready, the next step is analysis of each of these features. The question, 'how does each of the features vary with the emotion' needs to be answered here. Usually every feature doesn't contribute to the same extent to recognize different emotional states. Thus feature analysis and selection is an important step. The case of facial expression mentioned in this paper illustrates feature analysis and selection process in detail later.

D. Formulation of Rules

If-then rules are one of the most common forms of knowledge representation used in various domains. Systems employing such rules as the major representation paradigm are called rule based systems.

To design the rules for classifying emotions, all the relevant features needs to be studied in more detail to see its ability to distinguish between different target emotional states. Influential and useful features can be used to define rules. This approach remains broadly same across different modalities and is as follows:

1) Feature analysis has been done for each feature to see its ability to distinguish among the target emotional states, and accordingly useful features were shortlisted.

2) Rules are formed using each of these features for different target emotional states. A feature may yield one or more rules. Generally these rules have the form: if feature F1 has value less than or greater than T1 and feature F1 has value less than or greater than T2 then conclude emotion = e1. For each rule, the cut- off points T1 and T2 for a given emotion class is taken to be the approximate average of the value of that class with its immediate neighbor emotional class.

3) *Corresponding to each rule, we associate CF values for each emotional class. These values of CFs are decided on the conditions mentioned in Table 1.*

TABLE I. DEFINING CERTAINTY FACTOR (CF) FOR RULES

Range of the CF	*CF Values*	*Belief and Disbelief*	*Indicated by*
Greater than 0.2 and up to 0.4	0.3	High evidence	High Inter class distance
Greater than 0.1 and up to 0.2	0.2	Moderate evidence	Medium Inter class distance
Equal to 0.1	0.1	Low evidence	Low Inter class distance

This heuristic has been arrived at based on empirical studies of the various feature graphs and behavior of the CF calculus. There may be multiple rules associated with each feature. Multiple rules, when they fire simultaneously (based on values of different features) may saturate the values of CF associated with them. To minimize this possibility, we have chosen relatively lower range of CF values. Given our observation that most features do not provide a high degree of discrimination for any of the emotions, a high value did not appear justified for any individual feature. The chosen range also allows the CF value to climb steadily to a high range, when there are many features supporting an emotion. The rules may point to a specific emotional state or a set of emotional states. If the distance of an emotion with its neighboring emotion is found to be less than 5– 6% of the entire spread (overall range) for that features value, then these emotions are grouped as a subset. Allocation of the values of CF to these classes is done based on the following rules, derived based on analysis of the emotion profile.

High Interclass Distance: If the interclass distance of an emotional class (either singleton or non-singleton) with its neighbors (left side and right side) is more than 15% of the entire spread for that feature, then the chances of a confusion with the neighboring class is low and hence the CF value associated with this class for that feature is 0.3.

Medium Interclass Distance: If the interclass distance of a emotional class (either singleton or non-singleton) with its neighbors (left side and right side) is in between 6-15% of the entire spread for that feature, then the CF value associated with this class is 0.2.

Low Interclass Distance: If the interclass distance of a emotional class (either singleton or non-singleton) with its neighbors (left side and right side) is less than 6% of the entire spread for that feature, then the CF value associated with this class is 0.1.

The exercise is done for the modalities like facial expression and speech. The next section will discuss in detail one of the case scenario for facial expression beginning with databases to the rules formulation.

IV. CASE STUDY FOR FACIAL EXPRESSION

We illustrate the process with a concrete example of emotion recognition from facial expressions. We used standard database, Cohn-Kanade (CK) [18] of the static images, where individuals are constrained to look straight at the camera and they are photographed with single colored background and illumination conditions do not vary drastically. Therefore, preprocessing issues are not a concern here. We utilize 184 images from 57 subjects. We have 32 female and 25 male subjects for the emotional states of neutral, anger, happy, fear, sad and disgust.

A. Feature Extraction

We have used the geometric features for emotion recognition and defined the model as a point-based model. The frontal view face model [19] is composed of many elements like mouth, nose, eyes and brows that could be used for analysis (Fig. 1 and Table 2). We used a set of 18 points in the frontal view image and using these points we defined a total of 21 features (f3, f4, f5, f6., f7, f8, f9, f10, f11, f12, f13, f14, f15, f16, f17, f19, f20, f21, f22, f23, f24 as shown in Fig. 1, mostly in the form of inter-point distances. For example, the feature f3 is the distance between left eye outer corner, A to left eyebrow outer corner, E.

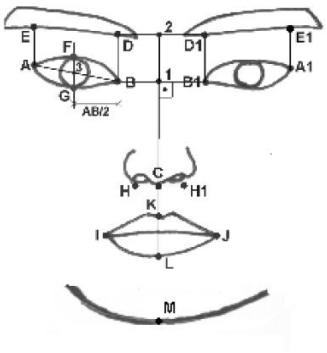


Fig.1. Facial points of frontal-view [19]

Similarly feature f4 (symmetrical to f3) is the distance between right eye outer corners, A1 to right eyebrow outer corner, E1. Each of these points has been extracted from the image. All the distances were computed. For example, mouth width is the distance between the tips of the lip corner. Similarly lip thickness, distance between left eye to left eyebrow, etc., were computed.

The distances are compiled into an output file (.xls) that is used for further analysis. All these distances were obtained for different emotions including the neutral state for all subjects. Facial expressions are often characterized by variation of a feature from its value in the neutral state, rather than its absolute value in a given state. Therefore, we standardize these features w.r.t their neutral value. These parameters were normalized in the following manner:

$$\textit{Normalized Value} = (\textit{Measured Value} - \textit{Neutral State Value}) / \textit{Neutral State Value} \quad (6)$$

B. Feature Analysis

As discussed earlier all features might not be useful in forming the rules. Individually each of these has to be analyzed. For example, the feature, lip distance (horizontal distance- f16 and vertical distance- f17) could be seen as varying with emotions (Fig. 2 and Fig. 3).

TABLE II. FEATURES OF THE FACIAL POINTS OF THE FRONTAL VIEW [19]

Features	Feature Description
f3	Distance AE
f4	Distance A1E1
f5	Distance 3F, 3 is the centre of AB (See Fig. 1)
f6	Distance 4F1, 4 is the centre of A1B1 (See Fig. 1)
f7	Distance 3G
f8	Distance 4G1
f9	Distance FG
f10	Distance F1G1
f11	Distance CK, C is 0.5HH1 (f0)
f12	Distance IB
f13	Distance JB1
f14	Distance CI
f15	Distance CJ
f16	Distance IJ
f17	Distance KL
f19	Image intensity in circle (r(0.5BB1), C(2)) above line (D, D1)
f20	Image intensity in circle (r(0.5BB1), C(2)) below line (D, D1)
f21	Image intensity in circle (r(0.5AB), C(A)) left from line (A, E)
f22	Image intensity in circle (r(0.5A1B1), C(A1)) right from line (A1, E1)
f23	Image intensity in the left half of the circle (r(0.5BB1), C(I))
f24	Image intensity in the right half of the circle (r(0.5BB1), C(J))
Total	21 Features

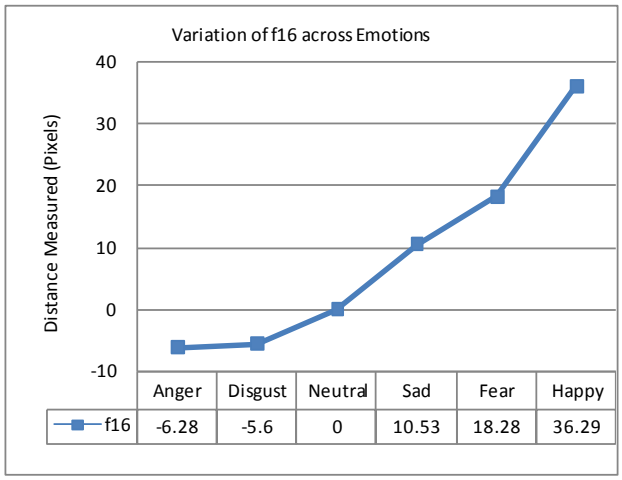


Fig.2. Variation of feature f16 (horizontal lip distance) across emotions

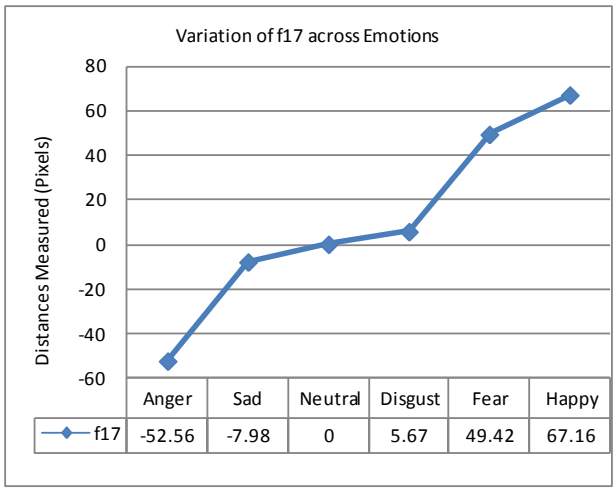


Fig.3. Variation of feature f17 (vertical lip distance) across emotions

We did these analyses using the individual features (f3 to f24) to see how each of these is varying across emotion. We found that eleven features (i.e. f3, f4, f9, f10, f11, f12, f13, f14, f15, f16, and f17) show more significance variation across the considered emotional states among all twenty one features. Also it is found that all the symmetrical pairs of features (like left eye vertical distance, f9 and right eye vertical distance, f10) do not always contribute to the same level to distinguish between the same set of emotions. The lip movement (horizontal lip distance, f16 and vertical lip distance, f17) provides good separation between 'happy', 'sad', 'fear' and 'neutral' emotions, but doesn't differentiate between 'anger' and 'disgust'. To validate these separations between emotional states, rules structure will be formed.

C. Formulation of Rules

As discussed earlier all features might not be useful in forming the rules. Individually each of these has to be analyzed. For example, the feature, lip distance (horizontal distance- f16 and vertical distance- f17) could be seen as varying with emotions (Fig. 2 and Fig. 3). From the trend of feature f16 (Fig. 2), it is seen that the emotions 'neutral', 'sad', 'fear' and 'happy' are distinguishable individually, whereas the emotions, 'disgust' and 'anger' are found to be close together (as the distances with its neighbour are found to be in the range of 5- 6% of the entire spread). Depending on the interclass distances of these classes CFs have been allocated (Table 1) and rules have been formed. For each rule (of the type if – then), the cut off point (i.e., upper limit and lower limit) belonging to the emotion class is taken to be the average of the value of that class with its immediate emotional class. For example, for 'sad' emotion the cut off points to be considered are 5 and 14, forming the singleton class and due to high inter class distances the CF values is to be considered as 0.3 (see Table 3). Similarly, the feature f17 also varies across emotions (Fig. 3). It is observed that 'neutral' along with 'disgust' is forming a non-singleton class while rest of the emotions is acting as singleton classes. Depending on distances between these classes, CFs has been allocated and rules have been formed. We found a total of five conditions each for the feature f16 and feature f17 to classify emotions. Examples of rules (Rule 1 and Rule 2) are shown below.

Example Rule 1: *Using dist_horizontal_lip (F16) for emotion identification*

(i) if (dist_horizontal_lip <= -3)
CFDis=0.2; CFAng=0.2;
(ii) if ((dist_horizontal_lip > -3) && (dist_horizontal_lip <= 5))
CFNeu=0.3;
(iii) if ((dist_horizontal_lip > 5) && (dist_horizontal_lip <= 14))
CFSad=0.3;
(iv) if ((dist_horizontal_lip > 14) && (dist_horizontal_lip <= 27))
CFFear=0.3;
(v) if (dist_horizontal_lip > 27)
CFHap=0.3;

Example Rule 2: *Using dist_vertical_lip (F17) for emotion identification*

(i) if (dist_vertical_lip < -30)
CFAng=0.3;
(ii) if ((dist_vertical_lip < -3) && (dist_vertical_lip >= -30))
CFSad=0.2;
(iii) if ((dist_vertical_lip < 27) && (dist_vertical_lip > -3))
CFNeu=0.3; CFDis=0.3;
(iv) if ((dist_vertical_lip >= 27) && (dist_vertical_lip < 58))
CFFear=0.3;
(v) if (dist_vertical_lip >= 58)
CFHap=0.3;

Such kind of exercise is done for each of the selected features extracted from the face. These features are f3, f4, f9, f10, f11, f12, f13, f14, f15, f16 and f17. Symmetrical pair of features like (f3, f4), (f9, f10), (f12, f13) and (f14 and f15) do not vary in the same way across different emotions and hence the resulting rules may differ. In the formulation of rules, we considered each of these features individually. Total of 11 rules have been formed for emotion identification using facial static images.

D. Recognizing Emotions from Facial Expressions using Rules

These rules have been tested on the database (CK database of facial expression) and final value of CF has been computed corresponding to each of the 6 emotional states. The emotion with the highest value of final CF is considered and counted against the expected emotion class for each image for all the subjects. For example, Table 3 shows the computed values of CF labelled as CF_SAD, CF_NEU, CF_ANG, CF_HAPPY, CF_FEAR and CF_DISGUST corresponding to all the six emotions - sad (S), neutral (N), anger (A), happy (H), fear (F) and disgust (D).

TABLE III. EXAMPLES OF COMPUTED VALUES OF CF USING RULES FROM FACE FOR FEMALE SUBJECT

Updated Value of CF computed using rules for respective emotion							
Subjects	**Actual Emotion**	**CF_ Sad**	**CF_ Neu**	**CF_ Ang**	**CF_ Happy**	**CF_ Fear**	**CF_ Disgust**
s1	S	**0.83**	0.56	0.30	0.36	0.72	0.00
s1	N	0.00	**0.97**	0.00	0.00	0.30	0.00
s1	A	0.10	0.37	**0.91**	0.20	0.50	0.51
s1	H	0.30	0.30	0.30	**0.82**	0.36	0.30
s1	F	0.78	0.37	0.00	0.20	**0.84**	0.00
s1	D	0.00	0.00	0.85	0.51	0.20	**0.87**

A row in this table indicates an input image of an individual subject in a particular emotional state (subjects labelled as 1). Each subject has been tested across emotions. Final outcome for the same is indicated in these CF values under the six columns labelled from CF_SAD to CF_DISGUST. For example, row 3 corresponds to subject-1 in 'angry' state; the table shows the maximum value of CF under the emotion class of 'anger' (0.91) showing correct identification. Similarly, the maximum value of CF for the subject-1 (row 6) is 0.87 and is for the target emotion of disgust. Though the value belonging to 'anger' is coming close to this value, we are considering the highest value of CF to identify the target emotion associated with the input image. Hence, the computed emotion matches with the 'predicted emotion' which is 'disgust' in this case and 'anger' in the previous case. Similarly computed value of CF has been analyzed for each of the emotions. The overall correctness of recognizing emotions using rule based approach in a unimodal system from facial expression is found to be 86.43% (i.e., out of 184 images, 159 images are correctly recognized). The recognition rates are found to be 80% (93 images correctly recognized out of 112) and 88.89% (64 images correctly recognized out of 72) for female and male subjects respectively.

V. WORKING WITH OUR OWN MULTIMODAL DATA

As seen from literature humans recognize emotion by fusing information from multiple sources: speech signals, facial expressions, gesture, bio-signals and others. Inadequacy of unimodal recognition systems provides the basis to go for multimodal recognition. We extend emotion recognition for multimodal data based on our rule based model. This model is based on preparing a set of rules derived from the individual modalities. The rules are mixed together independent of the modality into a single group. In order to test, we propose to include the data source as facial expressions with speech. The database used in the experiments consists of audio samples and static frontal images of different people (graduate students in the age group of 21 to 28 years). Total of 11 subjects participated in our experiment (5 female and 6 male). Each of these subjects was told to read a single sentence under four emotional states (anger (A), happy (H), sad (S) and neutral (N)). For the process of inducing the desired emotional state, individual subjects were shown a small video clipping of 2-3 minute corresponding to each of the four emotional categories. During this, facial expression was captured by the digital camera. The subjects chosen in our experiment don't wear 'glasses' and males don't have 'beard' on their face – this made the analysis easier. We have total of 20 images with utterances (5 each of 'anger', 'happy', 'sad', and 'neutral') of female and 23 images (6 each of 'anger', 'sad' and 'neutral' but 5 is of 'happy') with utterance of male subjects. The compiled set of rules for speech and facial expression was run against this dataset. We now discuss the results obtained, and compare with the performance of the same when using facial expression and speech alone.

a) Results using Facial Expression: Unimodal Approach

The average emotion recognition rate of the system using our own database is found to be 65% (for female subjects), 65.21% (for male subjects) and 67.44% overall. The emotion 'sad' is the best recognized and has 82% recognition rate overall. But this is not true with male subjects. 'Anger' is hard to distinguish from others and hence having the least accuracy.

b) Results using Speech: Unimodal Approach

The average emotion recognition rate of the system was found to be 55% (for female subjects), 62.5% (for male subjects) and 56.6% overall. It has been observed that the emotion 'happy' is hard to recognize both in female as well as in male subjects. The emotion 'sad' shows reasonably good recognition rate for male as well as female subjects.

c) Results when combining Facial Expression and Speech: Bimodal Approach

The average emotion recognition rate of the bimodal emotion recognition system (adding the two sets of rules together) using rules is found to be 75% (for female subjects), 65.21% (for male subjects) and 67.44% (overall). It has been observed that overall performance has increased by combining the inputs from speech signal and facial image in case of gender independent as well as gender dependent scenario.

VI. CONCLUSION

We presented a rule based approach for multimodal emotion recognition, which provides an elegant method for the design of multimodal recognition of emotion. We have formulated a multimodal recognition framework built around if-then rules using certainty factors to capture uncertainty of individual features. Multimodal emotion recognition performs better than unimodal emotion recognition system. Emotions such as 'anger' and 'sad', which was hard to recognize with facial expression yields better result when combined with speech modality. To the best of our knowledge, this approach has not been tried in the literature. This technique appears to be simple and effective for this problem. There are a number of avenues for extending this work. A more realistic evaluation with large data and more modalities is, perhaps, the most important. At present, we have used the Confirmation theory as used in MYCIN approach [12]. One of the major concerns against the use of certainty factor is that they have no sound theoretical basis; though, they often work well in practice. We allocated the values of CF to the emotional classes based on heuristic rules as defined in section III. These have been derived based on the analysis of the individual features across different emotions. In this work, we have ignored the possibility of having more than one emotional state at a time. We also would like to investigate alternative uncertainty models like the Dempster-Shafer Theory. Demspter Shafer theory provides more flexibility in assigning belief to various subsets of emotions.

The databases used for the expression analysis are all based on subjects who "performed" a series of different expressions. There is a significant difference between expressions of a spontaneous and of a deliberate nature. Without a database of spontaneous expressions, the expression analysis system cannot be robust enough. This database issue is common for all the modalities. The multimodal data fusion for emotion recognition remains an open challenge as several problems still persist, related to finding optimal features, integration and recognition. Completely automated multimodal emotion recognition system is still at the preliminary phase, shows very limited performance and is mostly restricted to the lab environment.

REFERENCES

[1] L.S. Chen, T.S. Huang, T. Miyasato, and R. Nakatsu "Multimodal Emotion/Expression Recognition", in Proceedings of the 3rd International Conference on Face and Gesture Recognition, pp.366-371, 1998.

[2] De Silva and Ng, "Bimodal Emotion Recognition", Automatic Face and Gesture Recognition, in IEEE International Conference, pp. 332 – 335, 2000.

[3] N. Sebe, I. Cohen, T. Gevers, and T.S. Huang, "Emotion Recognition Based On Joint Visual and Audio Cues", Pattern Recognition, International Conference on, vol. 1, pp. 1136–1139, 2006.

[4] Z. Zeng, Jilin Tu, Liu, Huang, Pianfetti, Roth and Levinson, "Audio-Visual Affect Recognition", IEEE Transactions on multimedia, 9 (2), pp. 424-428, 2007.

[5] B.V. Dasarathy, "Sensor Fusion Potential Exploitation Innovative Architectures and Illustrative Approaches," in Proceeding of IEEE vol. 85, pp. 24–38, 1997.

[6] C. Busso, Z. Deng, and S. Yildirim, "Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information", in

Proceedings of ACM 6th International Conference on Multimodal Interfaces, pp. 205-211, 2004.

[7] Corradini, A., Mehta, M., Bernsen, N. and J.-C. Martin. Multimodal input fusion in human-computer interaction on the example of the on-going nice project. In Proceedings of the NATO-ASI conference on Data Fusion for Situation Monitoring, Incident Detection, Alert and Response Management, Yerevan (Armenia), August 2003

[8] H. Liao. "Multimodal Fusion", Master's thesis, University of Cambridge, July 2002.

[9] S. Kettebekov and R. Sharma, "Understanding Gestures In Multimodal Human Computer Interaction", International Journal on Artificial Intelligence Tools, 9(2), pp. 205-223, 2000.

[10] R. Sharma, V. Pavlovic, and T. Huang. "Toward Multimodal Human Computer Interface", In Proceedings of the IEEE, 86(5), pp. 853-860, 1998.

[11] M. Sasikumar, S. Ramani, S.M. Raman, K.S.R. Anjaneyulu, and R. Chandrasekar, "Rule Based Expert Systems – A Practical Introduction", Narosa Publishers, 2007.

[12] E.H. Shortliffe. and B.G. Buchanan, "A Model of Inexact Reasoning in Medicine", Mathematical Biosciences, vol. 23, pp. 351-379, 1975.

[13] J. Gordon and E.H. Shortliffe, "The Dempster-Shafer Theory of Evidence", in [Buchanan and Shortliffe, 1984] pp. 272-292, 1984.

[14] C.V. Negoita, "Expert Systems and Fuzzy Systems", Benjamin/Cummings, 1985.

[15] J.A. Doyle, "Truth Maintenance System", Artificial Intelligence, vol. 12, pp. 231-272, 1979.

[16] R. Reiter, "A Logic for Default Reasoning", Artificial Intelligence, vol. 13, pp. 81-132, 1980.

[17] D. McDermott and J. Doyle, "Non-monotonic Logic I", Artificial Intelligence, vol. 13, pp. 41-72, 1980.

[18] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis", in Proceedings of the International Conference on Automatic Face and Gesture Recognition, pp. 46-53, 2000.

[19] M. Pantic and L.J.M. Rothkrantz, "Automatic Analysis Of Facial Expressions: The State Of The Art", IEEE Trans. Pattern Analysis and Machine Intelligence, 28(12), pp. 2037-2041, 2000.

[20] Preeti Khanna and M. Sasikumar, "Recognizing Emotions from Keyboard Stroke Pattern", International Journal of Computer Applications, 11(9): December, 2010.

[21] Preeti Khanna and M. Sasikumar,"Recognizing Emotions from Human Speech", Think Quest 2010, International Conference on "Contours of Computing Technology in association with Springer Publications, March 2010.

[22] Preeti Khanna and M. Sasikumar, "Application of Vector Quantization in Emotion Recognition from Human Speech", Springer Series in Communications in Computer and Information Science (CCIS), ICISTM – pp. 118-125, 2011.

[23] A. Batliner, B. Schuller, D. Seppi, S. Steidl, L. Devillers, L. Vidrascu, T. Vogt, V. Aharonson, and N. Amir, "The Automatic recognition of emotions in speech," in Emotion-Oriented Systems, Pt. 2, Springer-Verlag, pp. 71–99, 2011.

[24] S. V. Ioannou, A. T. Raouzaiou, V. A. Tzouvaras, T. P. Mailis, K. C.Karpouzis, and S. D. Kollias, "Emotion recognition through facial expression analysis based on a neurofuzzy network," in Neural Networks, vol. 18 (4), Elsevier, pp. 423–435, 2005.

The SNCD as a Metrics for Image Quality Assessment

Avid Roman-Gonzalez
TELECOM ParisTech, Departement TSI
Paris, France

***Abstract*—In our era, when we have a lot of instrument to capture digital images and they go more in more increasing the image resolution; the quality of the images become very important for different application, and the development tool to quality assessment is a current issue. In this paper, we propose to use the Symmetric Normalized Compression Distance (*SNCD*) as a metrics for the measurement of image quality, especially when we analyze residual errors. We also show performance comparisons of other metrics that we can found in the various research literatures and the *SNCD*. We also present an analysis about the performance of the *SNCD* depending to the type of distortion.**

Keywords—Quality metrics; NCD; SNCD; Kolmogorov complexity; image quality.

I. INTRODUCTION

The quality assessment of images is an issue very important since different automatic tools for signal processing were developed. The results given by these automatic tools can be affected if the image quality is not enough good. Thus, it is necessary an image analysis as is explained in [1].

In the literature and in the related works, we can see that many metrics have been developed within the full-reference approach to allow comparison and thus an assessment of the quality between an image and its reference. Some quality metrics to assess images using the full-reference approach have also been evaluated in [2], [3] and [4].

Each metrics evaluated in [2], [3] and [4] works better or worse in cases for specific distortions. One of the best known metrics is the *PSNR* (peak signal to noise ratio) even if some results may appear to be inconsistent. For instance, if an equal amount of additive noise is added to different sections of an image, we obtain different image quality results based on a visual assessment as shown in Figure 1. Here, however, both images have the same *PSNR = 35.29*.

The visual effect of the distortions depends of the section where the artifacts have been placed. For this reason, it is necessary to look for another quality metric more correlated to the human subjective evaluation. In this work, we propose to use the *SNCD* that will be explained in follow.

An application of Kolmogorov Complexity is to estimate the shared information between two objects given by their Normalized Information Distance (*NID*) [5]. The *NID* is proportional to the length of the shortest program that can calculate *x* given *y*.

Fig.1. Two images with same PSNR. The same amount of noise has been added to rectangular areas at the top (left) and at the bottom (right) of this image.

The normalized information distance is calculated as follows:

$$NID(x,y) = \frac{K(x,y) - \min\{K(x), K(y)\}}{\max\{K(x), K(y)\}} \qquad (1)$$

where *K(x)* is the Kolmogorov complexity of *x*, *K(y)* is the Kolmogorov complexity of *y*, *x* and *y* are two strings to be compared, and *K(x,y)* is the joint Kolmogorov complexity of *x* and *y*.

The *NID* result is a positive value *r* in the range of $0 \le r \le 1$, with $r = 0$ if the objects are identical, and $r = 1$ stands for the maximum distance between them. However, the *NID* is not computable and therefore we need a computable approximation. A well-known approach is the Normalized Compression Distance *NCD* defined by [5] and by [6] considering *K(x)* as the compressed version of *x*, and taking it as a lower limit of what can be achieved with the compressor *C*. Thus, the Normalized Compression Distance (*NCD*) can be defined as shown in the following equation:

$$NCD(x,y) = \frac{C(x,y) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}} \quad (2)$$

Where *C(x, y)* represents the size of compressed file obtained by the concatenation of *x* and *y*. We use this equation to estimate the *NID*.

The *NCD* can be calculated easily between two strings or two files *x* and *y*, and it shows how different these files are. We can use the *NCD* for various applications with different classes of data as a parameter-free approach [7], [8], [9] and [10]. The *NCD* can also be used to classify the data by unsupervised methods [6].

We analyzed the Normalized Compression Distance *(NCD)* that should be an approximation of the Normalized Information Distance *(NID)* in more detail [5]. The *NID* is a symmetric measure as the Kolmogorov Complexity $K(x, y) = K(y, x)$. However, we observed experimentally that the *NCD* is not symmetrical, $NCD(x, y) \neq NCD(y, x)$. Therefore, we use a Symmetric Normalized Compression Distance (*SNCD*) defined as the arithmetic mean of *NCD(x, y)* and *NCD(y, x)*. The *SNCD* is given by:

$$SNCD(x,y) = \frac{1}{2}\left[NCD(x,y) + NCD(y,x)\right] \quad (3)$$

$$SNCD(x,y) = \frac{C(x,y) + C(y,x) - 2 \times \min\{C(x), C(y)\}}{2 \times \max\{C(x), C(y)\}} \quad (4)$$

The structure of this paper is as follows: In Section II we present the image quality metrics that we use to compare with the *SNCD*. Section III presents the correlation coefficients that we use for the comparison step. In Section IV is shown a description of the database used for this work. Section V presents our results and analysis. Finally in Section VI we present the conclusions and discussion.

II. Metrics For Image Quality Assessment

Multimedia images are always subject to a variety of distortions and modifications during the process of compression, transmission, reproduction, etc.

It is important to measure and identify the quality and quality degradation in the data in order to have control and a chance to improve the quality of the images.

To evaluate the quality of images, some methods use measures of comparison against a reference. In that sense, we have three approaches [11]:

The "full-reference" (FR) approach

The full-reference method requires full access to the original image as a reference. It is based on the following philosophy:

Distorted Signal = Reference Signal + Error Signal

We assume that the reference signal has a perfect quality, and we quantify the error of visual perception.

The "non-reference" (NR) approach

The non-reference approach does not require any access to the original image, but the quality assessment without reference is a very difficult task. Several researchers have done some work for the evaluation of specific distortions.

The "reduced-reference" (RR) approach

The reduced-reference approach does not require full access to the original image but needs some partial information as references such as a set of extracted features.

The related research develops methods and algorithms that can automatically assess the quality of an image. [11] present a concept for quality-aware images. They use features extracted from the original image; the feature extraction is based on wavelet coefficients. [12] propose how to quantify lost image information and explore a relationship between image information and image quality. The authors of [13] investigated whether observers used structural cues to direct their fixation as they searched for simple embedded geometric targets at very low signal-to-noise ratios; the authors demonstrated that even in case of very noisy displays, observers do not search randomly, but in many cases they deploy their fixation to stimulus regions that resemble some aspect of the target in their local image features. [2] show an evaluation of different recent full reference image quality assessment methods, where they performed a subjective evaluation.

For comparison, in the present work, from the many existing metrics in the literature with a full reference approach, we use the *PSNR* and *SSIM* metrics that are also used and evaluated in [2] and [14].

The *PSNR* (Peak Signal-to-Noise Ratio) is given by:

$$PSNR = 10 \log_{10}\left(\frac{L^2}{MSE}\right) \quad (5)$$

Where *MSE* is the Mean Squared Error and *L* is the maximum dynamic range; for gray-scale images with 8 bits/pixel $L = 255$.

Another metrics is the *SSIM* (Structural Similarity Index) that has three independent components: luminance, contrast, and structure. The *SSIM* is given by:

$$SSIM = f(l(x,y), c(x,y), s(s,y)) \quad (6)$$

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (7)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (8)$$

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (9)$$

Where μ_x, σ_x and σ_{xy} represent the global mean, the standard deviation, and the cross-correlation. C_1, C_2 and C_3 are selectable constants.

III. Comparison Of Metrics

In order to compare the different metrics and the *SNCD*, we use three correlation coefficients. These correlation coefficients are calculated from the results obtained by a subjective evaluation of images of the database and the results obtained by the metrics. This subjective assessment was performed by a group of experts who evaluated the degree of distortion of each image in the database.

The correlation measures we will use are:

- *The Pearson correlation coefficient (PCC)* is an index that measures the linear relationship between two quantitative random variables. Unlike the covariance, Pearson correlation is independent of the scale of the measured variables. To calculate the PCC, we use the following MATLAB instruction: *corr(MOS, RG, 'type', 'Pearson')*, where *MOS* is the result for the subjective evaluation, and *RG* is the result using the image quality metrics.
- *The Spearman correlation coefficient (SCC)* is a measure of correlation (association or interdependence) between two continuous random variables. To calculate it, the data is sorted and replaced by their ordered indices. We used the following MATLAB instruction: *corr(MOS, RG, 'type', 'Spearman')*, where *MOS* is the result for the subjective evaluation, and *RG* is the result using the image quality metrics.
- *The Kendall correlation coefficient (KCC)* is another non-parametric correlation measure. To calculate de *KCC*, we used the following MATLAB: *corr(MOS, RG, 'type', 'Kendall')*, where *MOS* is the result for the subjective evaluation, and *RG* is the result using the image quality metrics.

IV. Database Descripction

To perform *SNCD* metrics experiments and to make appropriate comparisons, we use a database that has already been used by other researchers and is available on the Internet. The database that we use is the Cornell-A57 collection [15], consisting of three original images (*baby*, *harbor*, and *horse)* as shown in Figure 2 and which also includes distorted images. For each original image, we have six types of distortion:

- Quantization of the *LH* (*L* = Low and *H* = High) sub-bands of a 5-level discrete wavelet transform, where the sub-bands were quantized via uniform scalar quantization (*FLT*)
- Additive white Gaussian noise (*NOZ*)
- JPEG BaseLine compression (*JPG*)
- JPEG2000 compression without visual frequency weighting (*JP2*)
- JPEG2000 compression with the dynamic contrast-based quantization algorithm (*DCQ*)
- Blurring by a Gaussian filter (*BLR*)

For each type of distortion, we have 3 intensities; thus we have a database of 54 images (3 images × 6 distortion types × 3 distortion parameters).

Fig.2. Original images of Cornell-A57 database.

Each image has a size of 512×512 pixels; we can see that the *baby* picture and the *horse* picture contain a predominant object that we will use to analyze the behavior of our selected compression methods together with the existing metrics.

V. Analysis Of Results

To evaluate the performance of the *SNCD* as a quality metrics, we made various experiment and also we analyzed the error maps. The errors *E* between the original image and the distorted image, are the absolute difference values between the original image *X* and the distorted image *Y*, *E = abs (X - Y)*. In order to validate the error map importance, we calculate for:

- The *SNCD* comparing the original image *X* and the distorted image *Y*.
- The *SNCD* for the original image *X*, and the error map *E*.
- The *SNCD* for the error map *E* and the distorted image *Y*.

For the first tests, we calculated the quality measures of the images of the entire database, and compared them with the subjective evaluation using correlation coefficients explained above. The subjective evaluation was obtained from seven imaging experts by using a continuous rating system; greater values represent a greater distortion. The results are shown in Table 1.

TABLE I. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH THE ENTIRE DATABASE OF 54 IMAGES.

COMPLETE DATABASE			
	PCC	***SCC***	***KCC***
PSNR	0.6347	0.6189	0.4309
SSIM	0.7528	0.8066	0.6058
SNCD XY jpeg	0.0967	0.1501	0.1287
SNCD XE jpeg	0.2943	0.1860	0.1217
SNCD EY jpeg	0.1245	0.1273	0.1063
SNCD XY zip	0.0929	0.0448	0.0518
SNCD XE zip	0.295	0.0278	0.0154
SNCD EY zip	0.0196	0.0789	0.0686

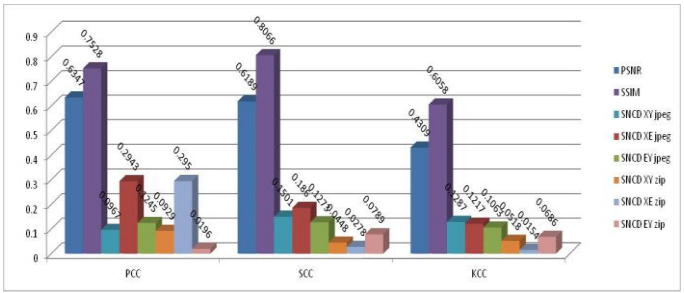


Fig.3. Summary results of Table 1.

We can see that the best results are obtained by the classical metrics; we obtain a Pearson correlation of 0.7528 using *SSIM* metrics, a Spearman correlation of 0.8066, and a Kendall correlation of 0.6058. The values obtained by the *SNCD* are really very low, indicating that it is not a good representation of the subjective assessment of quality; we obtained for the *SNCD* between the *X* image and the *E* map the following values: a Pearson correlation of 0.2943, a Spearman correlation of 0.1860, and a Kendall correlation of 0.1217 using a JPEG lossless compressor.

Another experiment we conducted was to sub-divide the database for each given parent image since, as mentioned above, the database contains two parent images with a predominant structure, and another parent image that does not have a predominant structure; then we could see how they behave with respect to the metrics. The results are shown in Tables 2 to 4.

TABLE II. RESULTS OF THE COMPARISON OF THE DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS FOR THE 18 *BABY* IMAGES.

BABY			
	PCC	***SCC***	***KCC***
PSNR	0.7786	0.6925	0.5686
SSIM	0.7559	0.7152	0.5556
SNCD XY jpeg	0.1983	0.3024	0.1895
SNCD XE jpeg	0.3102	0.2239	0.1111
SNCD EY jpeg	0.1503	0.5501	0.4510
SNCD XY zip	0.4109	0.4613	0.3595
SNCD XE zip	0.0975	0.0072	0.0458
SNCD EY zip	0.2105	0.3664	0.2680

TABLE III. RESULTS OF THE COMPARISON OF THE DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS FOR 18 *HARBOUR* IMAGES.

HARBOUR			
	PCC	***SCC***	***KCC***
PSNR	0.6137	0.7438	0.5461
SSIM	0.7375	0.8182	0.6382
SNCD XY jpeg	0.2312	0.0723	0.0066
SNCD XE jpeg	0.3629	0.2965	0.1645
SNCD EY jpeg	0.4642	0.1560	0.0724
SNCD XY zip	0.2773	0.0465	0.1118
SNCD XE zip	0.1945	0.0031	0.0461
SNCD EY zip	0.3626	0.1829	0.0855

TABLE IV. RESULTS OF THE COMPARISON OF THE DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS FOR 18 *HORSE* IMAGES.

HORSE			
	PCC	***SCC***	***KCC***
PSNR	0.7968	0.6863	0.4771
SSIM	0.7779	0.7936	0.5948
SNCD XY jpeg	0.3282	0.3230	0.2941
SNCD XE jpeg	0.2917	0.0423	0.0196
SNCD EY jpeg	0.0595	0.1538	0.0980
SNCD XY zip	0.3099	0.1950	0.1373
SNCD XE zip	0.0829	0.1373	0.1111
SNCD EY zip	0.1605	0.1889	0.1111

When we sub-divide the database into smaller databases for each parent image, we see that the traditional metrics for image quality still show a better performance (see Figs. 6.22 to 6.24). We obtain a Pearson correlation of 0.7786 for *baby* when using *PSNR*, a Spearman correlation of 0.7152 using *SSIM*, and a Kendall correlation of 0.5686 using *PSNR*. For the *harbour* image we obtain a Pearson correlation of 0.7375 when using *SSIM*, a Spearman correlation of 0.8182 using *SSIM*, and a Kendall correlation of 0.6382 using *SSIM*. Finally, for the *horse* image we obtain a Pearson correlation of 0.7968 when using *PSNR*, a Pearson correlation of 0.7936 using *SSIM*, and a Kendall correlation of 0.5948 using *SSIM*. We also see that the performance of the *SNCD* has improved somewhat, although is still not comparable with the classical metrics, but it has improved somewhat compared with the experiment of the complete database. For the *baby* image, we obtain a Pearson correlation of 0.4109 when using *SNCD XY*, a Spearman correlation of 0.5501 using *SNCD EY*, a Kendall correlation of 0.4510 using *SNCD EY*. For the *harbour* image, we obtain a Pearson correlation of 0.4642 when using *SNCD EY*, a Spearman correlation of 0.2965 using *SNCD XE*, and a Kendall correlation of 0.1645 using *SNCD XE*. Finally, for the *horse* image we obtain a Pearson correlation of 0.3282 when using *SNCD XY*, a Spearman correlation of 0.3230 using *SNCD XY*, and a Kendall correlation of 0.2941 using *SNCD XY*.

We could imagine that *SNCD* can improve the comparison performance for images with predominant structure, but experience shows that it is not.

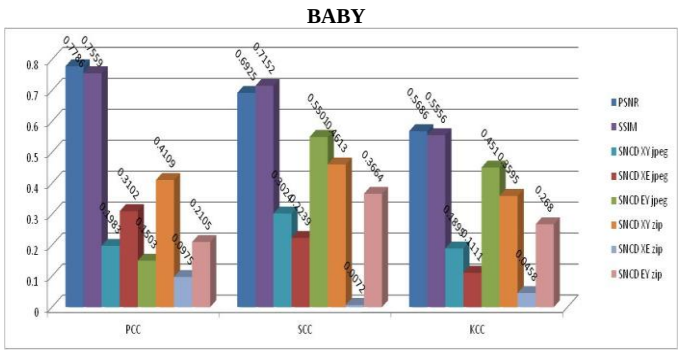


Fig.4. Summary results of Table 2.

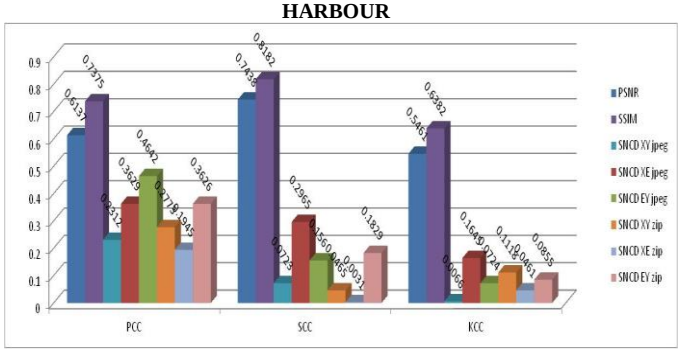


Fig.5. Summary results of Table 3.

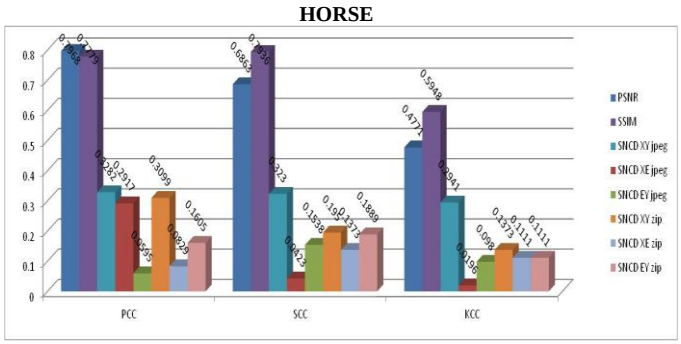


Fig.6. Summary results of Table 4.

Therefore, the next experiment to perform is to sub-divide the database according to the type of distortion. In this case, we have 6 types of distortion with 9 images for each one. The results are shown in Tables 5 to 10.

TABLE V. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH IMAGES DISTORTED BY BLURRING (9 IMAGES).

BLR DISTORTION			
	PCC	***SCC***	***KCC***
PSNR	0.5904	0.4667	0.3889
SSIM	0.9421	0.8000	0.6667
SNCD XY jpeg	0.8243	0.5167	0.3889
SNCD XE jpeg	0.7199	0.5000	0.3889
SNCD EY jpeg	0.5738	0.3833	0.2778
SNCD XY zip	0.3872	0.2833	0.2778
SNCD XE zip	0.5924	0.4333	0.2778
SNCD EY zip	0.6477	0.5333	0.4444

TABLE VI. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH IMAGES DISTORTED BY JPEG2000 + DCQ (9 IMAGES).

DCQ DISTORTION			
	PCC	***SCC***	***KCC***
PSNR	0.5637	0.5000	0.3889
SSIM	0.9369	0.9667	0.8889
SNCD XY jpeg	0.9472	0.8833	0.7778
SNCD XE jpeg	0.4522	0.3833	0.2222
SNCD EY jpeg	0.9115	0.8500	0.7222
SNCD XY zip	0.5940	0.7333	0.5556
SNCD XE zip	0.2456	0.2333	0.1667
SNCD EY zip	0.9051	0.9667	0.8889

TABLE VII. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH IMAGES DISTORTED BY A FLT ALLOCATION (9 IMAGES).

FLT DISTORTION			
	PCC	***SCC***	***KCC***
PSNR	0.9100	0.9000	0.7222
SSIM	0.8982	0.8500	0.6667
SNCD XY jpeg	0.4327	0.3333	0.2222
SNCD XE jpeg	0.9533	0.9167	0.7778
SNCD EY jpeg	0.9432	0.9000	0.8333
SNCD XY zip	0.4342	0.2667	0.1667
SNCD XE zip	0.9519	0.9500	0.8333
SNCD EY zip	0.9803	0.9667	0.8889

TABLE VIII. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH IMAGES DISTORTED BY JPEG2000 COMPRESSION (9 IMAGES).

JP2 DISTORTION			
	PCC	***SCC***	***KCC***
PSNR	0.7957	0.8000	0.6667
SSIM	0.8641	0.8167	0.6667
SNCD XY jpeg	0.6422	0.6833	0.5000
SNCD XE jpeg	0.7495	0.7000	0.5000
SNCD EY jpeg	0.7470	0.7000	0.5556
SNCD XY zip	0.1047	0.1500	0.1667
SNCD XE zip	0.6645	0.6333	0.5000
SNCD EY zip	0.6742	0.7167	0.5556

TABLE IX. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH IMAGES DISTORTED BY JPEG COMPRESSION (9 IMAGES).

JPG DISTORTION			
	PCC	***SCC***	***KCC***
PSNR	0.7008	0.6333	0.5000
SSIM	0.9178	0.9333	0.7778
SNCD XY jpeg	0.6659	0.7167	0.6111
SNCD XE jpeg	0.1015	0.4167	0.1667
SNCD EY jpeg	0.6852	0.7333	0.5556
SNCD XY zip	0.7225	0.7833	0.6111
SNCD XE zip	0.0300	0.0667	0.0556
SNCD EY zip	0.0163	0.0833	0.1111

TABLE X. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH IMAGES DISTORTED BY GAUSSIAN NOISE (9 IMAGES).

NOZ DISTORTION			
	PCC	***SCC***	***KCC***
PSNR	0.9340	0.9500	0.8333
SSIM	0.8834	0.9500	0.8333
SNCD XY jpeg	0.3986	0.2500	0.2222
SNCD XE jpeg	0.3254	0.2833	0.2222
SNCD EY jpeg	0.4414	0.5000	0.3889

SNCD XY zip	0.5715	0.3333	0.3333
SNCD XE zip	0.8552	0.8333	0.7222
SNCD EY zip	0.9194	0.9167	0.8333

BLR DISTORTION

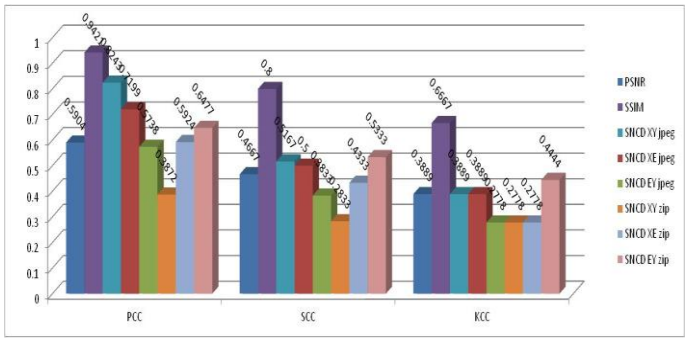


Fig.7. Summary results of Table 5.

DCQ DISTORTION

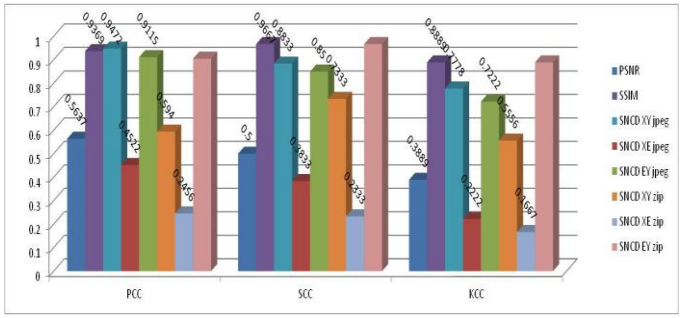


Fig.8. Summary results of Table 6.

FLT DISTORTION

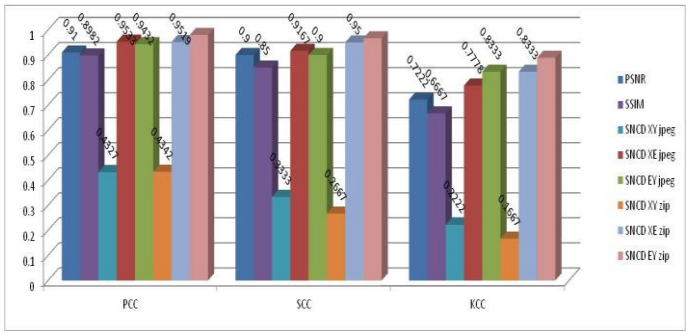


Fig.9. Summary results of Table 7.

JP2 DISTORTION

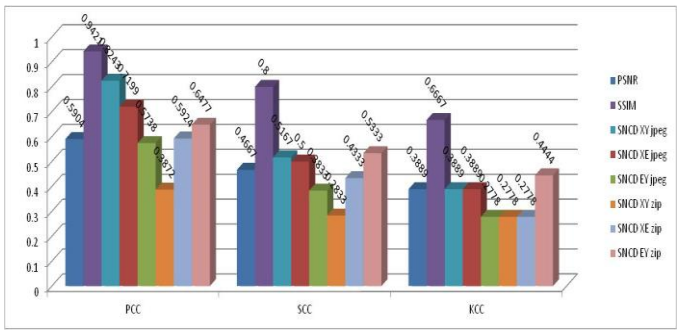


Fig.10. Summary results of Table 8.

JPG DISTORTION

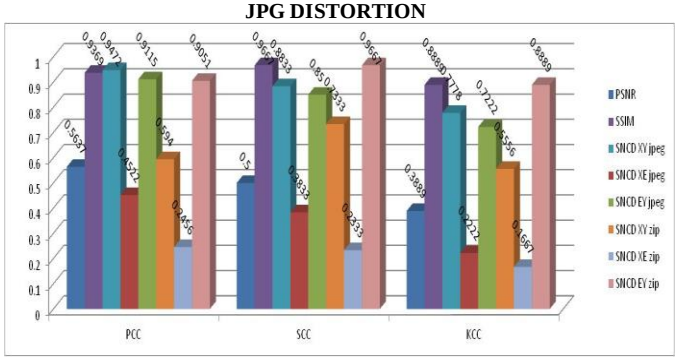


Fig.11. Summary results of Table 9.

NOZ DISTORTION

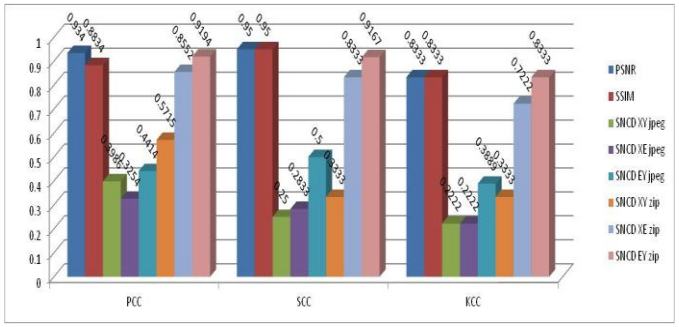


Fig.12. Summary results of Table 10.

- The results of this experiment grouped by the type of distortion are very interesting. We have encouraging results for the *SNCD*. The performance of the *SNCD* has improved considerably in all cases. It outperforms the traditional metrics *SSIM* and *PSNR* for the *DCQ* case and for the filtering case; however, for the remaining distortion cases, the obtained values are quite comparable to the classical metrics (see Figs. 7 to 12):
- For *BLR* distortion, we obtain of a Pearson correlation of 0.8243 when using *SNCD XY*, a Spearman correlation of 0.5167 using *SNCD XY*, and a Kendall correlation of 0.4444 using *SNCD EY*.
- For *DCQ* distortion we obtain a Pearson correlation of 0.9472 when using *SNCD XY*, a Spearman correlation of 0.9667 using *SNCD EY*, and a Kendall correlation of 0.8889 using *SNCD EY*.
- For *FLT* distortion we obtain a Pearson correlation of 0.9803 when using *SNCD EY*, a Spearman correlation of 0.9667 using *SNCD EY*, and a Kendall correlation of 0.8889 using *SNCD EY*.
- For JP2 distortion we obtain a Pearson correlation of 0.7495 when using *SNCD XE*, a Spearman correlation of 0.7167 using *SNCD EY*, and a Kendall correlation of 0.5556 using *SNCD EY*;
- For *JPG* distortion we obtain a Pearson correlation of 0.7225 when using *SNCD XY*, a Spearman correlation of 0.7833 using *SNCD XY*, and a Kendall correlation of 0.6111 using *SNCD XY*.

- For *NOZ* distortion we obtain a Pearson correlation of 0.9194 when using *SNCD EY*, a Spearman correlation of 0.9167 using *SNCD EY*, and a Kendall correlation of 0.8333 using *SNCD EY*.

For all distortions cases, the performance of *SNCD* deteriorates as the method is based on data compression, and therefore, cannot identify the compression distortions, but still shows very comparable values.

In the experiments where we sub-divided the database by type of distortion we have good results for *SNCD*. Why we do not have the same results when we work with the database sub-divided per parent image, or when working with the entire database? A reason may be that the *SNCD* method properly evaluates the distortion or quality of the images, but does not consider the magnitude of the type of distortion for the entire database. This means that for the subjective assessment, some kind of distortion is more influential than another. In contrast, during *SNCD* computation, the sequence of distortion types can be rearranged; however, the *SNCD* determines with good approximation the intensity of the type of distortion. This holds for all results shown in the different tables.

Fig.13. Distorted images of the Cornell-A57 database with the same or about the same *MSE*.

Another experiment is to have distorted images with the same or about the same mean squared error *MSE*. For this experiment, we take the original image of Figure 2 and create distorted images. We calculate the measure of quality of the images of the new database (9 distorted images for the original image shown in Figure 2; the distortions are: JPEG compression, JPEG2000 compression and Noise; all distortions with about the same *MSE* values between 3200 and 3400 and *PSNR* values between 22 and 24; these distorted images are shown in Figure 13) and compared them using the correlation coefficients explained above. The results are shown in Table 11.

TABLE XI. RESULTS OF THE COMPARISON OF DIFFERENT METRICS TO EVALUATE IMAGE QUALITY USING THE CORRELATION COEFFICIENTS WITH THE ENTIRE DATABASE OF 9 IMAGES.

COMPLETE DATABASE			
	PCC	***SCC***	***KCC***
PSNR	0.7722	0.9160	0.8003
SSIM	0.0576	0.0672	0.0572
SNCD XY jpeg	0.2219	0.1092	0.1143
SNCD XE jpeg	0.7333	0.6555	0.5145
SNCD EY jpeg	0.0536	0.4034	0.1715
SNCD XY zip	0.5946	0.6471	0.5145
SNCD XE zip	0.7496	0.7311	0.5717
SNCD EY zip	0.0345	0.2185	0.1143

VI. CONCLUSIONS AND DISCUSSION

The results obtained by the classical metrics, are better than the SNCD when we analyze the complete database.

When we sub-divide the database into smaller databases for each parent image, the traditional metrics still show a better performance, it was in contrast to our idea (we imagined that *SNCD* can improve the comparison performance for images with predominant structure).

The results for the experiment grouped by the type of distortion are very interesting. We have encouraging results for the *SNCD*. The *SNCD* outperforms the traditional metrics *SSIM* and *PSNR* for the *DCQ* case and for the filtering case; however, for the remaining distortion cases, the obtained values are quite comparable to the classical metrics. A reason for that may be that the *SNCD* method is based on compression techniques and cannot to evaluate distortions produced by compressors.

The *SNCD* method properly evaluates the distortion or quality of the images, but does not consider the magnitude of the type of distortion for the entire database. This means that for the subjective assessment, some kind of distortion is more influential than another. In contrast, during *SNCD* computation, the sequence of distortion types can be rearranged; however, the *SNCD* determines with good approximation the intensity of the type of distortion.

The *SNCD* as a metrics for assessing image quality is limited to a single type of distortion with different levels of intensity.

The researches in this topic must continue, finding a good metrics for image quality assessment responding to the visual evaluation is a important issue.

REFERENCES

[1] A. Roman-Gonzalez, “Digital Images Analysis”, Revista ECIPeru, vol. 9, N° 1, 2012, pp. 61-68.
[2] H. R. Sheikh, M. F. Sabir, A. C. Bovik, “A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms”, IEEE Transactions on Image Processing, vol. 15, N°. 11, 2006, pp. 3441-3451.
[3] I. Avcibas, B. Sankur, K. Sayood, “Statistical Evaluation of Image Quality”, Journal of Electronic Imaging, vol. 11, N° 2, 2002, pp. 206-223.
[4] I. Avcibas, “Image Quality Statistics and Their Use in Steganalysis and Compression”, PhD Thesis – *Bogazici University*, Istanbul/Istanbul Province, Turkey, 2001.
[5] M. Li and P. Vitányi, “The Similarity Metric”, IEEE Transaction on Information Theory, vol. 50, N° 12, 2004, pp. 3250-3264.

[6] R. Cilibrasi, P. M. B. Vitanyi; "Clustering by Compression", *IEEE Transaction on Information Theory*, vol. 51, N° 4, April 2005, pp 1523 - 1545.

[7] M. R. Quispe-Ayala, K. Asalde-Alvarez, A. Roman-Gonzalez, "Image Classification Using Data Compression Techniques"; 2010 IEEE 26th Convention of Electrical and Electronics Engineers in Israel – IEEEI 2010; Eilat – Israel; November 2010, pp. 349-353.

[8] E. Keogh, S. Lonardi, Ch. Ratanamahatana, "Towards Parameter-Free Data Mining", Department of Computer Science and Engineering, University of California, Riverside.

[9] B.J.L. Campana y E.J. Keogh, "A Compression Based Distance Measure for Texture", University of California, Riverside, EEUU 2010.

[10] A. Roman-Gonzalez, "Clasificación de Datos Basado en Compresión", Revista ECI¨Peru, vol. 9, N° 1, 2012, pp. 69-74.

[11] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. Yang, A. C. Bovik; "Quality-Aware Images"; *IEEE Transaction on Image Processing*.

[12] H. R. Sheikh, A. C. Bovik, "Image Information and Visual Quality", IEEE Transactions on Image Processing, vol. 15, N° 2, 2006, pp. 430-444.

[13] U. Rajashekar, A. C. Bovik, L. K. Cormack, "Visual Search in Noise: Revealing the Influence of Structural Cues by Gaze-contingent Classification Image Analysis", *Journal of Vision*, vol. 6, N° 4, 2006.

[14] Z. Wang, Q. Li, "Information Content Weighting for Perceptual Image Quality Assessment", IEEE Transactions on Image Processing, vol. 20, N° 5, 2011, pp. 1185-1198.

[15] E. P. Simoncelli, E. H. Adelson, "Nonseparable QMF Pyramids", Proc. SPIE, Visual Comm and Image Proc. IV, vol. 1199.3, 1989, pp. 1242-1246.

The Use of Software Project Management Tools in Saudi Arabia: An Exploratory Survey

Nouf AlMobarak, Rawan AlAbdulrahman, Shahad AlHarbi and Wea'am AlRashed
Software Engineering Department
King Saud University
Riyadh, Saudi Arabia

***Abstract*—This paper reports the results of an online survey study, which was conducted to investigate the use of software project management tools in Saudi Arabia. The survey provides insights of project management in the local context of Saudi Arabia from ten different companies which participated in this study. The aim is to explore and specify the project management tools used by software project management teams and their managers, to understand the supported features that might influence their selection. Moreover, the existence of the Arabic interface, the Hijri calendar and the Arabic documentation has been specially considered, due to the nature of the local context in dealing with the Hijri calendar and the prolific use of Arabic as the formal language in communication with clients in the public sector.**

Keywords—project management tools; survey; Hijri calendar; Arabic interface; software engineering; Arabic documentation.

I. INTRODUCTION

In recent years, the software industry has witnessed a proliferation in technology adoption within the software engineering process to deal with the increase in software complexity. In addition, the number of organizations is growing daily; which leads to many projects, tasks, risk analysis plans, scheduling and distinct resources that need to be monitored properly by the project managers. The necessitate for controlling those responsibilities accurately is to deliver a successful project; despite the fact of the impediments that the project managers or the team might face throughout the time when monitoring these simultaneously tasks (Pitch, Loch, & Meyer, 2002) [1].

Thus, a need for project management tools has emerged in which it can help the project's leaders in administering the business and the individuals with different assigned tasks. Moreover, it provides the chance to monitor the projects' progress and its flow in the organizations that runs many tasks across the region.

The Purpose of this paper is to emphasize the features that were essential to every project managed by different project managers in different organizations with distinct sector types. As well as, highlight the use of some special features in the local context that need to be supported by the applied project management tools.

This paper is structured into five sections : section 2 which represents a concise background of the study, section 3 illustrates the methodology of the conducted survey, its participants, materials and the way it was designed . Followed that , section 4 shows the results that were collected from the survey with its analysis. Section 5 represents the discussion that was made after observing and analyzing all the participants' responses for the most utilized tools in the local context along with any demand for the Arabic language in the tool's interface or documentation. And finally, section 6 wraps up and concludes all what has been presented in this paper.

II. BACKGROUND

Project's with many resources, phases and due dates can be hard and stressful, however, running multiple projects with different activities, timelines and a lot of deadlines can even be harder.

The project's leader might be unable to concentrate and follow-up each project's detail properly.

The planning and the design of the stages of each task would be proficient and well-organized when using a software technology tools which in the end can contribute in delivering a successful project (Attarzadeh & Ow, 2008) [2].

According to Alfaadel, Alawairdhi, and Al-Zyoud's study that aimed to investigate the most common reasons behind the success or failure of some IT projects. The results listed five reasons that lead to the project failure and one of those reasons was "poor planning and scheduling". The study was conducted in a form of a survey and was done by three hundred and eight project managers across Saudi Arabia (ALFAADEL, ALAWAIRDHI, & ALZYO, 2012)[3].

Furthermore, (Olateju, Abdul-azeez, & Alamutu, 2011) [4] stated in their study which was performed in Nigeria that the tools and the techniques of project management that were performed in the public institutions needs to be enhanced, due to its bad performance which causes the project's failure in many aspects.

Throughout the search that has been done on the project management tools it has been noticed that the market has presented a wide range of various tools. These tools provide a distinguished features and services that can fit with different industry's needs.

A comparative study was made by Mishra, A. and D. which presented the features of twenty popular project management tools side by side in a table to represent and compare their provided features (Mishra & Mishra, 2013) [5].

To the best of our knowledge, there was an absence of studies that showed and examined the use of project management tools in Saudi Arabia. So as a result, a survey in a form of a questionnaire was distributed among many project managers in Saudi Arabia who work in different organization (private sector / governmental sector). The main goal of this survey is to provide a good overview about the tools used by project managers in the local context and what the main preference features in their opinion that should be available in the tool. Finally, to gain information about the adoption of the Arabic language and Hijiri calendar format.

III. METHOD

A. Participants

The participants of this study were a total of sixteen project managers all were volunteers from different organizations across Saudi Arabia. Those participants represented many companies and fields. The respondents were four females and twelve males between the ages of twenty-five and forty with the majority of them in their thirties. In addition, Arabic was their native language, and their educational level varies between bachelor and master degree.

B. Materials

A web-survey software tool called " SurveyMonkey " was chosen to conduct and prepare this study, its main purpose is to help the individuals in creating and constructing a customized online survey that fit their needs. It also provides the ability to distribute it to the intended audience to reach the maximum possible number of participants. Then lastly, gather the data and analyze it.

Moreover, the survey was passed to the participants in several ways as some was sent to their email directly, while the others received it via the social network such as Twitter and Linkedin.

The participants were given one week period to complete the survey and provide their feedback. After that, the data were gathered and were prepared to be analyzed using the analyzing service that was provided by the "SurveyMonkey" web tool. As a final step, the last result which contains numbers and percentages were revised by two team members to guarantee the correctness of the outcome.

C. Design and procedure

The research design of this study was non–experimental, descriptive, and review-based as it focused on a research questionnaire that tries to identify, appraise, and select different high quality project management tools and if there are any relation that might be found and linked to the ones used in the local context from the project managers' point of view and their feedback that were provided earlier in the survey .

To be able to accomplish the previous goal, an online survey was distributed to several project managers whom were considered to be our intended audience.

The survey contains ten questions (see Appendix), it was organized in three parts; the first part requires the respondent to provide the name of his organization, this will let us decide the type of the organization whether it is a private or public sector.

The second part focused on the tool's data and its features, the reason for the inclusion of this part was to gain more information about the tools that was used by the participants, such as how many tools are used during managing the different aspects of the project, the name of the tool, the features that were important when trying to select a project management tool.

The third part primarily concentrated on the existence of Arabic language and its use in the organization as the questions mentioned the Arabic documentation usage, the support for the Arabic Interface in the provided tool and to what extent it is needed, then finally, the support of the Hijiri calendar and if it was used as the main date format when determining the deadlines for each and every project phase.

The survey's answers were a closed ended question (multiple choice) in which respondents were asked to select their answer or answers out of the choice list.

In some questions, if the answer wasn't included in the list, the respondent was provided with the ability to add his answer in a comment field.

The results were collected and analyzed automatically through the "SurveyMonkey's" analyzing feature. Afterwards, a group of two team members analyzed those outcome again manually using the calculator and excel sheet to ensure the results' accuracy.

The variable in this study was the use of project management tools by different governmental agencies in Saudi Arabia and some other local organizations in different sectors, such as medical, financial, and telecommunication sectors.

Since, the participants in the survey were project managers from different organizations, the most practical way to get their feedback in a relatively short period of time, is to send them the designed survey through their direct official email as well as contacting some of them personally through phone and social media such as Twitter and Linkedin.

So, as researchers, we can assert that the participants were the exact targeted and intended audience and the survey was taken seriously by them which led to the possibility of controlling the survey environment. Hence, the exclusion of some responses was not considered in this case (Zechmeister & Shaughnessy, 2011) [6].

The study was planned as follow, first, The project managers were contacted through the telephone, then the purpose of the study was briefly described to them at that time, they were asked about their native language, age and their level of education. After that, the respondents were requested to provide an email address to be able to send the survey's link to them.

For those who couldn't be reached by the telephone the answers for the previous questions were obtained from their personal account on the social networks (Twitter, Linkedin) where the survey's link was sent to them.

Afterwards a request letter within the sent email declared that the response will be needed for research purpose that is related to the use of different project management tools in their local organizations that they are working in. And, that the research was intended to recognize if there is any correlation between the features founded in different used tools and the nature of the organization's business that the project manager represent. It also has been noted that the information they will offer will be kept confidential and the collected responses will be used for analytical purpose only. Finally, the participants were then thanked for their cooperation and were provided with the hyperlink of the survey.

IV. RESULTS

The results presented below summarizes the main views and responses of different project managers. A total of 16 project managers responded to the survey, more than three-quarters, 81%, of participants were male project managers. These collected responses represent about 61.5% of the total project managers who were contacted.

The key information was gathered through the survey by determining the organization or governmental agency where the participated project manager stands for, so that the requirements for each and every organization can be easily understood. The Collected responses showed different organizations including, Ministry of Communication and Information Technology, local General Electric Healthcare, Zain, King Fahad Medical City, SADAD, Azian, EMC, International System Engineering, some telecommunication companies' vendors and local Nokia Siemens Networks. Responses from the governmental agency formed a small proportion of total responses about 12%.

The first question outlined that more than 90% respondents use at least one project management tool to help in planning, organizing, and managing project resources and tasks. Fig 1 shows the number of project management tools used by the participated project managers.

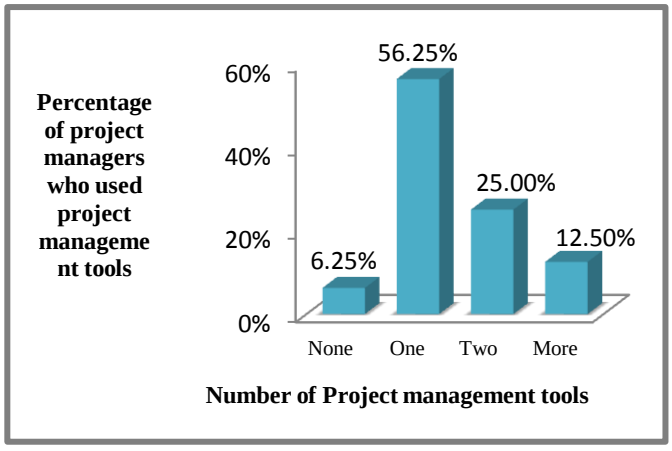


Fig.1. The number of project management tools used by the Project Managers

The result in Fig 2 summarizes the number of project managers who used some existing softwares.

The reported responses showed that these organizations depend on other tools rather than the ones that was mentioned as the main popular project management tools used in the most global organizations in their daily tasks. Additionally, the tools listed in the survey was evaluated as the most reliable tools used by international organizations in (Garmahis, 2009) [7] and (Gorbunova, 2011)[8]. As shown in Fig 2, Microsoft Project (Microsoft, 2013) [9] had the highest rank among the used tools by the participated project managers, more than ten project managers depend on this tool in managing their projects. Other project management software tools were used, such as special templates accessed through the organization implemented portal, Microsoft Project Server (Microsoft, Project Server 2013, 2013) [10] and Excel spreadsheets.

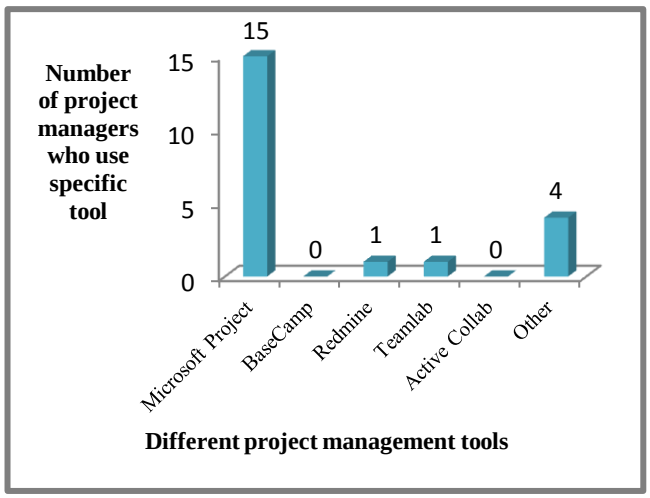


Fig.2. Project management tools used by different project managers in Saudi Arabia

Respondents were asked about their opinion on the most features that influenced their selection of any project management tool. To simplify this question and to limit our results; a list of some project management tool's features were picked and listed (Schwalbe, 2012) [11] as shown in the survey at Appendix A.

Most Project managers, 10 out of 16 were looking for the simplicity in choosing a specific tool, where the second most feature that was looked for was the ability to integrate the used project management tool with other unlike tools to combine project planning functions with added operational functions that does not exist in the integrated software tool, such as collaboration, document version control, and issue tracking activities (Duggan) [12].

While the third most important features evaluated equally by the participated project managers were time and resource estimations and tracking project progress. Moreover, less than half project managers, 6 out of 16, founded that estimating the cost feature considered as a significant feature to look for before selecting a tool. Some other features were selected by at most 5 project managers as important features, shown in Fig 3.

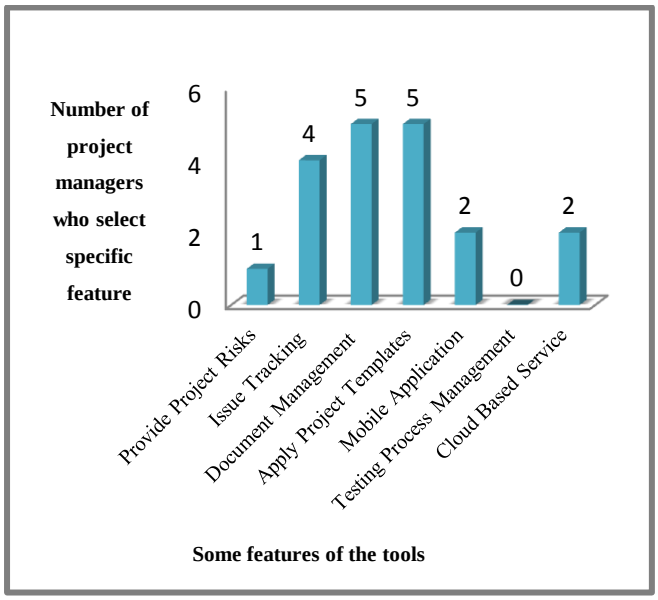


Fig.3. A sample of the tool's features selected by the Project Managers

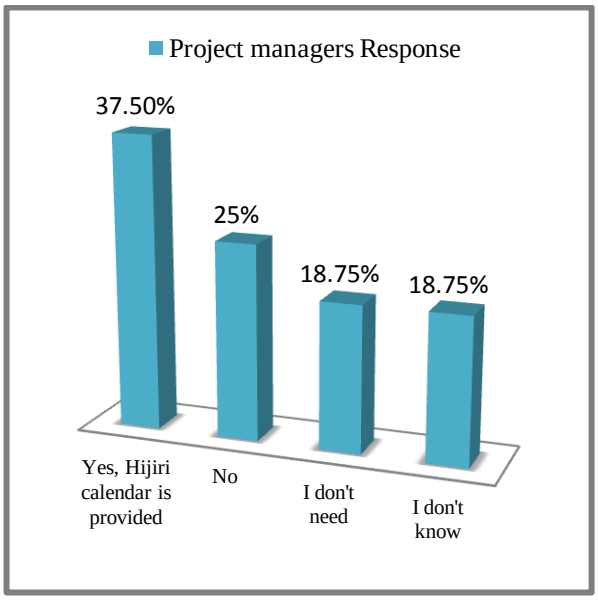


Fig.4. Support for Hijiri calendar in the tools used by the project managers.

A small percentage 18.75% of the used tools supported Arabic interface, which was unexpected in an Arabic environment. In addition, more than 80% stated that the Arabic interface wasn't supported or that they don't even know about its availability in their tool. In case of supporting this feature, the participants were asked if they were using this feature in their daily work. 100% of the project managers reported that they don't look for this kind of support for managing their daily project tasks as they didn't need to use the Arabic interface, even thought, it was supported by the used tool.

The participants have also been asked about the support of the Hijiri calendar. The results showed that the Hijiri calendar was supported by 37.50% of the tools that have been reported previously, whereas, 25% said that it wasn't supported. Moreover, 18.75% indicated that they don't need to use the Hijri calendar in their projects. At the same time, 18.75% responded that they don't know whether the tool support such feature or not. Fig 4 shows the project managers' responses about the support for a Hijiri calendar in their tool.

Furthermore, The project managers were asked if they need to use the Hijiri calendar in their project's due dates and deadlines in the case if this feature was supported. About 22% of the project managers do use the Hijiri calendar in their business while 78% don't, as shown in Fig 5.

As for the support of the Arabic documentation, 25% of the project managers reported that their tool provides the support for the Arabic documentation. While, 25% said that the Arabic documentation wasn't supported by the tool they use. Table 1 illustrates the support of the Arabic report.

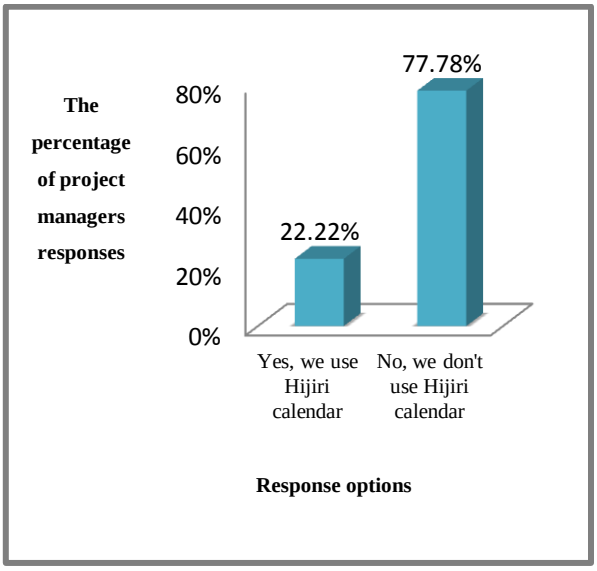


Fig.5. Use of the Hijiri calendar format in the organization's projects.

31.25% of the respondents didn't need the Arabic documentation in their organization and only 18.75% indicated that they don't know whether their tool support the Arabic report. If a need for an Arabic report has arisen and If the used tool doesn't support the Arabic documentation, participants were asked about the actions that need to be performed to overcome this issue. Twelve project managers have answered this question, eight said that they use manual translation whereas, one indicated that they used a translation tool while, the rest specified that they don't need Arabic documentation in their work.

TABLE I. SUPPORT FOR THE ARABIC DOCUMENTATION

Response options	Project management tools that support Arabic Documentation	
	Number of respondents	*Percentage*
Yes, the tool support Arabic documentation	4	25%
No, the tool doesn't support Arabic documentation	4	25%
I don't need	5	31.25%
I don't know	3	18.75%

V. DISCUSSION

Since the official language in Saudi Arabia is Arabic, it was expected to be used as the main written and spoken communication language in different private and governmental organizations.

However, according to our study, the results showed that despite having Arabic as the official language in the context of the study, project managers perceived that Arabic support of project management tools is not essential because software engineering phases and tasks are mainly conducted in English. Therefore, not supporting the Arabic Interface doesn't affect the manager's and/or the organization's decision when selecting a tool. Moreover, the Hijiri calendar wasn't the used format by most of the project managers in tracking the project's timeline.

Although, the official followed calendar format in the local context is Hijiri calendar. From the observations, it is noted that the official communication for projects is often conducted in Arabic; specifically with government and industry projects. It seems that the external documentation is often adapted from the internal documentation used by project managers in the local contexts. This highlights an opportunity for project management software tools to address this gap in supporting software engineering projects, particularly in the local context where documentation can be effectively generated and communicated in the format appropriate to all stakeholders and individuals involved.

During distributing the survey to gain the needed information from the largest possible audience, we have noticed that we gain a high response from managers which have been contacted in person. On the other hand, the response rate was lower for project managers who were contacted through the social network as twitter.

VI. CONCLUSION

A survey study was undertaken as part of looking for a dependency by some local organizations in different sectors and governmental agencies, on some of project management tools in administrating, managing and tracking several project phases.

The main purpose of the distributed survey was to find out, in case of using project management tools by the organizations, the key element and the players that lead to choose and use a precise project management tool. Also, to find and discover if there were any special requirements that need to be implemented in the tools that are used by the participant of an organization to cope with the local business needs.

We can conclude that, when analyzing the results it was found that the most dominant software tool used by project managers in Saudi Arabia was Microsoft project and a slight portion used Microsoft Excel Sheets to keep track of their everyday tasks. Additionally, the three top features that the project manager looks for in any tool were, in order, "Simplicity" then "Ability to integrate with other tools, as well as, provide time estimation feature" and "project progress & changes" along with "project resource estimation feature" which was rated evenly.

Moreover, The Arabic interface and the Arabic documentation weren't essential in the industry where the participants' work.

Finally, regardless of the fact that the Hijri calendar is the calendar format that is officially used in Saudi , the survey's result revealed the fact that it wasn't popular as it was expected to be and instead Georgian calendar format was followed.

ACKNOWLEDGMENT

This research paper would not have been possible without the interaction of many project managers, who consider the time to provide their responses in a relatively timely manner without any compensation. We would like also to extend our thanks and acknowledgment to Dr. Areej Alwabil for her dedicated effort in sharing with us her experience in project management, providing us with all the needed concepts, and allocating the time to guide us in writing this paper.

REFERENCES

[1] Pich, M. T., Loch, C. H., and De Meyer, A. (2002). " On uncertainty, ambiguity, and complexity in project management. Management Science ", vol. 48(8), pp.1008-1023.

[2] Attarzadeh, I., Ow, S. (2008). " Project Management Practices: The Criteria for Success or Failure ". Communications of the IBIMA, vol. 1, pp.234-241.

[3] Alfaadel, F., Alawairdhi, M., and Al-Zyoud, M. (2012, April). " Success and failure of IT projects: a study in Saudi Arabia ". In Proceedings of the 11th WSEAS international conference on Applied Computer and Applied Computational Science , pp.77-82.

[4] Olateju, O. I., Abdul-azeez, I. A., and Alamutu, S. A. (2011). " Project Management Practice In Nigerian Public Sector–An Empirical Study ". Australian Journal of Business and Management Research, vol. 1(8), pp. 01-07.

[5] Mishra, A. and Mishra, D., (2013). " Software project management tool : a brief comparative view ". ACM SIGSOFT Software Engineering Notes, vol. 38 (3), pp.01-04.

[6] Zechmeister,E. and Shaughnessy, J., (2011). " Research Methods in Psychology ". US: McGraw-Hill College, pp.161-175.

[7] Michael Garmahis. (2009). " Top 20 Project Management tools" . Available: http://garmahis.com/reviews/top-project-management-tools/. Last accessed 10th June 2013.

[8] Nina Gorbunova. (2011). " 10 Most Popular Online Project Management Solutions ". Available: http://www.famousbloggers.net/most-popular-online-project-management-solutions.html. Last accessed 10th June 2013.

[9] Microsoft. (2013). Microsoft Project. Available: http://office.microsoft.com/en-us/project/enterprise-project-portfolio-management-in-the-cloud-project-online-FX103802026.aspx. Last accessed 10th June 2013.

[10] Microsoft. (2013). Project Server 2013. Available: http://office.microsoft.com/en-us/project/project-portfolio-management-for-the-enterprise-project-server-FX103802061.aspx. Last accessed 10th June 2013.

[11] Schwalbe, K. (2012). "An Introduction to Project Management ". US: Schwalbe publishing, pp.11-30.

[12] Duggan, T. , " The Importance of Integrated Project Management Tools ". Available: http://www.ehow.com/info_8384297_importance-integrated-project-management-tools.html#ixzz2Us8hZB55. Last accessed 10th June 2013.

APPENDIX

PM Survey

1. The organization you work in

- ○ Zain
- ○ Mobily
- ○ STC
- ○ Sadad
- ○ National Guard Hospital
- ○ King Fahad Medical City
- ○ King Faisal Specialist Hospital
- ○ Ministry Of Health
- ○ GE Health Care

Other (please specify)

2. How many Project Management tools do you use ?

- ○ None
- ○ One
- ○ Two
- ○ more

(please specify)

3. What are the tools that you use ?

- ☐ Microsoft Project
- ☐ Freedcamp
- ☐ Dotproject.net
- ☐ BaseCamp
- ☐ Redmine
- ☐ Teamlab
- ☐ Clarizen
- ☐ Huddle
- ☐ Active Collab
- ☐ Assembla
- ☐ Other (please specify)

4. Which of the following MOST influenced your choice in Question 2 ?

- ☐ Simplicity
- ☐ Project resources estimation feature.
- ☐ Cost estimation feature.
- ☐ Ability to integrate with other tools.
- ☐ Time estimation feature.
- ☐ Provide project progress & changes.
- ☐ Provide project risks.
- ☐ Other (please specify)
- ☐ Issue tracking feature.
- ☐ Document Management feature.
- ☐ Apply project template from previous projects.
- ☐ Availability of Mobile application.
- ☐ Testing Process Management.
- ☐ Cloud Based service.

5. Does your tool support Arabic Interface ?

- ○ Yes
- ○ No
- ○ I don't need.
- ○ I don't know

6. If Yes , Do you use Arabic interface ?

- ○ Yes
- ○ No

7. Does it support Hijiri calendar?

- ○ Yes
- ○ No
- ○ I don't need
- ○ I don't know

8. If yes, do you use the Hijri calendar in your projects?

- ○ Yes
- ○ No

9. Does your tool support Arabic documentation ?

- ○ Yes
- ○ No
- ○ I don't need
- ○ I don't know

10. If No, What do you do in case you need Arabic reports ?

- ☐ Manually translation
- ☐ Using a translation tool.

Other (please specify)

Done

Communication in Veil: Enhanced Paradigm for ASCII Text Files

Khan Farhan Rafat
Dept. of Computer Science
International Islamic University
Islamabad, Pakistan

Muhammad Sher
Dept. of Computer Science
International Islamic University
Islamabad, Pakistan

***Abstract*—Digitization has a persuasive impact on information and communication technology (ICT) field which can be realized from the fact that today one seldom think to stand in long awaiting queue just to deposit utility bills, buy movie ticket, or dispatch private letters via post office etc. as these and other such similar activities are now preferably being done electronically over internet which has shattered the geographical boundaries and has tied the people across the world into a single logical unit called global village. The efficacy and precision with which electronic transactions are made is commendable and is one of the reasons why more and more people are switching over to e-commerce for their official and personal usage. Via social networking sites one can interact with family and friends at any time of his/her choice. The darker side of this comforting aspect, however, is that the contents sent on/off-line may be monitored for active or passive intervention by the antagonistic forces for their illicit motives ranging from but not only limited to password, ID and social security number theft to impersonation, compromising personal information, blackmailing etc. This necessitated the need to hide data or information of some significance in an oblivious manner in order to detract the enemy as regards its detection.**

This paper aims at evolving an avant-garde information hiding scheme for ASCII text files - a research area regarded as the most difficult in contrast to audio, video or image file formats for the said purpose.

Keywords— Embedded Secrets; Hide and Seek; Eccentric way of writing; ASCII Text Steganography; Communication in Veil; Stealth Communication

I. INTRODUCTION

The ease with which digital contents can be allegedly copied and distributed over the internet is the driving force behind exploration of new information hiding techniques for content's protection, verification, and its legitimate distribution. However, research that was primarily focused on protecting digital rights has inadvertently given way to steganography having Greek origin [1] which with the introduction of computers, has evolved into a science for in veil communication.

The name steganography first appeared in a manuscript entitled Steganographia written by Trithemus (1462-1516) and is a composition of two Greek words στεγανό-ς (Steganos) and γραφ-ειν (Graphos) which in English means covered writing [2]. The difference between cryptography and steganography is that the later hides the existence of secret data [3] whereas former concerns itself in making that data unintelligible [4].

A. Techniques for Steganography

Steganography can be realized through any of the following three ways [5]:

- **Insertion:** It involves direct embedding of secret information inside the body of Cover which results in an increased stego object file size. *One such old technique involves writing of data past end-of-file (EOF) mark that remains in veil to naive computer user.*
- **Substitution:** Information is hidden inside the cover by substituting secret information with cover's contents. The Stego Object may or may not retain actual Cover file size. *For example if 01001011 be the pixel elementthen a secret message bit 0 inserted at Least Significant Bit (LSB) position results in Stego byte as 01001010, where the change is so inconsequential that it easily deceives human eye.*
- **Cover Generation:** In this technique a Cover is generated based on the secret information that need covertly communication. *Spam Mimic is a freely available program that hides secret text message by generating random but meaningful text phrases.*

B. How Steganography Works

Steganography exploits limitations in Human's Auditory-Visual System (HAVS) that are briefly summarized as follows:

- **Inadequacy of Vision:** "Highest Resolution Perceivable Pixels: 28 Seconds Of Arc" as reported in [6] where subsequent discussion explicates on variety of physical factors as major hindrance for premier spatial frequencies being misconstrued by the human eye while eye's edifice and biological tests confer on an all-out professed frequency of around one cycle per arc minute (half arc-minute pixels) thereby failing to differentiate colors (where RGB Color ranges from 0(zero) to 2^24-1 in numbers).
- **Inexactitude of Auditory System:** The way head is contoured together with external ear (Pinna) enacts ethereal variations on entrant sounds according to its angle of inception in ear and can be apprehended as a filtering process. Human beings can hear voices having

frequency bounds 20 Hz to 20 KHz; both inclusive, however, age plays a significant role on the hearing threshold [7]. Further loud sounds tend to dominate modest ones and hence data may be superimposed on extremely low pitched noise that remain oblivious to human in presence of loud sounds.

C. Paper Plan

Rest of the paper is intended as follows: Section II briefly deliberates on ASCII text file format and the challenge it offers to researchers in devising text cover based steganographic schemes followed by related research in that area. Evaluation parameters for our proposed scheme are discussed in Section III while our choice of steganographic model is given in Section IV. Section V actually deliberates on our proposal. Quantified Test results and allied illustrations are given in Section VI. Section VII explains conjectural aspect of our proposed scheme while Section VIII explains its advantages and limitations. Future work follows in Section IX whereas Section X concludes the discussion.

II. Related Research

A. ASCII Text File Format

American Standard Code for Information Interchange abbreviated as ASCII, is a 7-bit code that facilitates text communication between different devices. One of the salient attributes distinguishing ASCII file format from other file formats like image, audio and video is that the former lacks auxiliary space that later exploits in projecting bounded contents. Another trait of Text files is that these are saved and presented for view in the manner the human beings are familiar with. The draw back with ASCII character codes is that changing its single bit results in code that may render a character or word as erroneous/misspelled thereby drawing immediate attention of the viewer. This, however, is also what makes ASCII text centric steganography a challenging task for researchers and hence is the prima facie of this research.

B. Literature Review

Steganography by virtue of being seamless in idiosyncrasy has emerged as a preferred choice for information hiding. Following discussion categorically expand on how text steganography is being used for covert communication.

1) English Language Specific Steganography: Following is the discussion on Steganographic schemes that exploits syntax and semantics of English language:

a) Acronym: Acronyms are contractions for comparatively long or frequently used words/phrases like As soon as possible which is abbreviated as ASAP. Author [8] suggested using acronyms together with corresponding words / phrases of English language to hide bits of secret information. The methodology works by arranging words/phrases in one of the two column table, the other column of which is populated with corresponding acronyms. The column containing words/phrases are labeled as "0" while acronyms are headed by label "1". Table 1 indicates one such arrangement.

TABLE I. Acronyms and Words/Phrases Arranged in a Two Column Table

ACRONYMS	WORDS
2L8	Too Late
ASAP	As Soon As Possible
C	See
CM	Call Me
F2F	Face to Face

Next, text cover composed of words/phrases and acronym is prepared. Secret information to be hidden inside the body of text cover is translated into bits. Text cover is then iterated to search for words/phrase or acronym matching those in the table till end of the message. Each time a word/phrase or acronym run into secret message bit (in sequence), it is examined with reference to column head (label) of the table. If secret message bit is 0 and the corresponding matching text in cover is word/phrase i.e., column labeled as 0, the cover text remains unchanged. However, if the secret message bit is 1 and the corresponding matching text in cover is word/phrase, the word/phrase in cover text is replaced by its corresponding acronym. In short, binary message bit 0 corresponds to having word/phrase in the Stego Object while binary message bit 1 corresponds to having acronym in place of words/phrases.

*b) Synonym:*Authorin [9]applied the aforesaid methodology on English language words that share same meaning/sense for the purpose of information hiding. Table 2 shows an arrangement of words having same sense/meaning and arranged in a two column table.

TABLE II. Different Words Sharing Same Meaning

WORDS	SYNONYMS
Big	Large
Chilly	Cool
Small	Little
Smart	Clever
Spaced	Stretched

c) Steganography through words that are spelled differently in British and American language: In [10] authorsextended aforementioned methodology on words that are spelled differently in British and American English - for information hiding purpose. Table 3 shows list of some of the words that are spelled differently in British and American language.

TABLE III. List of Word(s) Spelled Differently in British and American English

AMERICAN ENGLISH	BRITISH ENGLISH
Center	Centre
Criticize	Critise
Favorite	Favourite
Fulfill	Fulfil

*2) Miscellaneous Schemes for Text Steganography:*Following is a brief discussion on miscellaneous text-based steganographic schemes:

*a) Manipulation of Text Contents:*Authors in [11] suggested a number of eccentric proposals for data hiding in English Language text through modifications like making syntax errors, replacing words/phrases with their acronyms, artifact word format etc. as shown below:

- Inducing typographical errors – writing "there" in place of "there"
- Opting for "yr" in place of "your" and "TC" rather than "Take Care"
- Inserting extra carriage returns or segregating text into uneven paragraphs, or altering line or word space.
- Using annotating text e.g., :) that denotes 'pun'
- Use of bilingual text – "we always commit the same mistakes again, and 'je ne regrette rien'!".

*3) Use of Blank/Space Character:*In [12]authors proposed scheme that represents secret binary bit 1 with a single space while secret binary bit 0 depicts a double space. Following example illustrates the concept:

Example: Let 11010001 be the bits of our secret message bits and let "*A quick brown fox jumps over the lazy dog.*" be the cover text. Going in parallel with said scheme binary bit 0 will represent double spaces while no additional space is inserted for binary message bit 1. Resulting stego object after bit embedding will take the form*"A quick brownfoxjumpsoverthelazydog.", where black spaces denote a double space.*

*4) Steganography via Word Mapping Method:*In their attempt to increase capacity of data embedding by inserting 'spaces' based on word length of the cover authors [13] employed additional file form keeping locations of words targeted in hiding secret message bits. Table 4 shows substitution criteria for their proposed scheme.

5) Both Stego Object and index file are needed to retrieve hidden information at receiver's end. As comprehended from the said table, even and odd word lengths are compared with secret message bit pairs for inserting single or double space after identifying the word and location file updated accordingly.

TABLE IV. EVEN AND ODD WORD LENGTH

SECRET BIT PAIR	WORD SIZE	NUMBER OF BLANK SPACES AFTER
00	EVEN	TWO
01	EVEN	ONE
10	ODD	TWO
11	ODD	ONE

III. EVALUATION PARAMETERS

Perceptibility is the foremost requirement of any steganography system that does not involve sound. However, as regards its security nothing can precisely be said except to harden Wendy's efforts towards cracking or breaking a system by opting for information theoretically secure solutions consequent from the following deliberation:

Using Cachin [14] formal definition of security of steganographic system as:

$$D(P_c||P_s) \leq \varepsilon \qquad (1)$$

, one may infer that *perfect security* may be achieved whenever $\varepsilon = 0$, where P_c is the probability distribution of cover, P_s is that of stego object respectively over set of alphabets **A**,and $D(P_c||P_s)$ is computed using following:

$$D(P_c||P_s) = \sum_{a\epsilon A} P_c\,(a) log_2(\frac{Pc(a)}{Ps(a)}) \qquad (2)$$

Authors [15], however, differed on the aforesaid notion of perfect security by highlighting bound constraint concomitant with equation (2), and have elucidated with example that security is neither a measurable commodity nor can it be quantified.

IV. PREFERRED MODEL

Simmons [16] was first to propose a model based scenario for stealth communication where for Alice and Bob having agreed on a secret mode of communication before being sent to jail were kept in separate cells under supervision of Warden Wendy, must decide on their escape plan without raising suspicion. The two types of errors Warden Wendy may commit in such a scenario include:

- Type – I error:A hidden message gets detected by Wendy where in fact no message was sent.
- Type-II error:A secret message gets through Wendy as unnoticed.

Obviously we want our proposed scheme to maximize probability of occurrence of Type – II error.

Through literature review on model based steganography we found the one presented by [17] as close to what we have anticipated concerning evolution and implementation of our proposed solution. The same is illustrated in Figure 1 followed by its brief explanation.

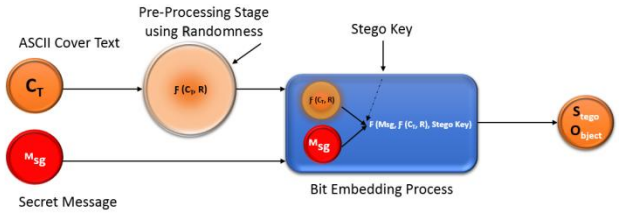


Fig. 1. Information Theoretically Secure Model for Steganography.

As apparent from Figure 1, the idea is to induce some sort of uncertainty in cover text (via some random process) and thereafter performing bit embedding over it, the theoretical substance of which is explicated in Section VII.

V. Proposed Steganographic Scheme

A. Objective

Of the three techniques used in steganography discussed earlier, our proposed scheme uses insertion for hiding secret message bits inside ASCII cover text file. However, to alleviate the overall effect of insertion on size of stego object we concentrated on increasing bit embedding capacity per byte in the cover text but without needing support of some additional file that must also be transferred along with stego object.

B. Preludes

To achieve aforesaid objective we extended our research on existing ASCII codes using Visual Basic 6 as our experimenting tool, and found eight such codes that can safely be inserted inside ASCII cover text file without raising perceptibility concern when the file gets opened with Microsoft Windows default application called 'Notepad'. For ease we shall refer to these ASCII codes as '**S.code**' in subsequent discussion. Further to retain insertion of additional bytes confined to reasonable range we considered total spaces in ASCII Text file as our threshold for byte insertion beyond which the scheme ceases to work. Another aspect of this selection was that change of words may not hamper the hidden bits where extra spaces not prefixed by any of the S.code(s), on account of malevolent activity, can easily be ignored at the time of bit extraction at receiving end.

Since $x = log_2^y \rightarrow y = 2^x$, hence $x = log_2^8 \rightarrow 8 = 2^x \; for \; x = 3$meaning thereby that each extra byte inserted in ASCII cover text using any of the S.code(s) can serve as a place holder for three bits of secret information.

The eight S.code(s) were first randomly shuffled and then assigned a tri-bit group (Range: 000 to 111) in sequence as shown in Table 5.

TABLE V. Default Allotment of Tri-bits to S.code(s)

Index	S.code	Tri-Bits	Re-Arranged
0	28	000	
1	29	001	
2	30	010	
3	129	011	
4	141	100	
5	143	101	
6	144	110	
7	157	111	

The tri-bits are later re-arranged on the basis of stego_key value where the starting point for tri-bit re-arrangement is obtained using following formula:

$$S_p \leftarrow (\sum_{i=0}^{31} Stego\_Key[i] * (i+1)) MOD \; 8 \qquad (3)$$

i.e. S_p holds the index value of the tri-bit groups starting from which the rest of the groups will be written in top down order against index 0 to 7 in 4th column of Table 5 for subsequent use. For example, let the starting point S_p be 6 then Table 5 will take the form shown in Table 6. As apparent S_p will vary from message to message.

TABLE VI. Key Dependent Allotment of Tri-bits to S.code(s)

Index	S.code	Tri-Bits	Re-Arranged Bits
0	143	000	110
1	29	001	111
2	157	010	000
3	129	011	001
4	141	100	010
5	28	101	011
6	144	110	100
7	30	111	101

C. Pre-Processing ASCII Cover Text

Shuffling of cover text content was not considered as feasible because being irrational it would create ambiguity, hence, we iterated through the cover text from start till end, searching for a 'space'. Upon finding it, a random number got generated using True Random Number Generator (TRNG) – the discussion of which is beyond the scope set for this research, in range 0 ~ 65537 and reducing it modulo 8. The outcome served as an index (Column 1st Table 6 refers) for selecting the specific S.code after which it was inserted before the said space.

D. Pre-Processing ASCII Cover Text

Since we have eight S.code(s) hence these can be arranged in 8! = 40320 unique ways (*Permutations*), an arrangement of which is shown in Table 7.

TABLE VII. 40320 Permutations for S.code(s)

R/C		28	29	30	129	141	143	144	157
1		141	144	29	143	30	129	157	28
2		30	143	129	144	28	29	157	141
3		29	141	143	157	144	28	30	129
4		144	157	129	141	143	30	29	28
5		143	30	157	28	141	144	129	29
6		129	144	28	29	157	143	141	30
...		...	...	...	...	...	...	...	...
40319		28	29	141	143	30	129	144	157
40320		141	28	30	129	29	157	143	144

E. Bit-Embedding Steps

1) Type / Select Secret Message and translate it into equivalent bits. Store message length and its corresponding file extension into four bytes (each) and translate those into equivalent binary bits. The 64-bits forms message header and is affixed before message bits.

2) Select 256-bit Stego key.

3) Process Stego Key as input through SHA-256 HASH algorithm [18].

4) Translate the outcome of Step 3 into equivalent binary bits and count the number of ON binary-bits. If the number equals or exceeds total number of binary message bits (inclusive of header bits) proceed to step 6.

5) Output of Step 3 (in place of initial Stego Key) serves as feedback to Step 3 followed by execution of Step 4.

6) Iterate cover text till end, searching for 'blank/space' by taking output of step 4 and processing one bit at a time. Mark the 'space' against ON HASH bit as secret message bit replacement position.

7) Compute a random row number for Table 7 using following equation:

$$\tau \leftarrow \left(\sum_{i=0}^{30} (Stego\ Key\ [i] * Stego\ Key\ [i + 1]) MOD\ 65537 \right) MOD\ 40320 + 1 \qquad (4)$$

8) Iterate cover text till end in search of 'space' marked for replacement (Step 6 refers), taking one tri-bit at a time. Locate S.code corresponding to those tri-bits in Table 6 that are to be treated as column head of Table 7. Replace S.code at cross section of row τ and column head of Table 7 with that affixed before the marked 'space'.

9) If no more tri-bits are to be processed, move to Step 12.

10) Increment row τ by 1. If it exceeds 40320, then set it to initial value of 1.

11) Repeat Step 8.

12) Terminate bit embedding process.

F. Bit-Extraction Steps

1) Select 256-bit Stego key.

2) Process Stego Key as input through SHA-256 HASH algorithm.

3) Translate the outcome of Step 2 into equivalent binary bits and count the number of ON binary-bits. If the number equals or exceeds 22 proceed to step 6.

4) Output from Step 3 (in place of initial Stego Key) serves as feedback to Step 2 followed by execution of Step 3.

5) Iterate stego object, searching for 'blank/space' by taking output bits, one bit at a time, of step 4. Mark the 'space' against ON HASH bit as pointer to hidden message bit.

6) Compute a random row number for Table 7 using equation (4).

7) Iterate stego 'space' characters marked as 'pointers' (Step 6 refers). Locate S.code affixed before marked 'space' in row τ of Table 7 and note down its corresponding column head. Search the column head thus for corresponding tri-bit group in 3rd column of Table 6 followed by their extraction and concatenation.

8) If first 22 bits gets processed, move to Step 11.

9) Increment row τ by 1. If it exceeds 40320, then set it to initial value of 1.

10) Of the 66 hidden bits thus obtained, the first 32 when converted into bytes gives hidden message length while the next its file extension.

11) Repeat Step 8 based on computed hidden message length vide Step 11.

12) Translate the extracted bits into bytes and save it in a file having extension as obtained in Step 11 which is the hidden message.

13) Terminate bit extraction process.

VI. Test Results

A. Perceptibility

Figures 2, 3 and 4 are screen shorts for cover text (extracted from: http://en.wikipedia.org/wiki/Steganography), pre-processed cover text and stego object respectively which are 100% identical in appearance:

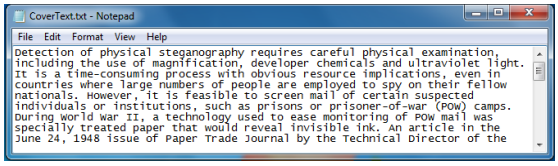


Fig. 2. ASCII Cover Text File

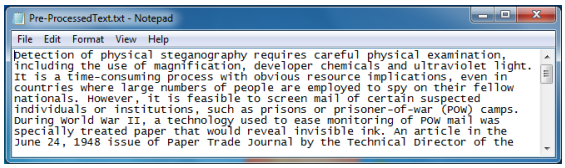


Fig. 3. Pre-Processed ASCII Cover Text File

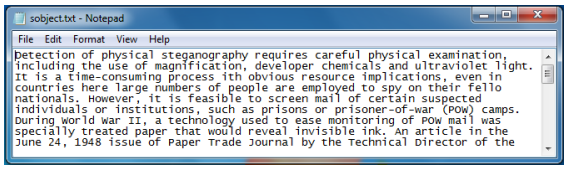


Fig. 4. Stego Object

B. File Lengths of Cover Text and Stego Object

Directory listing of the aforesaid files illustrated in Figure 5 does not show any difference unless viewed exclusively in terms of actual file size and that on persistent storage.

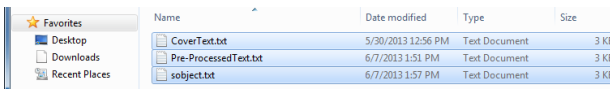


Fig. 5. Directory Listing for Cover, Pre-Processed Cover Text Files and Stego Object respectively

C. Quantified Similarity

The similarity between cover text, pre-processed text and stego object is quantified using Jro-Winkler distance [19][20]which was observed as 0.98071 while computed mean, variance and standard deviation are given in Table 8.

TABLE VIII. TABULATED MEAN, VARIENCE AND STANDARD DEVIATION

Computation	Cover Text	Pre-Processed Cover Text	Stego Object
Mean	0.09238	0.09368	0.09013
Variance	0.88189	1.18137	1.17942
Standard Deviation	0.02969	0.03437	0.03569

D. Graphical Illustration

Figures 6, 7 and 8 are graphical illustrations of probability distribution plots between ASCII cover text, pre-processed cover text and stego object respectively using MiniTab16 [21]:

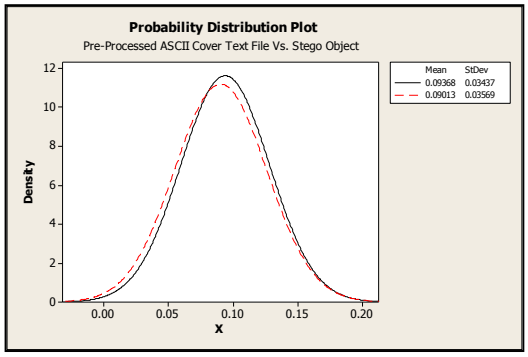


Fig. 6. Contrasting Probability Distribution Plots of Cover Text and Pre-Processed Cover Text

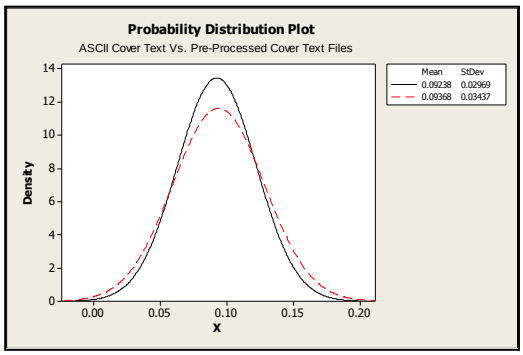


Fig. 7. Contrasting Probability Distribution Plots of Pre-Processed Cover Text File and Stego Object

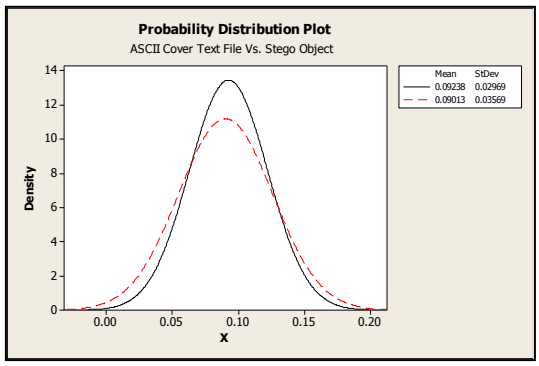


Fig. 8. Contrasting Probability Distribution Plots of ASCII Cover Text File and Stego Object

As apparent from preceding test results and graphical illustrations that Wendy will have a tough time to guess about the cover text that would have been used for information hiding purpose.

VII. THEORETICAL SUBSTANCE

Let δ be the cover text, and let α denote secret message bits that are to be hidden inside δ.Bit embedding involves searching for a blank/space in cover text and then replacing it with a double space to represent secret message bit '1' while for binary bit 0 no extra space is inserted. We may then represent bit embedding and extraction processes as follows:

Bit embedding: *if* $(\delta_j = \text{""})$ *then if*$(t = 1)\delta_j =$, *where* $t \in \alpha = \{1,0\}^*$, *j= 1, 2, 3 , … , total_ spaces &black-"", blue-"" &red-"" denotes blank/space characters respectively for clarity.*

Bit extraction: *if* $(\delta_j = \text{""})$ *thenif*$(\delta_{j+1} = \text{""})\{m \leftarrow 1; j \leftarrow j+1\}$ *else* $m \leftarrow 0; j \leftarrow j+1$, *where "j" denotes space counter which has an initial value of 1.*

From above it is obvious that Wendy can easily extract the hidden information by merely knowing the algorithm. Hence, there seems a daring need to induce some sort of uncertainty so that Wendy may not ascertain on actual cover text used in bit embedding process.

For incomprehensible steganography, stego object (**S**) must be an exact replica of cover text(**C**) – the chance of occurrence of which, however, is 1/100 as for remaining cases $\boldsymbol{S \neq C}$. Hence, security of steganographic system relates to the uncertainty involved in detection of actual cover used for whichintroduction of pre-processing stage, as in our case, will force Wendy to first confirm on the actual cover text used, then to figure out how to extract (key dependent) hidden bits from it and finally to recognize original message.

VIII. ADVANTAGES AND LIMITATIONS

Following are some of the pros and cons of our proposed data hiding scheme:

1) Advantages

- Key dependency.
- Information Theoretically Secure Solution.
- Imperceptibility of information hidden inside cover text.
- Increased bit embedding capacity per inserted byte.
- Prior information about hidden message's type and length facilitates in bit-extraction process.
- Doubles Wendy's effort towards cracking the system.

2) Limitations

- Increase in stego object's file size to a number equal to that of spaces in cover text file.
- Less 'Notepad' (default application for ASCII text files) opening the stego object through other applications may result in incomprehensible text.

IX. Future Work

Following are recommended as future work:

- Exploring new ideas to increase bit embedding capacity of text-cover.
- Compression of secret message.
- Adding another security layer via Key controlled encryption.
- Random, stego keydependent, bisectionof encrypted bitsjust before commencement of bit embedding processto illude comprehension about its starting point. OR

  Selecting random, stego keydependent, secret bit embedding starting point in text cover and then traversing cyclically just before that point.

X. Conclusion

Communicational ease that internet offers had its confronting impact when legitimate owners of information got deprived of their stake due to its illicit online copying and distribution. In addition, for many - it also appeared as a direct assault on their privacy and hence was reluctant to store, share and distribute digital contents over the internet. The situation necessitated the need to evolve copyright schemes to protect digital contents that inadvertently has given way for covert communication which is now being exploited in full especially in regions where cryptography is outlawed. With ever changing requirement for persistent storage and retrieval of data a variety of file formats like text, image, audio and video etc. have been evolved of which, excluding text file format, others offer Meta date (additional space) for storing secret information, and hence are the preferred choice for cover/carrier when it comes to information hiding. The fact of the matter, however, is that text format is the cheapest and most preferred choice for communication e.g. SMS texting, IP packets etc. and being least susceptible to be used as cover/carrier is an ideal choice for stealth communication.

This paper presented a novel stegokey dependent insertion-based steganographic scheme for ASCII cover text files. Salient characteristic of the scheme included increased bit embedding capacity per byte (i.e. in place of 1/8 bits per inserted byte, the embedding capacity has raised to 3/8 bits) in contrast to other prevalent text–cover centric steganographic schemes. Pre-processing of ASCII cover text before bit embedding has added an additional security layer which would double the amount of efforts Wendy may now need to get the hidden secret excavated out of the stego object.

References

[1] Chincholkar A.A. and Urkude D.A. (2012) Design and Implementation of Image Steganography. Journal of Signal and Image Processing, ISSN: 0976-8882 & E-ISSN: 0976-8890, Volume 3, Issue 3, pp.-111-113.

[2] F.A.P. Petitcolas, R.J. Anderson, and M.G. Kuhn, Information Hiding: A Survey, Proc. IEEE, 1999

[3] Donovan Artz, Digital Steganography: Hiding Data within Data, IEEE Internet Computing 1089-7801/ 01/$10.00, 2001 IEEE http://computer.org/internet/ MAY • JUNE 2001, p. 75-80

[4] Al.Jeeva,V.Palanisamy and K. Kanagaram, Comparative Analysis Of Performance Efficiency And Security Measures Of Some Encryption Algorithms, International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com, Vol. 2, Issue 3, May-Jun 2012, pp.3033-3037.

[5] Kipper, G.: Investigator's guide to steganography. Auerbach Publications, 2004 (240 p.), ISBN: 0849324335

[6] Michael F. Deering,The Limits of Human Vision, Sun Microsystems; http://www.swift.ac.uk/about/files/vision.pdf

[7] Ashraf Seleym and Dina Darwish, Real-time Covert Communications Channel for AudioSignals, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 3, September 2012, ISSN (Online): 1694-0814

[8] Shirali-Shahreza, M.H. Text Steganography in chat, 3rd IEEE/IFIP International Conference in Central Asia on Internet, (ICI 2007), 2007, pp. 1-5.

[9] Shirali-Shahreza, M.H, A New Synonym Text Steganography, International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP '08), 2008, pp. 1524-1526.

[10] Shirali-Shahreza, M., Text Steganography by Changing Words Spelling, 10th International Con-ference on Advanced Communication Technology (ICACT), 2008. 17-20 Feb. 2008, Volume: 3 Page(s): 1912 - 1913

[11] Mercan Topkara, Umut Topkara, Mikhail J. Atallah. Information Hiding through Errors: A Con-fusing Approach.2007. Internet: http://umut.topkara.org/papers/ToToAt_SPIE07.pdf, [July 12, 2012]

[12] Bender, W., Gruhl, D., Morimoto, N. & Lu, A. "Techniques for data hiding", IBM Systems Journal, Vol 35, pp. 313-336.1996.

[13] Souvik Bhattacharyya, Indradip Banerjee and Gautam Sanyal, A Novel Approach of Secure Text Based Steganography Model Using Word Mapping Method (WMM). International Journal of Computer and Information Engineering 4:2, p 96-102, 2010. Shirali-Shahreza, M.H. Text Steganography in chat, 3rd IEEE/IFIP International Conference in Central Asia on Internet, (ICI 2007), 2007, pp. 1-5.

[14] Caching, C. (1998). An information-theoretic model for Steganographic. Proc. of 2nd Workshop on Information Hiding, vol. 1525, Lecture Notes in ComputerScience.

[15] Khan Farhan Rafat, Muhammad Sher. On the Limits of Perfect Security for Steganography System. International Journal of Computer Science issues (IJCSI), Issue 4, May 2013.*Under publication*.www.ijcsi.org

[16] G. J. Simmons, "The prisoners' problem and the subliminal channel," inAdvances in Cryptology:Proceedings of Crypto 83(D. Chaum, ed.), pp. 51–67, Plenum Press, 1984.

[17] J.Z lner, H. Federrath, H. Klimant, A. Pfitzmann, R. Piotraschke, A. Westfeld, G.Wicke, G.Wolf. Modeling the security of steganographic systems. Proc. 2nd Workshop on Information Hiding, April 1998, Portland, LNCS 1525, Springer-Verlag, 1998, pp. 345-355.

[18] T Hansen - 2006, US Secure Hash Algorithms (SHA and HMAC-SHA), http://tools.ietf.org/html/rfc4634

[19] Jaro, M. A. 1989. Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. Journal of the American Statistical Association 84:414–420.

[20] Winkler, W. E. 1999. The state of record linkage and current research problems. Statistics of Income Division, Internal Revenue Service Publication R99/04.

[21] Minitab 16. http://www.facebook.com/Minitab [June 6, 2012]

Collaborative Spectrum Sensing under Suburban Environments

Aamir Zeb Shaikh
Department of Telecommunications Engineering,
NED University of Engineering & Technology,
Karachi 75270. Pakistan

Dr. Talat Altaf
Department of Electrical Engineering,
NED University of Engineering & Technology,
Karachi 75270. Pakistan

***Abstract*— Collaborative spectrum sensing for detection of white spaces helps in realizing reliable and efficient spectrum sensing algorithms, which results in efficient usage of primary spectrum in secondary fashion. Collaboration among cognitive radios improves probability of detecting a spectral hole as well as sensing time. Available literature, in this domain, uses Gudmundson's exponential correlation model for correlated lognormal shadowing under both urban and suburban environments. However, empirical measurements verify that the suburban environment can better be modeled through double exponential correlation model under suburban environments in comparison to Gudmundson's exponential correlation model. Collaboration among independent sensors provides diversity gains. Asymptotic detection probability for collaborating users under suburban environments using double exponential correlation model has been derived. Also, the Region of Convergence performance of collaborative detection is presented which agrees well with analytical derivations.**

***Keywords*— collaborative spectrum sensing; suburban environment; asymptotic analysis; cognitive radio; opportunistic access.**

I. INTRODUCTION

Cognitive radio is a revolutionary concept that aims to utilize licensed RF spectrum in an unlicensed/opportunistic fashion [1]. The successful opportunistic usage of spectrum requires information about unused spectrum bands termed as white spaces. This can be accomplished through Beacon-assisted techniques and spectrum sensing algorithms. In Beacon-assisted based methods [2], primary user / licensed user transmits a beacon signal to cognitive users (opportunistic users) regarding available white-space on particular time and frequency bands that is decoded by the secondary users for successful exploitation of those spectral holes (white spaces). In spectrum sensing based techniques [3]; cognitive radio users detect white spaces (either individually or collaboratively) and exploit the identified bands in opportunistic fashion.

Spectrum sensing methods can be distributed in three categories i.e. transmitter, receiver and interference-temperature based detection algorithms [3]. From these, transmitter detection based methods are a preferred way of sensing for presence/absence of spectral holes. These methods can be implemented through various techniques including coherent detection, feature detection and energy based detection algorithms [4]. Coherent sensor is an optimal linear detector for known primary signals in presence of white Gaussian noise [5]. However, detector implementation requires demodulation of received signals for achieving the optimal gains. Cognitive Radio operates in an opportunistic fashion that can operate in licensed as well as unlicensed bands and thus, it is quite difficult to demodulate each received signal before deciding in favor of presence/absence of a primary user. So, coherent sensor is generally not a preferred method of spectrum sensing. Cyclostationary feature based detector is an efficient and reliable method of spectrum sensing. These detectors compute Spectral Correlation function of received signals which serves as the signature of the particular signals. Thus, these detectors can easily distinguish between primary user signals, noise and other interfering sources (by using the features of corresponding signatures). However, these gains are achieved on the basis of exact licensed user information as well as received computational complexity.

On the other hand, energy sensing is a semi-blind method of detection of primary users [6]. Energy spectrum sensor computes the energy of the received signal samples and compares with a pre-set threshold. Setting of threshold requires noise information only. Computational simplicity makes these detectors a preferred choice for spectrum sensing cognitive radios. We use semi-blind algorithm i.e. energy detector for sensing the white spaces in our cognitive radio network under suburban environment.

Collaboration among secondary radios provides diversity gains. In [7] it is shown that under independent and identically distributed (i.i.d.) shadowing the detection probability can be significantly improved by exploiting user-collaboration. However, due to correlated shadowing the diversity gains are reduced [8].

In previous works, the detection probability (or missed-detection) for cognitive radio applications is computed using Gudmundson's exponential correlation model [9] under both urban and suburban environments. However, the results demonstrated by using real-time received signal measurement campaign performed by authors [10], analyze that the exponential correlation model best fits the autocorrelation function under urban environments but the results under suburban environments better follow double exponential model.

This paper considers the application of double exponential correlation model under suburban environments for cognitive radio applications. Asymptotic probability of detection is also derived. The simulation results verify the improvement in

detection probability under double exponential correlation model in comparison to Gudmundson's correlation model under suburban environment.

The paper is organized as follows. The proposed system model and a brief discussion on the double correlation model are presented in Section II. In addition to that, this section also includes the derivation of asymptotic probability of detection using double exponential model under suburban environments. Section III presents the evaluation of detection probability under given environment conditions. Section IV concludes the paper in addition to a brief discussion on future work.

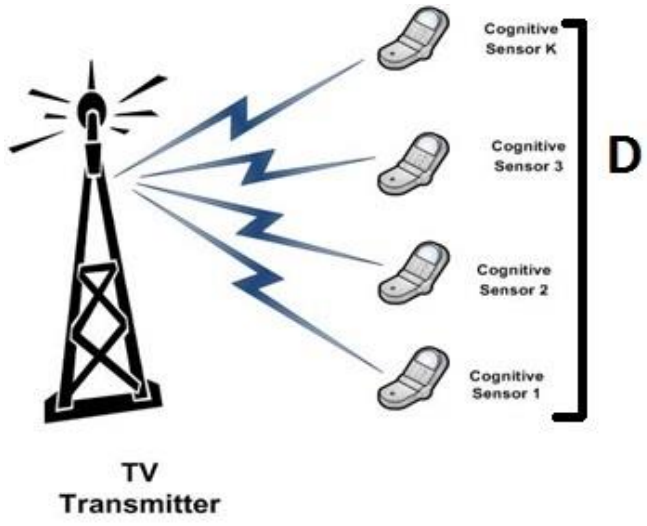


Fig. 1. Cognitive radio network, sensing TV Transmitter

II. System Model

The proposed spectrum sensing scenario in this paper includes a TV transmitter as primary/licensed radio network and the secondary radio network consists of a large number of cognitive sensors, detecting primary transmissions collaboratively, as shown in Figure [1]. The received signal energy from primary transmitter to cognitive sensor can be represented with y(n) that can be defined as a binary hypothesis testing rule:

$$y(n) = \begin{cases} \mathbf{n_i(n)} & ;H_0 \\ \mathbf{h_i(n)x_i(n)+n_i(n)} & ;H_1 \end{cases} \quad (1)$$

n = 0, 1,2,3,…, k shows sample number of received signal, $n_i(n)$ represents samples of AWGN, $n_i(n) \sim \mathcal{N}(0, \sigma_n^2)$ and $h_i(n)$ shows fading channel coefficients, $h_i(n) \sim \mathcal{N}(0, \sigma_h^2)$ and $x_i(n)$ represents the Primary User signal samples. H_0 hypothesis represents presence of a spectral hole while H_1 represents the presence of a primary user signal.

Probability of detection (P_D) represents the presence of a legitimate user (a.k.a. primary user), probability of missed-detection (P_{MD}) shows that the detector flags as the presence of a legitimate user while in actually it is absent and the false alarm probability (P_{FA}) represents the absence of a legitimate user, however, due to erroneous noise estimation or other errors the detector flags the presence of a primary user.

Thus, received signal vector y can be defined as a multivariate Gaussian random variable with following distribution:

$$\mathbf{y} \sim \begin{cases} \mathcal{N}(\boldsymbol{\mu}_0, \sigma^2 \boldsymbol{\Sigma}) & ;H_0 \\ \mathcal{N}(\boldsymbol{\mu}_1, \sigma^2 \boldsymbol{\Sigma}) & ;H_1 \end{cases} \quad (2)$$

Where μ_0 and $\mu_1 \in \mathbb{R}^{K\times 1}$ represent mean of received signal while $\sum \in \mathbb{R}^{K\times K}$ is representing the common covariance matrix of y under both null and alternative hypotheses i.e. H_0 and H_1. Mean and Covariance Matrices can be defined as:

$$\boldsymbol{\mu}_0 = \sigma_n^2 \times \mathbf{1}, \quad \boldsymbol{\mu}_1 = (P_S + \sigma_n^2) \times \mathbf{1} \quad (3)$$

σ_n^2 is the mean signal power of received signal, P_s is the mean signal power of the primary transmitter, received at cognitive sensor , '1' shows the vector of 1s, μ_0 is the mean, whereas common covariance matrix is defined as:

$$\boldsymbol{\Sigma} = \frac{P_S}{m}\boldsymbol{\Lambda} + \frac{\sigma_n^4}{m}\mathbf{I_K} \quad (4)$$

$m = \tau\varpi$, Where τ represents the sensing time and ϖ shows the bandwidth. Λ represents double exponential correlation covariance matrix with k x k measurements.

$$\boldsymbol{\Lambda}_{a,b} = \xi^{|a-b|}, a, b = 1, 2, 3, 4, \ldots\ldots k \quad (5)$$

A. Channel Model

We use double exponential model for suburban environments [10], which is given by following equation:

$$\xi = h\, e^{-d/d_{A1}} + (1-h)\, e^{-d/d_{A2}} \quad (6)$$

Where d_{A1}, d_{A2} are the short and long correlation distances, h shows the strength between short and long correlation distance and d represents the distance travelled by the cognitive user.

Considering the case of secondary users with one-dimensional distribution within a fixed distance D as in (i.e. $d = D/(k-1)$) where k represents the number of sensing.

B. Hypothesis Testing

It is assumed that the received signal power is less than the noise power as described in IEEE 802.22 [12]. The probability density function (PDF) of **y** under both null and alternate hypothesis is given by:

$$f_0(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{k}{2}}|\boldsymbol{\Sigma}_0|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}_0^{-1}(\mathbf{y}-\boldsymbol{\mu}_0)\} \quad (7)$$

$$f_1(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{k}{2}}|\boldsymbol{\Sigma}_1|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1}(\mathbf{y}-\boldsymbol{\mu}_1)\} \quad (8)$$

Thus, Likelihood Ratio test can be evaluated as:

$$\Upsilon(\mathbf{y}) = \frac{(\mathbf{x};H_1)}{(\mathbf{x};H_0)} = \frac{\frac{1}{(2\pi)^{\frac{k}{2}}|\boldsymbol{\Sigma}_1|^{\frac{1}{2}}}\exp\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_1^{-1}(\mathbf{y}-\boldsymbol{\mu}_1)\}}{\frac{1}{(2\pi)^{\frac{k}{2}}|\boldsymbol{\Sigma}_0|^{\frac{1}{2}}}\exp\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu}_0)^T\boldsymbol{\Sigma}_0^{-1}(\mathbf{y}-\boldsymbol{\mu}_0)\}} \quad (9)$$

Using Neyman-Pearson Lemma, the detection probability can be computed by [11,13]:

$$P_d = \mathbf{Q}\{\frac{\Delta}{\sigma}\sqrt{1^T\Lambda^{-1}1} + Q^{-1}(P_{fa})\} \quad (10)$$

Where $\Delta = \mu_1 - \mu_0$, P_{fa} represents the false alarm probability and missed-detection probability is shown by P_{md}=1-Pd.

From equation (10):

$$\sqrt{1^T\Lambda^{-1}1} = \sqrt{\frac{(1-\xi)k+2\xi}{1+\xi}} \quad (11)$$

For very large number of sensing, detection probability can be derived as:

$$\lim_{k\to\infty}\sqrt{\{\frac{(1-\xi)k+2\xi}{1+\xi}\}} \quad (12)$$

$$\lim_{x\to 0}\sqrt{\{1-ke^{-a_1x}-(1-k)e^{-a_2x}\}\{1+\frac{D}{x}\}} \quad (13)$$

$$\lim_{x\to 0}\sqrt{\{\frac{1-ke^{-a_1x}}{x}-\frac{(1-k)e^{-a_2x}}{x}\}\{x+D\}} \quad (14)$$

$$= \sqrt{\frac{D(ha_1-a_2-ha_2)+2}{2}} \quad (15)$$

Where $a_1 = \frac{1}{d_{A1}}$ and $a_2 = \frac{1}{d_{A2}}$.

III. Numerical Results

In this section, the detector performance in terms of ROC metric is examined under suburban environments. Furthermore, asymptotic behavior of collaborative spectrum sensing is also examined and compared with Gudmundson's exponential correlation model.

In addition, a cluster-based architecture is also proposed to improve the sensing performance under double correlation model.

It is assumed that a very large number of cognitive sensors, placed equidistantly, are sensing TV transmissions.

The parameters are numerical computations are given as: Δ = -5.19, σ =2.3, D=100, false alarm probability $P_{fa}=10^{-3}$, detection probability P_d = 0.9, short correlation distance d_{A1}=25 m, long correlation distance d_{A2}=200m, strength between short and long correlation distance =0.2, total distance travelled by the cognitive sensor d, with 20 number of sensing.

Figure [2] compares the performance of the proposed double correlation model with classic exponential correlation model.

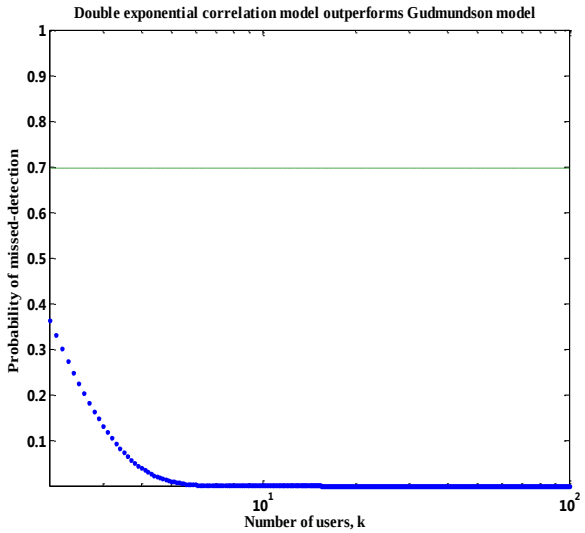


Fig. 2. Verifies the theoretical recommendation that double exponential Correlation model outperforms the Gudmundson's Exponential Correlation Model under suburban environments

C. Clustered Sensing:

Clustered sensing corresponds to the case where a number of sensors combine their sensing results for improved detection probability. In this case we consider twenty sensors to submit their results at a fusion center that combines the results of the sensors using OR based combination rule. The following equations can be used to determine detection and false alarm probabilities under suburban environment.

D.

$$Q_d = \sum_{i=m}^{n}\binom{n}{i}(p_d)^i(1-p_d)^{n-i} \quad (12)$$

$$Q_{fa} = \sum_{i=m}^{n}\binom{n}{i}(p_{fa})^i(1-p_{fa})^{n-i} \quad (13)$$

Once the probability of detection and false alarm are computed, the results are announced through control channel to all the secondary users so that the cognitive users may benefit from the results appropriately. The results in Figure [3] indicate that even two and three round of sensing results improves detection probability significantly.

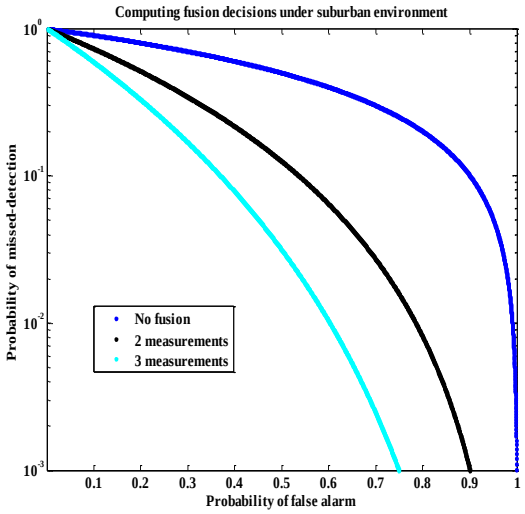


Fig. 3. The improvement in detection performance due to clustered sensing.

IV. Conclusion

Asymptotic performance of collaborative spectrum sensing algorithm is derived and analyzed, using double exponential correlation model under suburban environments. Numerical results have shown that our proposed channel model performs significantly better than classical Gudmundson's exponential correlation model. The proposed model will find wide applications especially in practical wireless systems as the missed-detection metric attains a minimal value very fast (i.e. with minimal number of secondary sensors) in comparison to exponential correlation model. Furthermore, to enhance the performance, a light-weight hard decision combining based strategy for cluster-based detection algorithms is also proposed. This strategy exploits cooperation among sensors to improve detection performance.

Operation of cognitive radio networks mainly in unlicensed RF spectrum poses a threat of interference to the secondary users. Thus, presently, we are studying the performance of spectrum sensors under external interference. The results to this study will indicate the performance metrics under realistic wireless systems.

Acknowledgment

Authors are thankful to Dr. Muhammad Imran Aslam at NED University of Engineering & Technology, Karachi for his useful discussions/suggestions.

References

[1] A. Ghasemi and E. S. Sousa, "Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing," Communications Letters, IEEE, vol. 11, pp. 34-36, 2007.

[2] A. Tajer and X. Wang, "Beacon-assisted spectrum access with cooperative cognitive transmitter and receiver," Mobile Computing, IEEE Transactions on, vol. 9, pp. 112-126, 2010.

[3] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," Communications Surveys & Tutorials, IEEE, vol. 11, pp. 116-130, 2009.

[4] D. Cabric, et al., "Implementation issues in spectrum sensing for cognitive radios," in Signals, systems and computers, 2004. Conference record of the thirty-eighth Asilomar conference on, 2004, pp. 772-776.

[5] J. G. Proakis, "Digital communications, 1995," ed: McGraw-Hill, New York.

[6] Y. Zeng, et al., "A review on spectrum sensing for cognitive radio: challenges and solutions," EURASIP Journal on Advances in Signal Processing, vol. 2010, p. 2, 2010.

[7] E. Visotsky, et al., "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on, 2005, pp. 338-345.

[8] A. Ghasemi and E. S. Sousa, "Opportunistic spectrum access in fading channels through collaborative sensing," Journal of Communications, vol. 2, pp. 71-82, 2007.

[9] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," Electronics letters, vol. 27, pp. 2145-2146, 1991.

[10] A. Algans, et al., "Experimental analysis of the joint statistical properties of azimuth spread, delay spread, and shadow fading," Selected Areas in Communications, IEEE Journal on, vol. 20, pp. 523-531, 2002.

[11] A. Ghasemi and E. S. Sousa, "Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing," Communications Letters, IEEE, vol. 11, pp. 34-36, 2007.

[12] IEEE 802.22 "Working Group on Wireless Regional Area Networks (WRAN)," IEEE Tech. Rep.

[13] A. Z. Shaikh and T. Altaf, "Performance Analysis of Correlated Multiple Antenna Spectrum Sensing Cognitive Radio," International Journal of Computer Applications, vol. 50, 2012.

On an internal multimodel control for nonlinear multivariable systems - A comparative study

Nahla Touati Karmani
Automatic Research Laboratory, LA.R.A, National Engineering School of Tunis (ENIT), Tunis El Manar University, Tunis, Tunisia

Dhaou Soudani
Automatic Research Laboratory, LA.R.A, National Engineering School of Tunis (ENIT), Tunis El Manar University, Tunis, Tunisia

Mongi Naceur
Automatic Research Laboratory, LA.R.A, National Engineering School of Tunis (ENIT), Tunis El Manar University, Tunis, Tunisia

Mohamed Benrejeb
Automatic Research Laboratory, LA.R.A, National Engineering School of Tunis (ENIT), Tunis El Manar University, Tunis, Tunisia

***Abstract*—An internal multimodel control designed for nonlinear multivariable systems, is proposed in this paper. This approach is based on the multi-modeling of nonlinear systems and the realization of a specific inversion of each model. A comparative study is presented between two structures of the internal multimodel control. The first one is based on switching models and the second on residues techniques as fusion method. The case of a second order nonlinear multivariable system shows the effectiveness of both structures.**

Keywords—internal multimodel control; nonlinear multivariable systems; inverse model; switching models; residues techniques.

I. INTRODUCTION

Among the control structure of physical systems, the Internal Model Control (IMC) is one of the powerful control laws for open-loop stable plants [6]. It provides an open-loop framework for checking closed-loop stability and also highlights the inherent performance limitations due to model uncertainties, non-minimum phase plant characteristics and actuator constraints.

The IMC systems are characterized by a control device consisting of the controller and a simulation model of the process. The internal model loop computes the difference between the process and its model which represents the effect of disturbances and model mismatch.

The proposed controller in this structure is the inverse of the selected model, if it is realizable. This task is the main problem of the IMC approach due to the difficulty of direct model inversion for the majority of physical systems, which gives a structure generally unrealizable [8].

The IMC design procedure is quite extensive and diverse. It has been developed in many forms; these include single-input-single-output (SISO) and multiple-input-multiple-output (MIMO) formulations, continuous and discrete-time design procedures, design procedures for unstable open-loop systems, for nonlinear systems and so forth.

Developments of nonlinear IMC (NIMC) have been proposed for continuous time systems and for a class of discrete-time systems [9].

In this context, several NIMC schemes have been proposed, such us proposing an analytical inverse of the nonlinear model based on a physical understanding of the plant [10], or by representing the nonlinear plant by a collection of local linear models. The global controller in this case is obtained by combining these controllers using a fuzzy system. This last scheme combines two approaches: the internal model and multimodel control.

The multimodel approach is extremely interesting when dealing with nonlinear complex systems. In this approach, the nonlinear plant is described by a combination of local linear models, each of which is valid in a particular operating region. First, local controllers for the local models are tuned. Next, the control actions of these local controllers are combined in the form of a global controller to be implemented on the nonlinear plant. Different methods are proposed to design the global controller such us switching between local controllers or by combining them (residues techniques).

Based on the NIMC and the multimodel representation of nonlinear systems, this paper proposes the Internal MultiModel Control (IMMC) designed for nonlinear multivariable systems. Two structures will be presented, the IMMC based on the switching and the residues techniques.

This paper is organized as follows: section II presents a brief review on the internal model control strategy. Section III is devoted to multimodel approach. Section IV presents the proposed structure for the internal multimodel control designed for nonlinear multivariable systems. In section V, an illustrative example is considered to show the effectiveness of the proposed structures and to compare them. Finally, concluding remarks are drawn in Section VI.

II. NONLINEAR MULTIVARIABLE INTERNAL MODEL CONTROL

Internal model control is considered as a robust control method for open-loop stable systems. The nonlinear internal model control structure designed for multivariable systems is composed of a stable nonlinear multivariable process model, and a feed-forward controller as shown in Fig. 1.

The controller is obtained as the inverse of the plant model [2].

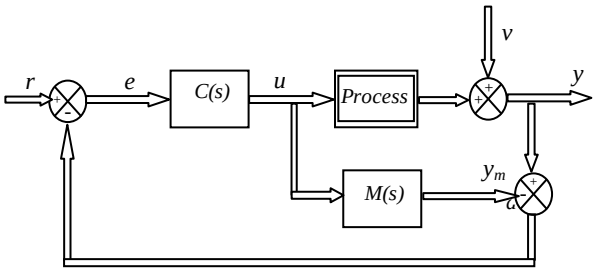


Fig.1. The internal model control structure

In the basic configuration of the IMC for multivariable systems, *M* is the chosen model for the process, *C* the controller and *v* a vector of disturbance affecting the system. The manipulated *l* input vector *u* is introduced to both the process and its model. The difference *d* between the process output vector *y* of dimension *m* and model output vector y_m , is compared to the reference vector *r* to generate the controller input vector *e* . In this paper, we propose a fully actuated system. So, *l=m*.

The model inversion is the main problem of the IMC since the direct model inversion, for physical systems, gives a structure generally unrealizable.

The development of a nonlinear general approach of the IMC can raise serious difficulties, because of the complexity of the nonlinear systems, the absence of the mathematical inversion methods for the nonlinear models and consequently the difficulty of the design of a nonlinear internal model controller [1]. Multivariable linear internal model control is based on transfer matrix models, while nonlinear systems are usually described by nonlinear state-space models.

By using the inversion method proposed in [2], based on the gain matrix to realize an inverse model for the nominal model, we obtain the IMC controller shown in Fig. 2.

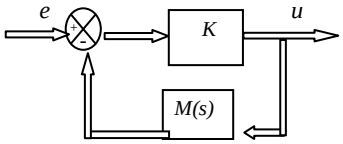


Fig.2. Generalized controller structure

III. MULTIMODEL CONTROL

The multimodel control represents a relatively new approach on nonlinear control strategies. It makes possible to represent nonlinear complex systems by using a set of simple linear models constituting a model's library, each of them is valid in a particular operating region. For each model, a suitable controller can be designed off-line [5].

The multimodel control can be perceived as a fusion of the partial command calculated from each model of this library or as a switching approach between these models.

The switching method consists on the choice of the nearest model to the process which leads to the least modeling error. The fusion method consists on evaluating the contribution of each partial command to the effective controller of the system.

It is based on calculating the residues which are frequently formulated by the geometric distance between the real output and the local models outputs [7]. The highest validity index is given to the system that gives the best estimation. In the opposite case the validity index is close to zero.

IV. MULTIVARIABLE INTERNAL MULTIMODEL CONTROL

In this paper, an internal multimodel control is proposed for nonlinear multivariable systems as shown in Fig. 3. The first step consists on describing the nonlinear system to control by a model's library. The analysis and control of these linear models are easy and they can be exploited for an IMC structure. Identification or linearization around various operating points or convex polytopic transformation can be used to define these models.

The multimodel system describing the nonlinear multivariable system should be stable. In the proposed structure, the M_i, $i=1...n$, are linear state-space models describing the multivariable system. For each model M_i, a controller matrix C_i is designed in order to ensure the nominal performances for the pair (M_i,C_i). C_i is the IMC controller obtained by using the inversion method as proposed in Fig. 2. It is integrated in the closed-loop configuration, as presented in Fig. 3. All local controllers are calculated instantly and simultaneously. *y* and *u* are the system output and input vectors. $\tilde{y}_i$, $\tilde{u}_i$ and $e_{i,,}$ $i=1...n$ present the output, input and controller input vectors respectively for each model M_i. *d* is the modeling error, K_i the gain matrices and *v* the validation index.

Two methods are proposed in this paper, the first one is based on the switching technique: the model with the smallest error is chosen. After this operation, the correspondent control input *u* is applied to the system. The second method is based on residues techniques.

The calculation of the controller to be applied to the real system depends, not only on the partial controllers, related to the library's models, but also of the validities resulting from the residues which represent errors between system and models outputs.

A. System description

Let's consider a nonlinear multivariable system described by the following state space equation

$$\begin{cases} \dot{x} = f(x,u,t) \\ y = g(x,u,t) \end{cases} \quad (1)$$

Where

- x is the system state vector, $x \epsilon R^m$;
- y is the output vector, $y \epsilon R^m$;
- And u is the input vector, $u \epsilon R^m$.

This system can be represented by a library of *n* linear models M_i , where each model can be described by the following state space equations

$$\begin{cases} \dot{x}_\iota = f_i(x_i, \tilde{u}_i, t) \\ \tilde{y}_i = g_i(x_i, \tilde{u}_i, t) \end{cases} \quad (2)$$

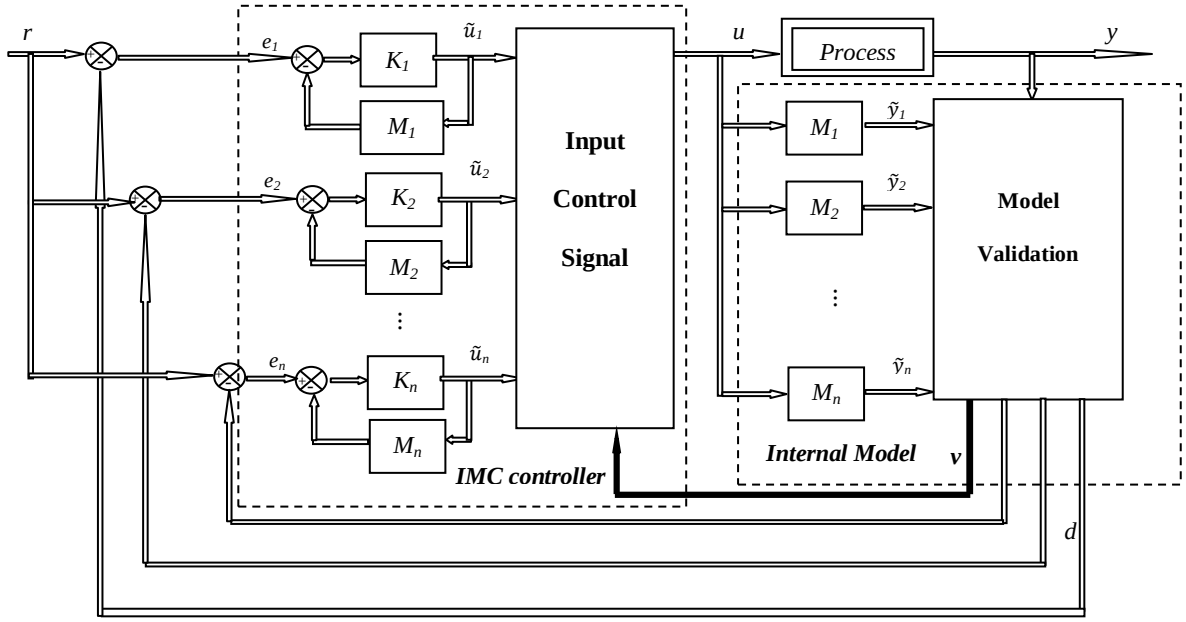


Fig.3. Proposed structure for internal multimodel control designed for nonlinear multivariable systems

For each model is associated a controller vector $\tilde{u}_i$ and an output vector$\tilde{y}_i$. The M_i can be represented by (A_i, B_i, C_i, D_i), where A_i, B_i, C_i and D_i are the state matrices.

B. The switching method

First of all, let's define the *n* distance vectors d_i, $d_i = \{d_{ji}\}_{j=1\dots m}$. d_i characterizes the difference between system outputs y and model outputs $\tilde{y}_i$, $y = \{y_i\}_{i=1\dots m}$ and $\tilde{y}_i = \{\tilde{y}_{ij}\}_{j=1\dots m}$.

$$d_{ji}(t) = y_j(t) - \tilde{y}_{ij}(t), i = 1 \dots n \text{ and } j = 1 \dots m \qquad (3)$$

Let's consider f_i such that

$$f_i(t) = \sum_{j=1}^{m} \left\| d_{ji}(t) \right\|, i = 1 \dots n \text{ and } j = 1 \dots m \qquad (4)$$

The controller correspondent to the model with the lowest f_i is the one to apply to the system.

C. The residues techniques

For the residues techniques, all controllers participate on calculating the final controller applied to the system, according to their validities $v_i, i = 1 \dots n$.

To quantify the contribution of the *n* partial command $\tilde{u}_i = \{\tilde{u}_{ij}\}_{j=1\dots m}$ to the global controller of the system, it's necessary to define validity values v_i, making it possible to give preemption, at every moment, to the nearest model at the detriment of the others.

$$v_i = \frac{\frac{1}{\|f_i\|}}{\frac{1}{\sum_{i=1}^{m} \frac{1}{\|f_i\|}}} \qquad (5)$$

The control law u_i to be applied to the system is given by the following expression.

$$u_i(t) = \sum_{j=1}^{n} v_j(t) \tilde{u}_{ij}(t), \; i = 1 \dots m \qquad (6)$$

This control law depends on the partial commands and the validity coefficients relating to the models M_i.

V. APPLICATION

In order to illustrate the proposed internal multimodel control for multivariable nonlinear systems, let's consider the following multivariable system with two inputs two outputs *(m=2)*.

$$\begin{cases} \dot{x}_1(t) = -2x_1(t) - \cos(x_1(t))\, x_2(t) + u_1(t) \\ \dot{x}_2(t) = \cos(x_1(t))\, x_1(t) - |x_2(t)| x_2(t) + 5u_2(t) \end{cases} \qquad (7)$$

The plant equations can be transformed to the following state space equations

$$\begin{cases} \dot{x}(t) = A\, x(t) + Bu(t) \\ y(t) = C\, x(t) + Du(t) \end{cases} \qquad (8)$$

Where

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, A = \begin{pmatrix} -2 & -\cos(x_1(t)) \\ \cos(x_1(t)) & -|x_2(t)| \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}, u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}, \; y = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } D = 0.$$

We propose to apply the convex polytopic transformation in order to define the models describing the system.

$$\dot{x} = \sum_{i=1}^{n} \mu_i (A_i x + Bu) \quad i = 1 \dots n \qquad (9)$$

A_i, B, C and D define the local models and the μ_i are the activation functions.

We suppose that the matrix $A = \{a_{ij}\}$ such us a_{ij} are bounded

$$a_{ij} \in [a_{ijmin}, a_{ijmax}], i, j = 1,2. \quad (10)$$

Let $\cos(x_1(t)) \in [-1,1]$ and $x_2 \in [0,10]$.

For the proposed system, 2^2 models (A_i, B, C, D) are proposed to describe the nonlinear multivariable system

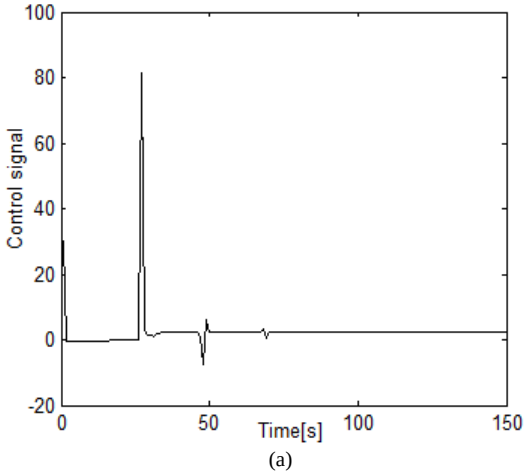
$$A_1 = \begin{pmatrix} -2 & -1 \\ 1 & -10 \end{pmatrix}, A_2 = \begin{pmatrix} -2 & -1 \\ 1 & 0 \end{pmatrix},$$
$$A_3 = \begin{pmatrix} -2 & 1 \\ -1 & -10 \end{pmatrix} \text{ and } A_4 = \begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix}.$$

The considered local models for this example are significantly different from each others. This allows seeing clearly the difference between the two proposed methods applied to the system.

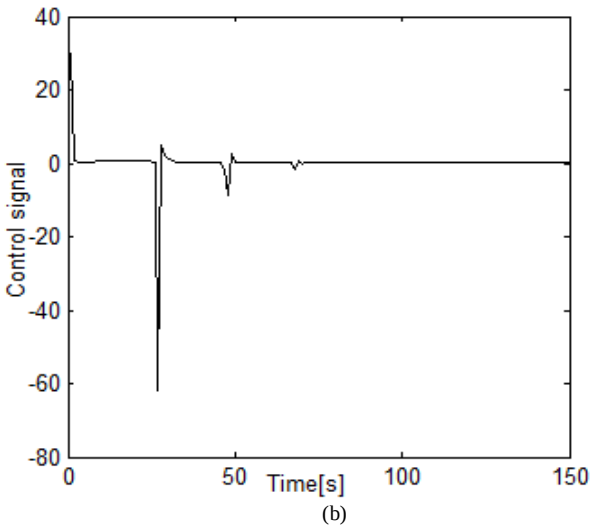
A. *Simulation results for the switching method*

Simulation results are shown in Fig. 4, 5 and 6 where the reference vector r is $r = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and the gain matrices K_i, *i=1…4* such that $K_i = \begin{pmatrix} 30 & 0 \\ 0 & 30 \end{pmatrix}$.

Fig. 4, 5 and 6 display simulation results for the switching method. Fig. 4 represents the internal multimodel controllers $u_1(t)$ and $u_2(t)$. System outputs $y_1(t)$ and $y_2(t)$ are displayed in Fig. 5. Control signals are elevated and varied from one moment to other. This is due to the switching between linear models leading to control signals variation. It can be shown a sluggish output responses (the system outputs reach the references after 80seconds) and an important overshoot. Fig. 6 displays the chosen models for the switching method. Their correspondent controllers are applied to the system. For this example, the switching method hasn't shown satisfactory results.

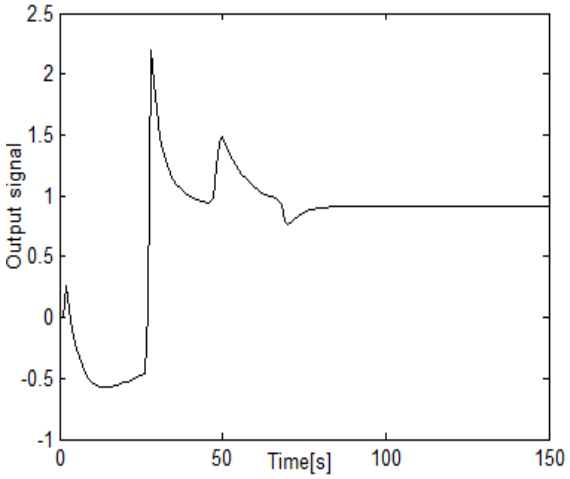


(a)

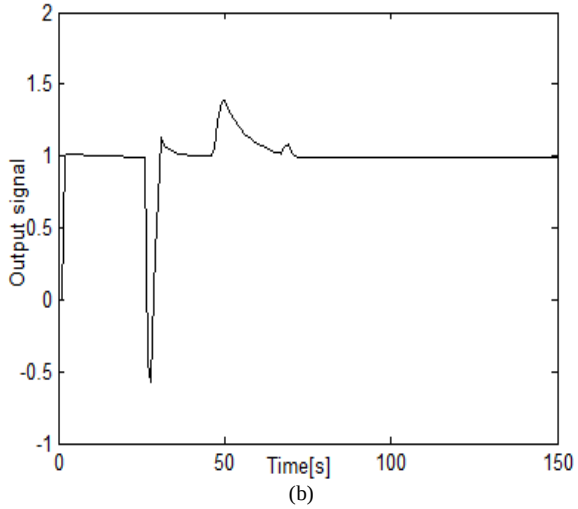


(b)

Fig.4. (a)Control signal u_1 (b) Control signal u_2



(a)



(b)

Fig.5. (a) Output signal y_1 (b) Output signal y_2

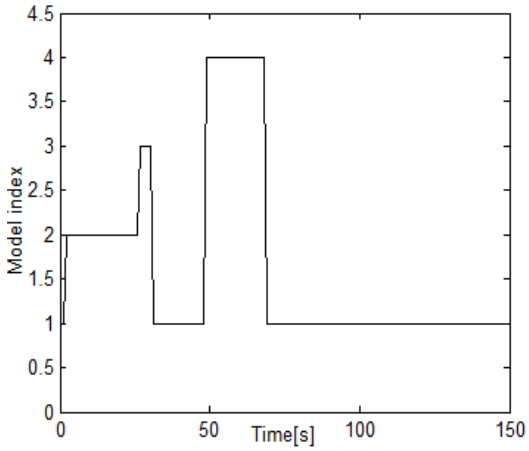


Fig.6. Chosen models for the switching method

B. Simulation results for the residues techniques

Simulation parameters for residues techniques are taken the same as the switching method. Fig. 7, 8 and 9 display simulation results for the residues techniques. Fig. 7 represents the evolution of the internal multimodel controllers $u_1(t)$ and $u_2(t)$. System outputs $y_1(t)$ and $y_2(t)$ are displayed in Fig. 8. The system outputs present an acceptable overshoot and are able to reach perfectly the references at the steady state. Fig. 9 shows model's validities.

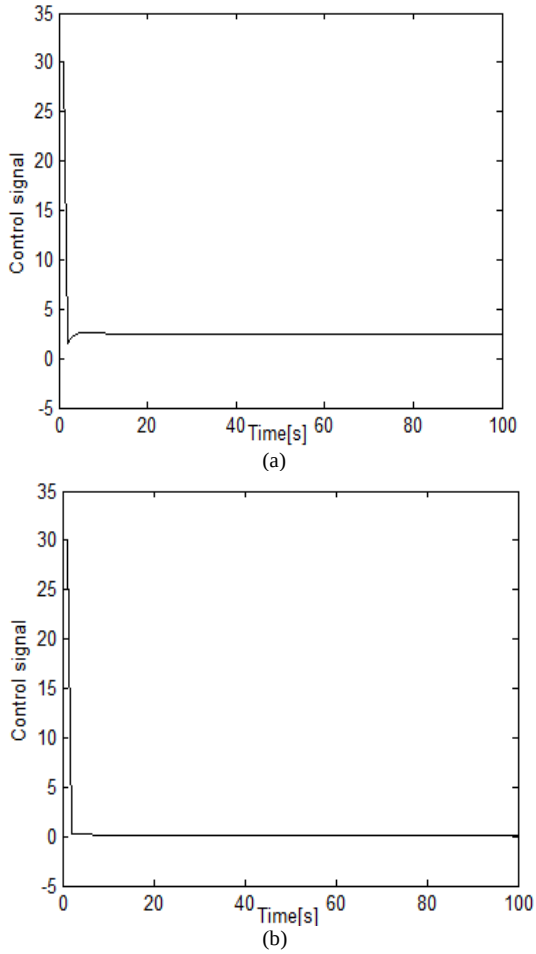


Fig.7. (A) Control signal u_1 (b) Control signal u_2

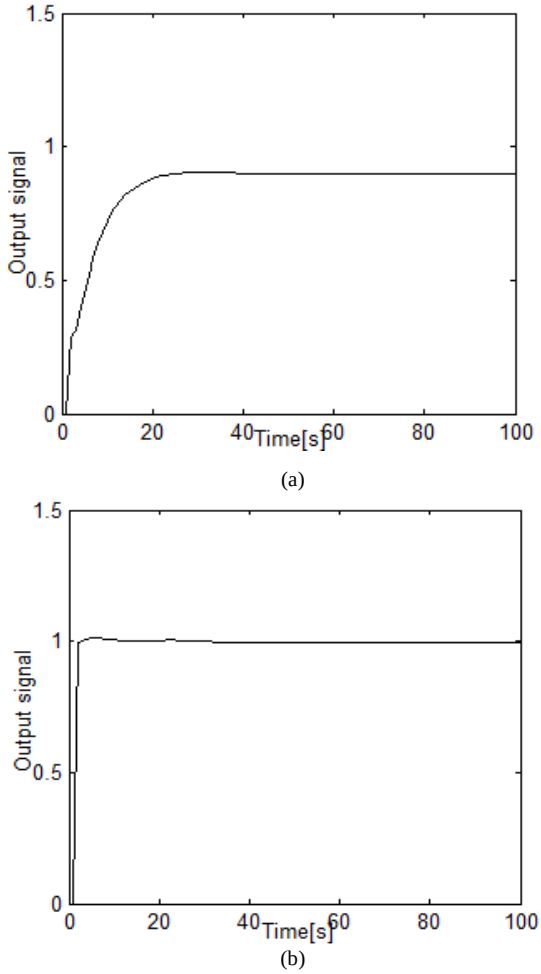


Fig.8. (a) Output signal y_1 (b) Output signal y_2

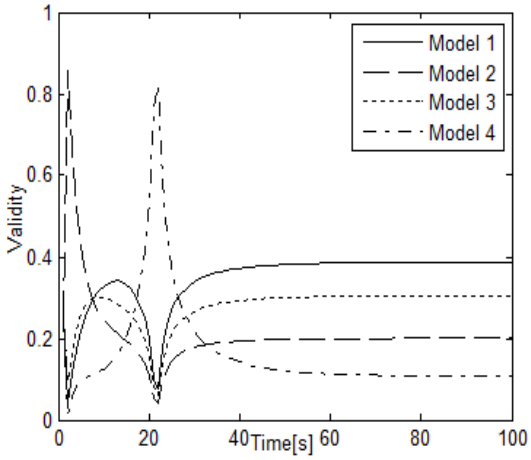


Fig.9. Validity indices

VI. CONCLUSION

In this paper, two structures for internal multimodel control were proposed for nonlinear multivariable systems.

The first structure is based on the determination of control signals by switching between different inverse models outputs. While the second one is based on using the residues techniques as fusion method to calculate the control signal.

The implementation of the second structure based on residues techniques is more complex and needs more computation time than the first one based on the switching method, but it is more efficient. In fact, by using the residues techniques, the system outputs reach quickly the imposed reference without important overshoot.

REFERENCES

[1] S. Bel Hadj Ali, "Commande par modèle interne de processus dynamiques", PhD, ENIT, Tunis, 2003.

[2] M. Benrejeb, M. Naceur , D. Soudani , "On an internal model controller based on the use of a specific inverse model", ACIDCA 2005, pp.623 626, Tozeur, 2005.

[3] N. Touati , D. Soudani., M. Naceur and M. Benrejeb , "On the internal model control of multivariable linear system", International Conference on Sciences and Techniques of Automatic control and computer engineering, STA'2011, Sousse, 2011.

[4] N. Touati , D. Soudani., M. Naceur and M. Benrejeb, "Internal Model Control of Coupled Tanks Process", 4th Conference of Applied Research and Technology Transfer CRATT'2012, Tunis, 2012.

[5] R. H. Nyström, J. M., Böling , J. M. Ramstedt , H. Toivonen and K. Häggblom, "Application of robust and multimodel control methods to an ill-conditioned distillation column", Original Research Article Journal of

[6] Process Control, Volume 12, Issue 1, January 2002, pp. 39-53.

[7] M. Morari and E. Zafiriou , "Robust Process Control", Ed. Prentice Hall, Englewood cliffs, N.J., 1989.

[8] M. A. Henson and D. E. Seborg, "Internal Model Control Strategy for Nonlinear Systems", A.I.Ch.E. Journal, vol. 37, 1991, pp. 1065-1081.

[9] J. Calvet and Y. Arkun, "Feedforward and feedback linearization of non-linear systems and its implementation using internal model control," Ind. Eng. Chem. Res., vol. 27, no. 10, pp. 327*346, 1988.

[10] C.G. Economou, M. Morari, and B. O. Palsson, "Internal model control 5. Extension to nonlinear systems,"Ind. Eng. Chem. Proc. Des. and Dev., vol. 25, pp. 403-409, 1986.

[11] C. E. Garcia and M. Morari,"Internal model control. 1. a unifying review and some results', Industrial Engineering Chemistry Process Design and Development", vol. 21, 1982, pp. 403–411.

Contribution of the Computer Technologies in the Teaching of Physics: Critical Review and Conception of an Interactive Simulation Software

Abdeljalil Métioui
Faculté des sciences de l'éducation
Université du Québec à Montréal
Montréal, Canada

Louis Trudel
Faculté d'éducation
Université d'Ottawa
Ottawa, Canada

***Abstract*—in the present research, we will synthesize the main research results about the development of interactive computer environments for physics teaching and learning. We will see that few types of software propose environments that take into account the user's erroneous representations in order to make him become aware of his mistakes. The majority of these softwares present modelling activities that are restricted to the automatic collection of experimental data and their analysis under graphical form. As a consequence, we will present the design of computer environments for the learning of the phenomenona of absorption and diffusion of light which will take into account the user's initial representations. The design of these environments is divided in five steps: (1) diagnostic of the user's initial representations; (2) confrontation of the user's initial representations by the simulation; (3) reconstruction by the user of his representations following the completion by the user of the simulation; (4) reconstruction of the user's representations following the presentation by the software of scientific information related to the case studied and (5) assessment of the current state of understanding of the user by the software.**

***Keywords*—critical review; physical; conception; interactive environments; representations**

I. Introduction

To facilitate the teaching and the learning of the basic physical concepts, the researchers resort more and more to the computer technologies to create interactive environments. To make easier the design of these environments, we carried on a literature review about the use of the computer technology in physics teaching and learning followed by their critical analysis, restraining our study to environments aiming at the modification of initial representations of the users, our main subject of study.

This review allowed us to identify the main difficulties and challenges linked to the building of such interactive environments, and also provided guidelines to conceive them, the main of which consists in using simulation in a step of conceptual change. Using these guidelines, we present the main stages of a step of conceptual change and the different roles the simulation plays in it. We illustrate the sequence of this step by taking into account the initial representations of the science preservice teachers in primary school concerning the properties of light. Finally, we will present he limits of our study and perspectives for future research.

II. Overview of the Problems With Respect to the Teaching and the Learning of the Physics

According to the organization of cooperation and economic development (OECD), the problems associated with the teaching and learning of sciences last in spite of multiple reforms of the ministerial programs [1]. The difficulties encountered are multiple: (1) the acquisition of the basics concepts of physics is not easy because the pupil must deconstruct the explanatory model to which he refers to explain a given phenomenon. In this regard, the Canadian Council on Learning [2] note that "the effort necessary to integrate scientific knowledge always more complex and often counter-intuitive can provoke the discouragement of the pupil and even divert him from the study of sciences"; (2) the formation of the teachers in didactics [3] and in sciences ([1], [4], [5]) is not sufficient; (3) the allocated time doesn't allow the teachers to study the construction of the concepts of physics through history, in order to analyze the difficulties met by famous scientists to study different natural and constructed phenomena; (4) the number of pupils in a classroom doesn't permit the teacher to take into account the conceptual difficulties of each pupil and (5) the laboratories as dispensed don't succeed to make the pupils see the interactions that exist between the theory and the practice, and unfortunately these laboratories are often reduced to a simple verification of the theory [6]. To overcome these difficulties, some researchers demonstrate that a teaching with the help of software allows the pupils to achieve meaningful learning: "Successful teaching practices have been implemented internationally in a small number of physics classrooms. These often involve strategically planned tutorials, concept cheks in lecture classes and increased opportunities for student discussion" [7]. Such a teaching requires smaller groups of pupils and a strong formation of the teachers, which is difficult to do in our present systems of education. Thus, according to several researchers, the use of technologies in teaching could offer promising avenues for the formation in sciences. What is the situation in the particular case of physics, central point of the present research?

III. Synthesis of Research Works about the Use of Computer Technologies in Physics Teaching

To facilitate the learning of physics, one resorts more and more to the computer technologies to develop activities of modelling and simulation: mechanical and kinematics ([7],

[8], [9], [10], [11], [12]), electric circuits and electromagnetism [13], [14], [15], [16], [17], [18]), optic ([19], [20], [21]) and acoustic phenomena [22]. Although there are a lot of researches on students' representations concerning mechanical or electrical phenomena, the situation is not the same concerning optics' phenomena where researches on students' representations are rather rare. Moreover, the majority of the activities in these researches are mainly concerned with the automated collection of experimental data and their graphic treatment by the user. In this regard, the automated collection refers to the process of data collection that is assured by a probe plugged to an external port of a computer permitting to take, for example, measures of the temperature or the intensity of the electric current. The data collected by such a process is transmitted thereafter and recorded, permitting their treatment by analysis software such as Regressi [11].

According to our literature review and like other research having achieved a critical review of works that conceives simulation software to support science teaching and learning ([23], [24]), one must note that there are not enough simulation software's that create environments allowing the user to detect his own mistakes explicitly and to favor their evolution toward more scientific representations, as it is the case of the works of Muller, Bewes, Sharma and Reimann [7] and those of Baser and Durmus [14], to mention only them. For example, Baser and Durmus have developed strategies centered on the conceptual conflict among preservice teachers of the primary order in the case of the working of simple electric circuits. In their strategy, the user had to complete at first a questionnaire that served to destabilize him on a conceptual level once he will have verified his answers with the help of the simulation. This stage is indispensable to create the first conditions that will make the user's erroneous conceptions evolve toward more scientific ones.

However, in most cases, the simulation doesn't take into account the user's representations and cannot help him to learn, as the research of Yeo, Loss, Zadnik, Harrison and Treagust [25] have shown. According to these authors, the use of simulated experience on the projectiles motion made the students interact superficially since they kept their intuitive representations on the motion which are erroneous. In the same perspective, Zhou, Brouwer, Nocente and Martin [26] showed that it is possible to create the conceptual change while resorting to experiences that will be followed by a discussion initiated by the teacher where students will argue their ideas following the simulations. But for that to happen, the teacher must have the necessary expertise to manage the erroneous representations of the students. The authors synthesize their results according to what follows: "This study demonstrates that computer-based applets, designed in the light of constructivism, can be helpful in fostering conceptual change/learning, but they should be used in a constructivist teaching environment to be more effective. The effectiveness of computer applets is a function of the applet design, the instructional environment in which they are used, and the teacher's readiness for using new instructional technology. Well-designed new media applications must be used in a constructivist teaching environment by enthusiastic teachers to be effective. University science teachers, normally without a pedagogical background, need to become familiar with the teaching suggestions from cognitive studies and the way that new media works in order to be effective teachers." (p. 47).

Besides, we identified in most analyzed works another problem that is associated to the research methodology used. Indeed, the simulated experimentations have not been achieved with an experimental group and a control group to show the relative merits of the simulation and the traditional laboratory.

With respect to this issue, Smetana and Bell [24] underline what follows: "[…], several studies did not include a comparison group. In these cases, the question remains whether students would have learned as much without the simulation, under the traditional instruction. In those studies that employed a variety of instructional interventions, the contribution of the simulation to student outcomes is uncertain without appropriate controls" (p. 1320).

The present research appears in the lineage of research works led by Baser and Durmus [14] in electricity and propose the conception of computer environments in the case of the optics to the teachers in formation and in service for the primary order.

IV. Conception of Computer Environments

The conception of computer environments for the learning of the phenomenon of the absorption and the diffusion of light will be structured in five stages, as illustrated in figure 1.

What do we mean however by the term "conception of computer environments"? We use it in the sense of Tchounikine [27] according to whom the term CEHL (Computer Environments for Human Learning) refers to the acts of imagining, of thinking, of elaborating, of representing a computer artifact taking into account the educational objectives pursued and the various constraints of the educational situation. The terms of realization or construction refer to making executable on a computer, i.e., to program it.

The first stage serves to evaluate the user's initial representations. In order to do so, the user will have to complete a multiple choices questionnaire already included in the software in which every question is formulated as statement and where the user should indicate, while justifying his choice, if it is true or false. The construction of the questionnaire is described at the end of this section. As an illustration, some questions relative to our theme of study are presented in table I.

In a second stage, the user will confront his anticipated answers with the results of the simulation. To this end, a window in the simulation will be at its disposal to validate himself his answer. It will incite him to throw into question his conceptual structure and he will probably want to change his explanations in the case where some of his anticipated answers revealed to be erroneous. If the user wishes it, the software will invite him to rephrase his explanations again (third stage).

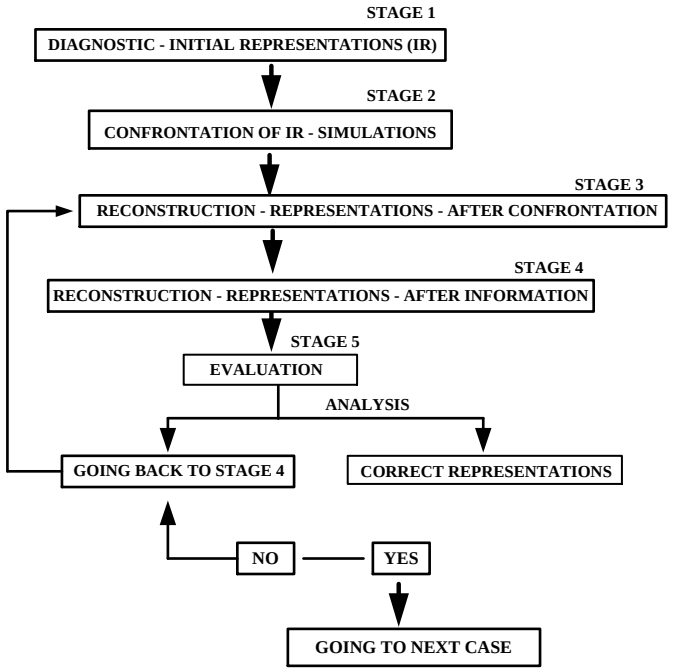


Fig.1. Theoretical schema of Interactive Simulation Software

TABLE I. EXAMPLE OF QUESTIONNAIRE TO MULTIPLE CHOICES - LIGHT PHENOMENA

Question 1: When one illuminates a mirror with a pocket-size lamp, light emitted by the bulb of the pocket-size lamp reaches the mirror and stops there. Justify your choice.	True ☐	False ☐
Question 2: A white sheet of paper is deposited on a table. When one illuminates it with a pocket-size lamp, light emitted by the bulb reaches the paper and stop there. Justify your choice.	True ☐	False ☐
Question 3: One places a yellow tennis ball on a table in a piece illuminated by the light of the day; the bullet will distribute light. Justify your choice.	True ☐	False ☐

According to Zhou [28], this phase of verification with the help of the simulation will facilitate the conceptual change: "The ability for students to visually compare the consequence of their predictions with the realistic process can be helpful in creating cognitive conflict and facilitating conceptual change". (p. 108). In a fourth stage, he will be invited again to rephrase his erroneous answers following the scientific information that will be offered to the user by the software. More precisely, the software will present him a synthesis of the theories, the laws and the scientific models whose acquisition is indispensable to explain the results obtained in the experimentation of the phenomenon and a window of documentation will be dedicated to this end.

The objective is to allow the user to acquire some basic notions with respect to the phenomenon under study, without revealing him however directly the proper scientific explanations.

At this stage, it is likely that he will throw into question his conceptual structure and will rebuild a new structure that, this time, will be correct on a scientific level. In a fifth stage, one will present him again the questionnaire that he has completed at the time of the first stage. For every question, he should evaluate different representations while indicating for each if it is true, partially true, incomplete or false, while justifying his choice (tables II, III and IV).

TABLE II. ASSESSMENT BY THE USER OF DIFFERENT REPRESENTATIONS QUESTION 1

What do you think of the following sentences about lighting a mirror with a pocket-size lamp? Justify your choice.				
	T[1]	**PT[2]**	**I[3]**	**F[4]**
Light emitted by the lamp reaches the mirror and stop there because the mirror absorbs light, it doesn't redirect it.	☐	☐	☐	☐
Light propagates itself in a straight way. If a mirror has a plane surface, light reaches it but doesn't reflect upon it. Light can reflect upon the mirror if it makes an angle with it.	☐	☐	☐	☐
The light reaching the mirror reflects upon it when the mirror has a flat surface.	☐	☐	☐	☐
Light will be reflected by the mirror on the wall in front of the source. The mirror doesn't absorb it.	☐	☐	☐	☐

[1]True; [2]Partially true; [3]Incomplete; [4]False

TABLE III. ASSESSMENT BY THE USER OF DIFFERENT REPRESENTATIONS QUESTION 2

What do you think of the following sentences about the lighting of a white sheet of paper deposited on a table? Justify your choice.				
	T[1]	**PT[2]**	**I[3]**	**F[4]**
Light emitted by the lamp reaches the paper and stops there because light stops where one illuminates.	☐	☐	☐	☐
Light emitted by the lamp reaches the paper and stops there. The white sheet of paper doesn't reflect light because it is an opaque body.	☐	☐	☐	☐
The two bodies are different: light is transparent, whereas the sheet of paper is opaque. It is for that reason that light reaches the paper and stops there.	☐	☐	☐	☐
Light doesn't reflect upon the paper since it is an opaque body.	☐	☐	☐	☐

[1]True; [2]Partially true; [3]Incomplete; [4]False

TABLE IV. ASSESSMENT BY THE USER OF DIFFERENT REPRESENTATIONS QUESTION 3

What do you think of the following sentences about a yellow tennis ball placed on a table in a piece illuminated by the light of the day? Justify your choice.				
	T[1]	**PT[2]**	**I[3]**	**F[4]**
The ball distributes a part of light; the light we see is the light that is reflected by the object (yellow). The object reflects a part of the light of the sun and it absorbs a part of it.	□	□	□	□
The ball doesn't distribute light because it is an opaque body.	□	□	□	□
The ball doesn't distribute light because the ball is not a source of light.	□	□	□	□
The ball doesn't distribute light because the color is not a source of light, it is rather what our eye discerns.	□	□	□	□

[1]True; [2]Partially true; [3]Incomplete; [4]False

In this stage, the user should not resort to the simulation, but rather to refer to his new knowledge and to write down in a window his assessment, as asked. In the case where the user didn't really understand the phenomenon, his answers risk to destabilize it, or even to sow a doubt.

The conception of such an interactive environment must be based upon the identification at first of the representations of the user with respect to the scientific phenomena being studied, here the properties of light. To this end, we analyzed the answers to a paper-pencil questionnaire distributed to a hundred-twenty (120) preservice teachers in a course on the didactics of the sciences. We also took into account the few research works on the representations of the teachers and pupils with respect to the properties of light ([29], [30], [31]). If the analyses done by the user are appropriated, our interactive software will invite him to complete a questionnaire of reinforcement that will ask new questions.

V. CONCLUSION AND LIMIT OF OUR SURVEY

Following our analysis of research works on the development of computer environments, very few pursue the goal of understanding what they really bring to correct the erroneous representations of the pupils, as reported in the international literature, and that the traditional teaching doesn't succeed in correcting, as several organisms, like the OECD [1], confirmed it.

Also, very few research works have tried to see how the multimedia resources could be used to take into account the mistakes of the learners in order to generate the conceptual conflict and to propose environments with which the user must interact in order to re-establish his conceptual balance and that, in the perspective of conceptual change advocated by Posner, Strike, Hewson and Gertzog [32]. Also, few researches worried about the users' requirements in order to use the computer tools adequately, like the graphic representation as a tool of acquisition of knowledge. Several researches showed that the pupils have serious difficulties to use such a graphic representation correctly [32]. The following passage, borrowed from Muller, Bewes, Sharma and Reimann [7] supports our conclusion: « […] limited research has been conducted on how resources like linear multimedia can be altered promote conceptual change. Multimedia research has investigated student learning of scientific topics […], but the issue of misconceptions has rarely been addressed. Studies have also typically been conducted in controlled laboratory environments, with learners who have little or no prior knowledge neither about the subject matter nor experience in the ways of knowing, learning and thinking in the domain. »

Finally, with regard to the second aspect of our research, we intend to pursue our work of conception, notably in the scientific notions that the interactive environment will present to the user (stage 4). For it, we are achieving a conceptual analysis in order to identify the most important scientific notions than the user should acquire. This analysis will give account of the erroneous theories constructed by the scientists during history and will serve to valorize the user's erroneous representations among others. As for the part of realization (according to Tchounikine) of the proposed environments, it will be developed once our work of conception will be advanced. To this topic, we are conscious that the realization of the computer part could bring us to review some elements of our conception for reasons bound to their programming. Thus, we think of resorting to other supports as the video [11] in the case of the experimentations that would be difficult to achieve on the screen of a computer.

REFERENCES

[1] Organization of cooperation and economic development (OCDE). Declining enrolment in science & technology Studies. Is it real? What are the causes? What can be done?, 2005.

[2] Conseil canadien sur l'apprentissage. Apprentissage informel de la science au Canada. Carnet du savoir, 2007.

[3] N. Coppens, G. Rebmann, and V. Munier. "Suivre l'évolution des conceptions des élèves en mécanique : développement et évaluation d'exercices informatisés," Didaskalia, vol. 35, 2009, pp. 37-58.

[4] A. Métioui and L. Trudel. "Quebec secondary physics teachers and modern science: The case of the concept of matter," The International Journal of Science in Society, vol. 3, no 1, 2012a, pp. 177-190.

[5] M. Walport. Science and mathematics secondary education for the 21st century. Report of the Science and Learning Expert Group, 2010. Innovation and Skills (England):

[6] http://interactive.bis.gov.uk/scienceandsociety/site/learning/files/2010/02/Science-and-Learning-Expert-Group-Report-Annexes-31.pdf

[7] N. Nonnon, and A. Métioui. "L'appropriation du processus de vérification expérimentale par des étudiants en formation des maîtres au secondaire : exemple de la diffusion de la chaleur dans un liquide," Res Academica, vol. 21, no 1, 2003, pp. 39-61.

[8] D. A. Muller, J. Bewes, M. D. Sharma and P. Reimann. "Saying the wrong thing: Improving learning with multimedia by including misconceptions," Journal of Computer Assisted Learning, vol. 24, no 2, 2008, pp. 144–155.

[9] M. Couture. "Realism in the design process and credibility of a simulation-based virtual laboratory," Journal of Computer Assisted Learning, vol. 20, no 1, 2004, pp. 40-49.

[10] Z. Smyrnaiou, and A. Weil-Barais. "Évaluation cognitive d'un logiciel de modélisation auprès d'élèves de collège," Didaskalia, vol. 27, 2005, pp. 133-149.

[11] P.-K. Tao and R.-F. Gunstone. "The process of conceptual change in force and motion during computer-supported physics instruction," Journal of research in science teaching, vol. 36,l no 7, 1999, pp. 858-882.

[12] L. Trudel and A. Métioui. "Effect of a video-based laboratory on the high school pupils' understanding of constant speed motion," International Journal of Advanced Computer Science and Applications, vol. 3, no 5, 2012, pp. 71-76.

[13] G.-G. Zhou, W. Brouwer, N. Nocente, and B. Martin. "Enhancing conceptual learning through computer-based applets: The effectiveness and implications," Journal of Interactive Learning Research, vol. 16, no 1, 2005, pp. 31-49.

[14] M. Baser. "Effects of conceptual change and traditional confirmatory simulations on preservice teachers' understanding of direct current circuits," Journal of Science Education and Technology, vol. 15, no 5, 2006, pp. 367-381.

[15] M. Baser and S. Durmus. "The effectiveness of computer supported versus real laboratory inquiry learning environments on the understanding of direct current electricity among pre-service elementary school teachers," Eurasia Journal of Mathematics, Science & Technology Education, vol. 6, no 1, 2010, pp. 47-61.

[16] R. H. Berube. Computer simulated experiments for electric circuits using electronics workbench Multisim, New Jersey: Prentice Hall, 2004.

[17] S. Labrique, D. Grenier, and F. "Labrique. Entre savoir et savoir-être, du rôle des TIC dans le processus d'apprentissage du génie électrique : compte rendu d'innovation," Didaskalia, vol. 24, 2004, pp. 104-131.

[18] M. Ronen, and M. Eliahu. "Simulation - a bridge between theory and reality: The case of electric circuits," Journal of Computer Assisted Learning, vol. 16, 2000, pp. 14-26.

[19] C. Vreman-de Olde, and T. de Jong. "Student-generated assignments about electrical circuits in a computer simulation," International Journal of Science Education, vol. 26, no 7, 2004, pp. 859-873.

[20] C. Buty. "Richesses et limites d'un « modèle matérialisé » informatisé en optique géométrique," Didaskalia, vol. 23, 2003, pp. 39-63.

[21] B. Eylon, M. Ronen, and U. Ganiel. "Computer simulations as tools for teaching and learning: Using a simulation environment in optics," Journal of Science Education and Technology, vol. 5, no 2, 1996, pp. 93-10.

[22] C. Viudez. "Expérimentation d'un environnement informatique dans le cadre de travaux pratiques sur l'interféromètre de Michelson en licence: compte rendu d'innovation," Didaskalia, vol. 23, 2003, pp. 147-161.

[23] A. Séjourné. "Éléments théoriques pour la conception d'un hypermédia en sciences physiques et pour l'analyse de l'activité des élèves : le cas des phénomènes sonores," Didaskalia, vol. 23, 2003, pp. 65-99.

[24] S.-W.-L. Lee, C.-C. Tsai, Y.T. Wu, M.J. Tsai, T.-C. Liu, F.K. Hwang, and C.Y. Chang. "Internet-based science learning: A review of journal publications," International Journal of Science Education, vol. 33, no 14, 2011, pp. 1893-1925.

[25] L.K. Smetana and R.L. Bell. "Computer simulations to support science instruction and learning: A critical review of the literature," International Journal of Science Education, vol. 34, no 9, 2012, pp. 1337-1370.

[26] S. Yeo, R. Loss, M. Zadnik, A. Harrison, and D.F. Treagust. "What do students really learn from interactive multimedia? A physics case study," American Journal of Physics, vol. 72, no 10, 2004, pp. 1351-1358.

[27] G.G. Zhou, W. Brouwer, N. Nocente, and B. Martin. "Enhancing conceptual learning through computer-based applets: The effectiveness and implications," Journal of Interactive learning Research, vol. 16, no 1, 2005, pp. 31-49.

[28] P. Tchounikine. Précis de recherche en ingénierie des EIAH, 2009. Récupéré le 18 août 2012 du site de l'auteur :

[29] http://membres-liglab.imag.fr/tchounikine/Precis.html

[30] G.G. Zhou. "Conceptual change in science: A process of argumentation," Eurasia Journal of Mathematics, Science & Technology Education, 6(2), 2010, pp. 101-110.

[31] A. Métioui and L. Trudel. "The model of the rectilinear propagation of light and the study of the variation of the size of a shadow," US-China Education Review, vol. 2, no 9, 2012b, pp. 173-186.

[32] [30] K. Ravanis, K. Zacharos, and A. Vellopoulou. "The formation of shadows: The case of the position of a light source in relevance to the shadow," Acta Didactica Napocensia, vol. 3, no 3, 2010, pp. 1-6.

[33] N.J. Selley. "Children's ideas on light and vision," International Journal of Science Education, vol. 18, no 6, 1996, pp. 713-723.

[34] G. Posner, K. Strike, P. Hewson, and W. Gertzog. "Accomodation of a scientific conception: Toward a theory of conceptual change," Science Education, vol. 66, no 2, 1982, pp. 211-228.

[35] W.M. Roth, and M.K. McGinn. "Graphing: a cognitive ability or a cultural practice?," Science Education, vol. 81, 1997, pp. 91-106.

VI. AUTHORS PROFILE

Abdeljalil Métioui is a professor at the Université du Québec à Montréal. He has a M.A. in physics and a Ph.D. in Didactics of Physics. He has taught Didactics of the Sciences at Université Laval (Quebec), Université de Moncton (New-Brunswick) and Université Sainte-Anne (Nova Scotia). He has directed research programs in science-teaching and technologies and published numerous articles, as well as given papers on students' and teachers' alternative frameworks in science. Recently he co-directed the publication of one book on the history of the sciences on the teaching.

Louis Trudel is professor at the faculty of education at University of Ottawa. He was awarded a Ph.D. in education at the Université du Québec à Montréal. He has published several papers on science education and the pedagogical applications of computer science to the field of education.

A Novel Permutation Based Approach for Effective and Efficient Representation of Face Images under Varying Illuminations

Shylaja S S*, K N Balasubramanya Murthy, S Natarajan
Department of Information Science and Engineering,
P E S Institute of Technology,
Bangalore, Karnataka, INDIA

***Abstract*—Paramount importance for an automated face recognition system is the ability to enhance discriminatory power with a low-dimensional feature representation. Keeping this as a focal point, we present a novel approach for face recognition by formulating the problem of face tagging in terms of permutation. Using a fundamental concept that, dominant pixels of a person will remain dominant under varying illuminations, we develop a Permutation Matrix (PM) based approach for representing face images. The proposed method is extensively evaluated on several benchmark databases under different exemplary evaluation protocols reported in the literature. Experimental results and comparative study with state-of-the-art methods suggest that the proposed approach provides a better representation of face, thereby achieving higher efficacy and lower error rates.**

Keywords—Biometrics; Face Recognition; Independent Component Analysis (ICA); Linear Discriminant Analysis (LDA);Locality Preserving Projections (LPP); Pattern Recognition; Permutation Matrix (PM); Singular Value Decomposition (SVD).

I. INTRODUCTION

Facial Recognition, one of the biometric techniques is a form of computer vision that uses faces to identify a person or verify a person's claimed identity. The main attraction for Face recognition research over other biometric techniques is that it is nonintrusive in nature. This has paved path to several applications like security, access control, surveillance system, human computer interface, recognition in galleries and consumer devices. Significant amount of work is carried out by Research Scientists and Engineers in the field of Face recognition and is still growing exponentially. In spite of these efforts, studies reveal [1,2] that there is no single state of art face recognition system. The main reason quoted is that applications in this area have varying requirements and constraints. Therefore there has been always growing interest for developing newer computational algorithms that can serve as a model for human face recognition function.

Recognition of faces primarily involves collection of descriptive measurements known as feature vectors extracted from each training set of images. These can be facial components, such as eyes, nose, and mouth and facial outline [30]. Feature extraction is performed to provide effective information that is useful for distinguishing between faces of different persons and to be stable with respect to the geometrical and photometrical variations.

Face Recognition has been tackled in two modes, one focusing on generic face recognition algorithms and other on specific face recognition algorithms. Generic face recognition algorithms should work well with databases having complex conditions like combination of pose, illumination and expressions variations. Some of the example methods experimented in this direction are Singular Value Decomposition (SVD), Linear Descriminant Analysis (LDA), Locality Preserving Projections (LPP), Hidden Markov models (HMM), Gaussian Mixture Model (GMM) [3 – 17]. These systems work on extracting face manifold in a subspace. On the other hand, specific face recognition method requires preprocessing of database to meet a specific condition, keeping other conditions invariant. Spherical representations have been used for modeling illumination variations or both illumination and pose variations in face images [18]. Mian et al. [19] have handled the expression problem using a fusion scheme in which three kinds of methods, spherical face representations, Automatic preprocessing of 3D Spherical Face Representation (SFR), Scale-Invariant Feature Transform (SIFT)-based face data matching and a modified iterative closest point(ICP) have been combined to achieve the final result. Their results have showed the potential of appearance-based methods for solving the expression problem [20].The models for illumination range from highly specular objects such as mirrors to models for matte objects. Most objects belong to the latter category, which is described by a Lambertian reflectance model. Early shape from shading approaches assumed a constant albedo field. But they have shown to be violating [21] at locations such as eyes and edges of mouth. For the human face, the Lambertian reflectance model with a varying albedo field has been shown to provide a reasonable approximation.

Pose variation is essentially a correspondence problem. With an assumption of Lambertian reflectance model, dense correspondences across poses are available with a rank-1 constraint implied. Recovering a 3D model from 2D images is a difficult task. Two types of approaches have been around: [22-25] one using model-based and the other using the image-based approach. Explicit [26, 27] knowledge of prior 3D models have been used in Model-based approaches. Image-based approaches have not used prior 3D models. In general, model-based approaches register the 2D face [28] image to 3D models that are given beforehand.

II. PERMUTATION MATRIX – AN OVERVIEW

PM is a square binary matrix that has exactly one entry 1 in each row and each column and 0's elsewhere. Each such matrix represents a specific permutation of m elements and, when used to multiply another matrix, can produce that permutation in the rows or columns of the other matrix [33].

PM is obtained by reordering the rows of an identity matrix I. A typical PM looks like

$$PM = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

It is obtained from 4X4 Identity matrix by cyclic permutation of the columns C1->C2->C3->C4->C1.

In LU factorization problem the PM has been very useful. In case of LU factorization without pivoting, A = LU, where L unit lower triangular, U upper triangular. This may not always exist even if A is nonsingular. In the case of LU factorization with row pivoting, A = PLU, where P is the PM, L is the unit lower triangular and U is the upper triangular matrices. This exists for all nonsingular A's.

The product PA is the matrix A with its rows reordered in exactly the way as the rows of I were reordered to produce P. Well written software for solving systems of linear equations chooses the pivot elements carefully to ameliorate the bad effects of round off error. The natural order of the rows gives way to an ordering that pays attention to the magnitudes of the elements available as pivots. There is a necessity to choose strong pivot elements to avoid growth of round off errors. Applying P reorders or permutes the rows of A so that the re-ordered matrix has a factorization

$$\boldsymbol{P^T A = LU}$$

III. PROPOSED METHOD

A. Introduction

The method proposed in this paper is based on generation of a PM. It captures the dominant values in the image matrix thus suppressing the less dominant values. The idea is that, between two samples of the same person images relativity is maintained in dominant places. Even if the variations intensity values do occur, it is usually found that the dominant values continue to be dominating. This is found after an exploration is done on the data set at hand. The data sets have several kinds of variations and in spite of these; variations found at leading positions are not much. The algorithm scans the columns for the leading and non-leading elements and assigns weights for them differentially. The PM thus generated is binarized. The advantage with such a matrix is that it is sparse in nature.

The mathematical representation is given below:

Let the input square matrix A of order n be represented as

$$A=\begin{bmatrix} a_{11} & a_{12.} \ldots\ldots & a_{1n} \\ a_{21} & a_{22} \ldots\ldots & a_{2n} \\ a_{n1} & a_{n2} \ldots\ldots & a_{nn} \end{bmatrix}$$

The notation A (i, j) denotes the element located i^{th} row and jth column of A.

B. Sequence Generation

Let P={p1,….,pk} be the set of permutation matrices corresponding to the gallery image set G. Each image in the gallery set is a square matrix with a resolution of nXn and k is the number of images in the gallery set. For each p € P, generate sequence vectors S={s1,….sk}, where each s € S is a sequence vector generated from non-zero entries of PM of the image. The value s generated is the max (column of p) for each column excluding earlier considered row and column values. Similarly, the permutation matrices and the sequences are generated for the set Q, the probe set. Let the set PQ={pq1,….pql} represent the sequence set of the probe image after quantizing the probe with application of permutation. Finally the tagged images consisting of the permutation sequences are compared.

The computation of similarity score Si(n) is carried out as follows: Given any probe q ∈ Q , it is first mapped it onto the feature space. Formally, we have x = F(q) ∈ Rd , where x is the feature vector extracted from the probe p through the feature extraction function F(), and Rd is the d-dimensional feature space. Similarly the feature vector of every image gi in gallery G is extracted through F(). Finally, the probe set Q is scored against the gallery G by comparing similarity scores Si(n) based on F(Q) for each qi ∈ Q and gn ∈ G. We assume that smaller similarity score implies closer match. Thus, the objective of a permutation based face recognition task is to determine the identity of q, i.e., ID(q), where ID(q) ∈ ID{g1, . . . ,gK } using the concept of PM.

For each PQ find the distance between gallery's permutation sequence and the probe sequence and the result is represented as set R={r1….rk}.We then compute d= min(R), the min distance between the probe and gallery.

If the d is below the thresholdϵ, the probe is accepted as a match otherwise marked as unknown. The image score is not only checked against the top score but also tested for other ranks where rank L represents the image match in first L scores. The one which yields few permutations will be marked as the matched one. The limiting case for the permutation in most cases is zero. The block diagram in Fig 1 depicts the overall process. The proposed method has been compared against SVD, LDA, LPP and ICA techniques that are being used for face recognition.

C. Pseudocode

```
colMax = [] #Our 1xn feature vector
for i = 0 to n: #Number of iterations
      maxEntry = -1
for r = i to n: #row
   for c = i to n: #column
      if maxEntry < A[r][c]: #Get index of column with
                              biggest entry.
         maxEntry = A[r][c]
         colMax[i] = c
        return colMax[] #resulting feature vector for
                         image, A.
```

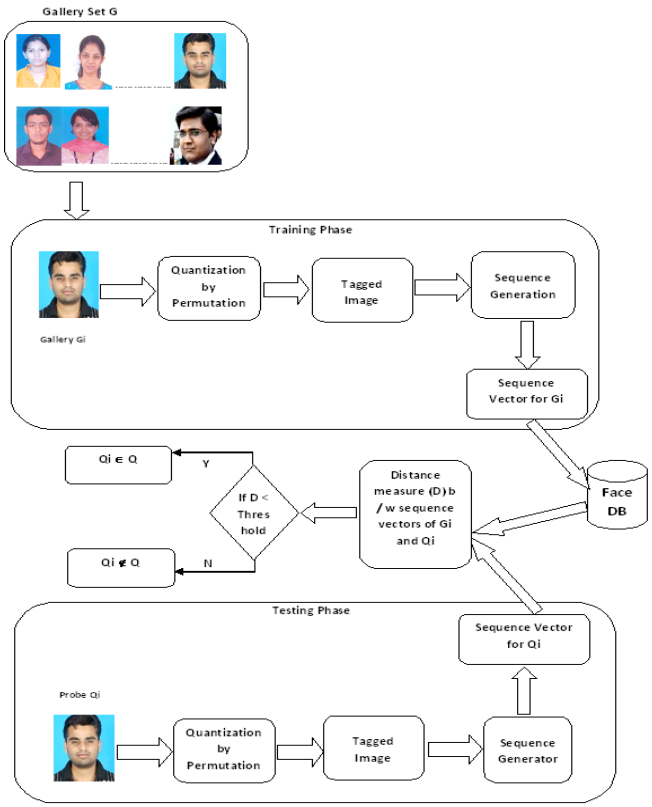


Fig.1. Block Diagram of the proposed method

D. Computational Complexity

The number of comparisons required for generating PM is:

$$\sum_{i=1}^{n-1}(n-i)$$

Where n is the order of the P matrix

$$\sum_{i=1}^{n-1}(n-i)=(n-1)+(n-2)+\ldots+2+1$$

$$=\frac{n(n-1)}{2}$$

$$=\frac{n^2-n}{2}$$

$$=\theta(n^2)$$

The reduction in number of comparisons per image using PM when compared to n2 algorithms is

$$n^2-\frac{n(n-1)}{2}$$

$$\text{Efficiency Factor (EF)}=\left[1-\left\{\frac{\sum_{i=1}^{n-1}(n-i)}{n^2}\right\}\right]X100 \qquad .....(4)$$

Now simplifying $\dfrac{\sum_{i=1}^{n-1}\left(1-\frac{i}{n}\right)}{n}$

$$\frac{\sum_{i=1}^{n-1}\left(1-\frac{i}{n}\right)}{n}=\frac{\left(1-\frac{1}{n}\right)+\left(1-\frac{2}{n}\right)+\left(1-\frac{3}{n}\right)\ldots++\left(1-\frac{n-1}{n}\right)}{n}$$

$$=\frac{(n-1)-\frac{1}{n}(n-1)\frac{n}{2}}{n}$$

$$=\frac{1}{2}-\frac{1}{2n} \qquad ....(5)$$

Now, substituting (5) in (4) we get,

$$=\left[1-\left(\frac{1}{2}-\frac{1}{2n}\right)\right]X100$$

$$=\left[\frac{1}{2}+\frac{1}{2n}\right]X100$$

From a computational complexity point of view, the PM method has a saving in cost of efficiency by a factor of 49.5% for a matrix size of 100X100. Higher the resolution of the image, higher will be the efficiency. The permutation sequence generated from PM has size n. The advantage of this method is that the sequence generation of an image doesn't need the knowledge of database images and is produced on the fly.

IV. Results And Discussions

A. Experimental Setup

Extensive experiments were carried out to investigate the efficacy of our proposed approach for face recognition. Essentially, six benchmark databases namely, ORL [35], Face 95 [37], Face96 [37], Yale [34], Pose [36] and Grimace [37] have been utilized for experimentations. Fig. 2 shows the selected representative subjects in the six databases. These databases incorporate several deviations from the ideal conditions, including pose, illumination, occlusion, and gesture.

Several standard evaluation protocols reported in the face recognition literature including statistical metrics, CMC curve and ROC plots have been adopted. We compare our algorithm with four state-of-the-art baseline techniques for face recognition namely SVD [29,3], LDA[6], LPP [31] and ICA [32]. It is appropriate to indicate that the developed approach has been shown to perform well for the cases of severe illumination and expression with little change in pose, scale and rotation.

Sample images from YALE database (Courtesy: Yale University [34])

Sample images from ORL database (Courtesy: AT & T Labs [35])

Sample images from POSE database (Courtesy: CMU [36])

Sample images from GRIMACE database (Courtesy: University of Essex [37])

Sample images from FACE95 database (Courtesy: University of Essex [37])

Sample images from FACE96 database (Courtesy: University of Essex [37])

Fig.2. Sample Face Images from Benchmark Databases

The following table 1 gives a summary of these databases.

TABLE I. Summary of Benchmark Database

Face Database	No. of Subjects	No. of Images	Variations
FACE 96	152	3040	Expression, Illumination and Pose
FACE 95	72	1440	Expression, Illumination and Pose
GRIMACE	20	360	Expression, Pose and Illumination
ORL	40	400	Pose and Expression
POSE	20	575	Pose and Expression
YALE	15	165	Expression and Illumination

Organization of database is one of the important steps in experimentation setup. In this direction, the databases are organized into two groups, gallery and probe. Thus, all experiments were performed with two experimentation protocols. The first protocol consisted of one mage per subject in the gallery with two different resolutions and the second protocol was similar to first except that, it contained five images per subject in the gallery set. . There was no overlap between the gallery and probe sets. Few subjects were used for unknown set. The size of gallery and probe for each of the databases under different experimental configurations is tabulated below. Original images were normalized and were resized to 100X100 pixels and 180X180 pixels for the all databases. Further all images were formatted with 256 gray levels per pixel.

TABLE II. Size of gallery, probe and unknown sets for different databases with five images per person in Gallery

Database Name	GSET 1	PSET 1	GSET 5	PSET 5	USET
FACE 96	144	2736	716	2164	240
FACE 95	66	1254	328	989	120
ORL	40	320	180	200	60
YALE	243	315	309	249	25
POSE	20	530	101	450	24
GRIMACE	18	85	120	220	20

B. Features Extraction

Feature extraction is carried out using four predominant existing approaches namely SVD, LDA, LPP, ICA and our newly proposed PM approach. The features are extracted from the gallery set and probe set. The face representations obtained from these techniques are illustrated below:

- SVD: Aims at identifying a lower dimensional space maximizing the variance among the data.

Fig.3. Sample Eigen Faces

- LPP: Aims at finding linear transformation-preserving local structure information of original data.

Fig.4. Sample Laplacian Faces

- LDA: Objective is to identify a lower dimensional space minimizing the interclass similarity while maximizing the intra-class similarity simultaneously.

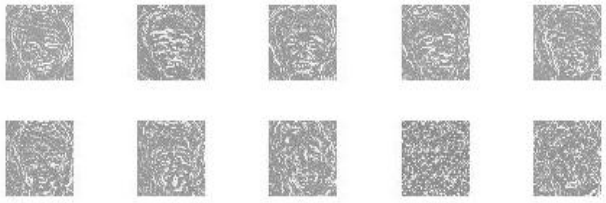


Fig.5. Sample Fisher Faces

- ICA: A function from an m-dimensional space to an n-dimensional space such that the transformed variables

give information on the data that is otherwise hidden in the large data set.

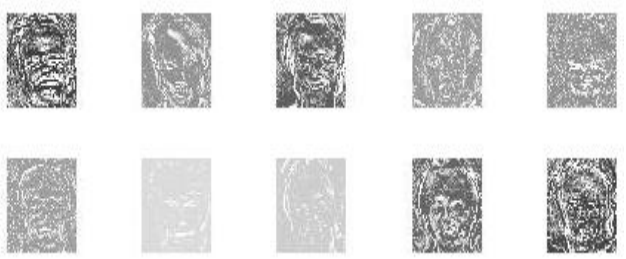


Fig.6. Sample ICA Faces

- Permutation: Aims at identifying dominant intensity values which are likely to be dominant even under variations in two samples of images.

Fig.7. Sample Permuted Faces

C. Recognition Accuracy

After having extracted features from the gallery images and probe images, recognition is possible with relatively little computational effort by comparing sequence vectors. The difference between probe sequence and gallery image sequence is found and the number of zeros is counted. More the number of zeros, closer is the match. This gives the similarity measure in our case. The experiments were carried out in Matlab 7.5 on a 32 bit Core 2 Duo, 2 GHz processor, with 8 GB RAM.

TABLE III. Recognition rates (%) of SVD, LDA, LPP, ICA and PM based methods on query images of 100 X 100 resolutions and with 1 image per subject in gallery

Database Name	SVD	LDA	LPP	ICA	PM
FACE 96	88.59	NA	70.33	71.44	89.97
FACE 95	86.45	NA	69.26	69.97	87.97
ORL	86.50	NA	83.50	76.50	88.00
YALE	87.20	NA	83.20	82.80	88.80
POSE	87.13	NA	76.58	79.32	88.40
GRIMACE	87.60	NA	81.60	85.20	88.80

The recognition rates we have been able to obtain with the different methods including PM are reported in table 3 to 6. Consistent results have been obtained using PM. The relative advantage of PM based over other techniques is quite evident from the recognition accuracy, error rates and other trivial metrics. The proposed PM approach copes well with the problem of illumination and expressions in the presence of small pose variations, achieving high recognition rates of 95.22%. The PM approach, however, shows slight degradation in recognition rates by 2% to 3% for the severe pose variations. It is closer to SVD in terms of accuracy and substantially performs better than LDA, LPP and ICA.

TABLE IV. Recognition rates (%) of SVD, LDA, LPP, ICA and PM based methods on query images of 180 X 180 resolutions and with 1 image per subject in gallery

Database Name	SVD	LDA	LPP	ICA	PM
FACE 96	90.34	NA	70.33	73.06	91.04
FACE 95	88.47	NA	71.59	71.08	89.99
ORL	89	NA	85	78	90.5
YALE	89.60	NA	85.20	84.00	91.20
POSE	88.61	NA	78.48	80.17	90.51
GRIMACE	88.80	NA	83.6	87.6	90.8

PM method is simpler, has higher recognition speed and smaller memory requirements. A key success of PM approach is the unique way of extracting relevant information. More precisely, it exploits only dominant features.

TABLE V. Recognition rates (%) of SVD, LDA, LPP, ICA and PM based methods on query images of 100 X 100 resolutions and with 5 images per subject in gallery

Database Name	SVD	LDA	LPP	ICA	PM
FACE 96	90.06	89.97	73.01	72.00	91.04
FACE 95	89.08	84.83	72.70	70.07	90.09
ORL	88.50	88.00	84.00	77.00	90.00
YALE	88.40	86.00	83.20	82.80	90.80
POSE	89.03	86.08	78.06	78.90	90.93
GRIMACE	89.20	86.00	83.60	86.80	91.20

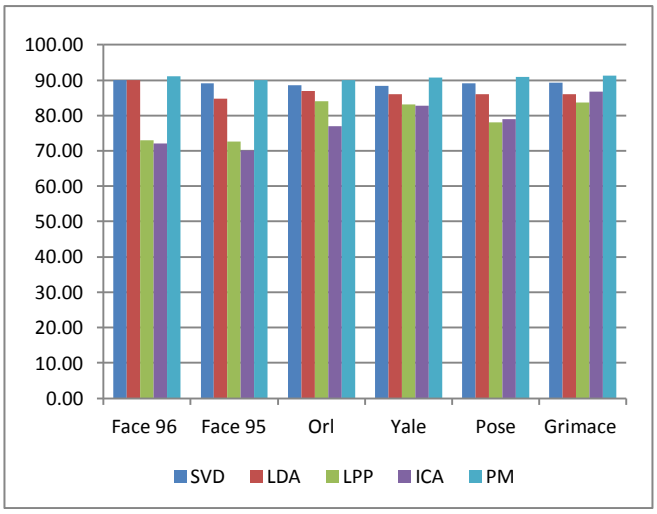


Fig.8. Performance of various methods with 5 images per person in gallery and each image with 100 X 100 resolution

TABLE VI. Recognition rates (%) of SVD, LDA, LPP, ICA and PM based methods on Query images of 180 X 180 resolutions and with 5 images per subject in gallery

Database Name	SVD	LDA	LPP	ICA	PM
FACE 96	94.36	91.77	72.92	74.72	95.22
FACE 95	90.80	88.27	74.92	73.10	91.81
ORL	92.00	89.50	86.50	80.50	94.00
YALE	92.80	89.6 0	89.20	86.00	94.00
POSE	92.62	89.24	81.22	79.54	93.67
GRIMACE	92.80	91.20	89.20	89.60	94.00

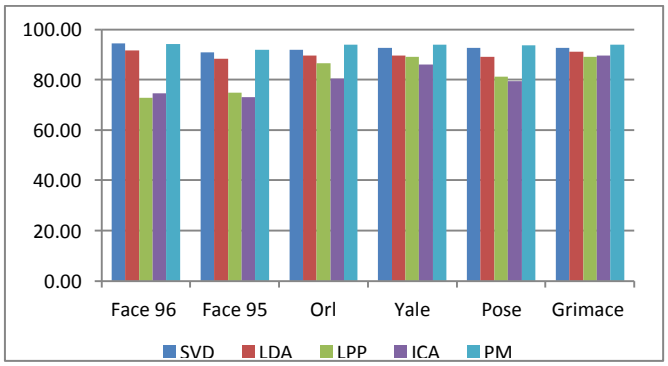


Fig.9. Performance of various methods with 5 images per person in gallery and each image with 180 X 180 resolution

D. Variation Of Gallery Set

Table 3 to 6 shows the results of varying the number of images per subject used in the gallery set. It is evident that LDA and LPP are more sensitive to the particular choice of the gallery set. In other words, both LDA and LPP suffer from Small Sample Size (SSS) problem which makes them sometimes perform poorer. Apart from this LPP's performance is sensitive to the values for the neighborhood size and the similarity matrix. Thus, PM based approach like SVD do not suffer from SSS problem.

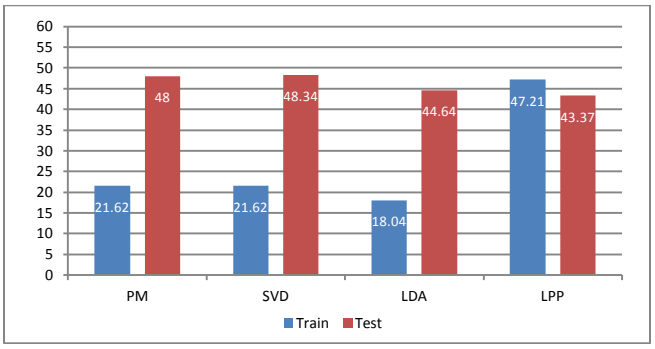


Fig.10. Test and Train Time for Different Methods on Face 95 Database

E. Train and Test Time Requirements

In this experiment, the time requirements of each method are discussed. Fig. 10 and 11 gives the runtime for training and testing phase for various methods on Face 95 and ORL databases respectively. It can be seen that PM, SVD and LDA have minimal time requirements in terms of training and testing. However, time requirements for LPP method are slightly higher because of its computational complexity.

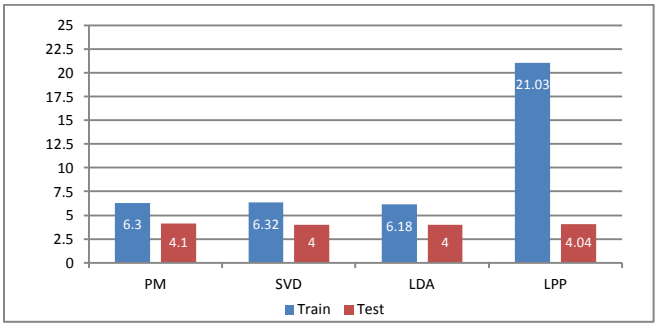


Fig.11. Test and Train Time for Different Methods on ORL Database

V. Performance Evaluation

A. CMC and ROC Plots

To measure respectively the performance of the proposed method for verification and identification tasks, we present both ROC curves and CMC curves for the results obtained in this paper.

The CMC curve for SVD, LDA, LPP, ICA and PM based face recognition on Grimace database is plotted in figure 12 (a) and the corresponding match score for rank 1 to 17 are tabulated in the table 8. Based on the CMC curve and tabular values, it is evident that PM has faster convergence with 94% at Rank 1 and 100% accuracy at Rank 9. Similarly, the CMC plot for the five methods on Face 95 is illustrated in figure 12 (b). The corresponding match score is tabulated in the table 9. The results obtained evaluate here, the PM converged with 100% at rank 9.

TABLE VIII. TOP 17 Rank Scores of PM, SVD, LDA, LPP and ICA on Grimace database

Method	1	3	5	7	9	11	15	17
PM	94.00	96.45	98	99.2	100	100	100	100
SVD	92.80	95	96.5	97.8	99	99.5	99.8	100
LDA	91.2	93.2	94	95.8	96.7	97.5	98	98.2
LPP	89.2	91.5	92	93.6	94.2	95.6	96.7	97.5
ICA	89.6	92	93.8	94.2	95	96.2	96.8	98

We plot CMC and tabulate results obtained on ORL database against these methods in figure 12 (c) and table 10 respectively. Here, the convergence is still faster with full accuracy at Rank 5. Figure 12 (d) and (e) illustrates the CMCs plotted against the POSE and YALE databases with respect to the five methods. Ranks scores for POSE and YALE have been tabulated in table 11 and 12 respectively. The convergence rate with 100% accuracy is always faster with PM method which is evident from the test conducted on several databases.

TABLE IX. TOP 17 Rank Scores of PM, SVD, LDA, LPP and ICA on FACE 95 database

Method	1	3	5	7	9	11	15	17
PM	91.81	96.2	97.5	99.2	100	100	100	100
SVD	90.80	93.6	94.5	95	96.2	97	97.92	98.2
LDA	88.27	91.92	93	94.2	95	95.8	96.2	97
LPP	74.92	78.5	80	82.92	84.2	85	86.5	88
ICA	73.10	76	78.2	82	83.8	84.92	87	89

TABLE X. TOP 17 Rank Scores of PM, SVD, LDA, LPP and ICA on ORL database

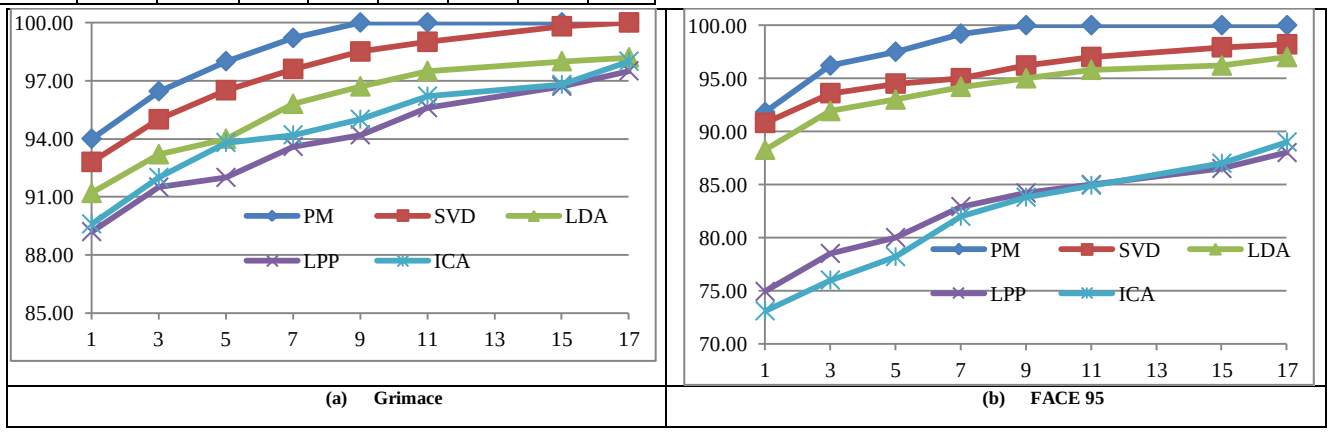
Method	1	3	5	7	9	11	15	17
PM	94.00	99	100	100	100	100	100	100
SVD	92.00	94.8	97.2	98.6	100	100	100	100
LDA	89.5	92.8	94.2	96	96.5	96.5	96.5	96.5
LPP	86.5	89	90	92.6	92.6	93	93	93
ICA	80.5	83	84.8	86	87.2	88	88.5	88.5

TABLE XI. TOP 17 Rank Scores of PM, SVD, LDA, LPP and ICA on POSE database

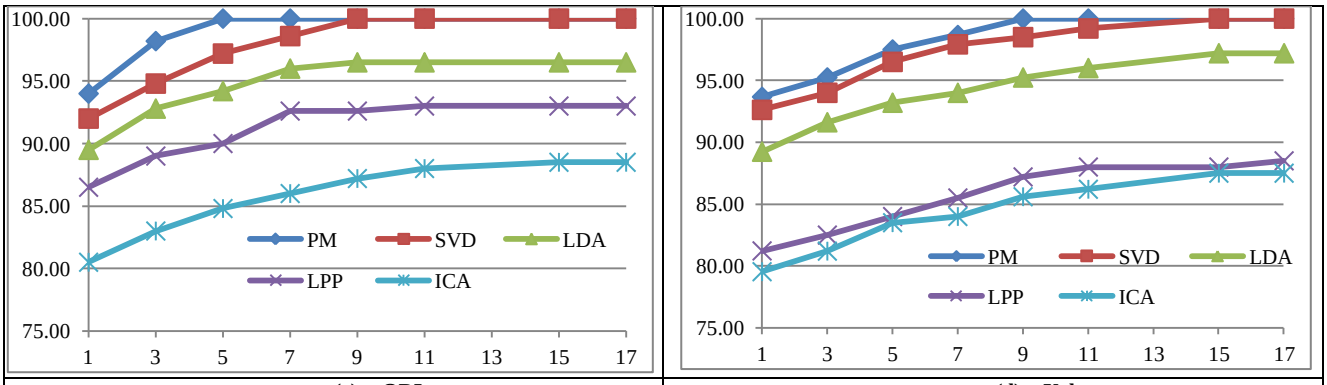
Method	1	3	5	7	9	11	15	17
PM	93.67	95.24	97.5	98.67	100	100	100	100
SVD	92.62	94	96.5	97.92	98.5	99.2	100	100
LDA	89.24	91.62	93.2	94	95.24	96	97.2	97.2
LPP	81.22	82.5	84	85.5	87.2	88	88	88.5
ICA	79.54	81.2	83.5	84	85.6	86.2	87.5	87.5

TABLE XII. TOP 17 Rank Scores of PM, SVD, LDA, LPP and ICA on YALE database

Method	1	3	5	7	9	11	15	17
PM	94.00	97.8	98.6	100	100	100	100	100
SVD	92.80	95	96.5	97.92	98.5	99.6	100	100
LDA	89.6	92.8	93.92	94.2	95.5	95.5	95.5	96
LPP	89.2	91.6	92.8	94.6	95.2	96	96	96
ICA	86	87.5	88.2	88.92	90	91.5	92.2	92.2



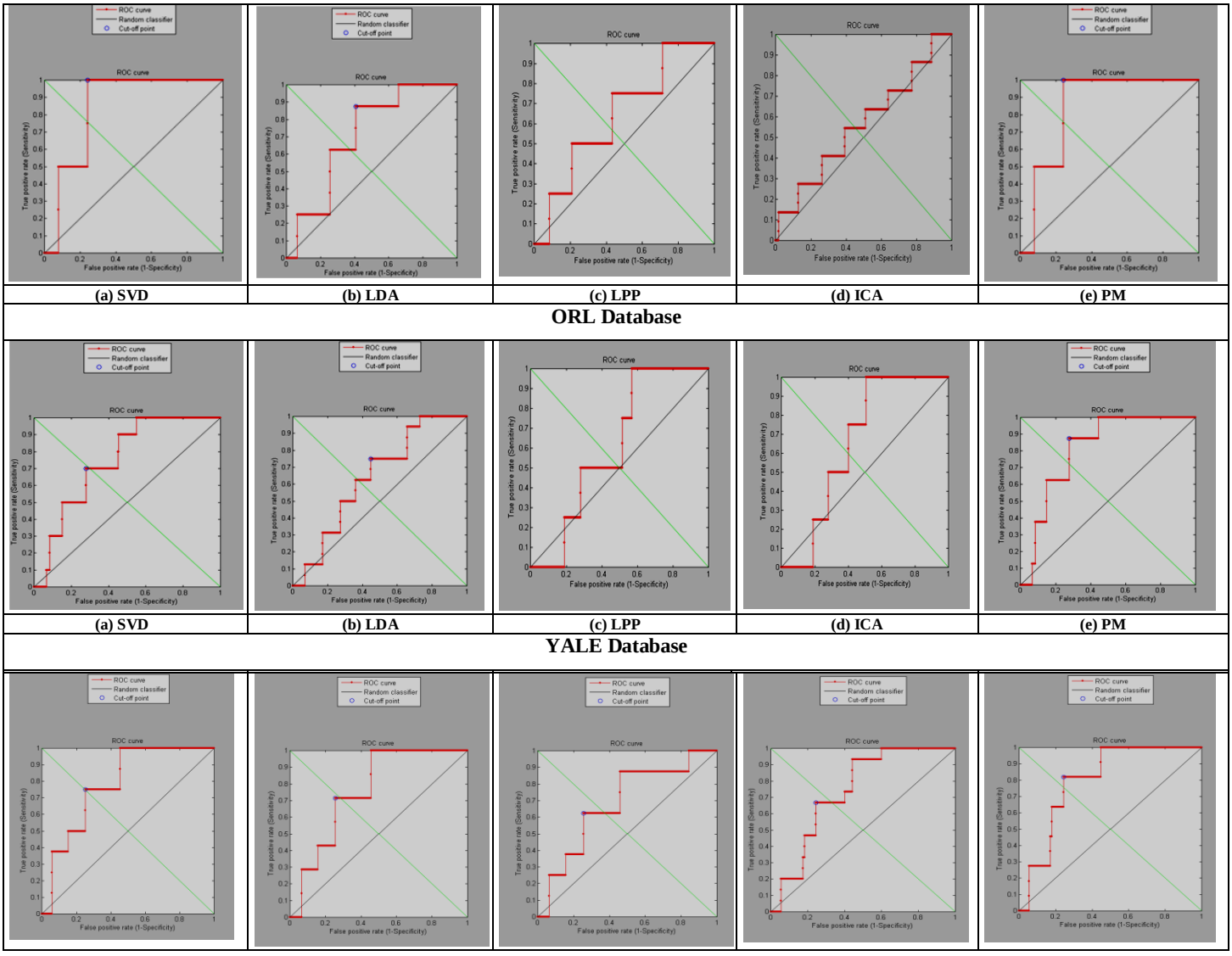
(a) Grimace **(b) FACE 95**



(c) ORL (d) Yale

Fig.12. CMC plots for SVD, LDA, LPP, ICA and PM based method on GRIMACE, FACE 95, ORL, POSE and YALE databases

The ROC plots for different methods are depicted in the Fig. 13. The area under ROC is statistically greater than 0.5 for PM, SVD, LDA and some cases for LPP. The blue circle indicates the cut-off point for best Sensitivity and Specificity. The cut-off point for PM varies from 0.9 to 1.0 which is an indicator of effectiveness of the method used. In all of the test cases the plot is well above the guess line and reaching to almost top left corner as desired on some databases. The cut-off point indicates the algorithm's strength in identifying true and false positives.



(a) SVD (b) LDA (c) LPP (d) ICA (e) PM

ORL Database

(a) SVD (b) LDA (c) LPP (d) ICA (e) PM

YALE Database

(a) SVD	(b) LDA	(c) LPP	(d) ICA	(e) PM
GRIMACE Database				

Fig.13. ROC Graphs predicting True Positive vs. False Positive Rate of SVD, LDA, LPP, ICA and PM

B. Contingency Table

Face Recognition is a pattern classification problem, in which the outcomes are labeled either as positive (p) or negative (n). There are four possible outcomes namely, if the outcome from a classification is p and the actual value is also p, then it is called a true positive (TP); however if the actual value is n then it is said to be a false positive (FP). Conversely, a true negative (TN) has occurred when both the classification outcome and the actual value are n, and false negative (FN) is when the classification outcome is n while the actual value is p. We define two experiments from P positive instances and N negative instances on Face95 and ORL databases. The four outcomes are depicted in a 2×2 contingency table or confusion matrix, as shown in Fig. 14.

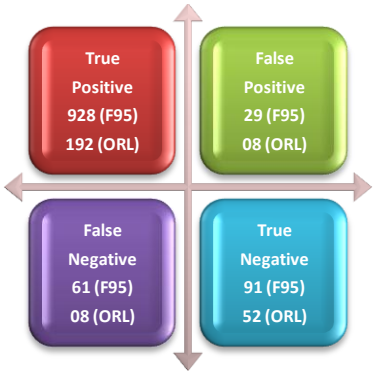


Fig.14. Contingency Table for Face 95 and ORL

The contingency table can derive several evaluation metrics. Some of them used in our experiments are:

- True Positive Rate / Recall / Sensitivity = TP/P = TP / (TP + FN)
- True Negative Rate / Specificity= TN / (FP + TN)
- False Positive Rate = FP/N = FP / (FP + TN)
- Accuracy = (TP + TN) / (P+N)
- Error Rate = (FN + FP) / (P+N)
- Precision / Positive Predictive Value (PPV) = TP / (TP + FP)
- Negative Predictive Value (NPV)= TN / (FN + TN)

TABLE XIII. Various Biometric Metrics Computed for PM on Face95 and ORL databases

Metrics	Face 95	ORL
TPR	0.938	0.96
FPR	0.241	0.133
TNR	0.75	0.86
PPV	0.97	0.96
NPV	0.59	0.86
ACC	0.918	0.938
ER	0.08	0.06

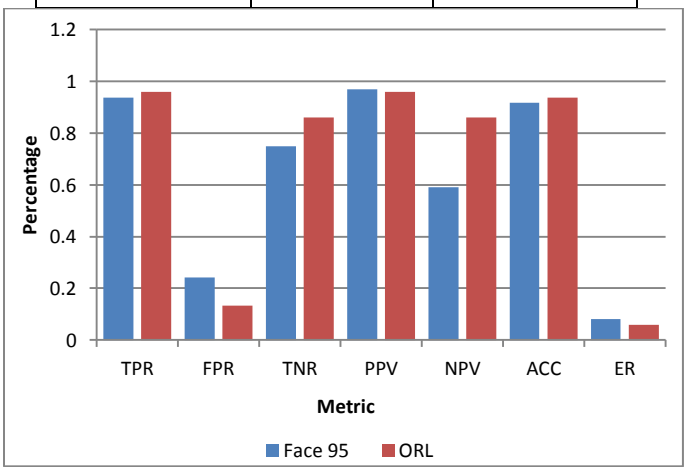


Fig.15. Graph indicating various metrics for PM on Face 95 and ORL Databases.

TABLE XIV. Trivial Biometric Metrics Computed for PM on Face95 and ORL databases

Metrics	Face 95	ORL
F-Score	0.95	0.94
Precision gain	1.08	1.26
Accuracy gain	1.02	1.26

Typically, a good biometric system must have low error rate, FPR and should have higher TPR, TNR, PPV, NVP, ACC. From the contingency table (Fig. 14), various metrics (Table 13 and Fig. 15), we can infer that the proposed technique satisfies these requirements. Apart from these metrics, many performance evaluation metrics like F-Score and Gain have been stated as standard metrics in many literatures. Table 14 depicts the score that we have been able to achieve for these metrics.

VI. Conclusion and Future Enhancements

In summary, PM based approach seems to offer satisfactory results for face recognition under varying illumination, expression and pose to certain extent. The results obtained have been verified on larger and benchmark databases. A more important problem that we did consider here is the reliable rejection of images of faces not contained in the data bases. There are many possible ways to improve the robustness of our scheme, for instance, by expanding the set of faces images in the gallery with all possible face variations. Considerable comparative analysis with the state of-the-art algorithms clearly reflects the relative potency of the proposed approach.

However, these results are absolute in terms of the experimental set up including the database selection and segregation, parameters chosen, etc.

The proposed PM approach reveals a number of interesting outcomes namely, faster convergence rate of recognition, reduction in feature vector size and saving in time complexity for generation and transmission of feature vector. Considering its efficacy at low-subspace dimensionality and its relative simplicity, the PM based approach is a highly effective subspace modeling technique for surveillance systems. The error rate is well below 0.1 which indicate the robustness of the algorithm proposed and is quite satisfactory.

Future work is suggested towards working on local permutation matrices instead of global holistic one. Yet another dimension would be to look at newer ways of generating permutations. The research can also be directed towards assigning differential weights to image matrix.

Acknowledgment

The authors would like to *acknowledge* the support provided by Mr. Kiran Kumar and Mr. Vinay, during the experimental setup.

REFERENCES

[1] Jones, M.J., "Face Recognition: Where we are and where to go from here", In the IEEE Transactions on Electronic, Information and Systems, Vol. 129 (2009), No. 5, pp. 770- 777, January 2009.

[2] Eds. W. Zhao and R. Chellappa, "Face Processing: Advanced Modeling and Methods", Academic Press, 2006.

[3] Chou-Hao Hsu, Chaur-Chin Chen, Nat. Tsing Hua Univ., Hsinchu, "SVD-Based Projection for Face Recognition", In the proceedings of IEEE International Conference on Electro/Information Technology, pp. 600 – 603, 2007.

[4] Yong Xu; Yingnan Zhao, "Comparison Study on SVD-Based Face Classification", In the proceedings of International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 343 - 346 2006.

[5] James, E.A.K, Annadurai, S, "Implementation of Incremental Linear Discriminant Analysis using Singular Value Decomposition for Face Recognition", First Advanced Computing, pp. 172 – 175, 2009.

[6] Steven Fernandes and Josemin Bala, "Performance Analysis of PCA-based and LDA-based Algorithms for Face Recognition," International Journal of Signal Processing Systems, Vol. 1, No. 1, pp. 1-6, 2013.

[7] Ruicong Zhi, Qiuqi Ruan, "Two-Dimensional Direct and Weighted Linear Discriminant Analysis for Face Recognition", In the Journal of Neurocomputing, Volume 71 Issue 16-18, pp.3607 – 3611, 2008.

[8] Alaa Eleyan, Hasan Demirel, "PCA an LDA based face recognition using feedforward neural network classifier", In the proceedings of Multimedia Content Representation, Classification and Security, LNCS, vol. 4105, pp. 199 – 206 , 2006.

[9] Jiazhong He; Minghui Du, "Face Recognition Based on Projection Map and SVD Method for One Training Image per Person", In the proceedings of International Conference on Computational Intelligence for Modeling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, pp. 20 – 24, 2005.

[10] Ben Niu, Qiang Yang, Simon Chi-Keung Shiu, Sankar K. Pal, "Two-dimensional Laplacianfaces Method for Face Recognition", In the Journal of Pattern Recognition Volume 41(10), pp. 3237-3243, 2008.

[11] Jiangfeng Chen Bo Li Baozong Yuan, "Face Recognition using Direct LPP Algorithm", In the proceedings of 9th Signal Processing, pp. 1457 – 1460, 2008.

[12] Nguyen Duc Thang, Tahir Rasheed, Young-Koo Lee, Sungyoung Lee, Tae-Seong Kim, "Content-Based Facial Image Retrieval using Constrained Independent Component Analysis", In the Journal of Information Sciences, pp. 3162-3174, 2011.

[13] Kailash J karande, Sanjay N Talbar, "Independent Component Analysis of Edge Information for Face Recognition", In the International Journal of Image Processing (IJIP), Volume 3, Issue 3, pp. 120 – 130, 2009.

[14] Jin Young Kim; Dae Young Ko; Seung You Na, "Implementation and enhancement of GMM face recognition systems using flatness measure", In the proceedings of 13th IEEE International Workshop on Robot and Human Interactive Communication, 247 – 251, 2005.

[15] Bredin, H.; Dehak, N.; Chollet, G., "GMM-based SVM for face recognition," , In the proceedings of 18th International Conference on Pattern Recognition, vol.3, pp.1111-1114,2006.

[16] Jen-Tzung Chien, Chin-Pin Liao, "Maximum Confidence Hidden Markov Modeling for Face Recognition", In the IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 606 – 616, 2008.

[17] Shih-Ming Huang, Jar-Ferr Yang, Shih-Cheng Chang, "Robust face recognition using subface hidden Markov models", In the Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), pp.1547 – 1550, 2010.

[18] Lei Zhang Samaras, D. "Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics", In the IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 351 – 363, 2006.

[19] Mian, A. S., Bennamoun, M., Owens, R., " An efficient multimodal 2D-3D hybrid approach to automatic face recognition", In the IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1927-1943, 2007.

[20] Yue Ming and Qiuqi Ruan, "Expression-robust 3D Face Recognition using Bending Invariant Correlative Features", In the proceedings of Informatica, pp. 231–238, 2011.

[21] R Basri and D W Jacobs: "Lambertian reflectance and linear subspaces", IEEE ", In the IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 25, No. 2, pp. 218-233, 2003.

[22] Shaohua Zhou Chellappa, R, "Illuminating light field: image-based face recognition across illuminations and poses", Proceedings. Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004.

[23] A S Georghiades, D J Kriegman and P N Belhumeur: " From Few to many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose," ", In the IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 23, No. 6, pp. 643-660, 2001.

[24] Demetri Terzopoulos, Yuencheng Lee and M. Alex O. Vasilescu, "Model-Based and Image-Based Methods for Facial Image Synthesis, Analysis and Recognition", in the Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition (FG'04), pp. 3 -8, 2004.

[25] Wenyi Zhao, Rama Chellappa, "Image-based Face Recognition: Issues and Methods", 2002.

[26] Ansari, A. Mahoor, M.H. Abdel-Mottaleb, M, "Normalized 3D to 2D model-based facial image synthesis for 2D model-based face recognition", In the proceedings of IEEE GCC Conference and Exhibition (GCC), pp. 178 – 181, 2011.

[27] Sang-Il Choi, Chong-Ho Choi, Nojun Kwak, "Face recognition based on 2D images under illumination and pose variations", In the Pattern Recognition Letters, Vol. 32, pp. 561-571, 2010.

[28] Jingu Heo Savvides, M. Rapid 3D face modeling using a frontal face and a profile face for accurate 2D pose synthesis", In the IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, 632 – 638, 2011.

[29] L. Zhao, W. Hu and L. Cui, "Face Recognition Feature Comparison Based SVD and FFT," *Journal of Signal and Information Processing*, Vol. 3 No. 2, 2012, pp. 259-262, 2012.

[30] Raj Babbar Sharma, P. K. Singh, "Enhancement in Image Recognition through Surveillance", In the International Journal of Computer Science and Management Research, Vol 1 Issue 2, pp. 254-258, 2012.

[31] Xiaofei He, Shuicheng Yan, Yuxiao Hu, Partha Niyogi, Hong-Jiang Zhang, "Face Recognition Using Laplacianfaces", In the IEEE

Transactions on Pattern Analysis and Machine Intelligence, pp. 328 – 340, 2005.

[32] http://www.postech.ac.kr/~seungjin/courses/dr/readings/ica.pdf

[33] http://en.wikipedia.org

[34] http://www.cvc.yale.edu/projects/yalefaces/yalefaces.html

[35] http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[36] http://vasc.ri.cmu.edu/idb/html/face/

[37] http://cswww.essex.ac.uk/mv/allfaces/index.html

Effective Performance of Information Retrieval by using Domain Based Crawler

Sk.Abdul Nabi[1]
Department of CSE
AVN Inst. Of Engg.& Tech.
Hyderabad, India

Dr. P. Premchand[2]
Dean, Faculty of Engineering
University college of Engineering
Osmania University, Hyderabad, India

***Abstract*—World Wide Web continuously introduces new capabilities and attracts many people [1]. It consists of more than 60 billion pages online. Due to this explosion in size, the information retrieval system or Search Engines are being upgraded day by day and it can be used to access the information effectively and efficiently. In this paper, we have addressed Domain Based Information Retrieval (DBIR) System. In this system we crawl the information from the web and added all links to the data base which are related to a specific domain. It simply ignores which are not related to that domain. Because of that we can save the Storage Space (SS) and Searching Time (ST) and as a result it improves the performance of the system.**

It is an extension of Effective Performance of Web Crawler (EPOW) System [2], in which it has two Crawler modules. The first one is Basic Crawler. It consists of multiple downloaders to achieve parallelization policy. The second one is Master Crawler, which is used to filter the URLs send by the Basic Crawler based on the Domain and sends back to the Basic Crawler to extract the related links. All these related links are collectively stored into the database under a unique domain name.

Keywords—Domain Based Information Retrieval (DBIR); Storage Space (SS); Searching Time (ST); Master Crawler; Basic Crawler; EPOW.

I. INTRODUCTION

The Web crawler [3] is a computer program that downloads data or information from World Wide Web for search engine. Web information is changed or updated rapidly without any information or notice. Web crawler searches the web for updated or new information. Web crawler [4, 5] is a software agent. It can be also called as Spider or Robots, which is the main component of a Search engine. Crawling the whole web is not possible because of its size and growth. Fig .1 represents how fast the amount of internet hosts increased in the last few years. When the crawler has finished with the current state of the network during the time host will have grown a lot larger and the documents which were indexed will have become outdated. Because of this reason web crawlers are being updated day by day and it becomes more popular in Information Retrieval Systems or Search Engines.

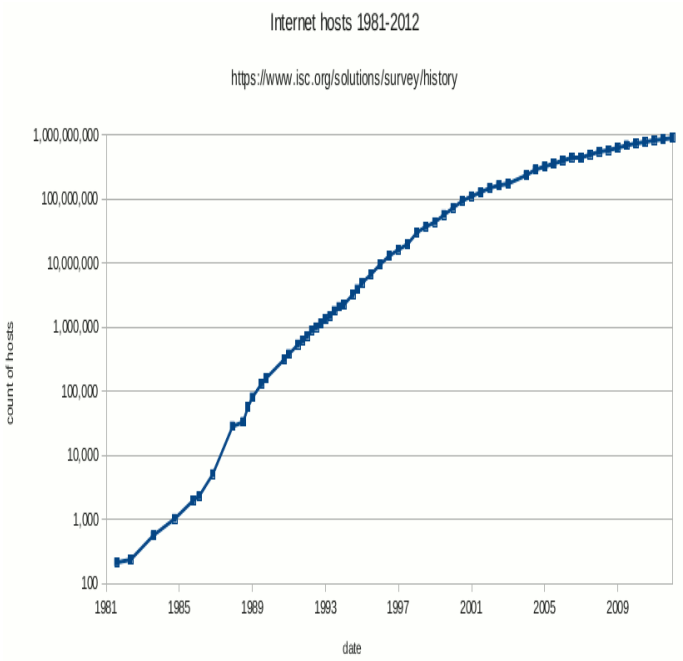


Fig. 1 Internet growth (Total no. of Hosts / Year) (taken from http://en.wikipedia.org/wiki/file)

The advent of the World Wide Web caused a dramatic increase in the usage of the Internet [2, 6]. The World Wide Web is a broadcast medium where a wide range of information can be obtained at a low cost. Information on the World Wide Web is important not only to individual users, but also to the organizations especially when the critical decision making is concerned. Most users obtain World Wide Web information using a combination of search engines [7] and browsers. However these two types of retrieval mechanisms do not necessarily produce all the information needed by the user. When the User tries to search the information hundreds of irrelevant documents return in response to a search query, only less than 18% of web pages are relevant to the user. To overcome this problem we need one effective search engine, which produces maximum relevant information in minimum time and at low cost.

Many academic and industrial researchers have looked at web searching technologies over the last few years, including storage, indexing, ranking techniques, crawling strategies and a significant amount of work on the structural analysis of the web [8]. Thus, highly efficient crawling systems are needed in order to download the hundreds of millions of web pages indexed by the major search engines. In fact, search engines battle against each other primarily based on the size and currency of their primary database, in addition to the quality and response time of their ranking function. Even Popular search engines, such as Google or AltaVista, presently cover up only restricted parts of the web, and a large amount of their data is several months out of date.

Applications involving search are everywhere in the world. The field of computer science which is most involved with R&D for search is Information Retrieval (IR). "[9, 10] Information Retrieval is a field concerned with the structure, analysis, organization, storage, searching, and retrieval of information." (Salton, 1968). General definition can be applied for many types of information and search applications. Primary focus of IR has been on text and documents since 1950's.

TABLE I. DIMENSIONS OF IR

Content	Applications	Tasks
Text	Web search	Ad hoc search
Images	Vertical search	Filtering
Video	Enterprise search	Classification
Scanned docs	Desktop search	Question answering
Audio	Forum search	
Music	P2P search	
	Literature search	

All search engines [11] available on the internet need to traverse web pages and their associated links to copy them and to index them into a web database. The programs associated with the page traversal and recovery is called crawlers. The main decisions associated with the crawlers algorithm are when to visit new sites that have been discovered through links. The parameters to balance are network usage and web server overload against index accuracy and completeness.

II. SCOPE & OBJECTIVES

This proposed system aims at creating a search engine which searches the information in domain form from the web by saving the storage space and searching time and as a result it improves the performances of the system.

The main objective of the system is as follows

- Reduces the Overhead of the user.
- Maintains the Freshness of the page.
- Provides Scalability and Portability.
- Provides High Performance.

III. RELATED WORK

An ultimate goal of web mining is to analyses the structure of Web. This can be achieved by the system that is capable of gathering more relevant information from the web. Web Crawler is defined as "a software component that iteratively collects information from the web, download pages and follows the linked URL's" [12]. Implementation of crawlers stared in early 1990's. But till the Google introduced its distributed crawling module in late 1990's, all previous crawlers were stand alone. Fig .2 represents the Process of a General Crawler. It approaches Breadth First Search or Horizontal search.

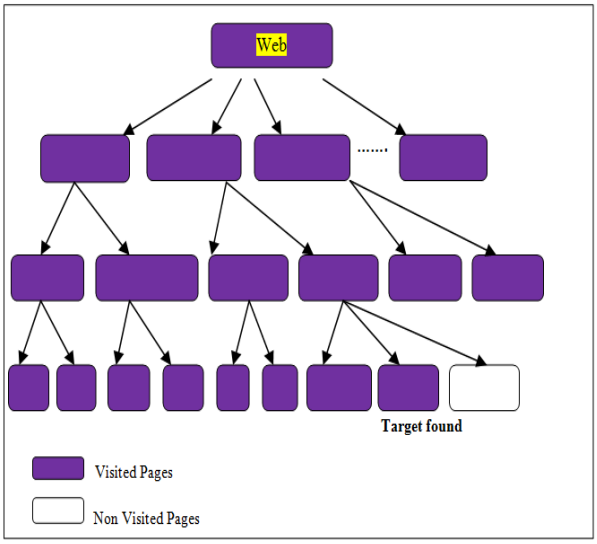


Fig. 2 General Crawler Process.

General crawler crawls all the pages from the web by using breadth first strategy. In this process, when we want to search particular information, it has to search in horizontal manner. Because of this method, it retrieves more non relevant information which is not required to that context. It takes longer time to access the actual information, thus it reduces the precision and recall.

Now days, a major challenge faced by the current web crawlers is to select the important pages for downloading. The crawler cannot download all the pages from the web due to large size. In our DBIR system, crawler will select the pages

and it visits important pages first by prioritizing the URLs in the Queue.

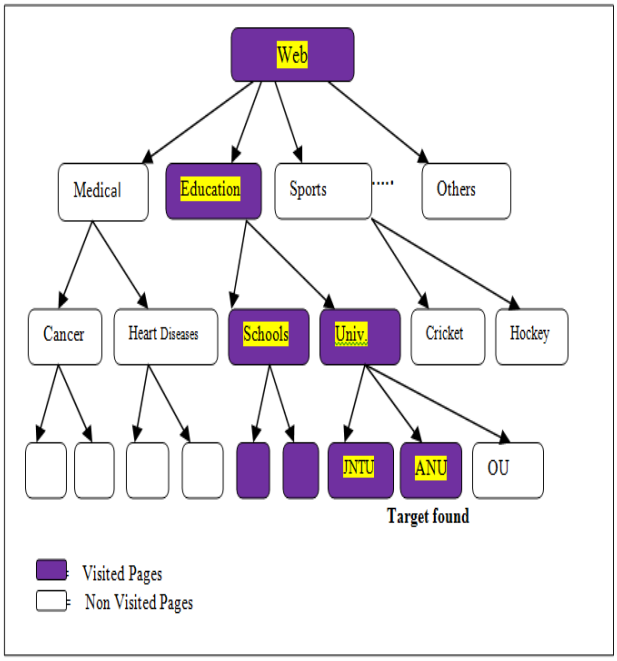


Fig. 3 DBIR Crawler Process.

In our DBIR System, It searches vertically based on the domain. In this process it skips or ignores all non relevant information thus we can improve the precision and recall.

Example to search ANU (Acharya Nagarjuna University) Text and domain name is given as Education in our DBIR System:

The Fig. 3 represents that searching is done in vertically according to the domain name is Education. In each level it skips non relevant pages, thus it improves the performance of the Searching process.

The fig.4 is the example for crawling the web pages which are interconnected through links.

IV. Domain Based Information Retrieval (DBIR) System

It is an extension of our earlier Effective performance of Web crawler (EPOW).In this proposed system we have added Ranking adaption with pattern matching (RAPM) algorithm in Master Crawler(Effective Web Crawler).

In this approach, it has two Crawlers. The first one is Basic Crawler and the second is Effective web crawler (i.e. master crawler). Basic Crawler consists of multiple downloaders. It fetches web pages (documents) from the World Wide Web when provided with a corresponding URL. Effective web crawler receives the URLs which is sent by the basic crawler and decides what page to request next based on the category (Domain) and issues a stream of requests (URLs) to the basic crawler (Downloader).The basic crawler downloads the requested pages and supplies them to the master crawler for analysis and storage.

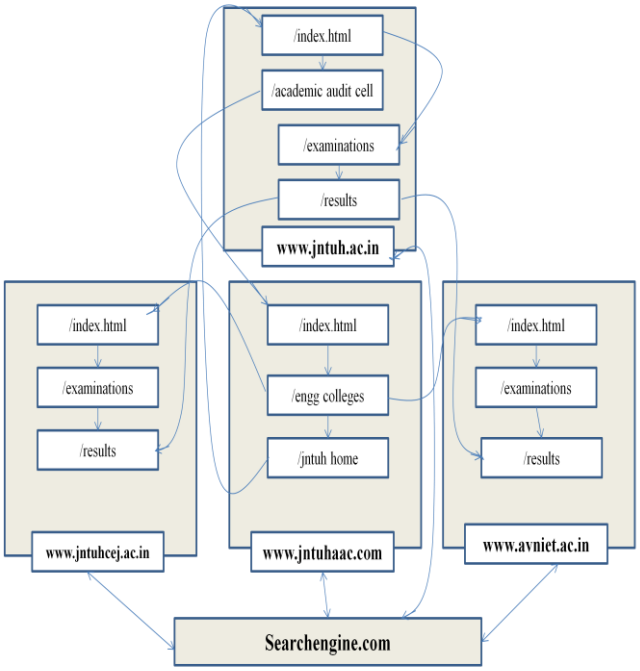


Fig. 4 Example of Crawling the Web

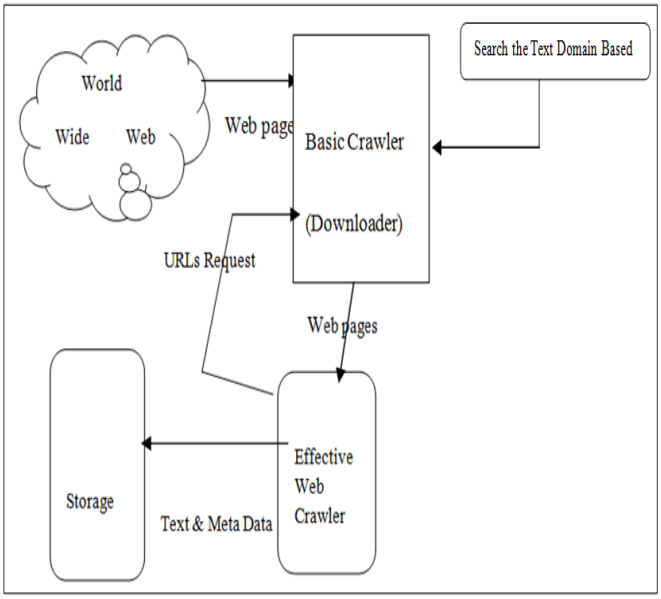


Fig. 5 Domain Based Information Retrieval (DBIR) System.

Master crawler collects all the related information and stored in the database by category wise. This process will be iteratively done till maximum relevant documents are fetched from the web at regular interval.

Once user enters the query, Master crawler analyzes the query and sends multiple URLs list which is relevant to the previous document. The web has a very dynamic nature [13, 14], and crawling a fraction of the web can take weeks or months. By the time a web crawler has finished its crawl, many events can have happened, including creations, updates and

deletions. The pages will remain outdated due to these modifications. The objective of this DBIR system is to keep the average freshness [15] of pages in its collection as high as possible, or to keep the average age of pages as low as possible. For that we have adapted optimal Revisit Policy. This method is for keeping average freshness high by ignoring the pages that change too often, and the optimal for keeping average age low is to access frequencies that monotonically increase with the rate of change of each page.

V. IMPLEMENTATION

A. *Flow of the Data in Domain Based Information Retrieval System*

To implement the crawler in real time we took a GUI application wherein the user has to login to use our DBIR system with a valid id and password. The User can get the id by signing-up into our application which will create a separate profile for the user storing his information, searched terms , sites visited which will provide privacy for each user. The user needs to add the sites in his profile which will be indexed by our DBIR crawlers and maintain the details in the user profile for searching the terms. Adding of sites has been allocated with category wise and admin will see in which category the site to be indexed by the system should. Once a site is indexed, all the details are crawled and maintained by the system in user profile. The system has admin for the application who can control all the activities of the user and system. Admin can upload the domain names, so that whatever indexes crawled related to that domain will be stored in separate storage area.

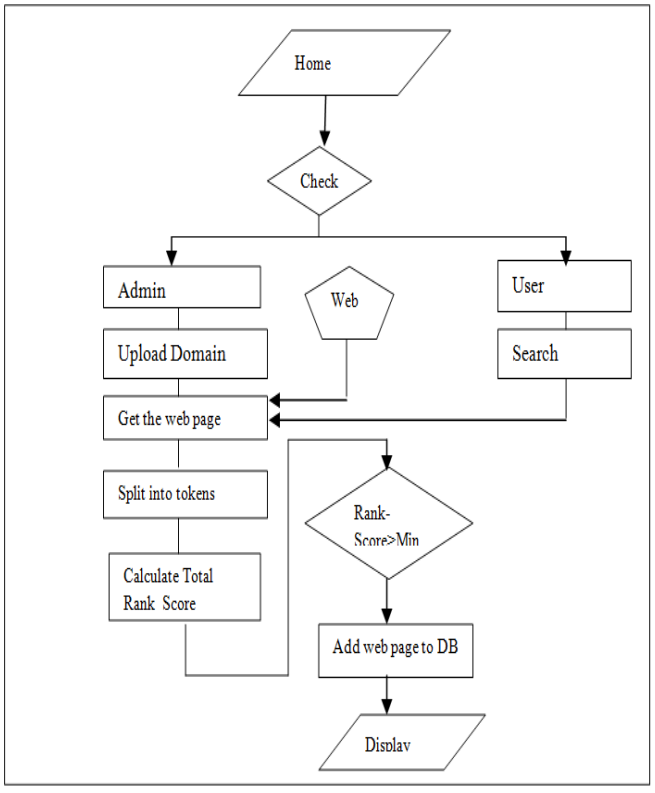


Fig. 6 Flow Chart for Domain Based Information Retrieval System (DBIR)

B. *Rank Adaption with pattern Matching Algorithm*

Input: A webpage (wp), Rank Table

Output: Total Rank of the webpage

Step 1: Initialize input Rank=0;

Step 2: Select the first term (t) and Corresponding rank from Rank table

Step 3: Webpage will be divided into tokens or words (Tokenizes)

Step 4: Compare the first term with webpage (wp).for each matching pattern, the Rank_Score will be added by 1

Step 5: Total Rank_Score = Rank * standard weightage of the page

Step 6: Select next term find the total Rank_Score of that term.

Step7: Repeat until all the terms are compared and find the overall Rank Score of the Webpage.

Step8: If Overall Rank_Score > Min Rank then the web page will be stored in Database.

Step 9: End.

VI. CONCLUSION & FUTURE DETECTIONS

Whenever we are searching for the content from the web, most of the crawlers are not returning the expected results and they give a list of sites to be searched by user which is related to the searched text but not required by the user. Thus it increases the user overhead and consumes time. This system decreases the overhead of the user by providing less in number and most relevant results based on the category of the searched term. To achieve this system uses DBIR Method, unlike other search engines, here two crawlers are used for crawling and indexing sites. The introduction of additional crawler with multiple downloaders increases the speed and performance of the crawling.

The front end of the system enables user to maintain his entire query and search history by providing a login system with Id and password. The validation process provides security and privacy to the users unlike other browsers. The system also has an admin who can control all the users and its activities. User or Admin has to add the sites to be crawled by this DBIR system. User can also share the search results with others using SMS and Email. Overall, the system is designed to decrease the burden on the internet surfers and provide an optimized GUI for safe and easy browsing.

Web is very large and dynamic. Searching the required relevant content from the web is always difficult. Grouping the collections from the web is always challenging. We need to gather from broad range of domains. This project validated limited no of collections. We need to validate a large number of collections from various domains. This Front end application can be implemented for the specific organizations or institutes to provide the privacy to the users to access the web. It can be used for maintaining their history and profile of their employees.

ACKNOWLEDGMENT

We would like to thank everyone who has motivated and supported us for preparing this manuscript.

REFERENCES

[1] Ricardo Baeza-Yates, Felipe Saint-Jean, and Carlos Castillo. Web structure, dynamics and page quality, In Proceedings of String Processing and Information Retrieval (SPIRE), Springer LNCS, 2002.

[2] Sk.Abdul Nabi and Dr. PremChand ,Effective Performance of Information Retrieval, International Journal of Web & Semantic Technology (IJWesT) Vol.3, No.2, April 2012.

[3] S S Vishwakarma , A Jain , A K Sachan, A Novel Web Crawler Algorithm on Query based Approach with Increases Efficiency,International Journal of Computer Applications , Volume 46– No.1, May 2012.

[4] Wenqing Yuan , Research on Prototype Framework of a Multi-Threading Web Crawler for E-Commerce, In Proceedings of the International Conference on Management and Service Science, MASS '09. IEEE Transactions, 2009.

[5] Heydon and Najork. Mercator: "A scalable, extensible Web crawler", World wide web2 (4), 1999.

[6] RM. Vidhyavathy, E.Ramaraj, N. Venkatesan, A Study Of Mining Web Intelligent Information, International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 9, September 2012.

[7] H.Vernon Leighton and J. Srivastava, Precision among www search services (Search Engines), 1997.

[8] Alexandros Ntoulas, Junghoo Cho, and Christopher Olston. What's new on the web?: the evolution of the web from a search engine perspective. In Proc. Of 13th conference on World Wide Web, May 2004.

[9] Kowalski, Gerald, Mark T Maybury, Information Retrieval Systems: Theory and Implementation, Kluwer Academic Press, 1997.

[10] Terrence A.Brooks. Web Search : how the web has changed information retrieval. Information Research, April 2003.

[11] H.Vernon Leighton and J. Srivastava. Precision among www search services (search Engines), 1997.

[12] S.Brin , L.Page , " The Anatomy of a large–Scale Hyper Textual Web Search Engine". In Proc. Of the 7th World Wide Web Conference, 1998.

[13] Recardo Baeza – Yates Carlos Castillo and Felipe Saint – Jean., Web Dynamics , Springer , 2004.

[14] Brain E. Brewington and George Cybenko, How Dynamics is the Web. In Proceedings of the Ninth International World – Wide Web Conference , May 2000.

[15] Junghoo Cho and Hector Garcia-Molina. Effective page refresh policies for web crawlers.ACM Transactions on Database Systems, 28(4), December 2003.

AUTHORS

Prof Shaik.Abdul Nabi1 is the Head of the Dept. of CSE, AVN Inst. Of Engg.& Tech, Hyderabad, AP, India. He completed his B.E (Computer Science) from Osmania University, A.P. He received his M.Tech. from JNTU Hyderabad campus and currently he is pursuing Ph.D. in the area of Web Mining from Acharya Nagarjuna University, Guntur, AP, India. He is a certified professional by Microsoft. His expertise areas are Data warehousing and Data Mining, Data Structures & UNIX Networking Programming.

Dr P.Premchand2 is a professor in the Department of Computer Science & Engineering, Osmania University, Hyderabad, A.P,India. He completed his ME (Computer Science) from Andhra University, A.P. He received Ph.D degree from Andhra University, A.P. He has guided many scholars towards the award of Ph.D degree from various Universities. He was a director of AICTE, New Delhi, during 1998-99. He also worked as the Head of the Dept of CSE , Additional Controller of Examinations and chairman of BOS, Faculty of Engineering, OU. Currently he is working as the Dean, Faculty of Engineering, Osmania University, Hyderabad, and AP.

An Intelligent Diagnostic System for Congenital Heart Defects

Amir Mohammad Amiri
Dept. Electrical and Electronic Engineering (DIEE)
University of Cagliari
Cagliari, Italy

Giuliano Armano
Dept. Electrical and Electronic Engineering (DIEE)
University of Cagliari
Cagliari, Italy

***Abstract*—congenital heart disease is the most common birth defect. The article describes detection and classification of congenital heart defect using classification and regressing trees. The ultimate goal of this research can decrease risk of cardiac arrest and mortality in compared with healthy children. The intelligent system proposed in three stages technique for automate diagnosis:(i) pre-processing(ii), feature extraction, and (iii) classification of congenital heart defects (CHD) using data mining tools. The intelligent diagnostic system has been validated with a representative dataset of 110 heart sound signals, taken from healthy and unhealthy medical cases. This system was evaluated in the test dataset with the following performance measurements global accuracy: 98.18%, sensitivity, 96.36% and specificity 100%. This results show the feasibility of classification based on optimized feature extraction and classifier. This paper follows the Association for the recommendations of the Advancement of Medical Instrumentation.**

Keywords—congenital heart defects; Heart murmurs; newborns; classification and regression trees;

I. INTRODUCTION

Physicians use the stethoscope as a device to listen to the acoustic signals which cannot be analyzed by the human ear. The interpretation can provide valuable information regarding the function of heart valves, and is capable of detecting of disorders, especially problems related to the valves. Various techniques have been developed to analyze and classification of the heart sound [1].

Physicians use a physical exam and special heart tests to diagnose congenital heart defects. They often find severe defects during pregnancy or soon after birth. Many congenital heart defects cause few or no signs and symptoms. They are often not diagnosed until children are older.

Congenital heart defects are characterized by anomalies in the structure of the heart and its related valves and vessels; such as holes in the heart, narrowed, leaky valves, malformed, missing vessels or heart chambers ventricular septal defect and coarctation of the aorta are typical examples of CHDs. These defects are the most common type of birth defect and typically present in infancy. In many cases, the cause of the defect in a particular infant is unknown; from an epidemiological standpoint, maternal smoking during pregnancy, genetics, and chromosomal abnormalities have been cited as factors contributing to CHD [2].

Congenital heart problems range from simple to complex. Some heart congenital defects in children can be watched by the child's physician and managed with medications, while others require surgery, sometimes as soon as in the first few hours after birth. A child may even ”grow out” of some of the simpler heart problems, such as patent ductus arteriosus (PDA) or atrial septal defect (ASD), as these defects may simply resolve on their own as the child grows. Other infants will have a combination of malformations and require several operations throughout their lifetime.

Newborns with heart congenital defects experience no symptoms. The heart defect may be diagnosed if the health care provider hears an abnormal sound, called a murmur.

Children with normal hearts also can have heart murmurs, called innocent or functional murmurs. A provider may suggest tests to rule out a heart defect.

Some babies born with a heart congenital defect can appear healthy at first and can be sent home with their family before their heart defect is detected. Congenital heart defects affect approximately 1 in 125 live births. Of these, 30% have extra-cardiac anomalies (such as tracheoesophageal fistula, anorectal anomalies), which might require surgery within the first year of life [3]. These babies are at risk for having serious problems within the first few days or weeks of life and often require emergency care.

Although a normal heart still orders an echocardiogram for reassurance, even though the cost of an echocardiogram is high. The result of this practice is a misallocation of healthcare funds, since echocardiograms are expensive. While it is clearly important to avoid that healthy newborn are sent for echocardiogram, it is also important to avoid that a newborn that has a pathological heart murmur is sent home without proper treatment [4].

Many studies performed on heart congenital defects are concerned with various stages of life, but our dataset included newborns from one day to 2 months after birth [5]. The classification at this stage is very important because pathological heart sounds in newborns are more difficult to diagnose [6]. We achieve a high accuracy result to discriminate CHD by using optimized features and CART [4-8].

In this paper, we present a method for automatically segmentation of phonocardiogram (PCG) data. The proposed methods utilize CART to identify whether a systolic and diastolic pathological murmur holds. Notably, features extraction was very effective to improve experimental results. Our results show an accuracy of 98.18%, which significantly

improves the current state-of-the-art on this specific problem. In fact, other relevant works report accuracies. Improvements have been obtained also in sensitivity and specificity.

II. METHODS

This phase involves pre-processing and features extraction of the signal that is certain characteristic properties of heart sound that are unique to the signal and are thus suitable for classification purposes. Figure 1 show stages used in the proposed diagnostic system.

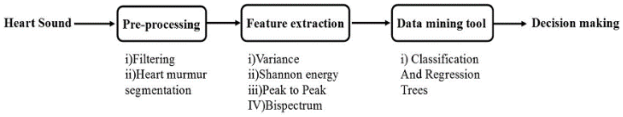


Fig.1. Stages used in an intelligent diagnostic system.

A. Pre-processing

Pre-processing occurs in two steps that will be described in the following order:

1) Filtering: The first step of signal processing is filtering heart sounds, with the goal of removing the unwanted noise. The recording of PCG usually has a sampling frequency higher than 8 kHz. In the event that the recording environment cannot be controlled enough, noise is coupled into the PCG. To avoid unpredictable effects brought by noise, filtering becomes important for later processing. Since the main spectrum of first and second (S1 and S2 respectively) heart sound occurs within the range of 200 Hz, the system filters the original heart sound using a 3rd order band-pass Butterworth filter, with cut-off frequencies at 50 Hz and 200 Hz. An electronic stethoscope has been used to record heart sounds, giving rise to a dataset at 44 kHz and converted to 4 kHz.

2) Segmentation: The second step of pre-processing is a segmentation method aimed at identifying the heart sound components S1 and S2 and timing interval between them. Although the detection can also be manual, we used to identify S1 and S2 with an automatic procedure. The segmentation method is based on the timing between high amplitude components. The fact that the time interval that occurs between S1 and S2 (systole) is always less than the one between S2 and S1 (diastole) is the basis for this process. Heart sound signals still have very complicated patterns, with numerous small spikes that have little impact on diagnosis but may influence the location of S1 and S2. Peak conditioning was performed for the obtained peaks using wavelet transform, which enabled the process of cycle detection. We used the Wavelet transform based on Complex Morlet Wavelet (CMW) for finding peak locations. CMW is a kind of Wavelet transform, which are a powerful tool in time-frequency analysis for PCG signals (see figure 2).

K-means uses an iterative method that minimizes the sum of distances from each object to its cluster centroid, over all clusters. Each class of heart murmurs contains distinctive information in time and frequency domains. This stage involves the extraction of each cardiac cycle of the PCG signal, the formed after the peak detection and conditioning stages. As systolic (S1-S2) and diastolic (S2-S1) murmurs occur within the time intervals that were calculated by the peak conditioning process, these time intervals were clustered into two clusters [9].

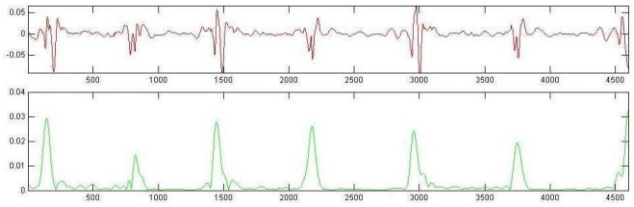


Fig.2. Peak detection: (top) original signal (bottom) coefficients line

Clusters 1 and 2 occur consecutively and indicate a single cardiac cycle. The smaller time interval of each cycle was then identified as systole while the other interval was identified as diastole. After peak detection and condition, cardiac cycle is identified by using k-means, a non-hierarchical clustering algorithm in which observation are divided in k mutually exclusive clusters. We extracted each single cycle of PCG signals using clustering (figure 3).

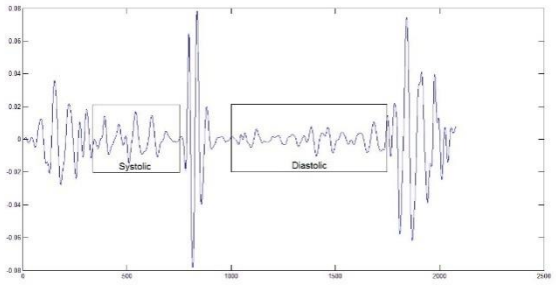


Fig.3. Sample of signal with systolic and diastolic murmur

B. Feature extraction

This phase is focused on extracting signal features that better highlight the properties of the PCG signal, with the goal of identifying those that are more suitable for classification purposes. It consists of two major steps: feature extraction and feature selection. In the former step we extract several features including Variance, Peak to Peak, Energy Shannon, Bispectrum.

The latter step (i.e. feature selection) was aimed at reducing the size of the feature vector. In particular, we used gains and variable importance in CART to measure the score each of variable. To calculate the importance score of a variable, CART looks at the improvement measure of each variable, in its role as a surrogate to the primary split. The values of these improvements are summed over each node and are scaled according to the best performing variable. In particular, the variable with the highest sum of improvements is scored 100, while other variables have lower scores. Variable importance scores (VIS) are summarized in Table 1. It lists all variables used and not used in the tree building process. A score is attached to each variable, and is based on

the improvement each variable makes as a surrogate to the primary splitting variable. Variable importance measure highlights variables whose significance is masked or hidden by other variables in the tree building process.

TABLE I. SCORE OF VARIABLE IMPORTANCE

No	Feature	Improvement	VIS
7	Variance	0.30724	100.0000
2	Peak to Peak	0.25342	89.0333
4	Shannon Energy	0.04802	53.5714
6	Wigner bispectrum	0.04802	53.3577

As Shannon energy and Wigner distributions have been very important for improving the classification, let us spend a few words on these techniques.

Shannon Energy: Shannon energy is another applicable method which we made use of. The calculation of the Average Shannon Energy is based on signal segments. Therefore, here we segment the data, 0.02 seconds and with a 0.01- second signal segment overlapping throughout the signal. The average Shannon energy is calculated as (see [10] for more information):

$$E = -x^2 \cdot \log x^2 \tag{1}$$

$$absolute\ value\ E = |x| \tag{2}$$

Wigner Bispectrum: Wigner high-order spectrum is an extension of Wigner-Ville distribution. It keeps the advantages of Wigner–Ville distribution and has also the advantages of High-Order Spectra. High-Order Spectra have been widely used in the non-gauss and non-stationary realm, which is quite applicable to PCG signals. In particular, by combining Wigner-Ville distribution, we could get the time-frequency characters at the same time. The study has proved that under low SNR circumstances, the Wigner bispectrum is better than Wigner-Ville distribution. The second order Spectra of Wigner-Ville Distribution of signal x (t) is finally defined as follows [1]:

$$W_{2x}(t, f_1, f_2) = \int_{\tau_1}\int_{\tau_2} x^*\left(t - \frac{1}{3}(\tau_1 + \tau_2)\right) x\left(t + \frac{1}{3}(2\tau_1 - \tau_2)\right) \exp(-j2\pi(f_1\tau_1 + f_2\tau_2))\, d\tau_1 d\tau_2 \tag{3}$$

C. Classification and Regression Trees

Classification of congenital heart defect is novel application of CART for clinical and physiological data. CART developed by Breiman et al. (1984), is a nonparametric statistical method that creates binary decision trees. It is a step-by-step process in which a decision trees are constructed by either splitting each node on the tree in two daughter nodes.

The realistic objective of partitioning is to find partition s of the data such that terminal nodes are as such homogeneous as possible. The quantitative measure of node homogeneity is called impurity function. The simplest idealization of the impurity function is the number of patients who meet an objective criteria divided by the total number of patients in the node. Ratios close to 0 or 1 are considered more pure.

To partition a node, CART examines all possible splits of the explanatory variables. In general, the number of splits for ordinal or continuous variables is 1 minus the number of distinctly observed values. A potential split is judged by its reduction of the impurity function for both daughter nodes it creates. The partitioning iteratively continues by splitting each node in two daughter nodes and continues until the tree is saturated that is, until no further partitions can be found [11].

The decision tree for predicting heart murmurs is reported in figure 4. We start at the top of the tree and follow different branches, depending on conditions involving the predictor variables. Trees with multiple layers of splits may be conceptualized as describing interactions between predictor variables. Once we arrive at an end-point of the tree, we used 8 nodes and variables classified in two classes (classes 0 and 1 were normal and pathological murmurs respectively [12]).

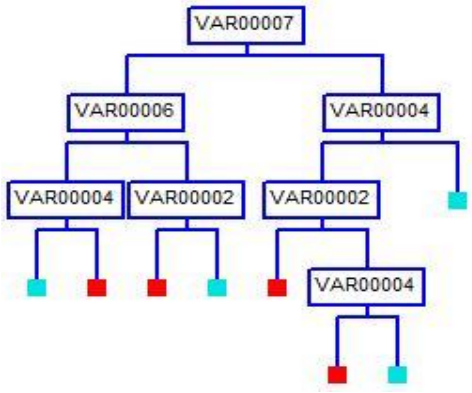


Fig.4. Illustration of decision tree structure.

We calculated the likelihood ratio (LR) to obtain sensitivity and specificity on a tree, defined as follows:

$$LR+ = \frac{\text{sensitivity}}{1-\text{specificity}} \tag{4}$$

$$LR- = \frac{1-\text{sensitivity}}{\text{specificity}} \tag{5}$$

The interpretation of likelihood ratios is intuitive: the larger the positive likelihood ratio, the greater the likelihood of heart disease; the smaller the negative likelihood ratio, the lesser the likelihood of congenital heart defects.

III. EXPERIMENTAL AND RESULTS

In this study, a biomedical system based on variance, peak to peak, Shannon energy and bispectrum was developed in order to diagnose two different heart sounds. A total of 110 heart sounds (normal and pathological) were studied. In this section, we present the results of the application of the above

proposed classification and regression trees technique. K-fold cross validation (K=8) has been used as training and test strategy.

Classification results of the CART are displayed by confusion matrix. Results are shown in Table 2 in the form of a confusion matrix, together with percentage classification accuracy. It can be seen that out of 55 normal signals, 53 were correctly classified as normal, and 2 were misclassified as pathological. Similarly, out of 55 pathological signals, they were correctly classified as pathological without misclassification. A detailed analysis of the misclassified example showed that it was in fact very difficult to classify, even by human experts.

TABLE II. CLASSIFICATION RESULT OF CONGENITAL HEART DEFECTS IN NEWBORNS

Actual Group	Normal	Pathological	Percent Correct
Normal	53 (96.36%)	2 (3.64%)	96.36%
Pathological	0 (0.00%)	55 (100%)	100%
Average/Overall	110		98.18%

Summarizing, 98.18% accuracy, 96.36% sensitivity and 100% specificity were obtained by CART, when used to distinguish between the 110 normal and pathological heart murmurs in newborns.

Let us point out that, for this system, both high sensitivity and specificity are important. In particular, high sensitivity reduces the number of newborns with normal (innocent) murmurs who are identified as pathological murmur and sent to echocardiogram for further testing. More importantly, high specificity reduces the number of newborns with pathological murmurs that are identified as innocent murmurs and have been released with a potentially deadly heart condition. For each fold, learning has been performed in two steps: growing and pruning. It is worth noting that pruning has been performed provided that decision tree error curve did not trespass the threshold of 1%.

The CART decision tree error curve archived automated growing of a too large tree, followed by automated pruning to find the right-sized tree [13]. The rationale for the growing/pruning process is illustrated in the error curve (fig. 5).

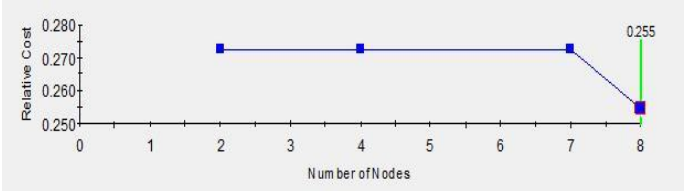


Fig.5. CART decision tree Error curve.

Figure 6 shows a curve which outlines the relationship between classification errors and tree size. The scale is always between 0 and 1, so it is called a relative error curve. A tree with a relative error of 0 or nearly 0 is usually too good to be true. The proposed model shows excellent performance for application of diagnosis of congenital heart defects. In a Receiver Operating Characteristic (ROC) curve for a binary classification problem, the true positive rate (Sensitivity) is reported as function of the false positive rate (1-Specificity) for different cut-off points. ROC curve is reported in Figure 6 for normal and pathological murmurs.

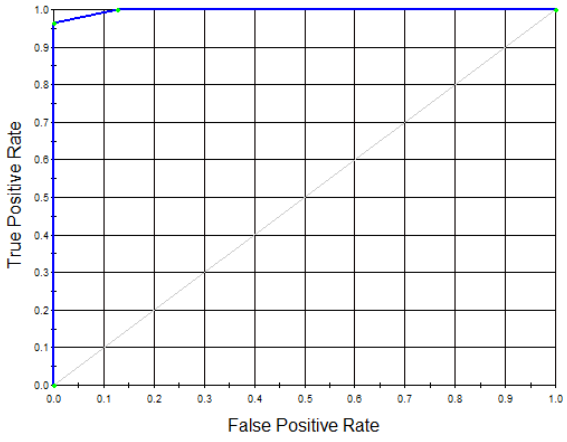


Fig.6. ROC curve of normal and pathological murmurs.

A predictive model with perfect performance has an area under ROC curve equal to 1. We obtained, on average, an accuracy of 0.99 the ROC curve highlights the excellent performance of CART to discriminate of heart murmurs.

IV. CONCLUSIONS

In this paper the system which is able to differentiate between normal and congenital heart defects using intelligent techniques is discussed. The intelligent diagnostic system proposes novelties in both segmentation of heart sounds and application of CART. The intelligent diagnostic system not only helps in accurate detection it is also useful for the physician who is in charge to help newborns saving lives of many case of abnormality. This technology is for high-volume screening of newborns suspected of having a heart disease. The software system proposed in this work can be considered the first release of a diagnostic tool able to support physicians in their diagnostic task.

Further work is under way to improve feature extraction and classification and also the diagnostic system can be can be saved for future use on other data.

REFERENCES

[1] C. Ahlstrom, P. Hult, P. Rask, J.E Karlsson, E. Nylander, U. Dahlstro, ”Feature extraction for systolic heart murmur classification”, Annals of Biomedical Engineering, Vol.34, No. 11, November 2006, pp.1666-1677.

[2] B. Roman, “Nourishing Little Hearts: Nutritional Implications for Congenital Heart Defects”, Vol. XXXV, No 8, August 2011, pp. 11-35.

[3] Michelle C White, “Anesthetic Implications of Congenital Heart Disease for Children Undergoing non-cardiac Surgery”, Pediatrics, Elsevier 2009, pp.504-509.

[4] S. L. Strunic, F. Rios-Gutierrez, R. Alba-Flores, G. Nordehn, S. Bums. “Detection and Classification of Cardiac Murmurs using Segmentation Techniques and Artificial Neural Networks”, Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2007, pp. 397-404.

[5] R. Shandas, L. Valdes-Cruz and Roop L. Mahajan. ”Artificial Neural

Network Based Method of Screening Heart Murmurs in Children" , Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231.2001, pp. 251-260.

[6] E. Loukis and M. Maragoudakis "Heart murmur identification using random forest in assistive environments", PETRA'10, June 23 - 25, Samos, Greece, 2010, pp. 1-6.

[7] F. Rios-Gutierrez, R. Alba-Flores, S. Strunic "Recognition and Classification of Cardiac Murmurs using ANN and Segmentation" , Electrical Communications and Computers (CONIELECOMP), 22nd International Conference on, IEEE2012.

[8] A.M. Amiri and G. Armano Diagnosis and classification of systolic murmur in newborns, The 10th IASTED international Conference on Signal processing, pattern recognition and applications, Innsbruck Austria, 2013, pp.480-484.

[9] C. N. Gupta, R. Palaniappan, S. Swaminathan, and S. M. Krishnan, "Neural network classification of homomorphic segmented heart sounds", Applied Soft Computing 2007, 7 pp. 286-297.

[10] Lantian Song and Fengqin Yu. "The Time-Frequency Analysis of Abnormal ECG Signals", Springer-Verlag Berlin Heidelberg 2010, pp. 60-66.

[11] A. Kumar Banerjee, N.Arora, U.S.N Murty, "Classification and Regression Tree (CART) Analysis for Deriving Variable Importance of Parameters Influencing Average Flexibility of CaMK Kinase Family", Electronic Journal of Biology, Vol. 4(1), 2008, pp.27-33.

[12] M. Rabinoff, C.M.R. Kitchen, I.A. Cook, and A.F. Leuchter, "Evaluation of Quantitative EEG by Classification and Regression Trees to Characterize Responders to Antidepressant and Placebo Treatment", The Open Medical Informatics Journal, 5, 2011, pp. 1-8.

[13] Steinberg, D., and P. Colla. 1995,"CART: Tree-structured non-parametric data analysis", San Diego, Calif., U.S.A.: Salford Systems.

POSIX.1 conformance for Android Applications

TayyabaNafees
Department of computer engineering
National University of science &Technology, H-12,
Islamabad, Pakistan

Prof. Dr. Shoab Ahmad Khan
Department of computer engineering
National University of science &Technology, H-12,
Islamabad, Pakistan

***Abstract*—Android operating system is designed for use in mobile computing by The Open Handset Alliance. Android market has hundreds of thousands of Android applications and these applications are restricted only to the mobiles. This restriction is mainly because of portability and compatibility issues of Android operating system. So need of employing these countless Android applications on any POISX Desktop operating system without disturbing the internal structure of application is very desirable. Thus we need to resolve these standardization and portability concerns by using POSIX standards (Portable Operating System Interface). The concepts of POSIX conformance for Android applications provide full-scale portability services and Android applications reusability for any POSIX desktop operating system. So Android applications will become usable for all POISX desktop users. This research theme introduces POSIX.1 Android thin layer model that simply provides the POSIX conformance for Android applications. It is using the POSIX.1 APIs for Android applications, which maintains the compatibility between the POISX Desktop operating systems and Android applications. We have analyzed our research work by implementation of the different applications in standard POSIX environment and, have verified its results. The results of POSIX.1 model clearly showed that it will boost up the Android applications market revenue up to 100% plus add real-time standardization and reusability.**

Keywords—Portable Operating System Interface (POSIX); Application Programming Interface (API); Operating system (OS); User interface (UI)

I. INTRODUCTION

Android is open source mobile OS .It is particularly adapted by various manufacturers and modified based on their own taste for its openness [1]. Currently, Android cell phones are becoming more sophisticated by providing functionalities that once expected from laptop and/or desktop computing systems. [2] For example, using cell phone, callers can now interact with system using spoken language, brows internet, exchange emails, chat online and social network medias, use navigation systems, etc. Mobile computing is real time computing. But mobile computing did not compete with Desktop OS because the Desktop users are still large in number plus it becomes the necessary need of user thus Mobile OS companies are still trying hard to make their space in the Desktop OS environment. Android is most famous and open source mobile operating system. It covers nearly 60% of mobile market but even Android OS (operating system) had the compatibility limitations. Therefore the need of standardization and portability is very essential. Android applications have multiple dependences so this limits the Android application utilization. One of the best possible solutions for catering these limitations is POSIX.

POSIX is an international standard with an exact definition and a set of assertions, which can be used to verify compliance. A conforming POSIX application can move from system to system with a very high confidence of low maintenance and correct operation. If you want software to run on the largest possible set of hardware and operating systems, POSIX is the way to go. [3]

POSIX conformance for android Application is the basic aim of this research in, which multiple Android applications are, used as sample input with the POSIX Application Programming Interface (API) standards. The research agenda based on the POSIX.1 thin layer model, which gives the POSIX conformance for Android applications. This POSIX.1 thin layer model hierarchy is:

- Selection of POSIX standard for Android Applications (POSIX.1).
- Need of POSIX.1 binding language.
- Conversion of sample examples of Android in POSIX binding language and test it
- Establishment of template for Android applications (POSIX thin layer model for Android applications)

A. The Problem Statement

Android is open source OS introduce by Google. Android is still developing. In the design of today's computing systems it is becoming increasingly important to design software with an open system architecture utilizing industry adopted standards. The need to develop open systems is driven by these major factors.

- Inefficient usage of manpower: First, gone are the days where a single developer can implement the entire system from scratch. Software development programs are continuously growing in scale, requiring teams of increasing size.
- Portability problem: Secondly, software does not operate in isolation; it must co-exist with the vast amount of commercially available software and can be run on available OS.
- Maintainability problems: The lifecycle of a software application is typically long requiring numerous modifications and updates as new features are added.
- Need of standardization: Lastly the biggest problem facing in these days is implementation of standards because portability and maintainability only fruitful when software developer follows the standards.

But Android performance is not enough, In addition, performance-analyzing environment has not been developed yet, and then its performance cannot be discussed well. Android OS addresses multiples challenges of today's software development process like interoperability, portability and compatibility issues. The major question is here, is Android application market is usable for all OS. Android applications standardization is major dilemma for Android market. Android applications for all OS are core idea of this research. But HOW is big question here. Thus Android applications need the openly published standard interfaces for competing these hybrids issues in Android OS. We are applying the Android applications standardizations by using the POISX.POSIX is based on UNIX, a well-established technology. POSIX defines a standard way for an application to interface to the operating system. [4] POSIX, the Portable Operating System Interface .The goal of POSIX is the source-code portability of applications: it means transform an application from one operating system to another by simple conversion. This Thin layer model of POSIX.1 provides the portability for Android applications that can be run on any operating system.

*1) Android Application Portability:*There are more than 600,000 apps and games available on Google Play store. [8] But the sorrowful act is limitation of these 600,00 apps only for the Android OS. All of this work need conformance for any operating system according to the users and developers need. Because developers are also looking to employ Android in a variety of other embedded systems that have traditionally relied on the benefits of true real-time operating systems performance, boot-up time, real-time response, reliability, and no hidden maintenance costs.

2) Is Android POSIX Compliance?: Android is considering a partial POSIX compliance. Limited POSIX threads (pthreads) library is implemented in Android Bionic library. It provides built-in support for pthreads, but implementation is very restricted. So Android applications conformance is very inspiring, which never has done yet.

3)Earliest Idea Invention of POSIX Compliance for Android: Android used the non standard Bionic library which restricted the android applications to only for android OS. So best into our knowledge this proposed model first time in the history trying to merge the mobile OS Android applications with desktop POSIX OS. All this innovation has been done under the umbrella of POSIX.1 that means standardization and consistency.

II. BACKGROUND

Basically android is not POSIX compliant but some time it called partially POSIX compliant so this work is very restraining in lecture. Till now there is no such thing implemented for any MOBILE Operation System especially for Android. There are some software's like blue stack that provides the portability for Android applications but the concept of standardization is not applied like POSIX there and secondly all these type of software's work like application file run and exit but not gives the compatibility with underlying machine OS. Hence there is no implementation related work. Now this chapter explains the brief history of Android OS, application development framework for Android and POSIX its standards and APIs.

A. What POSIX Is:

POSIX is a standard to allow applications to be source-code portable from one system to another. On a system conforming to a particular version of POSIX (as measured by the test suite it passed), you should be able to just compile and run those applications, which use the POSIX (and only the POSIX) functions. POSIX basically dependent on:

- A Compilation System: A compiler, basically. Real live POSIX systems are supposed to support a standard language. For this purpose the compiling language is C. for getting the POSIX support in any application each system has a variety of way of compiling code, for each occurrence. [1]
- Headers: A set of headers that defines the POSIX interface supported on the particular system. [1] #include<stdio.h> was used header file in given example
- Libraries: Libraries are pre-compiled, vendor-supplied objects that implement the POSIX functionality for any one. The libraries are linked into the application when it is built, or in the case of dynamically shared libraries, when user runs the program. [1]
- A Run-Time System: Once user has built the program, the run-time, or operating system, allows him/her to run the application. [1]

POSIX.1 on the other hand, is not considered to be basic functionality that all systems need in order to be useful (regardless of my personal opinion). Therefore, POSIX.1 is structured as a set of well-defined options that a vendor can support, or not. The only parts of POSIX.1 that aren't optional are some additions to the basic POSIX.1 signal mechanism. POSIX.1 options. [1]

B. Android

Android is s a software stack for mobile devices, which includes an operating system, a middleware and key applications. Android SDK provides the tools and APIs necessary to develop an application using JAVA (which is a popular language amongst the developers. Currently, Android is the most popular operating system out of the several Linux based mobile operating systems (e.g.,Maemo) [4].

- Linux Kernel: Android is based on Linux but is not Linux. The kernel of Android relies on Linux version 2.6 for core system services such as security, memory management, process management, network stack, and driver model. [6]
- Libraries: The surface manager of Android library takes care of the display of the system and OpenGL is an open-source utility, which takes care of graphics of the system. [6]
- Android Runtime: The development language that is used in this section is Java. The core Libraries of

Android is very powerful, yet simple and familiar development platform as it is very similar to Linux.

- Dalvik Virtual Machine (DVM): Dalvik virtual machine focuses on two of the most important issues of mobile system: limited space, and limited power. DVM converts all the files into smaller and more optimized (.dex) suitable for mobile systems [6].
- Application Framework and Application: The activity manager is responsible to keep track of life cycle of any application. [7]

III. IMPLEMENTATION OF ANDROID POSIX.1 THIN LAYER MODEL

A. Contribution

*1) Proposed ANDROID POSIX.1 Thin layer Model:*POSIX, the Portable Operating System Interface .The goal of POSIX is the source-code portability of applications: it means transform an application from one operating system to another by simple conversion. This goal is unattainable since most applications, especially the real-world ones, require more operating system support than you can find in any particular standard. The above unfeasible objective is now achievable through POSIX. POSIX is called useful." Useful," here, means "an aid to portability," and this brings us to the goal of POSIX: source-code portability of applications. The main intention of this work is that it will provide portability for the Android real world applications. But after the development of this thin layer model of POSIX.1.Android applications will become portable (POSIX compliance) and can be run on any operating system. This model provides the benefit to users as well as Android developers by increasing the number of users of android applications and reduces the developer time and cost because of portability and equivalence.

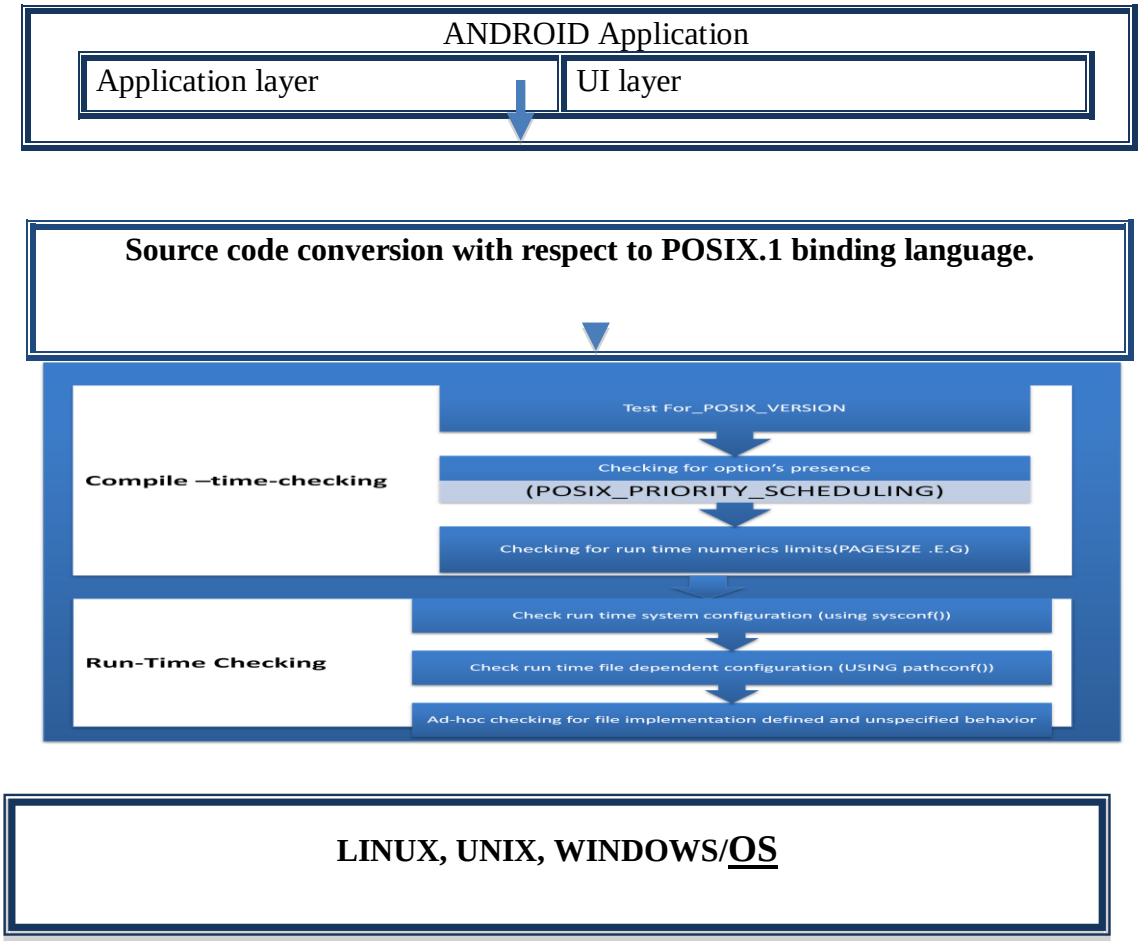


Fig.1. Proposed POSIX.1 thin layer model

2) The POSIX Development Environment: POSIX provides portability at the source level. This means that you transport your source program to the target machine, compile it with the Standard C compiler using conforming headers, and

link it with the standard libraries. The system vendor provides the compiler, the libraries, and headers. Strictly speaking, these are all black boxes and you do not need to know how they work. For POSIX .1 thin layer model implementation we used the following specifications.

TABLE I. DEVELOPMENT ENVIRONMENT SPECIFICATIONS

Specifications	POSIX environment	Android Environment
OS	Macintosh	Macintosh
IDE	Xcode	Eclipse INDIDGO
Language	C	Java
Complier	gcc version 4.2.1	Java complier

3) List of Android Applications: For this model we start the implementation from very simple to the complex one like multithreading [29].

Hello world

- Timer
- Text file creator, save data on it and display the text on the terminal
- Multithreading example [29]

The reasons of start test from very simple Android application to complex one Android application are:

- Is Android application will be POSIX compliant is a question itself. So we implement the very first sample example in both environment then we move forward that why it is part of our research work.
- The User interface means graphical user interface of POISX is not very supportive for android applications
- There is no such engine or converter that convert the whole application layer of Android application.
- All the gcc complier is not POSIX and all the desktop OS are not POISX compliant
- All the implement applications are very simple in Android environment but POSIX APIs are limited in numbers. Even for hello example POSIX standard C language have specified code.
- File creator and multithreading is very important example because it used very frequent OS calls. The IEEE Std 1003.1b-1993(pp.103) also used these examples for implementation.

TABLE II. TIMER APPLICATION FRAMEWORK

Timer Android application layer	POISX conformance Timer Android example

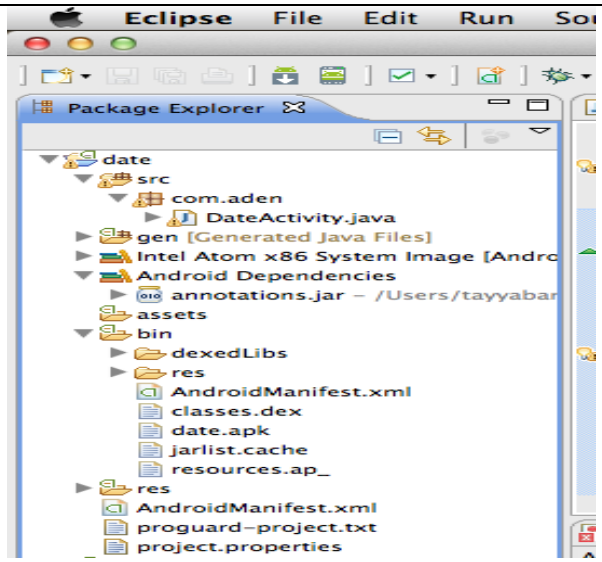
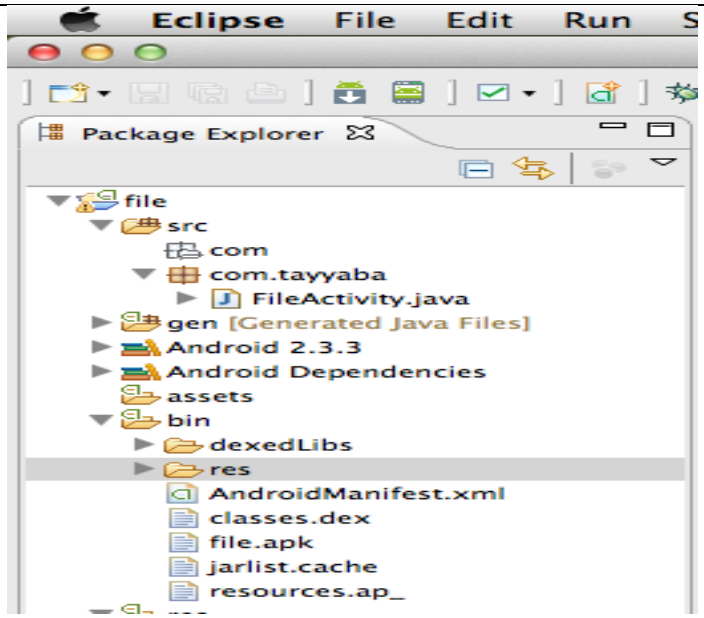
	#define _POSIXSOURCE 1
	#include<stdio.h>
	#include<time.h>
	main(argc,argv)
	intargc;
	char **argv;
	{
	struct tm *tmptr;
	time_t timer;
	timer = time(NULL);
	tmptr = localtime(&timer);
	printf("The current time is:\n%s",
	ctime(&timer));
	if (tmptr ->tm_isdst)
	printf("Daylight savings time\n");
	else
	printf("Standard time\n");
	exit(0);
	}

TABLE III. FILE CREATOR APPLICATION FRAMEWORK

File creator Android application layer	POISX conformance file creatorAndroid example

	#define _POSIX_SOURCE 1
	#include<stdio.h>
	int main(){
	FILE *fp;
	charch;
	int c;
	fp=fopen("data.txt","w");
	printf("\nTHIS DATA WRITTEN TO A FILE:");
	while((ch=getchar())!=EOF)
	putc(ch,fp);
	fclose(fp);
	fp=fopen("data.txt","r");
	c = fgetc(fp) ;
	while (c!= EOF)
	{putchar(c);
	c = fgetc(fp);
	printf("\nTHIS DATA WRITTEN TO A FILE:"+c);
	}fclose(fp);}

4) Android Application Template for POSIX.1: This template would be change according to the Application or need of the developer. But the #define _POSIX_SOURCE 1 is compulsory part of any application. [29]

TABLE IV. ANDROID POSIX.1 APPLICATION TEMPLATE

Template	Description
/* Feature test switches */ #define _POSIX_SOURCE 1	define the _POSIX_SOURCE macro to enable the POSIX symbols and disable all unspecified symbols.
/* System headers */	Each Standard C or POSIX function has one or more headers that must be included to define the symbols used by that function.
/* Local headers */	Most projects have at least one project header. These define common data structures and symbols that are used in many files.
/* Macros */	Define all of your macros here.
/* File scope variables */	These are variable that are shared by several functions in the same file.
/* External variables */	This is the list of variables defined in other modules and used in this module.
/* External functions */	There should be a prototype for each user-written external function that you use.
/* Structures and unions */	Define all of the structures that are used only in this file. Any structure that is used in multiple files should be in a local header file.
/* Signal catching functions */	Place signal catching functions in one place. Signals are an unusual calling mechanism and often hard to debug. Unless you point it out clearly in your source code, it may not be obvious that something is a signal catching function.
/* Functions */	Define functions here.
/* Main */	There is a main() function in this file

a) Used Some Core Portable Functions: The fgetc(), getc() and getchar() Functions are very portable. For example in file creation, deletion and read data from it .the application used these functions for reading data from created file.

```
c = getc(fp) ;
while (c!= EOF)
{
putchar(c);
c = getc(fp);
```

The call fgetc(stream) returns the next character from stream. If stream is at end-of-file, EOF is returned.The getc() function is the same as fgetc() except it may be implemented as a macro. These functions are very portable. So through these portable functions we are able to write a portable calls like for reading a data from text file char *fgets(char *s, int n, FILE *stream);

b) Opening and Closing File Functions: The fopen() function is used to connect a file with a stream:

fp=fopen("data.txt","w");
Create text file with name data and
w Create new file for writing. If a file with this name already exists, its contents are lost.

Some systems make a distinction between text files and binary files. While there is no such distinction in POSIX, a 'b' may be appended to the mode string to indicate binary. The b does not do anything but may be useful for compatibility with non-POSIX systems. If you are creating a binary file, include the b to make your program more portable. Most systems that do not support the b option will just ignore it.

Upon success, the fopen() function returns a pointer to a file descriptor. This pointer is used only as an argument to other functions. Do not attempt to manipulate the object it points at. If the open fails, fopen() returns a null pointer.

When you are finished with a file, you should close it. The call fclose(stream) will complete any pending processing, release system resources, and end access to the file. If there are no errors, fclose() returns zero. It returns EOF if any errors are detected.

```
int main(){
   FILE *fp;
   charch;
    int c;
fp=fopen("data.txt","w");
printf("\nTHIS DATA WRITTEN TO A FILE:");
while((ch=getchar())!=EOF)
putc(ch,fp);
fclose(fp);
```

5) Sample examples code matching with Android Application template for POSIX.1:

TABLE V. MATCHING OF POSIX.1 COMPLIANT ANDROID APPLICATION WITH POSXI.1 TEMPLATE

Template	Text file creator example
/* Feature test switches */ #define _POSIX_SOURCE 1	#define _POSIX_SOURCE 1
/* System headers */	#include<stdio.h>

/* Main */
/* Functions */
	int main(){
	fp=fopen("data.txt","w");
	printf("\nTHIS DATA WRITTEN TO A FILE:");
	while((ch=getchar())!=EOF)
	putc(ch,fp);
	fclose(fp);
	fp=fopen("data.txt","r");
	c = fgetc(fp) ;
	while (c!= EOF)
	{putchar(c);c = fgetc(fp);
	printf("\nTHIS DATA WRITTEN TO A FILE:"+c);} fclose(fp);}

/* File scope variables */	FILE *fp; char ch; int c;

/* External functions */	fclose(fp);
	putchar(c);
	fopen("data.txt","r");

Template	Timer example
/* Feature test switches */ #define _POSIX_SOURCE 1	#define _POSIX SOURCE 1
/* System headers */	#include<stdlib.h> #include<stdio.h> #include<time.h>

/* Main */
/* External functions */
	main (argc,argv)
	{ struct tm *tmptr;
	timer = time(NULL);
	tmptr = localtime(&timer);
	printf("The current time is:\n%s",
	ctime(&timer));
	if (tmptr ->tm_isdst)
	printf("Daylight savings time\n");
	elseprintf("Standard time\n");
	exit(EXIT_SUCCESS); }

/* File scope variables */	intargc; char **argv;
/* Structures and unions */	struct tm *tmptr; time_t timer;

6) Sample Examples:

TABLE VI. TESTED SAMPLE ANDROID APPLICATIONS CODE AND OUTPUT COMPARISON

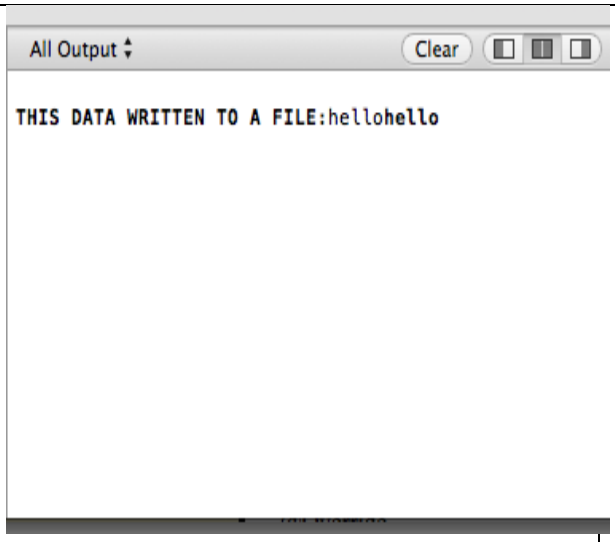
Sample Example Applications Comparison Table	
Text File creator POSIX Conformance Android application	File creator Android application
#define _POSIX_SOURCE 1	packagecom.tayyaba;

```
#include<stdio.h>
int main(){
   FILE *fp;
charch;
int c;
fp=fopen("data.txt","w");
printf("\nTHIS DATA WRITTEN TO A FILE:");
while((ch=getchar())!=EOF)
putc(ch,fp);
fclose(fp);
   //char *fgets(char *s, int n, FILE *stream);


fp=fopen("data.txt","r");
c = fgetc(fp) ;
while (c!= EOF)
   {
                        putchar(c);
                        c = fgetc(fp);
printf("\nTHIS DATA WRITTEN TO A FILE:"+c);
   }

fclose(fp);
}
```

```
importjava.io.BufferedReader;
importjava.io.FileNotFoundException;
importjava.io.IOException;
importjava.io.InputStream;
importjava.io.InputStreamReader;
importjava.io.OutputStreamWriter;
importandroid.app.Activity;
importandroid.content.Context;
importandroid.os.Bundle;
importandroid.util.Log;
importandroid.view.View;
importandroid.widget.EditText;
importandroid.widget.TextView;
importandroid.widget.Toast;
public class FileActivity extends Activity {private static final String
TAG = FileActivity.class.getName();
            private static final String FILENAME =
"myFileTayyaba.txt";
      @Override
public void onCreate(Bundle savedInstanceState) {
super.onCreate(savedInstanceState);
setContentView(R.layout.main);}
public void SaveText(View view){
          // EditText ET = (EditText)findViewById(R.id.editText1);
                        EditText ET =
(EditText)findViewById(R.id.editText1);
                        String textToSaveString =
ET.getText().toString()
                        //String textToSaveString = "Hello
Android tayyaba";
            writeToFile(textToSaveString);
                        String textFromFileString=
readFromFile();
                        if (
textToSaveString.equals(textFromFileString) )
            Toast.makeText(getApplicationContext(), "both string are
equal", Toast.LENGTH_SHORT).show();
                        else
            Toast.makeText(getApplicationContext(), "there is a
problem", Toast.LENGTH_SHORT).show();
                        Toast.makeText(this,"Text Saved
!",Toast.LENGTH_LONG).show();}
            private void writeToFile(String data) {
try {
OutputStreamWriteroutputStreamWriter = new
OutputStreamWriter(openFileOutput(FILENAME,
Context.MODE_PRIVATE));
outputStreamWriter.write(data);
Log.e(TAG, "File write : ");
outputStreamWriter.close();}
catch (IOException e) {
Log.e(TAG, "File write failed: " + e.toString());}          }
            private String readFromFile() {
       String ret = "";
try { InputStreaminputStream = openFileInput(FILENAME);
if ( inputStream != null ) {
            InputStreamReaderinputStreamReader = new
```

```
InputStreamReader(inputStream);
        BufferedReaderbufferedReader = new
BufferedReader(inputStreamReader);
        String receiveString = "";
        StringBuilderstringBuilder = new StringBuilder();
        while ( (receiveString = bufferedReader.readLine()) !=
null ) {
stringBuilder.append(receiveString);}inputStream.close();
        ret = stringBuilder.toString();
        TextViewtv = (TextView)findViewById(R.id.textView1);
tv.setText("text file data:"+ret);
        Log.e(TAG, "Can  read file: " + ret.toString());
} }  catch (FileNotFoundException e) {
        Log.e(TAG, "File not found: " + e.toString());}catch
(IOException e) {Log.e(TAG, "Can not read file: " + e.toString());
}return ret;}}
```

Output



IV. Analysis And Results

Essentially we are trying to provide the standardization (through POSIX) and portability of Android applications on multiple operating systems. Because a well-structured program is portable among the different programmers who may maintain it. Placing program elements in a consistent order makes finding things easier. [3]

- Portability: POSIX .1 thin layer model is initiation point for Android application portability to different operating systems.
- Reusability: POSIX .1 thin layer model provides the reusability of the Android applications on multiple operating systems.
- Standardization: POSIX .1 thin layer model is a standard way of transformation of application with damaging the application internal structure.
- Diversity: POSIX .1 thin layer model gives the diversity to the Android application market.

A. Quantified Feasibility Analysis

At this time Android covers the 53% of the Smartphone market share as shown in figure in 1. [23]. But we turn into 100% by introducing POSIX.1 Thin layer Model. It provides the viability for Android mobile users as well as developers. After implementation of this POSIX.1 Thin layer Model the Android applications can run on any operating system so the Android covers the 100% market, which means the revenue according to figure 9 it would be double. The statistical result is given in the table.1. This model focuses on the Android mobile users and Android developer through reusability and standardization

Fig.2. Worldwide Smartphone Sales to End User by Operating source: Gartner (February 2013)[23]

The Android applications market revenue increasing very rapidly [24]. Like in figure.1 Android market growth is 861.5%, which is tremendous. But question is here Why are restricting Android market only to the Mobile OS.

Global Mobile Applications Store Ranking in 2010 and 2009
(Ranking by Revenue in Millions of U.S. Dollars)

2010 Rank	Store	2009 Revenue	2009 Share	2010 Revenue	2010 Share	Year-Over-Year Growth
1	Apple App Store	$769	92.8%	$1,782	82.7%	131.9%
2	BlackBerry App World	$36	4.3%	$165	7.7%	360.3%
3	Noka Ovi Store	$13	1.5%	$105	4.9%	719.4%
4	Google Android Market	$11	1.3%	$102	4.7%	861.5%
	Total	$828	100.0%	$2,155	100.0%	160.2%

Source: IHS Screen Digest February 2011

Fig.3. Android market share [24]

The market of the desktop OS is very large as shown in the below table.1.but if we merge both these markets only for the Android the result is very magnificent in the form of revue however also in the form of manpower reduction which shown in the table.8.

TABLE VII. MARKET SHARE ANALYSIS FOR ANDROID DEVELOPERS [25]

POSIX Compliant operating systems	Market share of desktop Operating system	Android market share	Total market share for developers
Windows 7	44.55%	53.1%	97.65%
Windows XP	38.99%	53.1%	92.09%
Windows Vista	5.17%	53.1%	58.27%
Mac OS X 10.8	2.61%	53.1%	55.71%
Windows 8	2.67%	53.1%	55.77%
Mac OS X 10.6	1.97%	53.1%	55.07%
Mac OS X 10.7	1.93%	53.1%	55.03%

TABLE VIII. COMPARATIVE ANALYSIS OF APPLICATIONS DEVELOPMENT TIME AND MANPOWER FOR ANDROID DEVELOPERS

Application type	Application development manpower	Application development time	POSIX application development time	POSIX application development manpower
Entertainment	6-7 developers	30 days (min) 120 days (max)	3 developers	15 days (min) 60 days (max)
Lifestyle	2-3 developers	20 days (min) 60 days (max)	1 developer	10 days (min) 30 days (max)
Productivity	10-11 developers	30 days (min) 120 days (max)	5-6 developers	15 days (min) 60 days (max)
Libraries & Demo	10-11 developers	30 days (min) 120 days (max)	5-6 developers	15 days (min) 60 days (max)

1) Resulting Impact Factor for Android Developers: The feasibility study of the POSIX.1 thin layer Model clearly revealed a lot of benefits for developers.

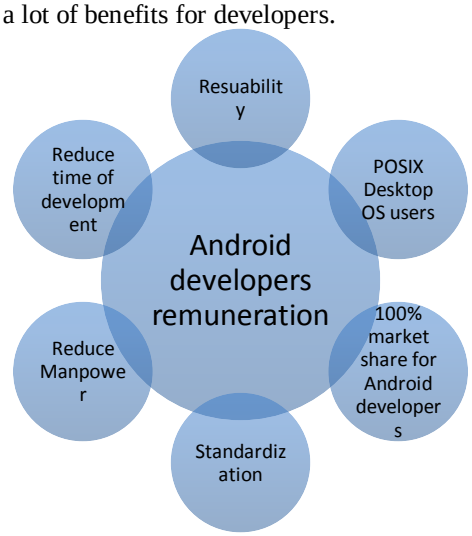


Fig.4. POSIX.1 thin layer Model resulting impact factors

V. LIMITATIONS

We have faced multiple limitations related to POSIX as well as related to the Android applications.

- POSIX have a list of standards and some of these standards are not still verified with IEEE. Secondly POISX bonding languages are very extinct so POISX programming is very difficult tasks. With passing each day POSIX standards are modified very

frequently. These abrupt changes in standards becomes the developer life miserable.

- A lot of Android applications are GUI dependent and POISX .1 support very limited GUI features so need of GUI functions in POISX .1.

There is no standard tool or engine for language conversion from java to standard C.

VI. Future Work

Till now there is only Application layer (code) implementation through this model but the need of implementation of UI layer is very stimulating and tempting. Although XML code conversion is very difficult and C library limitation for interface. The Hardware acceleration for Graphics subsystem is also in require for completing implementation.

Second option is related to making the Android OS POSIX compliant. This is not an easy task as there are a lot of limitations of Android hardware restriction, Android devices variety plus DVM [32] but the proposed model is one of best solution for all these limitations. The small size of usage hardware is obvious but implementation is not impossible by executing the Standard C library instead of using C/C++ for code conversion. [33] This is only the idea but achievement needs more attention.

In the below model we are try to introduce a new layer, which called the Java POISX APIs layer. This layer simply converts the all java APIs into POISX APIs but still in java language .so DVM consider it as java command and convert it into dex. Format.

The Android Runtime consists of the Dalvik virtual machine and the Java core libraries. The Dalvik virtual machine is an interpreter for byte code that has been transformed from Java byte code to Dalvik byte code. [30]

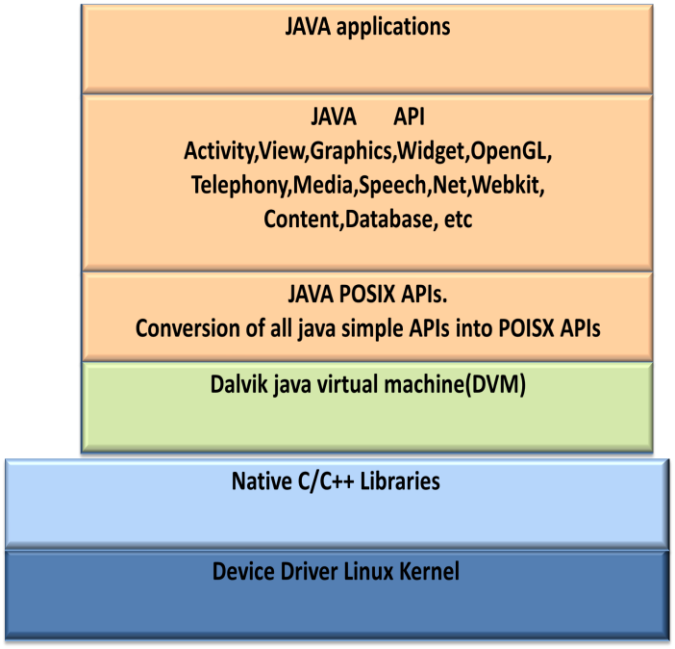


Fig.5. proposed Android OS POSIX compliant model

Dalvik itself is compiled to native code whereas the core libraries are written in Java, thus interpreted by Dalvik. It means conversion of java to C/C++ done here. But we are try to convert the Android OS POISX .One solution is the POSIX package. This package provides access to the POSIX API from Java. However essential question is that where put this POISX APIs library for Java?

As shown in above figure 7 we put the java POSIX APIs layer that Basically DVM do the conversion of java applications to .dex format means conversion of java to C/C++ .so DVM has not problem if there is any JAVA API so if we put the JAVA POSIX API [31] layer which convert the alljava simple APIs to POISX APIs but still in the java language. So DVM very easily do it conversion because DVM consider it a java API

VII. Conclusion

Currently, Android is the most popular operating system out of the several Linux based mobile operating systems (e.g.,Maemo) [4].POSIX .1 thin layer model assigns the Android applications to a wider marketplace without restricting to them with only mobile computing. In this work, the main theme of research provides the portability to the Android Application with POSIX.1 standard. In summary introducing this thin layer POSIX.1 model expands the market for Android applications and adds real-time capability and higher reusability

Acknowledgment

This research work is done under the supervision of Prof. Dr. Shoab Ahmad Khan. The Integrated development environments are Eclipse INDOGO for Android Application with Android virtual device (AVD) Intel Atom x86 with API level 10 and Xcode with gcc complier version is mention in previous Fig.1 at Macintosh OS.

References

[1] IEEE/ANSI Std 1003.1: Information Technology-- (POSIX®)--Part 1: System Application: Program Interface (API) [C Language], includes (1003.1a, 1003.1b, and 1003.1c). 1996.

[2] Bill O. Gallmeister,POSIX. 4: Programming for the Real World, 1995.ppt.4, 19-20,22,23

[3] Donald A. Lewine, POSIX Programmer's Guide Writing Portable UNIX Programs with the POSIX.1 Standard, 1991,pp.16-17, 25

[4] E. Oliver, A Survey of Platforms for Mobile Networks Research. Mobile Computing and Communications Review, December 2008, pp. 56-63.

[5] Hassan Reza, and Narayana Mazumder, A Secure Software Architecture for Mobile Computing (2012 IEEE)

[6] Android. http://code.google.com/android/

[7] Canalys Report. http://www.canalys.org

[8] http://www.android.com/about/

[9] http://developer.android.com/about/versions/index.html

[10] POSIX. 1: ISO/IEC 9945-1:1990 IEEE Std. 1003.1-1990

[11] Donald A. Lewine, POSIX Programmer's Guide Writing Portable UNIX Programs with the POSIX.1 Standard, 1991,pp.31-36

[12] http://developer.android.com/sdk/eclipse-adt.html[Aug.20, 2011].

[13] http://www.ibm.com/developerworks/opensource/library/os-android-devel/[Aug.10, 2011].

[14] http://en.wikipedia.org/wiki/Android_ (operating_system) [Aug.10, 2011]

[15] http://en.wikipedia.org/wiki/Android_Market [Aug.20.2011].

[16] http://developer.android.com/guide/basics/what-is-android.html[Aug.20, 2011].

[17] Barra, Hugo (10 May 2011). "Android: momentum, mobile and more at Google I/O".The Official Google Blog. Retrieved 10 May 2011.

[18] http://www.google.com/support/androidmarket/developer/bin/answer.py?answer=113475[Aug.6, 2011]

[19] http://www.android.com/about/[Aug, 6, 2011]

[20] Lawson, Stephen (17 March 2009). "Android Market Needs More Filters, T-Mobile Says". PC World.

[21] http://www.gartner.com/newsroom/id/2335616

[22] http://appleinsider.com/articles/11/02/18/rim_nokia_and_googles_android_battle_for_apples_ios_scraps_as_app_market_sales_grow_to_2_2_billion.html

[23] http://www.netmarketshare.com

[24] IEEE Std 1003.1b-1993 (Formerly known as IEEE P1003.4) (Includes IEEE Std 1003.1-1990)

[25] Kolin Paul, Tapas Kumar Kundu "Android on Mobile Devices: An Energy Perspective," 10th IEEE International Conference on Computer and Information Technology, 2010.

[26] Kyosuke Nagata,Saneyasu Yamaguchi "An Android Application Launch Analyzing System"

[27] IEEE Portable Applications Standards Committee, P1003.13: Infonnaiion Technology - Siandardized Applications Environment Profile - POSIX Real-time Application Support (AEP) (Draft 5) (Feb 1992).

[28] Namseung Lee, Sung-Soo Lim, "A Whole Layer Performance Analysis Method for Android Platforms", (2011 IEEE).

[29] Java POSIX APIs,

[30] http://bmsi.com/java/posix/posix-1.2.2/doc/index.html.Accessed March 2012

[31] Leonid , Aubrey-Derrick , Hans-Gunther , Ahmet Camtepe and Sahin Albayrak," Developing and Benchmarking Native Linux Applications on Android," Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Volume 7, pp 381-392, 2009.

[32] E. Cooper and R. Draves, "C threads". TR CMU-CS-88- 154, Carnegie Melloii University, Dept. of CS (1988).

Classification of Image Database Using Independent Principal Component Analysis

H. B. Kekre
Professor,
Computer Engineering
Mukesh Patel School of Technology,
Management and Engineering,
NMIMS University, Vileparle(w)
Mumbai 400–056, India

Tanuja K. Sarode
Associate Professor,
Computer Engineering,
Thadomal Shahani Engineering
College, Bandra(W),
Mumbai 400-050, India

Jagruti K. Save
Ph.D. Scholar, MPSTME,
NMIMS University,
Associate Professor,
Fr. C. Rodrigues College of
Engineering, Bandra(W),
Mumbai 400-050, India

***Abstract*—the paper presents a modified approach of Principal Component Analysis (PCA) for an automatic classification of image database. Principal components are the distinctive or peculiar features of an image. PCA also holds information regarding the structure of data. PCA can be applied to all training images of different classes together forming universal subspace or to an individual image forming an object subspace. But if PCA is applied independently on the different classes of objects, the main direction will be different for them. Thus, they can be used to construct a classifier which uses them to make decisions regarding the class. Also the dimension reduction of feature vector is possible. Initially training image set is chosen for each class. PCA, using eigen vector decomposition, is applied to an individual class forming an individual and independent eigenspace for that class. If there are n classes of training images, we get n eigenspaces. The dimension of eigenspace depends upon the number of selected eigen vectors. Each training image is projected on the corresponding eigenspace giving its feature vector. Thus n sets of training feature vectors are produced. In testing phase, new image is projected on all eigenspaces forming n feature vectors. These feature vectors are compared with training feature vectors in corresponding eigenspace. Feature vector nearest to new image in each eigenspace is found out. Classification of new image is accomplished by comparing the distances between the nearest feature vectors and training image feature vector in each eigenspace. Two distance criteria such as Euclidean and Manhattan distance are used. The system is tested on COIL-100 database. Performance is tested and tabulated for different sizes of training image database.**

Keywords—Image classification; Principal Component Analysis (PCA); Eigen value; Eigen vector; Variance; Nearest neighbor classifier; Orthogonal transform; Feature vector; Covariance matrix.

I. INTRODUCTION

Principal Component Analysis (PCA) is the most popular and the oldest multivariate statistical technique [1]. PCA was invented in 1901 by Karl Pearson [2], who formulated the analysis as "Finding lines and planes of closest fit to systems of points in space". The focus was on geometric optimization. Later it was re-invented by Harold Hotelling in 1933[3]. In image analysis, the term Hotelling transformation is often used for a principal component projection. PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Since patterns in data can be hard to find in data of high dimension, where the luxury of graphical representation is not available, PCA is a powerful tool for analysing data.

Generally classification of images is a two step process, feature vector generation followed by a nearest-neighbor classifier [4][5]. Classification accuracy depends on many factors. One major factor is the extraction of features to represent the image. Feature extraction is a special form of dimensionality reduction. PCA is a method used to reduce the number of features used to represent the data. The benefits of this dimensionality reduction include providing a simpler representation of the data, reduction in memory, and faster classification. PCA transforms the original variables to a new set of variables, that are uncorrelated and ordered such that the first few retains most of the information present in the data.[6]. These uncorrelated components are called principal components (PC) and are estimated from the eigenvectors of the covariance or correlation matrix of the original variables. PCA has been widely used for image processing applications such as face recognition[7][8][9][10], palm print recognition[11], image compression[12][13], image fusion[14], image enhancement[15] [16] ,object recognition[17] etc.

This research paper is structured as follows: Section II explains the generation of principal components by PCA method. Section III describes the methodologies of the system. Section IV presents the results. Finally section V describes the conclusions and proposing some possible future work followed by references.

II. PRINCIPAL COMPONENT ANALYSIS

Principle Components Analysis (PCA) is a well-known method to identify statistical trends in data. It projects the data from a higher dimension to a lower dimensional manifold such that the error incurred by reconstructing the data in the higher dimension is minimized. As shown in fig. 1, given a set of points in Euclidean space, the first principal component Z_1 corresponds to a line that passes through the multidimensional mean and minimizes the sum of squares of the distances of the points from the line. The second principal component Z_2 corresponds to the same concept after all correlation with the first principal component has been subtracted from the points. Principal components are a series of linear least square fits to a sample, each orthogonal to all previous. The Principle

Components reveal important information about the dispersion of the original data set.

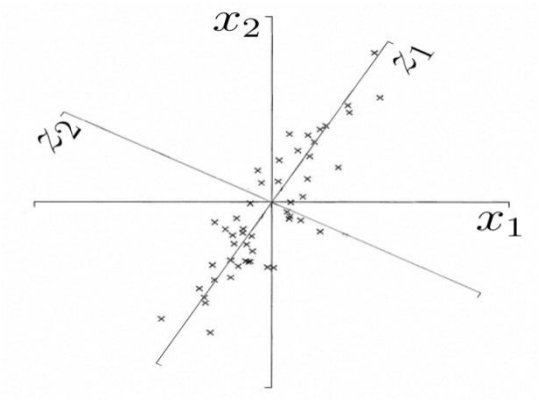


Fig.1. Principal components Z_1 and Z_2

PCA is based upon eigenvector decomposition of a covariance matrix. For multivariate data, covariance is a measure of the relationship between different variables, or dimensions of the data set. The general steps of PCA[18][19] are as follows:

1) Acquire data.

2) Subtract mean from the data.

3) Generate the covariance matrix[20]. Important property of covariance matrix is that it is square, real, and symmetric. This means that there always exists n real eigenvalues for an n×n covariance matrix.

4) Calculate the eigenvalues and eigenvectors of the covariance matrix. The first Principle Component is the eigenvector of the covariance matrix with the largest eigenvalue. It represents the most significant relationship between the data dimensions.

5) Compute the cumulative energy content for each eigenvector.

6) By ordering the eigenvectors in the order of descending eigenvalues (largest first), create an ordered orthogonal basis.

7) Use this basis to transform input data vector. Instead of using all the eigenvectors of the covariance matrix, the data can be represented in terms of only a few basis vectors of the orthogonal basis.

III. Proposed Algorithm

From the image database, some images are used for training and the remaining images are used for testing. The algorithm used to generate the feature vector for each training image is given below.

A. Feature vector generation for training images

Consider there are 'M' training images in each class and there are such 'N' classes. All images are converted into gray scale images.

For each class do the steps 1 to 10.

Step 1:Find the average image of that class. Refer equation 1.

$$I_{avg}(x,y) = \frac{1}{M}\sum_{i=1}^{M} I_i(x,y) \tag{1}$$

Step 2: Find zero mean images by subtracting the average image from each image as given in equation 2.

$$I_i = I_i - I_{avg} \quad for \quad i = 1...M \tag{2}$$

Step 3: Convert zero mean images into one dimensional vector by arranging the columns of an image one below the other as shown in fig.2.

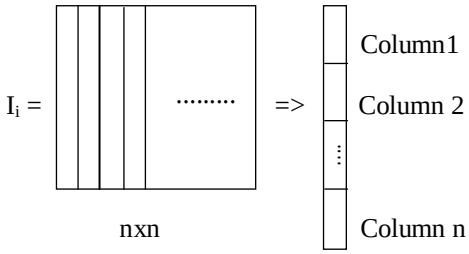


Fig.2. Conversion of an image into vector

Step 4: Organize these vectors say ϕ_i (for i=1 to M) in matrix form as shown below in fig.3.

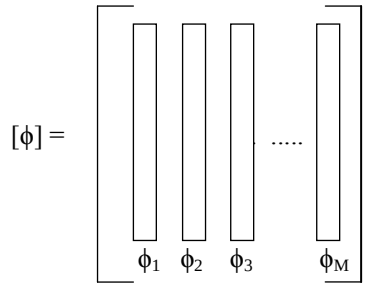


Fig.3. Organization of vectors

Step 5: Calculate the covariance matrix A(M x M) using the formula (refer equation 3).

$$A = 1/M\{[\phi]^t[\phi]\} \tag{3}$$

Step 6: Calculate the eigen values (λ_1 to λ_M) and eigen vectors (X_1 to X_M) by solving equation 4. Eigen vectors are ordered according to the corresponding eigen values from high to low.

$$[A - \lambda I]X = 0 \tag{4}$$

Step 7: Construct eigen images as given in equation 5. (Exclude eigen vector corresponding to lowest eigen value as it is extremely small comparatively)

$$[\phi]X_i = F_i$$
$$for\ i = 1..M-1 \quad (5)$$

Step 8. Convert each vector F_i into 2 dimensional eigen image as shown below in fig. 4.

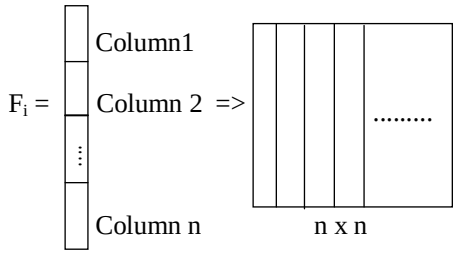


Fig.4. . Conversion of a vector to an image

Step 9: Calculate the cumulative energy μ for each F_i image.

Step 10: Calculate the feature vector V_i for each training image I_i of the class as in equation 6

$$V_i = \begin{bmatrix} w_{1i} \\ w_{2i} \\ \vdots \\ w_{(M-1)i} \end{bmatrix} \quad (6)$$
$$for\ i = 1..M$$

where the coefficient w_{ji} is calculated as given in equation 7

$$w_{ji} = \frac{1}{\mu_j} \sum_{x,y} I_i(x,y) F_j(x,y) \quad (7)$$
$$for\ j = 1...M-1,\ i = 1..M$$

After applying this entire procedure for all 'N' classes, we get the average image of each class, M-1 eigen images for each class and one feature vector of size (M-1)x1 for each training image of that class. Since total training images are MxN, we get 'MN' training feature vectors.

B. Feature vector generation for testing images

Each testing image is converted into gray scale image. The procedure used to generate feature vector for testing image I_{test} is given below(step 1 and step 2). This procedure is repeated for all other testing images.

For each training class do

Step1: Find zero mean test image as given in equation 8.

$$I_{ztest} = I_{test} - I_{avg} \quad (8)$$

Where I_{avg} is the average image of that class.

Step2: Calculate the feature vector V_{test} (refer equation 9)

$$V_{test} = \begin{bmatrix} w_{test1} \\ w_{test2} \\ \vdots \\ w_{testM-1} \end{bmatrix} \quad (9)$$

where each coefficient w_{testj} is given in equation 10.

$$w_{testj} = \frac{1}{\mu_j} \sum_{x,y} I_{ztest}(x,y) F_j(x,y) \quad (10)$$

Since there are 'N' classes, we get 'N' feature vectors for single test image. Each feature vector size is (M-1)x1.

C. Classification of testing image

When we apply the algorithms explained in section A, we get training feature vector set V containing MxN column feature vectors. Column vector V_{ij} denotes the feature vector for i^{th} training image of j^{th} class.

After applying the algorithm from section B we get feature vector set V_{test} (for single testing image) containing N column vectors. Each vector V_{testj} denotes the feature vector of testing image on j^{th} class. Procedure to classify the given testing image is given below:

```
For j=1 to N(Number of classes)
    For i=1 to M (Number of training images)
        D(i,j)=Euclidean_dist(V_ij,V_testj)
    End
End
For j=1 to N
    d_min(j)= Minimum value in j^th column of
              matrix D
End
J = Index of the minimum value of vector d_min
```

The given testing image is assigned to J^{th} class. This procedure is executed for all testing images. Manhattan distance criteria are also used to find the distance between training and testing feature vector.

IV. RESULTS

The implementation of the proposed technique is done in MATLAB 7.0 using a computer with Intel Core i5, CPU (2.50GHz and 6 GB RAM). The proposed technique is tested on the COIL-100[21] image database. Columbia Object Image Library (COIL 100) is a database of color images of 100 objects. The objects have a wide variety of complex geometric and reflectance characteristics.

The objects were placed on a motorized turntable against a black background. The turntable was rotated through 360° to

vary object pose with respect to a fixed color camera. Images of the objects were taken at pose intervals of 5 degrees. This corresponds to 72 poses per object. Thus this database has total 7200 images. Fig. 5 shows the sample images of coil database objects. Fig. 6 shows the sample of different images in a particular class.

Fig.5. Sample images of objects in coil database

Fig.6. Sample images in one class of coil database

Initially for training, 10 images per class are used. These images are orderly selected from database at an angle of 35º,70º,105º,140º,....so on. Remaining 62 images per class are used for testing purpose. Thus, out of 7200 images in database, 1000 images are used for training and 6200 images are used for testing purpose. Two distance criteria such as Euclidean distance (ED) and Manhattan distance (MD) are used for finding the difference between feature vectors. Results obtained are given in table I. Then the number of training images increased from 10 to 18. These images are orderly selected from database at an angle 0º,20º,40º,60º,....so on. Remaining 54 images per class are used for testing purpose. Thus out of 7200 images in database, 1800 images are used for training and 5400 images are used for testing purpose. Table II shows the performance in the case.

TABLE I. NUMBER OF CORRECTLY CLASSIFIED IMAGES (OUT OF 62 PER CLASS, 100 CLASSES) (1000 TRAINING IMAGES AND 6200 TESTING IMAGES)

Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD
1	19	26	21	36	35	41	33	31	61	17	18	81	28	28
2	34	34	22	16	21	42	30	31	62	12	13	82	28	30
3	35	34	23	34	35	43	23	25	63	28	32	83	25	32
4	42	38	24	6	10	44	34	32	64	18	19	84	28	32
5	27	28	25	26	23	45	26	29	65	12	13	85	40	38
6	27	27	26	21	26	46	33	31	66	11	13	86	9	11
7	24	10	27	26	27	47	31	28	67	11	19	87	35	37
8	28	25	28	40	43	48	37	37	68	31	32	88	29	28
9	33	31	29	29	35	49	17	22	69	22	21	89	26	36
10	38	44	30	11	9	50	9	9	70	13	9	90	28	31
11	30	35	31	15	20	51	25	24	71	20	24	91	28	29
12	27	33	32	36	35	52	41	40	72	34	32	92	23	21
13	20	18	33	21	24	53	36	35	73	32	32	93	19	18
14	31	30	34	30	31	54	22	27	74	41	42	94	12	24
15	31	32	35	32	32	55	21	24	75	37	36	95	23	21
16	26	26	36	19	21	56	43	37	76	22	26	96	23	26
17	41	39	37	23	24	57	40	39	77	39	35	97	34	36
18	13	16	38	28	30	58	12	14	78	31	31	98	26	26
19	28	27	39	23	22	59	19	24	79	23	25	99	20	25
20	39	38	40	44	42	60	42	42	80	41	40	100	38	41

Note : Numbers in green indicate Euclidean distance gives better performance than Manhattan distance for that class. Numbers in yellow indicate Manhattan distance gives better performance comparatively for that class. No color indicates identical performance.

Observations: Out of 100 object classes, in 36 object classes, Euclidean distance gives better performance. In 53 object classes Manhattan distance outperforms Euclidean distance and in 11 object classes, both the distance criteria equally performed.

TABLE II. NUMBER OF CORRECTLY CLASSIFIED IMAGES (OUT OF 54 PER CLASS, 100 CLASSES) (1800 TRAINING IMAGES AND 5400 TESTING IMAGES)

Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD
1	29	43	21	41	42	41	41	40	61	32	33	81	42	48
2	31	42	22	24	27	42	45	43	62	41	41	82	46	45
3	48	46	23	46	47	43	44	41	63	49	43	83	37	34
4	44	36	24	22	25	44	44	48	64	28	28	84	23	25
5	33	29	25	27	30	45	38	45	65	22	29	85	46	47
6	37	34	26	33	35	46	42	41	66	29	28	86	19	14
7	34	22	27	44	44	47	39	34	67	20	33	87	24	33
8	36	33	28	48	48	48	40	40	68	42	42	88	39	41
9	38	37	29	37	43	49	28	32	69	27	29	89	38	46
10	42	44	30	16	12	50	24	22	70	13	13	90	39	43
11	38	38	31	31	38	51	36	37	71	39	40	91	35	32
12	30	31	32	32	26	52	46	44	72	36	38	92	25	27
13	29	32	33	37	38	53	45	44	73	43	44	93	47	47
14	37	43	34	48	41	54	37	34	74	47	48	94	22	29
15	38	38	35	26	31	55	36	37	75	47	48	95	0	0
16	39	40	36	44	47	56	31	31	76	34	34	96	35	37
17	38	37	37	40	41	57	46	45	77	40	41	97	39	42
18	23	23	38	44	47	58	28	27	78	44	44	98	25	29
19	45	45	39	32	34	59	29	30	79	42	44	99	32	33
20	43	41	40	51	48	60	44	42	80	49	47	100	47	35

Note : Numbers in green indicate Euclidean distance gives better performance than Manhattan distance for that class. Numbers in yellow indicate Manhattan distance gives better performance comparatively for that class. No color indicates identical performance.

Observations: Out of 100 object classes, in 33 object classes, Euclidean distance gives better performance. In 51 object classes Manhattan distance outperforms Euclidean distance and in 16 object classes, both the distance criteria equally performed.

Then the number of training images increased from 18 to 24. These images are orderly selected from database at an angle 0º,15º,30º,45º,....so on. Remaining 48 images per class are used for testing purpose. Thus out of 7200 images in database, 2400 images are used for training and 4800 images are used for testing purpose. Table III shows the performance in the case. Table IV shows the overall performance.

TABLE III. NUMBER OF CORRECTLY CLASSIFIED IMAGES (OUT OF 48 PER CLASS, 100 CLASSES) (2400 TRAINING IMAGES AND 4800 TESTING IMAGES)

Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD	Coil database Object number	ED	MD
1	33	36	21	47	47	41	45	43	61	27	20	81	44	44
2	41	41	22	23	25	42	48	46	62	31	32	82	45	42
3	48	45	23	48	47	43	48	46	63	47	47	83	40	29
4	46	41	24	9	10	44	47	47	64	19	12	84	25	27
5	15	14	25	18	23	45	34	44	65	25	27	85	48	47
6	45	44	26	37	37	46	42	43	66	27	24	86	39	30
7	2	0	27	44	41	47	28	22	67	24	29	87	22	30
8	42	43	28	48	48	48	46	47	68	47	46	88	38	35
9	36	37	29	30	25	49	38	37	69	34	35	89	41	45
10	43	42	30	18	17	50	15	15	70	19	16	90	45	44
11	38	35	31	42	34	51	40	39	71	34	34	91	39	39
12	28	24	32	24	17	52	48	48	72	44	43	92	32	30
13	22	23	33	42	45	53	48	47	73	46	46	93	30	30
14	46	46	34	42	39	54	40	41	74	47	47	94	11	15
15	47	48	35	29	34	55	41	37	75	45	45	95	10	8
16	43	41	36	46	46	56	31	37	76	38	39	96	37	38
17	46	47	37	47	47	57	46	46	77	44	42	97	45	46
18	26	24	38	48	48	58	33	28	78	48	48	98	23	29
19	46	48	39	41	40	59	32	33	79	41	41	99	31	30
20	47	45	40	48	48	60	48	46	80	48	48	100	46	47

Note : Numbers in green indicate Euclidean distance gives better performance than Manhattan distance for that class. Numbers in yellow indicate Manhattan distance gives better performance comparatively for that class. No color indicates identical performance.

Observations : Out of 100 object classes, in 45 object classes, Euclidean distance gives better performance. In 31 object classes Manhattan distance outperforms Euclidean distance and in 24 object classes, both the distance criteria equally performed.

TABLE IV. NUMBER OF CORRECTLY CLASSIFIED IMAGES AND ACCURACY

Distance criteria	Training images : 10 per class		Training images : 18 per class		Training images : 24 per class	
	Number of Correctly classified images (out of 6200)	*Accuracy*	*Number of Correctly classified images (out of 5400)*	*Accuracy*	*Number of Correctly classified images (out of 4800)*	*Accuracy*
Euclidean distance (ED)	2700	43.55	3602	66.70%	**3695**	**76.98%**
Manhattan distance (MD)	**2799**	**45.15**	**3699**	**67.94%**	3640	75.83%

Observations: In the first case, only 10 images per class are used for training and 62 images per class are used for testing, i.e. around 14% of data is used for training, so the accuracy is below 50%. In the second case, 18 images per class are used for training and 54 images per class are used for testing, i.e. 25% of data is used for training, so the accuracy is now increased above 65%. Manhattan distance criteria gives better overall accuracy comparatively in both the cases. In the third case when out of 72 images per class, 24 images per class are used for training and 48 images per class are used for testing, i.e. around 33% of data is used for training, the accuracy is increased to above 75%. Euclidean distance criteria give better overall accuracy in comparison.

For different training database size, the accuracy for each class is calculated. Fig. 3,4 and 5 shows the number of classes in the different ranges of accuracy for different sizes of training databases.

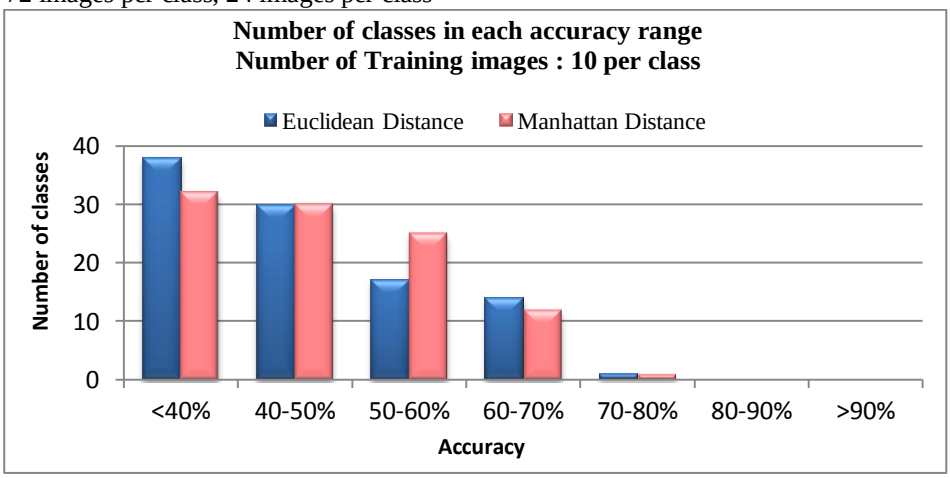


Fig.7. Accuracy performance of 100 classes when 1000 training images are used

Observations: When around 14% of data is used for training purpose, no object class gives more than 80% accuracy. Only one object class gives accuracy of 71%. Most of object classes give accuracy below 60%.

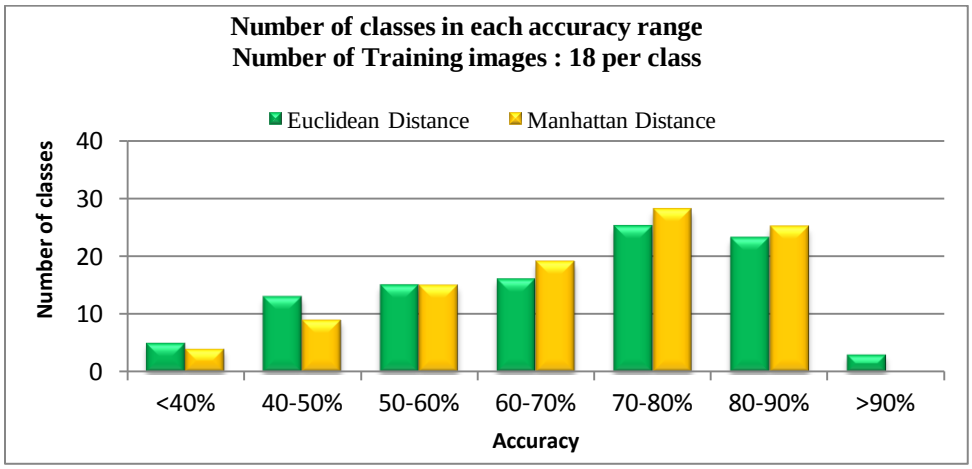


Fig.8. Accuracy performance of 100 classes when 1800 training images are used.

Observations: When 25% of data is used for training purpose, most of object classes give accuracy more than 60%. 51 object classes with Euclidean distance and 53 object classes with Manhattan distance give more than 70% accuracy. 3 classes give more than 90% accuracy with Euclidean distance.

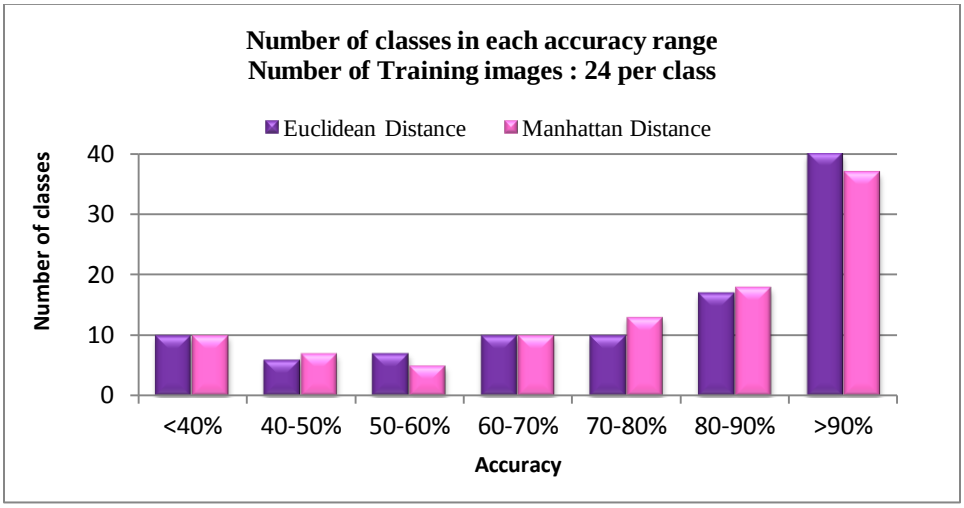


Fig.9. Accuracy performance of 100 classes when 2400 training images are used.

Observations: When around 33% of data is used for training purpose, most of object classes give more than 70% accuracy and around 40 classes give more than 90% accuracy. With Euclidean distance, 57 classes and with Manhattan distance, 55 classes give more than 80% accuracy.

V. CONCLUSIONS

The paper presents the application of PCA for an automatic classification of image database. Database used is COIL-100. It is a very large database containing 100 classes, each of 72 images, so total 7200 images. Each image is of size 128x128. If PCA is directly applied to an image for dimension reduction, then it would be computationally very intensive to find the eigen vectors of covariance matrix of size 128x128. In classification of data, generally training data is organized as columns of matrix and then PCA is applied to that matrix. But with such a large database, even if only 10 images per class are used for training purpose, the size of covariance matrix becomes 1000x1000. So to reduce the computational complexity, independent PCA is proposed and tested. In this technique the size of covariance matrix is n x n if 'n' images per class are used for training purpose. Experiments performed with three sizes of training database such as 10 images per class (13.88%), 18 images per class (25%) and 24 images per class (33.33%). When the training database is increased from around 14% to 25% and then to around 33%, the overall classification accuracy increases from 45% to 68% to 76%. Manhattan distance criterion gives overall better performance in comparison with Euclidean distance criterion when the size of training database is small. When 24 images per class are used for training, 40 object classes with Euclidean distance and 37 object classes with Manhattan distance give more than 90% accuracy. In this paper the technique is applied on grayscale image. It can be extended to all three planes of color image and combine the results.

References

[1] H. Abdi and L.J. Williams, "Principal component analysis, " Wiley Interdisciplinary Reviews: Computational Statistics, Vol.2, Issue 4, July 2010

[2] K.Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space," Philosophical Magazine **2** (11),1901,pp. 559–572.

[3] H.Hotelling,"Analysis of a complex of statistical variables into principal components," Journal of Educational Psychology, Vol 24(6), Sep 1933, 417-441.

[4] H.B.Kekre, T.K.Sarode and J.K.Save, "Image Classification in Transform Domain," International Journal of Computer Science and Information Security (IJCSIS), Vol.10,No.3, Pages 91-97, Mar 2012.

[5] H.B.Kekre, T.K.Sarode and J.K.Save, " Performance Comparison of Wavelet Transforms, Generated from Orthogonal Transforms, in Classification of Image Database, " International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), Volume 2, Issue 12, December 2012 ISSN: 2277 128X, pp. 10-18

[6] I.T.Jolliffe," Principal Component Analysis," Springer series in statistic, 2nd edition, New York:Springer; 2002.

[7] S.K.Dandpat and S.Meher, " Performance Improvement for Face Recognition Using PCA and Two-Dimensional PCA, " IEEE International Conference on Computer Communication and Informatics (ICCCI), Jan. 2013, pp. 1-5

[8] H. M. Hasan, Waleed A. AL.Jouhar and M. A. Alwan, "Face Recognition Using Improved FFT Based Radon by PSO and PCA Techniques, " International Journal of Image Processing (IJIP), Vol. 6, Issue : 1, 2012, pp. 26-37

[9] H.B.Kekre, S. D. Thepade, and A. Maloo," Eigenvectors of Covariance Matrix using Row Mean and Column Mean Sequences for Face Recognition, " International Journal of Biometrics and Bioinformatics (IJBB), Vol. 4, Issue :2, pp. 42-50

[10] H. B. Kekre, and K.Shah," Performance Comparison of Kekre's Transform with PCA and Other Conventional Orthogonal Transforms for Face Recognition, " Proceedings of the Second International Conference on Emerging Trends in Engineering and Technology (ICETET), 2009, pp. 873-879

[11] H.Li and L.Wang, "Palmprint Recognition Using Dual-Tree Complex Wavelet Transform and Compressed Sensing , " IEEE International Conference on Measurement, Information and Control (MIC), May 2012,Vol.2, pp.563-567

[12] Mohammad Bonab and Mostafa Bonab, "A New Technique for Image Compression Using PCA , " International Journal of Computer Science & Communication Networks,Mar. 2012,Vol. 2, No. 1, pp. 111-116

[13] V.Gaidhane, V.Singh and M.Kumar, "Image Compression Using PCA and Improved Technique with MLP Neural Network, " IEEE International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom), Oct 2010, pp. 106-110

[14] Nirosha Joshitha J, R. Medona Selin, "Image Fusion using PCA in Multifeature Based Palmprint Recognition, " International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Vol. 2, Issue : 2, May 2012, pp. 226-230

[15] Qiang Guo and Ya Chen, "Image Resolution Enhancement Using PCA Based Post Filtering, " Research Journal of Applied Sciences, Engineering and Technology, Maxwell Scientific Organization ,Vol.5, No. 2, Jan 2013, pp. 602-605

[16] Khan, M.A., Khan, A., Mahmood, T., Abbas, M. and Muhammad, N.," Fingerprint image enhancement using Principal Component Analysis (PCA) filters, " IEEE International Conference on Information and Emerging Technologies (ICIET), 978-1-4244-8003-6/10/June 2010, pp.1-6

[17] M. Asunción Vicente, Cesar Fernández, Oscar Reinoso, and Luis Payá, "3D Object Recognition from Appearance: PCA Versus ICA Approaches, " Image Analysis and Recognition Lecture Notes in Computer Science , Springer Publication, Volume 3211, 2004, pp 547-555

[18] L. C. Paul1, A. A. Suman, N. Sultan, "Methodological Analysis of Principal Component Analysis (PCA) Method," International Journal of Computational Engineering & Management, IJCEM, Vol. 16 Issue 2, March 2013, ISSN (Online): 2230-7893,pp. 32-38

[19] I. S.Bajwa,and S.I.Hyder,"PCA based Image Classification of Single-layered Cloud Types," IEEE -2005 International Conference on Emerging Technologies,September 17-18, Islamabad,pp. 365-369

[20] Smith, Lindsay I. "A tutorial on principal components analysis." Cornell University, USA 51 (2002): 52.

[21] Nene S. A., Nayar S. K. and Murase H.: Columbia Object Image Library (COIL-100), Technical Report CUCS-006-96 (1996)

AUTHORS PROFILE

Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engineering. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970 He has worked as Faculty of Electrical Engineering and then HOD Computer Science and Engg. at IIT Bombay. For 13 years he was working as a professor and head in the Department of Computer Engg. at Thadomal Shahani Engineering. College, Mumbai. Now he is Senior Professor at MPSTME, SVKM's NMIMS University. He has guided 17 Ph.Ds, more than 100 M.E./M.Tech and several B.E./ B.Tech projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 450 papers in National /International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and Life Member of ISTE Recently twelve students working under his guidance have received best paper awards and six research scholars have beenconferred Ph. D. Degree by NMIMS University. Currently 7 research scholars are pursuing Ph.D. program under his guidance.

Tanuja K. Sarode has Received Bsc. (Mathematics) from Mumbai University in 1996, Bsc.Tech.(Computer Technology) from Mumbai University in 1999, M.E. (Computer Engineering) from Mumbai University in 2004, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 10 years of experience in teaching. Currently working as Associate Professor in Dept. of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. She is life member of IETE, ISTE, member of International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT), Singapore. Her areas of interest are Image Processing, Signal Processing and Computer Graphics. She has more than 100 papers in National /International Conferences/journal to her credit.

Jagruti K. Save has received B.E. (Computer Engg.) from Mumbai University in 1996, M.E. (Computer Engineering) from Mumbai University in 2004, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 10 years of experience in teaching. Currently working as Associate Professor in Dept. of Computer Engineering at Fr. Conceicao Rodrigues College of Engg., Bandra, Mumbai. Her areas of interest are Image Processing, Neural Networks, Fuzzy systems, Data base management and Computer Vision. She has 8 papers in National /International Conferences/journal to her credit.

Mining Frequent Itemsets from Online Data Streams: Comparative Study

HebaTallah Mohamed Nabil
Faculty of Computers and Information Fayoum University
Fayoum, Egypt

Ahmed Sharaf Eldin
Faculty of Computers and Information Helwan University
Cairo, Egypt

Mohamed Abd El-Fattah Belal
Faculty of Computers and Information Helwan University
Cairo, Egypt

***Abstract*—Online mining of data streams poses many new challenges more than mining static databases. In addition to the one-scan nature, the unbounded memory requirement, the high data arrival rate of data streams and the combinatorial explosion of itemsets exacerbate the mining task. The high complexity of the frequent itemsets mining problem hinders the application of the stream mining techniques. In this review, we present a comparative study among almost all, as we are acquainted, the algorithms for mining frequent itemsets from online data streams. All those techniques immolate with the accuracy of the results due to the relatively limited storage, leading, at all times, to approximated results.**

***Keywords*—Data mining; frequent itemsets; data stream; sliding window model; landmark model; fading model.**

I. INTRODUCTION

Recently, the data generation rates in some data sources become faster than ever before. Examples include network traffic analysis, Web click stream mining, network intrusion detection, sensor networks, web logs, and on-line transaction analysis. This rapid generation of continuous streams of information has challenged our storage, computation and communication capabilities in computing systems. Systems, models and techniques have been proposed and developed over the past few years to address those challenges [Gaber M. et al., 2005].

In [Babcock B. et al., 2002] and [Lin C.H. et al., 2005], the Data Stream Model is characterized by that, some or all of the input data that are to be operated on are not available for random access from disk or memory, but rather arrive as one or more *continuous data streams*.

Data streams differ from the conventional stored relation model in several ways: *1) Continuity:* Data continuously arrive at a high rate. *2) Expiration:* Data can be read only once. *3) Infinity:* The total amount of data is unbounded. These characteristics lead respectively to the following challenges [Zhu Y. and Shasha D. 2002] in mining data streams: 1) Limited memory space. 2) Each item in a stream could be examined only once. 3) The mining result should be generated as fast as possible.

The infinite nature of these data sources is a serious obstacle to the use of most of the traditional methods since available computing resources are limited. One of the first effects is the need to process data as they arrive. The amount of previously happened events is usually overwhelming, so they can be either dropped after processing or archived separately in secondary storage. In the first case access to past data is obviously impossible whereas in the second case the cost for data retrieval is likely to be acceptable only for some "ad hoc" queries, especially when several scan of past data are needed to obtain just one single result [Silvestri C., 2006]. In the process of mining frequent itemset, traditional methods for static data usually read the database more than once. However due to the consideration of performance and storage constraints, on-line data stream mining algorithms are restricted to make only one pass over the data. Thus, traditional methods cannot be directly applied to data stream mining [Pauray S. and Tsai M., 2009].

II. BACKGROUND

According to [Li H. F. et al, 2006], data streams are further classified into: *1) offline data streams:* which characterized by discontinuity or regular bulk arrivals [Manku G. and Motwani R., 2002], such as a bulk addition of new transactions as in a data warehouse system, and *2) online data streams:* which characterized by real-time updated data that come one by one in time, such as a continuously generated transaction as in a network monitoring system.

A *transaction data stream* is a sequence of incoming transactions and an excerpt of the stream is called a *window*. A window,*W*, can be (1) either *time-based* or *count-based*, and (2) either a *landmark window* or a *sliding window*. *W* is time-based if *W* consists of a sequence of fixed-length time units, where a variable number of transactions may arrive within each time unit. *W* is count-based if *W* is composed of a sequence of batches, where each batch consists of an equal number of transactions. *W* is a landmark window if $W = [T1, T2, \ldots, T\tau]$; *W* is a sliding window if $W = [T\tau-w+1, \ldots, T\tau]$, where each Ti is a time unit or a batch, $T1$ and $T\tau$ are the *oldest* and the *current* time unit or batch, and *w* is the number of time units or batches in the sliding window, depending on whether *W* is time-based or count-based. Note that a count-based window can also be captured by a time-based window by assuming that a uniform number of transactions arrive within each time unit.

An itemset *X* is a Frequent Itemset (FI) in *W*, if $\sup(X) \geq \sigma$, where σ ($0 \leq \sigma \leq 1$) is a user-specified *minimum support threshold*. In the process of mining data streams, it is necessary to keep not only the FIs, but also the infrequent itemsets that are promising to be frequent later, since an infrequent itemset may become frequent later in the stream. Therefore, many of the existing approximate mining algorithms used a *relaxed* minimum support threshold (also

called a *user-specified error parameter*), ϵ, where $0 \leq \epsilon \leq \sigma \leq 1$, to obtain an extra set of itemsets that are potential to become frequent later.

There are many algorithms for mining frequent itemsets from data streams; according to [Pauray S. and Tsai M., 2009], all those algorithms are fallen into one of the following data stream mining models: *1) Landmark model, 2) Fading model and 3) Sliding window model.*

III. Approaches Of Mining Fis From Data Streams

A. Landmark model

Which considers all the data from a specified point of time (usually the time the system starts), to the current time. All the data considered are treated equally. In this model, knowledge discovery is performed based on the values between a specific timestamp called landmark and the present. See Figure 1.

In [Cormode. G., 2007], an algorithm called Lossy Counting is presented. It produces an approximate set of FIs over the entire history of a stream. The stream is divided into a sequence of buckets and each bucket consists of $B = \lceil 1/\epsilon \rceil$ transactions. It processes a batch of transactions arriving on the stream at a time, where each batch contains β buckets of transactions. The idea of maximum possible error is used to maintain all the possible frequent itemsets. Although the output is approximate, the error is guaranteed not to exceed a user-specified threshold.

According to [Cormode. G., 2007], this method attempts to use the available space as fully as possible. As, for each new transaction, it generates all the subsets, and stores them in a compact trie-based structure. When the space is full, it uses a pruning algorithm based on frequent items algorithms to delete the least frequent itemsets, and track the error in the estimated counts of each item.

In [Yu J. X. et al, 2004], an algorithm called FDPM is derived from the Chernoff bound, to approximate a set of FIs over a landmark window. Suppose that there is a sequence of N observations and consider the first n ($n << N$) observations as independent coin flips such that Pr(head) = p and Pr(tail) = $1 - p$. Let r be the number of heads. Then, the expectation of r is np. The Chernoff bound states, for any $\gamma > 0$:

$$\Pr(|r - np| \geq np\gamma) \leq 2e^{\wedge}((-np\gamma^2)/2) \quad (1)$$

After applying some substitutions and derivations:

$$\Pr\{ p - \epsilon \leq \bar{r} \leq p + \epsilon \} \geq (1-\delta) \quad (2)$$

The underlying idea of the FDPM algorithm is explained as follows. First, a memory bound, $n_0 \approx (2 + 2 \ln(2/\delta))/\sigma$, is also derived. Given a probability parameter, δ, and an integer, k. The batch size, B, is given as $k \cdot n_0$. Then, for each batch of B transactions, FDPM employs an existing non-streaming FI mining algorithm to compute all itemsets whose support in the current batch is no less than $(\sigma - \epsilon_B)$, where $\epsilon_B = \sqrt{(2\sigma \ln(2/\delta))/B}$. The set of itemsets computed are then merged with the set of itemsets obtained so far for the stream. If the total number of itemsets kept for the stream is larger than $c \cdot n_0$, where c is an empirically determined float number, then all itemsets whose support is less than $(\sigma - \epsilon_N)$ are pruned, where N is the number of transactions received so far in the stream and $\epsilon_N = \sqrt{(2\sigma \ln(2/\delta))/N}$. Finally, FDPM outputs those itemsets whose frequency is no less than σ N.

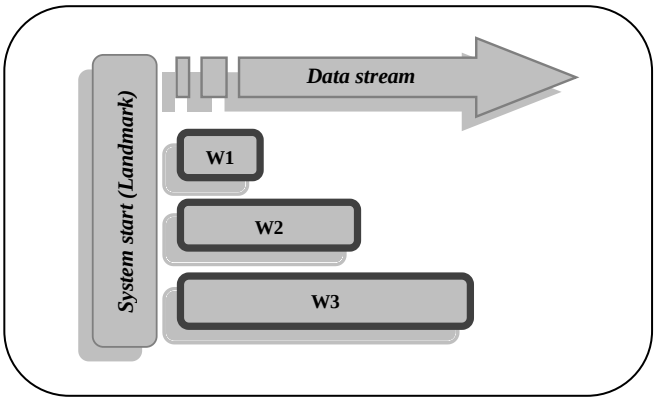


Fig.1. The landmark model

It solves the problem of huge number of sub-FIs problem, by first using a constant lowered minimum support threshold to compute a set of potential FIs and then using a gradually increasing lowered minimum support threshold to control the total number of sub-FIs kept in memory, but with a drawback of producing false negatives.

In [Jin R. and Agrawal. G., 2005], an algorithm called StreamMining is proposed. It is built on the idea of [Karp R.M. et al, 2003] to determine frequent items (or 1-itemsets). In [Karp R.M. et al, 2003] a two pass algorithm was presented for this purpose, which requires only $(1/\theta)$ memory, where θ is the desired support level. Their first pass computes a superset of frequent items, and the second pass eliminates any false positives. StreamMining algorithm addressed three major challenges in applying their idea for frequent itemset mining in a streaming environment. First, it developed a method for finding frequent k-itemsets, while still keeping the memory requirements limited. Second, it developed a way to have a bound on the superset computed after the first pass. Third, it developed a data structure and a number of other implementation optimizations to support efficient execution. This data structure called *TreeHash*, which implements a prefix tree using a hash table. It has the compactness of a prefix tree and allows easy deletions like a hash table. It also uses a relaxed minimum support threshold ϵ, like almost all the mining algorithms for data streams, so the memory requirements increase proportional to $1/\epsilon$. So, this algorithm should had to compute k-itemsets approximately after the first pass, without requiring any out-of-core or large summary structure, and ensure a provable bound on the accuracy of the results after the first pass on the dataset; because in streaming environments, second pass on the dataset is usually not feasible. Therefore, it is important that the set K computed above does not contain many false positives. It was different with [53][2002] in the space requirements. As, for finding frequent items, it takes $O(1/\theta)$ space, while [Manku G., and Motwani R., 2002] requires $O((1/\theta) \log (\theta N))$ space. As [Manku G., and Motwani R., 2002] requires an out-of-core data structure, while it used an in-core data structure. It also has deterministic bounds on the accuracy. One exception is

datasets with the average length of an itemset is quite large. In such case, some additional knowledge of maximal frequent itemsets helps efficiency of our algorithms.

In [Liu X. et al, 2005], an algorithm called FP-DS is presented. It uses a Frequent Pattern structure similar to the FP-DS tree in [Han J. et al, 2000]. The user can obtain current frequent itemsets online continuously without pattern-delay. Compared with the existing related algorithms, the FP-DS algorithm is especially suitable for the mining of long frequent items. It is unnecessary to enumerate every subset on transactions, nor produce a lot of frequent candidate items. The FP-DS tree stores the potential frequent itemsets. It does not need to store all subsets of itemsets independently. It reduces the storage capacity of itemsets and moreover, the itemsets are put in the order of the descending sequence of support of global 1-itemset. The more frequently the items appear the closer to the root of the tree. Such a compression tree has a higher compression ratio.

In the Landmark model, all FIs are outputted, although they are approximated; in other words, the data stream from system start to the existing point is scanned for mining (considering historical data, not only recent data). The support count is computed from the entire data set between the landmark and the current time. But, it isn't aware of time (time unconscious) and therefore it cannot distinguish between new data and old ones. In other words, it losses the time information the itemsets mined.

B. Fading model

That is called the ***Fading*** model in [Chang J. and Lee W., 2004], the ***Damped*** model according to [Zhu Y. and Shasha D., 2002] or ***Time-titled*** model according to [Chen, Y. et al, 2002] and [Pauray S. and Tsai M., 2009]; which is considered a variation of the *landmark* model. It also considers data from the start of streams up to the current moment, but the time period is divided into multiple time slots or assigned different weights to transaction such that new ones have higher weights than old ones. In other words time slots in recent time period are assigned at a fine granularity, while those in ancient time period are assigned at a coarse granularity. In this model, recent sliding windows are more important than previous ones. See Figure 2.

In [Chang J.H. and Lee WS., 2003], an algorithm called estDec is proposed. It uses a decay rate, d ($0 < d < 1$), to diminish the effect of old transactions on the mining result. As a new transaction comes in, the frequency of an old itemset is discounted by a factor of d. Thus, the set of FIs returned is called recent FIs. estDec algorithm adopts the mechanism in [Hidber C., 1999] to estimate the frequency of the itemsets. For example, let the decay rate and the support count of itemset X be d and v, respectively. As a new transaction containing X arrives, the new support count of X is equal to v×d+1. Obviously, when d equals 1, the time-fading model becomes the landmark model. Assume that the stream has received τ transactions, $\langle Y1, Y2, \ldots, Y\tau \rangle$. The decayed total number of transactions, Nτ, and the decayed frequency of an itemset, freqτ (X), are defined as follows:

$$N_\tau = d^{\tau-1} + d^{\tau-2} + \cdots + d^1 + 1 = \frac{1-d^\tau}{1-d} \qquad (3)$$

$$freq_\tau(X) = d^{\tau-1} \times w_1(X) + d^{\tau-2} \times w_2(X) + \cdots + d^1 \times w_{\tau-1}(X) + 1 \times w_\tau(X) \quad (4)$$

$$\text{where } w_i(X) = \begin{cases} 1 \text{ if } X \subseteq Y_i, \\ 0 \text{ otherwise} \end{cases}$$

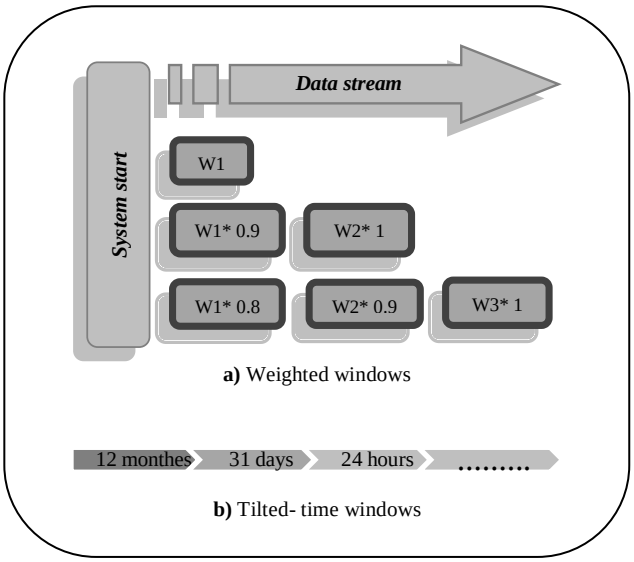


Fig.2. a) Fading model, b) Tilted time window

In [Giannella, J. et al, 2003], an algorithm called FP-Streaming is proposed. It proposed an FP-stream structure [Han J. et al, 2000] based algorithm, to mine frequent itemsets at multiple time granularities by a novel titled-time windows technique of [Chen, Y. et al, 2002]. Frequent patterns are maintained under a tilted-time window framework in order to answer time-sensitive queries. The frequency of an itemset is kept at a finer granularity for more recent time frames and at a coarser granularity for older time frames. For example, we may keep the frequency of an FI in the last hour, the last 2 h, the last 4 h, and so on.

The count of each itemset is asymmetrically distributed into multiple time slots such that the recent time period is assigned more time slots than the past. It is suitable for people to mine the recent data at a fine granularity while mining the long-term data at a coarse granularity. It computes a set of sub-FIs at the relaxed minimum support threshold, ϵ, over each batch of incoming transactions by using the FI mining algorithm, FP-growth [Han J. et al, 2000]. Two parameters, the minimum support count σ and the maximum support error ε where $\sigma \geq \varepsilon$, are used to classify all the itemsets into three categories: 1) Frequent: Support count is greater than and equal to σ. 2) Sub-frequent: Support count falls in $[\varepsilon, \sigma]$. 3) Infrequent: Support count is smaller than ε. Next, only frequent and sub-frequent itemsets are stored and organized in the FP-stream.

[Cohen E. and Strauss M., 2003] and [Chang J. and Lee W., 2004] have also provided variations of decay functions, like in [56][2003], under the time-fading model.

The fading model was proposed to overcome the limitation of time unconscious in the landmark model. It diminishes the effect of the old and obsolete information of a data stream on the mining result. In other words, it considers the effect of old

transactions in some way. It can assign different weights to transactions such that new ones have higher weights than old ones, these weights are decreasing as time passes by.

C. Sliding window model

In [Chang J.H and Lee WS., 2003], an algorithm called estWin is presented. At which the itemsets generated are maintained in a prefix tree structure, D. An itemset, X, in D has the following three fields: $\overline{freq}(X)$, err(X) and tid(X), where $\overline{freq}(X)$ is assigned as the frequency of X in the current window since X is inserted into D, err(X) is assigned as an upper bound for the frequency of X in the current window before X is inserted into D, and tid(X) is the ID of the transaction being processed, for X is inserted into D. For each incoming transaction Y with an ID tid_τ , estWin increments the computed frequency of each subset of Y in D. We prune an itemset X and all X's supersets if (1) $tid(X) \leq tid_1$ and $\overline{freq}(X) < \lceil \epsilon N \rceil$, or (2) $tid(X) > tid_1$ and $\overline{freq}(X) < \lceil \epsilon(N - (tid(X) - tid_1)) \rceil$. For each expiring transaction of the sliding window, those itemsets in D that are subsets of the transaction are traversed. For each itemset, X, being visited, if $tid(X) \leq tid_1$, $\overline{freq}(X)$ is decreased by 1; otherwise, no change is made since the itemset is inserted by a transaction that comes later than the expiring transaction. Then, pruning is performed on X as described before. Finally, for each itemset, X, in D, estWin outputs X as an FI if (1) $tid(X) \leq tid_1$ and $\overline{freq}(X) \geq \sigma.N$, or (2) $tid(X) > tid_1$ and $(\overline{freq}(X) + err(X)) \geq \sigma N$.

In [Chang JH. and Lee WS., 2004], an algorithm called estWin Lossy-counting-based is also presented as a similar method to [Chang J.H and Lee WS., 2003] based on the estimation mechanism of [Manku G. and Motwani R., 2002].

In [Lin C.H. et al, 2005], the first *time-sensitive sliding-window* was proposed, which regards a fixed time period as the basic unit for mining. At which the transaction data stream $TDS = T_1, T_2, \ldots, T_N$ is a continuous sequence of transactions, where N is the transaction identifier of latest incoming transaction T_N.

A transaction $T = (TU_{id}, T_{id}, \text{itemset})$, where TU_{id} is the identifier of the time unit, and T_{id} is the identifier of the transaction. A time-sensitive sliding window (TimeSW) in the transaction data stream is a window that slides forward for every time unit (TU). Each time unit TU_i consists of a variable number, $|TU_i|$, of transactions, and $|TU_i|$ is also called the size of the time unit. Hence, the current time-sensitive sliding window with w time units is $TimeSW_{N\_w+1} = [TU_{N\_w+1}, TU_{N\_w+2}, \ldots, TU_N]$, where N_{N_w+1} is the id of time unit of current TimeSW, and N is the TU_{id} of latest time unit TU_N. The window at each slide has a fixed number, w, of time units. The value $w = |TU_{N\_w+1}| + |TU_{N\_w+2}| + \ldots\ldots + |TU_N|$ is called the size of the time-sensitive sliding window and denoted as |TimeSW|.

It doesn't use a relaxed minimum support threshold, like almost all data stream mining algorithms; instead it maintained a data structure named the *discounting table* (DT) to retain the frequent itemsets with their support counts in the individual basic TUs of the current window.

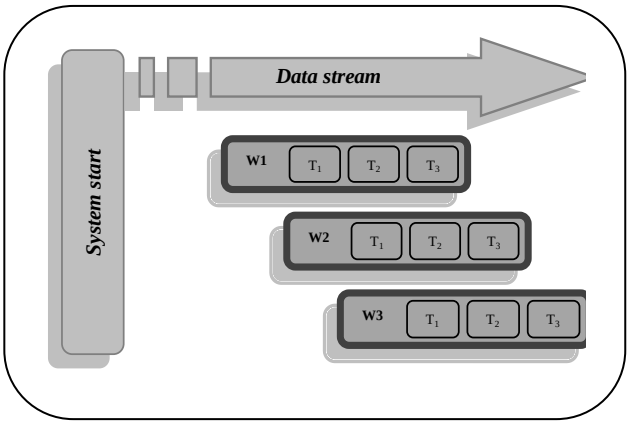


Fig.3. The sliding window model

Moreover, a data structure named the *Potentially Frequent-itemset Pool* (PFP) is used to keep the frequent itemsets in W_i and the frequent ones in TU_i. it includes the itemsets that are frequent in TU_i but not frequent in W_{i-1} in PFP because they are possibly frequent in W_i. Only the summary information derived from W_{i-1} is provided for mining frequent itemsets in W_i. It provides two alternatives to determine the frequent itemsets for output, having trade-off between accuracy and efficiency: *1) Recall-oriented and 2) Precision-oriented.*

In [Li H. F. et al, 2006] and [Li H. F. and Lee S. Y., 2009], an algorithm called MFI-TransSW is presented. It used an effective bit-sequence representation of items to reduce the time and memory needed to slide the windows. For each item X in the current SW, a bit-sequence with w bits, denoted as Bit(X), is constructed. If an item X is in the i-th transaction of current SW, the i-th bit of Bit(X) is set to be 1; otherwise, it is set to be 0. It consists of three phases: *1) window initialization phase:* when the number of transactions generated so far in a transaction data stream is less than or equal to a user-predefined sliding window size w, each item in the new incoming transaction is transformed into its bit-sequence representation; *2) window sliding phase:* after the sliding window TransSW becomes full, a new incoming transaction is appended to the sliding window, and the oldest transaction is removed from the window. The bitwise left shift operation is used to remove the aged transaction from the set of items in the current sliding window. After sliding the window, an item X in the current transaction-sensitive sliding window is dropped if and only if sup(X)TransSW = 0.; *3) frequent itemsets generation phase:* It is performed only when the up-to-date set of frequent itemsets is requested. a level-wise method is used to generate the set of candidate itemsets CI_k (candidate itemsets with k items) from the pre-known frequent itemsets FI_{k-1} (frequent itemsets with k-1 items) according to the Apriori property [Agrawal R. and Srikant R., 1994]. Then, the bitwise AND operation is used to compute the support (the number of bit 1) of these candidates in order to find the frequent k-itemsets FI_k.

In [Cheng J. et al, 2006], an algorithm called MineSW Algorithm is proposed. It is progressively increasing minimum support function. See preliminaries of time-sensitive sliding window model in the description of [Lin C.H. et al, 2005]. By

contrast with other algorithms which uses an error parameter, ϵ , to control the mining accuracy, which leads to a dilemma. It tackles this problem by considering $\epsilon = r.\sigma$ as a relaxed MST , where r ($0 \leq r \leq 1$) is the *relaxation rate*, to mine the set of FIs over each time unit t in the sliding window, allowing to increase ϵ at the expense of slightly degraded accuracy, but significantly improves the mining efficiency and saves memory usage. When an itemset is retained in the window longer, its minimum support required to approach the minimum support of an FI. Thus, the number of potential FIs to be maintained is greatly reduced.

In [Kun Li. Et al, 2008], an algorithm called BFI Steam is presented. It is a Bounded FIs algorithm, which maintains all accurate frequent itemsets from sliding windows by monitoring the boundary between frequent itemsets and infrequent itemsets; it restricts the update process on a small part of the tree. Mining all frequent itemsets with accurate frequencies is just to traverse the tree. It has no candidate generation and it uses a prefix-tree based structure, called BFI-tree, to maintain all frequent itemsets in the sliding window. The BFI-tree is a prefix-tree based data structure and is derived from CET structure in the Moment algorithm [Chi Y. et al, 2004]. It uses a count-based sliding window with fixed size of N, which always contains recent N transactions. BFI-tree monitors the boundary movements to efficiently maintain the selected part of infrequent itemsets. If a node status changes, either from infrequent to frequent or vice versa, it must come through the boundary and result in boundary movements. Boundary movements may cause recursive updates, which will be restricted in a small sub-tree. It may also cause creating new nodes, which need an additional scan on all transactions in the sliding window to compute their frequencies. In order to return accurate frequent itemsets, all transactions in the sliding window must be maintained in a highly compact structure. However, the boundary is stable at most time, which means the update cost is very small. BFI-tree uses the Apriori property [Agrawal R. and Srikant R., 1994] in construction and updates to prune infrequent nodes. So, 1) all the children of an infrequent node should be pruned and consequently all infrequent nodes are leaves in BFI-tree, 2) some children of a frequent node may be infrequent and should be pruned.

In [Ren J. D. and Li K., 2008], an algorithm called MRFI-SW is presented to mine Recent FIs with SW, which uses a transaction-sensitive sliding window. The transactions are denoted with a special representation, which can denote the number and the order of items that are contained in the transactions. Using Apriori property, frequent itemsets can be mined through processing the representation's information. In this representation, for each itemset X which is contained in transactions in current sliding window is constructed as a sequence. The length of the sequence is w, where w is the number of transactions in transaction-sensitive sliding window. Each entry is the form of (bit, order), denoted as R(x). If item X is in the i-th transaction in current sliding window, the i-th entry of R(X)_bit is set to be 1 and the order of items in a transaction can get from R(X)_order, otherwise the R(X) is set to be 0 (R(X)_bit=R(X)_order=0). The process of creating this sequence of items for transaction in current sliding window is called bit-order transform. For example, let <T1, (acd)>, <T2, (bce)>, <T3, (abce)>, and <T4, (be)>, and the size of sliding window be 3, hence, SW1=[T1, T2, T3] and SW2=[T2, T3, T4]. In SW1, because item **a** appears in T1 and T3 and is the first item in both transactions, so, R(**a**) is <(1, 1), 0, (1, 1)>. Similarly, R(**c**)=<(1, 2), (1, 2), (1, 3)>, R(**d**)=<(1, 3), 0, 0>, R(**b**)=<0, (1, 1), (1, 2)>, and R(**e**)=<0, (1, 3), (1, 4)>.

In [Naganth E.R. and Dhanaseelan F. R., 2008] a *graph structure* is proposed to capture the contents of transactions in a sliding window. The graph structure captures the contents of transactions in each batch of streaming data. Transaction items are arranged according to some canonical order, which can be specified by the user prior to the graph construction or the mining process. Whenever a new batch of transactions flows in, it appends to this list at each node its frequency count in the current batch. In other words, the last entry of the list at node X shows the frequency count of X in the current batch. When the next batch of transactions comes in, the list is shifted forward. The last entry shifts and becomes the second – last entry; this leaves room for the newest batch. It uses a pointer to indicate the last update at each node. If the pointer points to the previous entry in the list of frequency counts at a node X, then this indicates that X has just been visited at the update of the last batch. On the other hand, if the pointer points to a much earlier entry in the list at a node Y, then this indicates that Y has not been visited since then and that the frequent counts of Y for the entries in between should be 0s. Since this graph structure is constructed independent of the minimum support threshold, every transaction in the current window is captured. Once such a tree is constructed, we can mine frequent itemsets from it in a fashion similar to FP-growth [Han J. et al, 2000] using the user supplied σ. Since items are consistently arranged according to some canonical order, one can guarantee the inclusion of all frequent items using just upward traversals, leading to exact mining results. There is also no worry about possible omission ordoubly counting of items during the mining process. So, mining is delayed until it is needed to avoid lots of unnecessary computation.

In [Pauray S. and Tsai M., 2009], an algorithm called WSW Algorithm is proposed. It is a weighted SW framework is proposed. See preliminaries of time-sensitive sliding window model in the description of [Lin C.H. et al, 2005]. In the traditional sliding window model, only one window is considered for mining at each time point. WSW proposed a flexible framework for continuous query processing in data streams. The time interval for periodical queries is defined to be the size of a window. In traditional sliding window model, the size of a window is usually defined to be a given number of transactions, say T. Although only the latest T transactions are considered, the time to cover these T transactions may be long. If we ignore the significance of data at different time intervals, the effectiveness of the mining result may decrease. So the WSW model has the following two new features: (1) window size is defined by time, not the number of transactions, the purpose is to avoid the case where intervals that cover T transactions at different time points may vary dramatically. (2) number of windows considered for mining is specified by the user. Moreover, the user can assign different

weights to different windows according to the importance of data in each section. For example, the data near to the current moment may be more influential in the mining, and could be given a higher weight. The weights of windows affect the determination of frequent items. Even if the total support count of an item is large, if its support count in the window with a high weight is very low, it may not become a frequent item. Thus the consideration of weights for windows is reasonable and significant. So, the mining result would be closer to user's requirements.

By data characteristics, an improved algorithm, WSW-Imp, is explored to further reduce the time of deciding whether a candidate itemset is frequent or not. Experimental results show that the performance of WSW-Imp significantly outperforms that of WSW over weighted sliding windows.

In [Li H. F. and Lee S. Y., 2009], an algorithm called MFI-Time-SW is proposed. It works in a SW environment with a time-sensitive SW. See preliminaries of time-sensitive sliding window model in the description of [Lin C.H. et al, 2005]. The major differences between MFI-TransSW []Li H. F. et al, 2006] and [Li H. F. and Lee S. Y., 2009]and MFI-TimeSW are the following: 1) *Unit of data processing:* each time unit contains variable number of transactions. 2) *Bit-sequence transformation of a time unit:* For each item X in the current time-sensitive stream sliding window $TimeSW_{N\_w+1}$, a bitsequence with $|TimeSW_{N\_w+1}|$ bits, denoted as $Bit\ (X)_{N-w+1}^{TimeSW}$, is constructed. Similarly, if the item X is in the i-th transaction of $TimeSW_{N\_w+1}$, the ith bit of $Bit\ (X)_{N-w+1}^{TimeSW}$, is set to be 1. Otherwise, it is set to be 0. 3) *Number of sliding transactions:* In the window sliding phase of MFI-TimeSW algorithm, after the oldest time unit TU_{N_w+1} is removed from the current sliding window, a new time unit TU_{N+1} is appended to the window. If the aged time unit TU_{N_w+1} contains d transactions, MFI-TimeSW performs d times of bitwise left shift operation on the current sliding window. After that, MFI-TimeSW uses that same pruning method Item-Prune to improve the memory usage in mining process. 4) *Dynamic frequent threshold of itemsets:* the value of frequent threshold is s * |TimeSW| is a dynamic value, where $|TimeSW| = |TU_{N\_w+1}| + |TU_{N\_w+2}| + \ldots. + |TU_N|$.

The sliding window model captures recent pattern changes and trends, by utilizing only the latest transactions for mining. As in certain applications, users can only be interested in the data recently arriving within a fixed time period. Sliding window model ignores the important fact of that Itemsets are changing their frequencies according to certain time intervals. In other words the sliding window model ignores the history of the itemsets' frequencies.

IV. Storage, Time And Accuracy Tradeoff

TABLE 1 presents the important comparative parameters to distinguish among the state of the art algorithms which are as the following: *1) The mining model, 2) The accuracy of the results, 3) The processing strategy, 4) The units of processing and 5) The data to be processed.*

According to Table 1, the first and last parameters are closely related, at which the landmark and fading models process the whole data stream from the system start. On the other hand, the sliding window model processes only a recent portion of the data stream.

All the algorithms of the landmark and fading models using the count based windows, except the algorithm of [Giannella, J. et al, 2003], which uses the time based window. In the Sliding window model, all the algorithms using the count base window, except the algorithms of [Lin C.H. et al, 2005], [Cheng J. et al, 2006], [Pauray S. and Tsai M., 2009] and part of [Li H. F. and Lee S. Y., 2009], which considered a fixed time period as the basic unit for processing (i.e time based window).

The count based windows are easy for programmers to deal with and not easy for people to specify. By contrast, in time based windows, it is natural for people to specify a time period as the basic unit, but it is more difficult to deal with windows with variable sizes in terms of bytes. In the performance perspective, no difference between using the time based or the count based windows [Arasu A. and Widom J., 2003].

The algorithms of [Manku G., and Motwani R., 2002], [Jin R. and Agrawal. G., 2005], [Chang J.H. and Lee WS., 2003], [Chang J.H and Lee WS., 2003], [Chang JH. and Lee WS., 2004], [Kun Li. et al, 2008], [Ren J. D. and Li K., 2008] and [Pauray S. and Tsai M., 2009] are using a tuple processing mechanism, at which processing is done transaction by transaction; the rest of the algorithms are

TABLE I. COMPARATIVE SUMMARY AMONG THE SATE-OF-THE-ART ALGORITHMS

Data Stream Mining Model	Data Stream Mining Algorithm	Approximate/Exact Results	Tuple/ Batch Processing	Time/ Count base (transactions)	All / Recent Transactions
Landmark Model	Lossy Counting [Manku G.,Motwani R., 2002]	Approximate/ False Positives (FP)	Tuple	Count	All
	FDPM [Yu J. X. et al,2004]	Approx./False Negatives (FN)	Batch	Count	All
	StreamMining [Jin R. ,Agrawal. G.,2005]	Approx. / FP	Tuple	Count	All
	FP-DS [Liu X. et al,2005]	Approx. / FP	Batch	Count	All
Fading Model	estDec [Chang J.H.,Lee WS.,2003]	Approx./ FP	Tuple	Count	All (recent is more important)
	FP-streaming [Giannella, J. et al,2003]	Approx. / FP	Batch	Time	All (recent is more important)
Sliding Window Model (SW)	estWin [Chang J.H,Lee WS.,2003]	Approx. / FP	Tuple	Count	Recent
	estWin_Lossy_Counting based [Chang JH.,Lee WS.,2004]	Approx. / FP	Tuple	Count	Recent
	[Lin C.H. et al,2005]	Approx. / FP or FN	Batch	Time	Recent
	MFI-TransSW [Li H. F.et al,2006],[Li H. F.,Lee S. Y.,2009]	Approx. / FP	Batch	Count	Recent
	Mine-SW [Cheng J et al,2006]	Approx. / FN	Batch	Time	Recent
	BFI Steam [Kun Li. Et al,2008]	Exact	Tuple	Count	Recent
	MRFI-SW [Ren J. D.,Li K.,2008]	Exact	Tuple	Count	Recent
	Graph Structure _SW [Naganth E.R.,Dhanaseelan F. R.,2008]	Exact	Batch	Count	Recent
	WSW [Pauray S.,Tsai M.,2009]	Approx. / FP	Tuple	Time	Recent
	MFI-Time-SW [Li H. F.,Lee S. Y.,2009]	Approx. / FP	Batch	Time	Recent

Using batch processing. Processing each transaction against the entire stream in most cases is less efficient than processing a batch of transactions against the entire stream. In general, batch processing is more suitable for high speed data streams [Cheng J. et al, 2008]. See Figure 4.

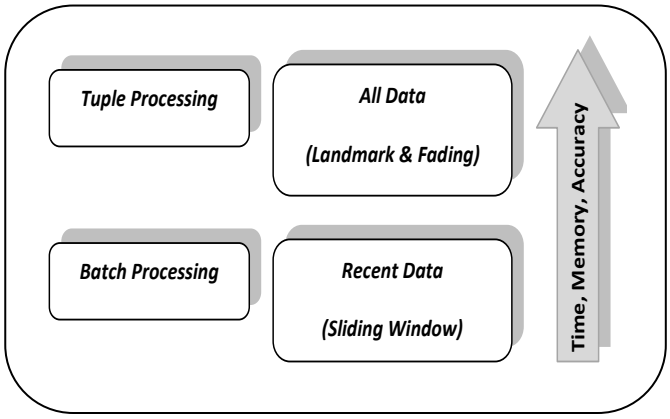


Fig.4. Time, memory and accuracy tradeoff

[Chang J.H. and Lee WS., 2003] favors recent itemsets by diminishing exponentially the effect of old transactions; but estimating itemsets' frequencies from their subsets leading to a propagated error. And [Chang J.H. and Lee WS., 2003] partitions the window according to a logarithmic scale with the recent frequency of an itemset recorded at a finer granularity using a tilted-time window, which answers time-sensitive queries at the expense of more than one record kept for each itemset leading to a very large structure over time which in turn degrade the mining process. All the variations of the landmark and fading models have the limitation of providing approximate answers for long-term data and adjust their storage requirement based on the available space.

All the algorithms are producing approximate results, except [Kun Li. et al, 2008], [Ren J. D. and Li K., 2008] and [Naganth E.R. and Dhanaseelan F. R., 2008], which produce exact results but for only a recent portion of the data stream (i.e they are under the sliding window model). All the approximate algorithms adopt false-positive approaches, except [Yu J. X. et al, 2004] and an option in [Lin C.H. et al, 2005], which adopts a false-negative approach. The false-positive approach uses a relaxed minimum support threshold, ϵ, to reduce the number of false-positives, so obtaining a more accurate result. However, to obtain a more accurate result, a smaller value of ϵ is to be set, leading to a larger set of sub-FIs to be maintained, consuming large amount of memory. The false-negative approach also uses a relaxed minimum support threshold, ϵ; however, its' use lowers the probability of discarding an infrequent itemset that may become frequent later. The error bound in the computed support and the possible false mining results of *most* of the false-positive approaches are implied by the following equation according the derivation of [Cheng J. et al, 2008]:

Support error bound = err $(X)/N_\tau$,

False results = $\{X \mid \text{freq}(X) < \sigma N \leq (\overline{\text{freq}}(X) + \text{err}(X))\}$

Almost all the online mining algorithms do the mining process during the entering of the stream, by constructing, filling and extracting FIs in parallel to the data stream entering; except [Li H. F. et al, 2006] and [Li H. F. and Lee S. Y., 2009] and [Naganth E.R. and Dhanaseelan F. R., 2008], which actually extract FIs from the filled data structure (i.e done during the data stream entering) only when it is requested by the user; which is a more effective strategy.

Any data stream mining algorithm aims to enforce correct and accurate results, minimized consumption of memory, fully utilized CPU and minimized time for processing. Correctness here refers to mining only true frequent itemsets (i.e no false positives, nor false negatives), and accuracy refers to mining exact or approximated frequencies for the itemsets (leading to true or nearest to true frequent itemsets).

Exact mining requires keeping track of all itemsets in the window and their actual frequency, because any infrequent itemset may become frequent later in the stream. However, the number of all itemsets is $O(2^{|I|})$ (that is, given a set of items I, the possible number of itemsets can be up to $2^{|I|} - 1$) making exact mining computationally difficult, in terms of both CPU and memory.

Also, using a relaxed minimum support threshold,ϵ, to control the quality of the approximation of the mining result leads to a dilemma. The smaller the value of ϵ, the more accurate is the approximation but the greater is the number of sub-FIs generated, which requires both more memory space and more CPU processing power. However, if ϵ approaches σ, more false-positive answers will be included in the result, since all sub-FIs whose computed frequency is at least $(\sigma-\epsilon)N \approx 0$ are outputted while the computed frequency of the sub-FIs can be less than their actual frequency by as much as $\sigma.N$.

Almost all the approximation algorithms produce false-positive results, at which, the set of sub-FIs kept is often too large in order to obtain a highly accurate answer. Thus, throughput is decreased and memory consumption is increased due to the processing of a large number of sub-FIs. But for the approximation algorithms which produce false negative results [Yu J. X. et al, 2004], it infringes the correctness. Accordingly, there is a direct proportion between the accuracy and memory consumption; more accurate results needs more memory usage leading to increased processing time.

V. CONCLUSION

According to the continues high flow of data streams and the relatively limited resources of CPU and storage, the process of mining frequent itemsets is chained to mining approximated frequencies, even it has guarantee on error bounds. The accuracy of the resulting FIs directly proportional with the memory usage; high accuracy needs high memory usage.

Therefore the algorithms of mining FIs from data streams which are using the Landmark model, Fading model (at which all the data stream is mined) or the Sliding Window model (at which only recent data is mined); leading to approximated results certainly.

VI. FUTURE WORK

We can focus in future to get exact mining results with regard to the available storage. Traditional data mining algorithms do not produce any results that show the change of the results over time. Dynamics of data streams using changes in the knowledge structures generated would benefit many temporal-based analysis applications.

REFERENCES

[1] Agrawal R. and Srikant R. (1994). Fast Algorithms for Mining Association Rules. In Proceedings of the 20th International Conference on Very Large Data Bases, pp. 487-499.

[2] Arasu A. and Widom J. (2003). Resource Sharing in Continuous Sliding-Window Aggregates. ACM SIGMOD Record, Volume 32 Issue 2.

[3] Babcock B., Babu S., Datar M., Motwani R., and Widom J. (June 2002) Models and issues in data stream systems. In PODS'02, Madison, WI.

[4] Cormode. G. (2007). Fundamentals of Analyzing and Mining Data Streams. WORKSHOP ON DATA STREAM ANALYSIS – San Leucio, Italy - March, 15-16.

[5] Chang J. and Lee W.(2004). Decaying Obsolete Information in Finding Recent Frequent Itemsets over Data Stream. IEICE Transaction on Information and Systems, Vol. E87-D, No. 6.

[6] Chen, Y., Dong, G.; Han, J.; Wah, B.W.; and Wang, J.(2002). Multidimensional regression analysis of time-series data streams. In Proc. 2002 Int. Conf. Very Large Data Bases (VLDB'02), 323.334.

[7] Chang J.H. and Lee WS. (2003). Finding recent frequent itemsets adaptively over online data streams. In:Getoor L, Senator T, Domingos P, Faloutsos C (eds) Proceedings of the Ninth ACM SIGKDD international conference on knowledge discovery and data mining, Washington, DC, pp 487–492.

[8] Cohen E. and Strauss M. (2003). Maintaining Time Decaying Stream Aggregates. Proc. of ACM PODS Symp.

[9] Chang J.H and Lee WS. (2003). estWin: adaptively monitoring the recent change of frequent itemsets over online data streams. In: Proceedings of the 2003 ACM CIKM international conference on information and knowledge management, New Orleans, Louisiana, USA, November 2003, pp 536–539.

[10] Chang JH. and Lee WS. (2004). A sliding window method for finding recently frequent itemsets over online data streams. J Inf Sci Eng 20(4):753–762.

[11] Cheng J, Ke Y. and Ng W. (2006). Maintaining frequent itemsets over high-speed data streams. In: Ng WK, Kitsuregawa M, Li J, Chang K (eds) Proceedings of the 10th Pacific-asia Conference on knowledge discovery and data mining, Singapore, April pp 462–467.

[12] Chi Y, Wang H, Yu P, Muntz R (2004). Moment: maintaining closed frequent itemsets over a stream sliding window. In: Proceedings of the 4th IEEE international conference on data mining, Brighton, UK, , pp 59–66.

[13] Cheng J, Ke Y and Ng W. (2008). A survey on algorithms for mining frequent itemsets over data streams. Knowl Inf Syst, 16:1–27 DOI 10.1007/s10115-007-0092-4.

[14] Gaber M., Zaslavsky A. and Krishnaswamy Sh. (June 2005). Mining Data Streams: A Review. SIGMOD Record, Vol. 34, No. 2.

[15] Giannella, J. Han, J. Pei, X. Yan, and P.S. Yu. (2003). Mining Frequent Patterns in Data Streams at Multiple Time Granularities. in H. Kargupta, A. Joshi, K. Sivakumar, and Y. Yesha (eds.), Next Generation Data Mining, AAAI/MIT.

[16] Han J., Pei J. and Yin Y. (2000). Mining frequent patterns without candidate generation. In Proceedings of ACM SIGMOD, 1–12.

[17] Hidber C. (1999). Online association rule mining. In: Delis A, Faloutsos C, Ghandeharizadeh S (eds) Proceedings of the ACM SIGMOD international conference on management of data, Philadelphia, Pennsylvania, pp 145–156.

[18] Jin R. and Agrawal. G. (2005). An algorithm for in-core frequent itemset mining on streaming data. To appear in ICDM'05.

[19] [19] Karp R.M., Shenker S., and Papadimitriou C.H. (2003). A simple algorithm for finding frequent elements in streams and bags. ACM Transactions on Database Systems (TODS), 28(1):51–55.

[20] Kun Li., Yong-yan Wang, Manzoor Ellahi, Hong-an Wang. (2008). Mining Recent Frequent Itemsets in Data Streams. IEEE 2008.

[21] Li H. F., Chin-Chuan Ho, Man-Kwan Shan, and Suh-Yin Lee (2006). Efficient Maintenance and Mining of Frequent Itemsets over Online Data Streams with a Sliding Window. In IEEE SMC.

[22] Li H. F. and Lee S. Y. (2009). Mining frequent itemsets over data streams using efficient window sliding techniques. Sceince Direct, Expert Systems with Applications 36 (2009) 1466–1477.

[23] Lin C.H., Chiu D.Y., Wu Y.H. and Chen A.L.P. (2005). Mining Frequent Itemsets from Data Streams with a Time-Sensitive Sliding Window. In Proceedings of 2005 SIAM International Conference on Data Mining.

[24] Liu X., Xu H., Dong Y., Wang Y. and Qian J.(2005). Dynamically Mining Frequent Patterns over Online Data Streams. Springer-Verlag Berlin Heidelberg. Y. Pan et al. (Eds.): ISPA 2005, LNCS 3758, pp. 645 – 654.

[25] Manku G., and Motwani R(2002). Approximate frequency counts over data streams. In Proc. 2002 Int. Conf. Very Large Data Bases (VLDB'02), 346.357.

[26] Naganth E.R. and Dhanaseelan F. R.(2008). Efficient Graph Structure for the Mining of Frequent Itemsets from data Streams. IJCSES International Journal of Computer Sciences and Engineering Systems, Vol.1, No.4.

[27] Pauray S. and Tsai M. (2009). Mining frequent itemsets in data streams using the weighted sliding window model. Elsevier, Expert Systems with Applications.

[28] Ren J. D. and Li K. (2008). Online data Stream mining of recent frequent Itemsets Based On Sliding Window Model. Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, 12-15.

[29] Silvestri C. (2006). Distributed and Stream Data Mining Algorithms for Frequent Pattern Discovery. Universit`a Ca' Foscari di Venezia, Dipartimento di Informatica, Dottorato di Ricerca in Informatica. Ph.D. Thesis: TD-2006-4.

[30] Yu J. X., Chong Z., Lu H., and Zhou A.(2004). False Positive or False Negative: Mining Frequent Itemsets from High Speed Transactional Data Streams. In Proceedings of the 30th International Conference on Very Large Data Bases, pp. 204-215.Zhu Y. and Shasha D. (2002). StatStream: Statistical Monitoring of Thousands of Data Streams in Real Time. In Proceedings of the 28th International Conference on Very Large Data Bases, pp. 358-369.

SVM Classification of Urban High-Resolution Imagery Using Composite Kernels and Contour Information

Aissam Bekkari, Mostafa El yassa, Soufiane Idbraim,
Driss Mammass and Azeddine Elhassouny
IRF – SIC laboratory,
Faculty of sciences
Agadir, Morocco

Danielle Ducrot
Cesbio laboratory,
Toulouse, France

***Abstract*—The classification of remote sensing images has done great forward taking into account the image's availability with different resolutions, as well as an abundance of very efficient classification algorithms. A number of works have shown promising results by the fusion of spatial and spectral information using Support Vector Machines (SVM) which are a group of supervised classification algorithms that have been recently used in the remote sensing field, however the addition of contour information to both spectral and spatial information still less explored.**

For this purpose we propose a methodology exploiting the properties of Mercer's kernels to construct a family of composite kernels that easily combine multi-spectral features and Haralick texture features as data source. The composite kernel that gives the best results will be used to introduce contour information in the classification process.

The proposed approach was tested on common scenes of urban imagery. The three different kernels tested allow a significant improvement of the classification performances and a flexibility to balance between the spatial and spectral information in the classifier. The experimental results indicate a global accuracy value of 93.52%, the addition of contour information, described by the Fourier descriptors, Hough transform and Zernike moments, allows increasing the obtained global accuracy by 1.61% which is very promising.

Keywords—SVM; Contour information; Composite Kernels; Haralick feature; Satellite image; Spectral and spatial information; GLCM; Fourier descriptors; Hough transform ; Zernike moments.

I. INTRODUCTION

The rich spectral information available in remotely sensed images allows the possibility to distinguish between spectrally similar materials [1]. However, supervised classification of satellite images (which assumes prior knowledge in the form of class labels for some spectral signatures) is a very challenging task due to the generally unfavourable ratio between the (large) number of spectral bands and the (limited) number of training samples available *a priori*, which results in the Hughes phenomenon.

The application of originally developed methods for the classification of lower dimensional data sets (such as multispectral images) generally provides poor results when applied to satellite images, particularly in the case of small training sets [2].

The classification of such images is similar to that of other image types, it follows the same principle, and it is a method of analysis of data that aims to separate the image into several classes in order to gather the data in homogeneous subsets, which show common characteristics. It aims to assign to each pixel of the image a label which represents a theme in the real study area (e.g. vegetation, water, built, etc) [3].

Several classification algorithms have been developed since the first satellite image was acquired in 1972 [4-6]. Among the most popular and widely used is the maximum likelihood classifier [7]. It is a parametric approach that assumes the class signature in normal distribution. Although this assumption is generally valid, it is invalid for classes consisting of several subclasses or classes that have different spectral features [8]. To overcome this problem, some non-parametric classification techniques such as artificial neural networks, decision trees and Support Vector Machines (SVM) have been recently introduced.

SVM is a group of advanced machine learning algorithms that have seen increased use in land cover studies [9, 10]. One of the theoretical advantages of the SVM over other algorithms (decision trees and neural networks) is that it is designed to search for an optimal solution to a classification problem whereas decision trees and neural networks are designed to find a solution, which may or may not be optimal.

This theoretical advantage has been demonstrated in a number studies where SVM generally produced more accurate results than decision trees and neural networks [7, 11]. SVMs have been used recently to map urban areas at different scales with different remotely sensed data. High or medium spatial resolution images (e.g., IKONOS, QUICKBIRD, LANDSAT (TM)/ (ETM+), SPOT) have been widely employed on urban land use classification for individual cities, building extraction, road extraction and other man-made objects extraction [12, 13].

On the other hand, the consideration of the spatial aspect in classification remains very important. For this case, Haralick has described methods for measuring texture in gray-scale images, and statistics for quantifying those textures. It is the hypothesis of this research that Haralick's Texture Features and

statistics as defined for gray-scale images can be modified to incorporate spectral information, and that these Spectral Texture Features will provide useful information about the image. It is shown that texture features can be used to classify general classes of materials, and that Spectral Texture Features in particular provide a clearer classification of land cover types than purely spectral methods alone.

As well as the contour information is concerned, survey approaches were developed for pattern recognition. The three most used methods are the Fourier descriptors (FD) classically used to shape recognition and template matching; the Hough transform (HT) which has become a standard tool in computer vision field. It allows the detection of lines, circles or ellipses in a traditional way; it can also be extended to the description of more complex object cases. The third method is the Zernike Moments (ZM) used to extract invariant shapes descriptors to some general linear transformations for the images classification.

This work presents the way adopted in our experiments to incorporate contour information into classification process. We have found that the use of this contour information with both spectral and spatial information allows increasing the accuracy obtained using only spectral and spatial information.

The proposed method consists into combining spatial, spectral and contour information to obtain a better classification. So we have started with the extraction of spatial information (Haralick texture features) [14], and the contour information (Fourier descriptors, Hough transform and Zernike moments). Then, we have used these descriptors combined with spectral values as entry of the SVM classifier. We have exploited the properties of Mercer's kernels to construct a family of composite kernels that easily combine spatial and spectral information. The three different composite kernels tested demonstrate enhanced classification accuracy compared to approaches that take into account only the spectral information, and a flexibility to balance between the spatial and spectral information in the classifier.

An extended version of the composite kernel that gives the best results will be used to introduce contour information in the classification process. The result obtained is compared with the same composite kernel using only spectral and spatial information to measure the contribution of contour information in the classification's overall accuracies.

This paper is organized as follows. In the second section, we will discuss the extraction of spectral, spatial and contour information especially the Grey-Level Co-occurrence Matrix (GLCM), Haralick texture features, Hough transform and Zernike moments used in experimentations. In section 3, we will give outlines on the used classifier: Support Vector Machines (SVM). Section 4 will describe the three different composite kernels used in experimentations. In section 5, the experimentations and results would be presented as well as the numerical evaluation. Finally, conclusions and future research lines would be provided in section 6.

II. Extraction of Information

A. Spectral Information

The most used classification methods for the remote-sensing data consider especially the spectral dimension. First attempts to analyze urban area used existing methodologies and techniques developed for land remote sensing, based on signal modeling. Each pixel-vector is regarded as a vector of attributes which will be directly employed as an entry of the classifier.

The traditional approach for classifying remote-sensing data may be summed up as: from the original data set, a feature reduction/selection step is performed according to the classes in consideration, and then classification is carried out using these extracted features. In our work, the step of a feature reduction/selection can be skipped considering that we have used multispectral images such as IKONOS, QUICKBIRD.

According to Fauvel [15] this allows a good classification based on the spectral signature of each area. However, this does not take in account the spatial information represented by the various structures in the image.

B. Spatial Information

Information in a remote sensed image can be deduced based on their textures. A human analyst is able to distinguish man-made features from natural features in an image based on the 'regularity' of the data. Straight lines and regular repetitions of features hint at man-made objects. This spatial information is useful in distinguishing the different field in the remote sensed image.

Many approaches were developed for texture analysis. According to the processing algorithms, three major categories, namely, structural, spectral, and statistical methods are common ways for texture analysis.

Many researches have been conducted on the use of Gabor filter banks [16] and co-occurrence matrices [17] for the spatial/spectral classification of multispectral data. Other researches have been conducted with mathematical morphology concepts. Palmason et al. [18] and Fauvel et al. [15] suggest an extraction method of morphological profiles. These profiles are computed on the first principal components of hyperspectral images. Plaza [19] uses also mathematical morphology to extract the endmembers of a hyperspectral image. Some other works [20] combine spectral classification with spatial segmentation based on watershed method.

In [21-23], the authors compare different spatial features in unsupervised classification of hyperspectral images; the studies used Gabor filter banks, co-occurrence matrices, Texture spectra and morphological profiles. The results obtained showed that the haralick features extracted from the co-occurrence matrices give the best performance in classification accuracies.

The GLCM method, proposed by Haralick [24, 25], involves two steps to generate spatial features.

First, the spatial information of a digital image is extracted by a co-occurrence matrix calculated on a pixel neighbourhood (pixel window) defined by a moving window of a given size. Such a matrix contains frequencies of any combination of gray levels occurring between pixel pairs separated by a specific distance and angular relationship within the window. The second step is to compute statistics from the gray level co-occurrence matrix to describe the spatial information according to the relative position of the matrix elements.

Even small, a co-occurrence matrix represents a substantial amount of data that is not easy to handle. This is why Haralick uses these matrices to develop a number of spatial indices that are easier to interpret.

Haralick assumed that the texture information is contained in the co-occurrence matrix, and texture features are calculated from it. A large number of textural features have been proposed starting with the original fourteen features described by Haralick et al [25], however only some of these features are in wide use. Wezska et al [26] used four of Haralick features. Conners and Harlow [27] use five features. Peng Gong and al. [28] show that these features are much correlated with each other. The authors have used the FORTRAN package TEXTRAN for the spatial feature extraction. The analysis was made on the near-infrared band (0.79-0.89/µm) with a quantization level of 16.

The interpixel distance was kept constant to 1, and the four main orientations were averaged. The window sizes used were 3x3, 5x5, and 7x7 pixels. Preliminary tests made with larger window sizes did not give satisfactory results. Ten texture features were first generated on a 5x5 pixel window. The three less correlated features were then selected to complete the study. The Fig.1. Represents the Correlation Matrix of the 16 Spatial Features.

	ASM	CON	COR	COV	DIS	ETP	HO	IDM	MA	VAR	AVE	ENT	STD	BW	DD	CS
ASM	*	-0.51	-0.34	-0.33	-0.74	-0.90	0.87	0.91	0.96	-0.44	-0.15	-0.83	-0.59	0.00	-0.02	-0.04
CON		*	0.21	0.72	0.92	0.73	-0.75	-0.64	-0.53	0.90	0.17	0.59	0.87	0.02	-0.03	0.06
COR			*	0.60	0.21	0.41	-0.19	-0.17	-0.28	0.47	0.15	0.40	0.55	0.01	-0.37	-0.15
COV				*	0.61	0.54	-0.46	-0.37	-0.33	0.95	0.16	0.45	0.87	0.00	-0.24	-0.06
DIS					*	0.91	-0.94	-0.87	-0.77	0.80	0.19	0.77	0.86	0.01	0.02	0.07
ETP						*	-0.96	-0.94	-0.90	0.67	0.20	0.88	0.81	0.00	-0.03	0.03
HO							*	0.98	0.90	-0.62	-0.18	-0.83	-0.75	0.00	-0.06	-0.08
IDM								*	0.94	-0.52	-0.16	-0.83	-0.65	0.00	-0.08	-0.08
MA									*	-0.45	-0.15	-0.82	-0.58	0 .00	-0.04	-0.05
VAR										*	0.18	0.54	0.93	0.01	-0.16	-0.01
AVE											*	0.20	0.21	0.02	-0.11	-0.05
ENT												*	0.72	0.01	-0.03	0.02
STD													*	0.01	-0.18	-0.02
BW														*	0.21	0.12
DD															*	0.41
CS																*

Fig. 1. The Correlation Matrix of the 16 Spatial Features.

In this work, we have chosen the five features used by Conners and Harlow, which are some of the most commonly used spatial measures and the three less correlated (Fig.1.); we have found that these five sufficed to give good results in classification [29].

These five features are: homogeneity (E), contrast (C), correlation (Cor), entropy (H) and local homogeneity (LH), and co-occurrence matrices are calculated for four directions: 0°, 45°, 90°and 135° degrees.

Let us recall their definitions considering a co-occurrence matrix *M*:

$$E = \sum_i \sum_j (M(i,j))^2 \quad (1)$$

$$C = \sum_{k=0}^{m-1} k^2 \sum_{|i-j|=k} M(i,j) \quad (2)$$

Where *m* is the dimension of the co-occurrence matrix *M*.

$$Cor = \frac{1}{\sigma_i \sigma_j} \sum_i \sum_j (i-\mu_i)(j-\mu_j) M(i,j) \quad (3)$$

Where μ_i and σ_i are the horizontal mean and the variance, μ_j and σ_j are the vertical statistics.

$$H = \sum_i \sum_j M(i,j) \log(M(i,j)) \quad (4)$$

$$LH = \sum_i \sum_j \frac{M(i,j)}{1+(i-j)^2} \quad (5)$$

Each texture measure can create a new band that can be incorporated with spectral features for classification purposes.

C. Contour Information

Fourier descriptors are classical methods to shape recognition and they have grown into a general method to

encode various shape signatures. Previous experiments have used Fourier descriptors to smooth out fine details of a shape. Also, using the portion of Fourier descriptors to reconstruct an image that smooths out the sharp edges and fine details found in the original shape. Filtering an image with Fourier descriptors provides a simple technique of contour smoothing.

Fourier description of an edge is also used for template matching. Since all the Fourier descriptors except the first one do not depend on the location of the edge within the plane, this provides a convenient method of classifying objects using template matching of an object's contour. A set of Fourier descriptors is computed for a known object. Ignoring the first component of the descriptors, the other Fourier descriptors are compared against the Fourier descriptors of unknown objects. The known object, whose Fourier descriptors are the most similar to the unknown object's Fourier descriptors, is the object the unknown object is classified to. They can also be used for calculation of region area, location of centroid, and computation of second-order moments.

On the other hand, in the detection of specific elements, There are algorithms that, so as to identify these basic forms, attempt to follow the contours to finally bind criteria more or less complex to trace the desired shape. Another approach to this problem is to try to accumulate evidences on a particular form existence, such as a line, a circle or an ellipse. It is this approach that has been adopted in the Hough transform. In recent decades, it has become a standard tool in computer vision field. It allows the detection of lines, circles or ellipses in a traditional way. It can also be extended to description of more complex objects cases.

Moreover, the methods of images representation by moments are among the first to have been applied in pattern recognition. The main motivation is to extract invariant shapes descriptors to some general linear transformations for the images classification. Since the initial work of H. Ming-Kuel [30] in 1962 on invariants derived from the image geometric moments, several approaches have been proposed. Most of these defined moments are expressed as radial moments of the image's circular harmonic functions. The image's Zernike Moments (MZ) were introduced by M.R. Teague [31]. He proposed to use complex polynomials of Zernike orthogonal within the unit circle. These methods are distinguished by the used radial kernel form, which is more or less appropriate to the extraction of invariant descriptors to flat similarities.

In the following we will introduce briefly the Fourier descriptors, the Hough transform and Zernike moments used in experiment to describe the contour information.

1) Fourier Descriptors

The Fourier Descriptors (FD) have been frequently used as features for image processing, remote sensing, shape recognition and classification.

The use of FDs for pattern recognition tasks started in the early sixties by Cosgriff [32] and Fritzsche [33]. A set of orthogonal FDs represent each pattern for the purpose of classification. The recognition system was independent of the character size and orientation. Furthermore, FDs were used as features for recognition systems for both handwritten characters [34] and numerals [35]. Granlund [34] used a small number of lower-order descriptors for the classification system. Those descriptors were insensitive to translation, rotation and dilation. Because of the small computational power available at that time the system could not be examined to give the suitable number of descriptors. The classification system was applied to a small number of characters. Nevertheless the system was able to produce a very good recognition rate of 98%.

Zhan and Roskies [35] started computing the FDs by translating the contour of handwritten numeral into a change of angle curve. A large number of Fourier coefficients are produced. For each coefficient two kinds of FDs are computed, the harmonic amplitude and the phase angle. Those pair of FDs is invariant under translation, rotation and change of size of the original handwritten numeral. All the FDs pairs fully describe the original signature.

Fourier descriptors were also used to describe open curves in an online character recognition system [36]. The one pixel thick strokes were taken online using a tablet. Then twenty FDs were computed and used for classification.

In remote sensing field the FDs were applied to the feature of the regions on the data for geometrical matching of the remote sensing images. It makes possible to monitor natural and artificial changes in land cover precisely.

The discreet Fourier function for a periodic polynomial function $f(t)$ is,

$$F(k) = \frac{1}{N}\sum_{t=0}^{N-1} f(t)\exp(-j2\pi kt/N), \tag{6}$$

Where N is the total number of points along the $f(t)$

$$F(k) = \frac{1}{N}\sum_{t=0}^{N-1} f(t)\{\cos(2\pi kt/N) + j\sin(2\pi kt/N)\} \tag{7}$$

$$with \quad k = 0,1,2,\ldots,N-1$$

The Fourier coefficients are

$$a_k = \frac{2}{N}\sum_{t=1}^{N-1} f(t)\cos\left(\frac{2\pi kt}{N}\right)$$
$$b_k = \frac{2}{N}\sum_{t=1}^{N-1} f(t)\sin\left(\frac{2\pi kt}{N}\right) \tag{8}$$

As said before the commonly used FDs are the harmonic amplitude A_k and the phase angle θ_k of the Fourier coefficients a_k and b_k above,

$$A_k = \sqrt{a_k^2 + b_k^2}$$
$$\theta_k = \tan^{-1}\frac{b_k}{a_k} \tag{9}$$

The harmonic amplitude A_k is a pure shape feature and doesn't contain information about the position or the orientation of the numeral but on the other hand the phase angle θ_k has those two features.

The fixed length feature vector would be

$$\{A_k, \theta_k\}_{k=1}^{M}, \quad (10)$$

Where M is a fixed integer number.

The original polynomial could be reconstructed from its FDs by using the following equation

$$f(t) \approx A_o + \sum_{k=1}^{N-1} a_k \cos\frac{2\pi kt}{N} + b_k \sin\frac{2\pi kt}{N} \quad (11)$$

$$with\ \ t = 0,1,2,....N-1$$

Where, A_o is the DC component of the function, and has no effect on the shape description.

2) Hough Transform

The Hough Transform (HT) is considered as a very powerful tool for detecting predefined features (i.e. lines, ellipses…) in images and has been used for more than three decades in the areas of image processing, pattern recognition and computer vision. Its main advantages are its insensitivity to noise and its capability to extract lines even in areas with pixel absence (pixel gaps) [37-39].

The Hough technique is particularly useful for computing a global description of a feature(s) (where the number of solution classes need not to be known *a priori*), given (possibly noisy) local measurements. The motivating idea behind the Hough technique for line detection is that each input measurement (*e.g.* coordinate point) indicates its contribution to a globally consistent solution (*e.g.* the physical line which gave rise to that image point).

As a simple example, consider the common problem of fitting a set of line segments to a set of discrete image points (*e.g.* pixel locations output from an edge detector). Fig.2. shows some possible solutions to this problem. Here the lack of *a priori* knowledge about the number of desired line segments (and the ambiguity about what constitutes a line segment) render this problem under-constrained.

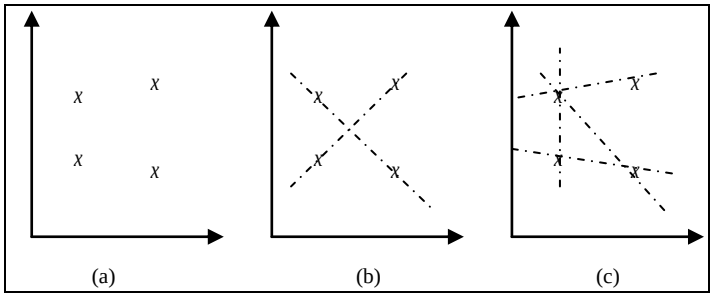


Fig. 2. **(a)** Coordinate points, when (**b)** and (**c)** Possible straight line fittings.

We can analytically describe a line segment in a number of forms. However, a convenient equation for describing a set of lines uses parametric or normal notion as follow:

$$\rho = x\cos\theta + y\sin\theta \quad (12)$$

Where ρ is the length of a normal from the origin to this line and θ is the orientation of ρ with respect to the X-axis. (Fig.3.) For any point (x,y) on this line, ρ and θ are constant.

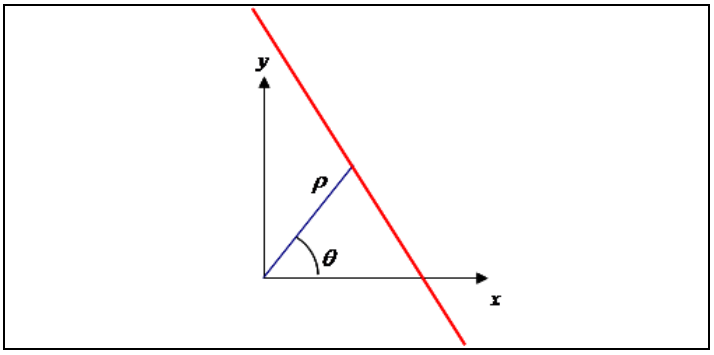


Fig. 3. Parametric description of a straight line (ρ, θ)

In an image analysis context, the coordinates of the point(s) of edge segments (*i.e.* (x_i,y_i)) in the image are known and therefore serve as constants in the parametric line equation, while ρ and θ are the unknown variables we seek. We plot the possible (ρ, θ) values defined by each (x_i,y_i) points in Cartesian image space map to curves (*i.e.* sinusoids) in the polar Hough parameter space.

This *point-to-curve* transformation is the Hough transformation for straight lines. When viewed in Hough parameter space, points which are collinear in the cartesian image space become readily apparent as they yield curves which intersect at a common (ρ, θ) point.

The transform is implemented by quantizing the Hough parameter space into finite intervals or *accumulator cells*. As the algorithm runs, each (x_i,y_i) is transformed into a discretized (ρ, θ) curve and the accumulator cells which lie along this curve are incremented. Resulting peaks in the accumulator array represent strong evidence that a corresponding straight line exists in the image.

3) Zernike Moments

The extraction of features from an image by the method of moments is one of the techniques commonly used. It obviously gives the amount of information which is encoded in the image [40]. A moment is an overall description of the distribution of pixels within an image. Each time a given order gives different information of other times on the image [41, 42]. The central moments of order p, q are given by the following expressions:

$$\mu_{pq} = \sum_x \sum_y \left(x-\bar{x}\right)^p \left(y-\bar{y}\right)^q I(x,y)$$

$$\text{With: } m_{pq} = \sum_X \sum_Y X^p\, Y^q I(x,y) \quad (13)$$

$$\bar{x} = \frac{m_{10}}{m_{00}} \text{ and } \bar{y} = \frac{m_{01}}{m_{00}}$$

Where $I(x, y)$ is the gray level of the pixel x, y. The central moments are given as following [39, 40]:

$$\mu_{pq} = \sum_x \sum_y \left(x-\bar{x}\right)^p \left(y-\bar{y}\right)^q I(x,y) \quad (14)$$

The normalized central moments are given by

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}{}^{\gamma}} \qquad \text{with} \quad \gamma = \frac{p+q}{2} + 1 \tag{15}$$

Hu moments are defined as a set of moment invariants [43], but are not orthogonal. The most interesting moments are orthogonal that can be obtained through the Zernike polynomials. The Zernike moments do not change the orientation, the scale and the translation. They remain robust to noise and to minor variations of the forms [44]. There is no redundant information because their bases are orthogonal. An image is best described by a small set of Zernike moments than any other type of moments such as geometric moments, Legendre, rotational or complex moments [45]. The Zernike moments are build using a set of complex polynomials which form a complete orthogonal set on the unit disk. For an image *f*, the Zernike moments are defined as follows [45]:

$$Z_n = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V^*_{nm}(\rho,\theta) \tag{16}$$

Where *m* and *n* define the order of the moment. Knowing that

$$V_{mn}(\rho,\theta) = R_{nm}(\rho) e^{+jm\theta} \tag{17}$$

Where $R_{nm}(\rho)$ is the radial polynomial Zernike. The latter can be described by:

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} r^{n-2s} \tag{18}$$

Where *n* and *m* are integers (their values are even integers). These moments can be used as a tool for comparing two classes by calculating the distance denoted by *d* between the vectors of Zernike moments of each class. If we are interested in comparing one class to multiple classes, the most similar image corresponds to that which is characterized by a smallest distance *d*.

III. SVM CLASSIFICATION

In this section we will briefly describe the general mathematical formulation of SVMs introduced by Vapnik [46, 47]. Starting from the linearly separable case in which the optimal hyperplanes are introduced. Then, the classification problem is modified to handle non-linearly separable data. At the end of this section, a brief description of multiclass strategies would be given.

A. Linear SVM

For a two-class problem in a *n*-dimensional space R^n, we assume that *l* training samples $x_i \in R^n$, are available with their corresponding labels $yi = \pm 1$, $S = \{(xi, yi) \mid i \in [1, l]\}$.

The SVM method consists of finding the hyperplane that maximizes the margin, i.e., the distance to the closest training data points for both classes [48]. Noting $w \in R^n$ as the normal vector of the hyperplane and $b \in R$ as the bias, the hyperplane H_p is defined as:

$$\langle w, x \rangle + b = 0, \forall x \in H_p \tag{19}$$

Where $\langle w, x \rangle$ is the inner product between *w* and *x*. If $x \notin Hp$ then $f(x) = \langle w, x \rangle + b$ is the distance of *x* to *Hp*. The sign of *f* corresponds to decision function *y = sgn (f(x))*.

Finally, the optimal hyperplane has to maximize the margin: $2/\|w\|$. This is equivalent to minimize $\|w\|/2$ and leads to the following quadratic optimization problem:

$$\min\left[\frac{\|w\|^2}{2}\right] \quad \text{subject to } y_i(\langle w, x_i \rangle + b) \geq 1 \ \ \forall i \in [1, l] \tag{20}$$

For non-linearly separable data, the optimal parameters (w, b) are found by solving:

$$\min\left[\frac{\|w\|^2}{2} + C\sum_{i=1}^{l} \xi_i\right] \quad \text{subject to } y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0 \ \ \forall i \in [1, l] \tag{21}$$

Where the constant *C* control the amount of penalty and ξ_i are slack variables which are introduced to deal with misclassified samples (Fig.4.). This optimization task can be solved through its Lagrangian dual problem:

$$\max_{\alpha} \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \quad \text{subject to } 0 \leq \alpha_i \leq C \quad \forall i \in [1, l] \quad \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{22}$$

Finally:

$$w = \sum_{i=1}^{l} \alpha_i y_i x_i \tag{23}$$

The solution vector is a linear combination of some samples of the training set, whose α_i is non-zero, called Support Vectors. The hyperplane decision function can thus be written as:

$$y_u = \text{sgn}\left(\sum_{i=1}^{l} y_i \alpha_i \langle x_u, x_i \rangle + b\right) \tag{24}$$

Where x_u is an unseen sample.

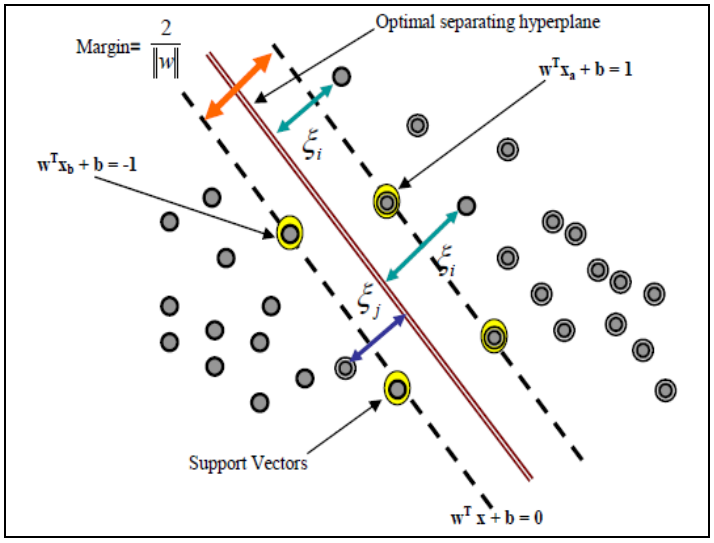


Fig. 4. Classification of a non-linearly separable case by SVMs. There is one non separable feature vector in each class.

B. Non-Linear SVM

Using the Kernel Method, we can generalize SVMs to non-linear decision functions. By this technique, the classification capability is improved. The idea is as following. Via a non-linear mapping Φ, data are mapped onto a higher dimensional space F (Fig.5.):

$$\Phi : R^n \to F \qquad x \alpha \ \Phi(x) \tag{25}$$

The SVM algorithm can now be simply considered with the following training samples: $\Phi(S) = \{(\Phi(x_i), y_i) / i \in [1, l]\}$. It leads to a new version of the hyperplane decision function where the scalar product is now: $\langle \Phi(x_i), \Phi(x_j) \rangle$. Hopefully, for some kernels function k, the extra computational cost is reduced to:

$$\langle \Phi(x_i), \Phi(x_j) \rangle = k(x_i, x_j) \tag{26}$$

The kernel function k should fulfill Mercers' conditions.

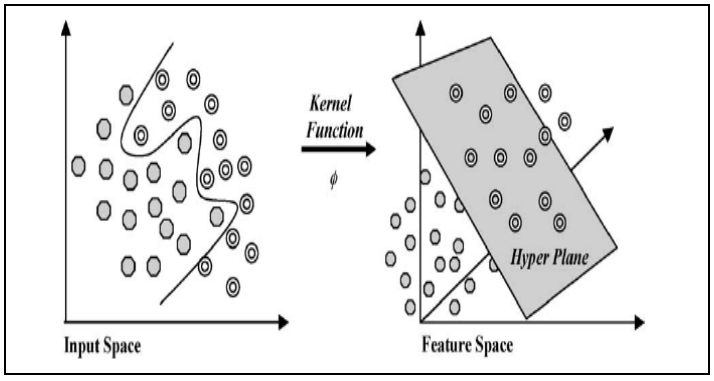


Fig. 5. Mapping the Input Space into a High Dimensional Feature Space with a kernel function

With the use of kernels, it is possible to work implicitly in F while all the computations are done in the input space. The classical kernels used in remote sensing are the polynomial kernel and the Gaussian radial basis function:

$$k_{poly}(x_i, x_j) = [(x_i \cdot x_j) + 1]^p \tag{27}$$

$$k_{gauss}(x_i, x_j) = \exp[-\gamma \|x_i - x_j\|^2] \tag{28}$$

C. Multiclass SVMs

SVMs are designed to solve binary problems where the class labels can only take two values: ±1. For a remote sensing application, several classes are usually of interest. Various approaches have been proposed to address this problem [49]. They usually combine a set of binary classifiers. Two main approaches were originally proposed for a k-classes problem.

- ***One versus the Rest***: k binary classifiers are applied on each class against the others. Each sample is assigned to the class with the maximum output.
- ***Pairwise Classification***: $k(k-1)/2$ binary classifiers are applied on each pair of classes. Each sample is assigned to the class getting the highest number of votes. A vote for a given class is defined as a classifier assigning the pattern to that class.

IV. Composite Kernels

In the following section, we will be dealing with three different kernel approaches that not only allow joining spectral and textural information for multispectral image classification, but also introducing the contour information by using an extended kernel version [50, 51].

A. The Stacked Features Approach

The most commonly adopted approach in multispectral image classification is to exploit the spectral content of a pixel (x_i). However, performance can be improved by including both spectral and spatial information in the classifier. This is usually done by means of the 'stacked' approach, in which feature vectors are built from the concatenation of spectral and spatial features.

Note that if the chosen mapping Φ is a transformation of the concatenation $x_i \equiv \{x_{i\text{-}spect}, x_{i\text{-}spa}\}$, then the corresponding 'stacked' kernel matrix is:

$$k_{\{Spect,Spa\}} \equiv k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \tag{29}$$

Which does not include explicit cross relations between $x_{i\text{-}spa}$ and $x_{i\text{-}spect}$.

Including the contour information is also possible by means of the 'stacked' approach; the feature vectors will be built from the concatenation of spectral, spatial and contour features:

$x_i \equiv \{x_{i\text{-}spect}, x_{i\text{-}spa}, x_{i\text{-}cont}\}$.

The corresponding 'stacked' kernel matrix $k_{\{Spect,Spa,Cont\}}$ remains the same in (29).

B. The Direct Summation Kernel

A simple composite kernel combining spectral and textural information naturally comes from the concatenation of nonlinear transformations of $x_{i\text{-}spa}$ and $x_{i\text{-}spect}$. Let us assume two nonlinear transformations $\varphi_1(.)$ and $\varphi_2(.)$ into Hilbert spaces H_1 and H_2, respectively. Then, the following transformation can be constructed:

$$\Phi(x_i) = \{\varphi_1(x_{i-spect}), \varphi_2(x_{i-spa})\} \tag{30}$$

And the corresponding dot product can be easily computed as follows:

$$\begin{aligned} k(x_i, x_j) &= \langle \Phi(x_i), \Phi(x_j) \rangle \\ &= \langle \{\varphi_1(x_{i-spect}), \varphi_2(x_{i-spa})\}, \{\varphi_1(x_{j-spect}), \varphi_2(x_{j-spa})\} \rangle \\ &= k_{spect}(x_{i-spect}, x_{j-spect}) + k_{spa}(x_{i-spa}, x_{j-spa}) \end{aligned} \tag{31}$$

In the same way, we can exploit the Mercer's properties to generalize this formulation in order to have a summation of multiple kernels:

$$k(x_i, x_j) = \sum_{m=1}^{p} k_m(x_i^m, x_j^m) \tag{32}$$

So to use spectral, spatial and contour information we take the case of $p=3$, then we will have:

$$\begin{aligned} k(x_i, x_j) = k_{spect}(x_{i-spect}, x_{j-spect}) + k_{spa}(x_{i-spa}, x_{j-spa}) + \\ k_{Cont}(x_{i-Cont}, x_{j-Cont}) \end{aligned} \tag{33}$$

C. The Weighted Summation Kernel

By exploiting properties of Mercer's kernels, a composite kernel that balances the spatial and spectral content in (28) can also be created, as follows:

$$k(x_i, x_j) = \mu k_{spect}(x_{i-spect}, x_{j-spect}) + (1-\mu) k_{spa}(x_{i-spa}, x_{j-spa}) \tag{34}$$

Where μ is a positive real-valued free parameter ($0 < \mu < 1$), which is tuned in the training process and constitutes a tradeoff between the spatial and spectral information to classify a given pixel.

This composite kernel allows us to introduce *a priori* knowledge in the classifier by designing specific μ profiles per class, and also allows us to extract some information from the best tuned μ parameter.

A generalization of the weighted summation to multiple kernels is possible by using "Linear combination methods", and we can linearly parameterize the combination function:

$$k_\mu(x_i, x_j) = \sum_{m=1}^{p} \mu_m k_m(x_i^m, x_j^m) \tag{35}$$

Where μ denotes the kernel weights. Different versions of this approach differ in the way they put restrictions on μ: the linear sum $(i.e., \mu \in \Re^p)$, the conic sum $(i.e., \mu \in \Re_+^p)$, or the convex sum $(i.e., \mu \in \Re_+^p \text{ and } \sum_{\mathrm{m}}^{\mathrm{p}} \mu_{\mathrm{m}} = 1)$. As can be seen, the conic sum is a special case of the linear sum and the convex sum is a special case of the conic sum. The conic and convex sums have two advantages over the linear sum in terms of interpretability.

First, when we have positive kernel weights, we can extract the relative importance of the combined kernels by looking at them. Second, when we restrict the kernel weights to be nonnegative, this corresponds to scaling the feature spaces and using the concatenation of them as the combined feature representation:

$$\Phi_\mu(x) = \begin{pmatrix} \sqrt{\mu_1}\Phi_1(x^1) \\ \sqrt{\mu_2}\Phi_2(x^2) \\ . \\ . \\ \sqrt{\mu_p}\Phi_p(x^p) \end{pmatrix} \tag{36}$$

And the dot product in the combined feature space gives the combined kernel:

$$\begin{aligned} \langle \Phi_\mu(x_i), \Phi_\mu(x_j) \rangle &= \begin{pmatrix} \sqrt{\mu_1}\Phi_1(x_i^1) \\ \sqrt{\mu_2}\Phi_2(x_i^2) \\ . \\ . \\ \sqrt{\mu_p}\Phi_p(x_i^p) \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \sqrt{\mu_1}\Phi_1(x_j^1) \\ \sqrt{\mu_2}\Phi_2(x_j^2) \\ . \\ . \\ \sqrt{\mu_p}\Phi_p(x_j^p) \end{pmatrix} \\ &= \sum_{m=1}^{p} \mu_m k_m(x_i^m, x_j^m) \end{aligned} \tag{37}$$

The combination parameters can also be restricted using extra constraints, such as the l_p-norm on the kernel weights or trace restriction on the combined kernel matrix, in addition to their domain definitions. For example, the l_1-norm promotes sparsity on the kernel level, which can be interpreted as feature selection when the kernels use different feature subsets.

So to use spectral, spatial and contour information we take the case of $p=3$, then we will have:

$$\begin{aligned} k(x_i, x_j) &= \mu_1 k_{spect}(x_{i-spect}, x_{j-spect}) + \\ &\quad \mu_2 k_{spa}(x_{i-spa}, x_{j-spa}) + \mu_3 k_{Cont}(x_{i-Cont}, x_{j-Cont}) \\ \text{with} \quad & \mu_m \in \Re_+ \text{ and } \sum_{\mathrm{m}=1}^{3} \mu_{\mathrm{m}} = 1 \end{aligned} \tag{38}$$

D. The Computational Complexity

The computational complexity of a multiple kernel learning (MKL) algorithm mainly depends on its training method (i.e., whether it is one-step or two-step) and the computational complexity of its base learner.

One-step methods using fixed rules and heuristics generally do not spend much time to find the combination function parameters, and the overall complexity is determined by the complexity of the base learner to a large extent. One-step methods that use optimization approaches to learn combination parameters have high computational complexity, due to the fact that they are generally modeled as a semi-definite programming (SDP) problem, a quadratically constrained quadratic programming (QCQP) problem, or a second-order cone programming (SOCP) problem. These problems are much harder to solve than a quadratic programming (QP) problem used in the case of the canonical SVM.

Two-step methods update the combination function parameters and the base learner parameters in an alternating manner. The combination function parameters are generally updated by solving an optimization problem or using a closed-form update rule. Updating the base learner parameters usually requires training a kernel-based learner using the combined kernel. For example, they can be modeled as a semi-infinite linear programming (SILP) problem, which uses a generic linear programming (LP) solver and a canonical SVM solver in the inner loop.

Note that solving the minimization problem in all kinds of composite kernels requires the same number of constraints as in the conventional SVM algorithm, and thus no additional computational efforts are induced in the presented approaches.

V. Experimantal Results

In this section, we are going to evaluate the proposed approach by using two high resolution satellite images with different resolutions representing the scene of urban areas.

A. Data

The first image used in classification is a subset of high resolution QUICKBIRD satellite image, with a high spatial resolution of 2.4 m per pixel. It represents urban scene areas. We dispose of four spectral bands: blue, green, red and near infrared. We can see in Fig.7. (a) a representation of this subset.

The second image is a subset of high resolution IKONOS satellite image. It has also four spectral bands: red, blue, green and near infrared, with a high spatial resolution of 4.1 m per pixel. This subset of the image is represented in Fig.8. (a).

We will have two files containing the extracted features for each image, “TrainFile.dat” and “TestFile.dat” respectively for learning and for classification, and divided on six classes as described in the following Table I.

B. Comparing Composite Kernels

Our experiments are divided on two stages (Fig.6. and Fig.9.). The first one concerns the studies of composite kernels proposed in section 4 using only spectral and spatial information. In the second stage we will use an extended version of the composite kernel that gave the best performance in the first stage, to introduce contour information in addition to spectral and spatial information.

TABLE I. Differents Classes

Class N°	Class name	Train samples	
		Image 1	*Image 2*
1	Asphalt	1 592	1 386
2	Green area	2 252	480
3	Tree	880	196
4	Soil	176	813
5	Building	4 217	920
6	Shadow	1 280	336
Total		10 397	4 131

So as we can see in Fig.6., that represents the first experience, we have developed a two step classification process: the first one is the extraction of the spatial and spectral features, so we compute Grey Level Co-occurrence Matrix (GLCM) to extract Haralick texture features that we have added to spectral information. The second step is the SVM classification; a supervised kernel learning algorithm widely used. We have selected SVMlight with composite kernels, which is an implementation of Support Vector Machines (SVMs) in C language [52].

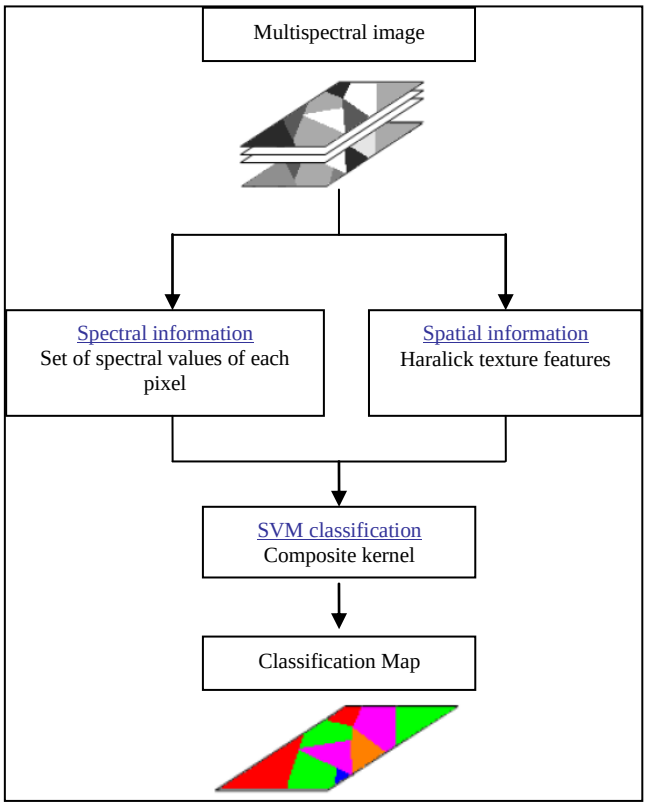


Fig. 6. A representative illustration of the first stage of the proposed workflow

To join spatial and spectral information, we have used three different kernel approaches as presented in section 4; named the stacked features approach in (29), the direct summation kernel in (31) and the weighted summation kernel in (34).

In the case of the weighted summation kernel, μ was varied with a step of 0.1 in the range [0, 1]. For simplicity and for illustrative purposes, μ was the same for all classes in our experiments. The penalization factor in the SVM was tuned in the range $C = \{10^{-1} \ldots 10^{7}\}$.

We have used the Gaussian RBF kernel (28) (with $\sigma = \{10^{-1} \ldots 10^{3}\}$) for the two kernels. k_{spect} uses a spectral information while k_{spa} uses Haralick features.

The classification map presented on (b) in Fig.7. and Fig.8., is obtained when the classification is performed using the stacked features approach (29). When the classification is performed using the direct summation kernel (31), we obtain the corresponding classification map which is presented on (c) in Fig.7. and Fig.8.. A visual analysis of classification maps shows those areas more homogeneous for the maps obtained using the direct summation kernel than those obtained by using the stacked features approach.

The fusion of the spectral and the spatial features using the weighted summation kernel give us the classification map presented on (d) in Fig.7. and Fig.8.. We can see that the classes are more connected and also we have got less misclassified pixels in the result compared to the other approaches.

Table II lists the accuracy estimates and kappa coefficient of the classification results, all models are compared numerically (overall accuracy, kappa coefficient).

Table III and Table IV presents respectively the confusion matrix results for SVM classification using the weighted summation kernel (34) based on spectral and spatial information, for both images used in experiments.

TABLE II. OVERALL ACCURACY (%) AND KAPPA COEFFICIENT OF CLASSIFIED IMAGES

Methods	Image 1		Image 2	
	Overall accuracy	***Kappa coefficient***	***Overall accuracy***	***Kappa coefficient***
SVM using only spectral information	87.56%	0.87	88.79%	0.88
The stacked features approach	94.13%	0.93	92.13%	0.91
The direct summation kernel	94.26%	0.93	92.38 %	0.92
The weighted summation kernel	94.48%	0.93	92.55%	0.92

TABLE III. CONFUSION MATRIX RESULTS (%) FOR SVM CLASSIFICATION USING THE WEIGHTED SUMMATION KERNEL FOR IMAGE 1. GLOBAL ACCURACY = 94.48%

Class name	Asphalt	Green area	Tree	Soil	Building	Shadow
Asphalt	93,66	1,41	1,91	1,01	1,63	0,38
Green area	1,13	94,99	0,00	1,08	1,54	1,26
Tree	0,28	1,07	92,82	2,50	0,82	2,51
Soil	4,84	0,95	0,00	93,87	0,34	0,00
Building	0,01	1,16	2,69	0,47	95,67	0,00
Shadow	0,08	0,42	2,58	1,07	0,00	95,85

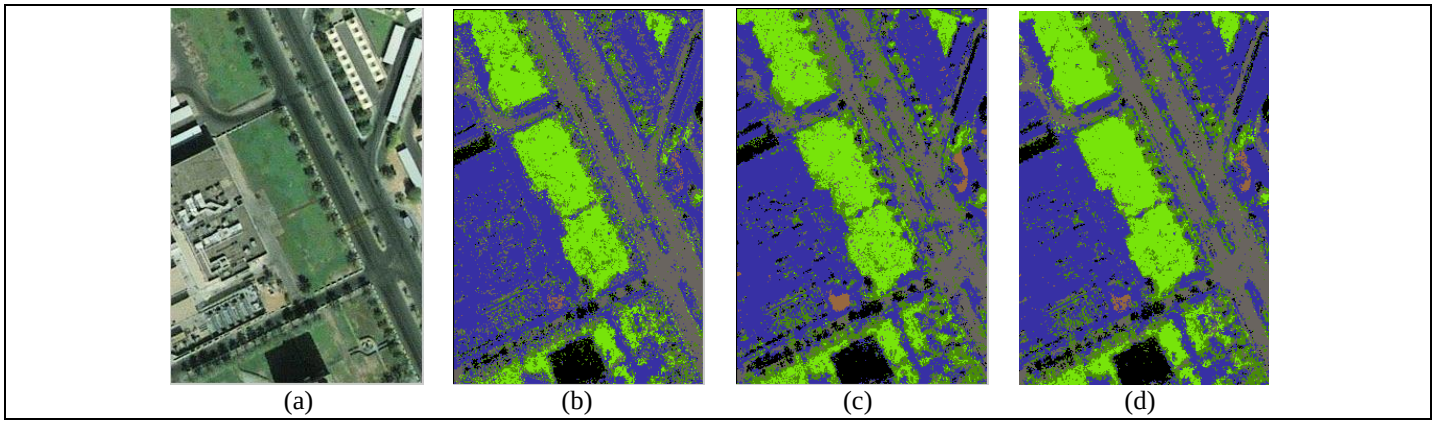


Fig. 7. (a) Original image 1, (b) Classification Map obtained using the stacked features approach, (c) Classification Map obtained using the direct summation kernel , (d) Classification Map obtained using the weighted summation kernel.

TABLE IV. CONFUSION MATRIX RESULTS (%) FOR SVM CLASSIFICATION USING THE WEIGHTED SUMMATION KERNEL FOR IMAGE 2. GLOBAL ACCURACY = 92.55%

Class name	Asphalt	Green area	Tree	Soil	Building	Shadow
Asphalt	89,36	2,04	1,92	1,50	3,32	1,86
Green area	5,13	92,21	0,00	1,03	1,54	0,09
Tree	1,18	1,52	93,15	1,92	0,03	2,20
Soil	1,75	1,13	0,64	93,04	3,44	0,00
Building	1,96	2,78	2,72	0,87	91,67	0,00
Shadow	0,62	0,32	1,57	1,64	0,00	95,85

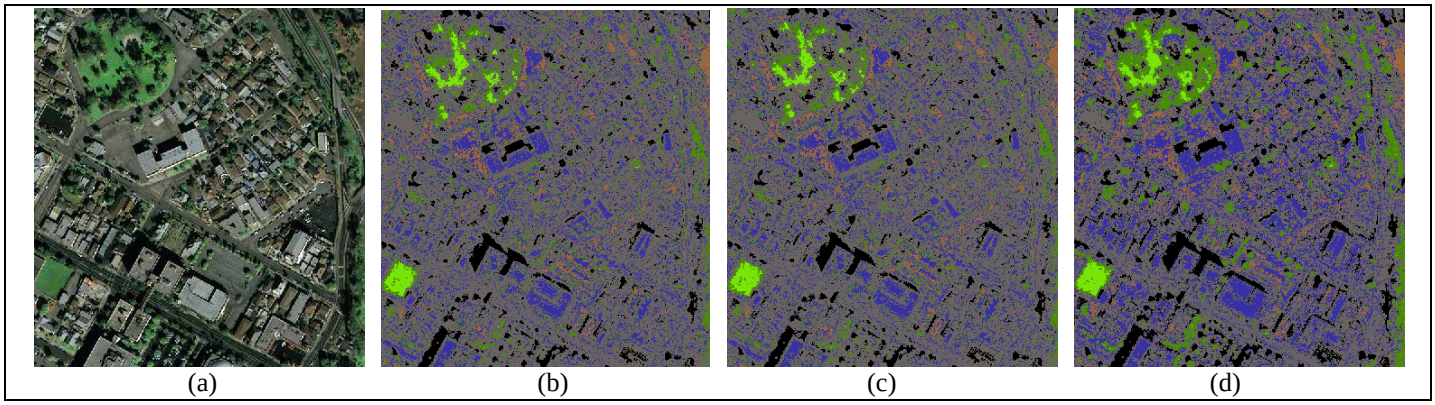


Fig. 8. (a) Original image 2, (b) Classification Map obtained using the stacked features approach, (c) Classification Map obtained using the direct summation kernel , (d) Classification Map obtained using the weighted summation kernel.

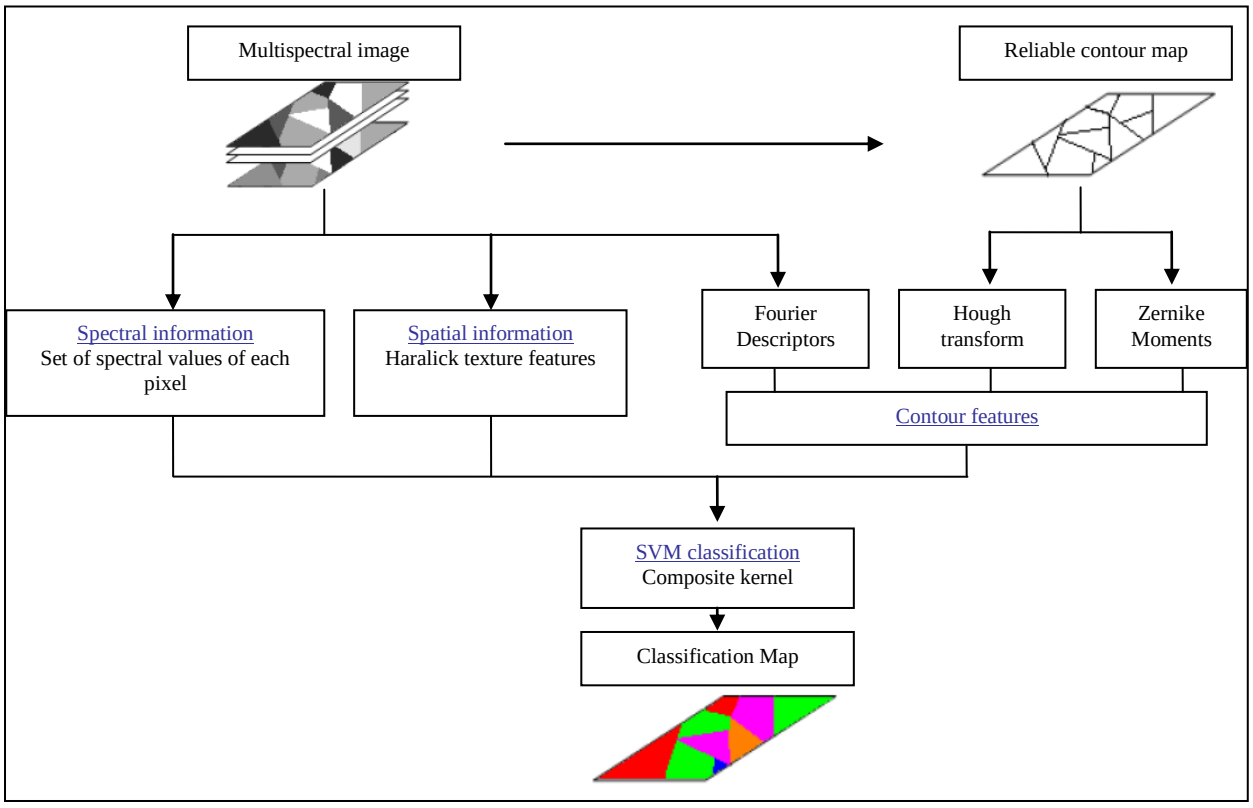


Fig. 9. A representative illustration of the second stage of the proposed workflow

C. Introducing Contour Information

In the second stage (represented by Fig.9.) we have started, like the first stage, with the extraction of the spectral and spatial features, so we have computed Grey Level Co-occurrence Matrix (GLCM) to extract Haralick texture features that we have added to spectral information. But, before the

SVM classification, we have an additional step that consists on building a reliable contour map from which we have extracted contour descriptors specially Hough transform and Zernike moments, while Fourier descriptors are extracted directly from the original image.

1) Edge Detector Choice

Generally the edge detectors can be grouped into three major categories: the first one is the Early vision edge detectors (Gradient operators, e.g. the detectors of Sobel and Kirsch). The second category is Optimal detectors (e.g. the Canny algorithm, etc.). The third category is the Operators using parametric fitting models (e.g. the detectors of Haralick, Nalwa-Binford, Nayar, Meer and Georgescu, etc) [53].

The edge detection process is greatly eased if, instead the original images, «edge enhanced» ones are used. This inevitably leads to the use of some edge detectors from the second category.

In the present work, we have chosen to use Canny edge detector. John Canny has treated edge detection as signal processing problem and aimed to design the «optimal» edge detector. He formally has specified an objective function to be optimized and used this to design the operator.

The objective function was designed to achieve the following optimization constrains [54]:

- Maximize the signal to noise ratio in order to provide good detection.
- Achieve good localization to accurately mark edges.
- Minimize the number of responses to a single edge (non-edges are not marked).

2) Building a Reliable Contour Map

The Canny method finds edges by looking for local maxima of the gradient of the image. The gradient is calculated using the derivative of a Gaussian filter. The method uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. This method is therefore less likely than the others to be fooled by noise, and more likely to detect true weak edges.

For simplicity and for illustrative purposes, we have used edge function in Matlab to extract contour map with the Canny method, and we have specified a scalar for thresh, this scalar value is used for the high threshold and 0.4*thresh is used for the low threshold. This scalar was varied with a step of 0.1 in the range [0, 1]. The Fig.10. Represents two values of threshold used for the first image.

For the choice of thresholds of the image contours that gives us a reliable contour map which will be used later in the classification process, we have adopted two measures proposed by Wiedemann [55], which are used for the evaluation of extraction methods roads from satellite images, these two measures are defined as follows:

Completeness = length of the reference contour in accordance with the extracted contour / length of the reference contour

Exactness = length of the extracted contour in accordance with the reference contour / length of the extracted contour.

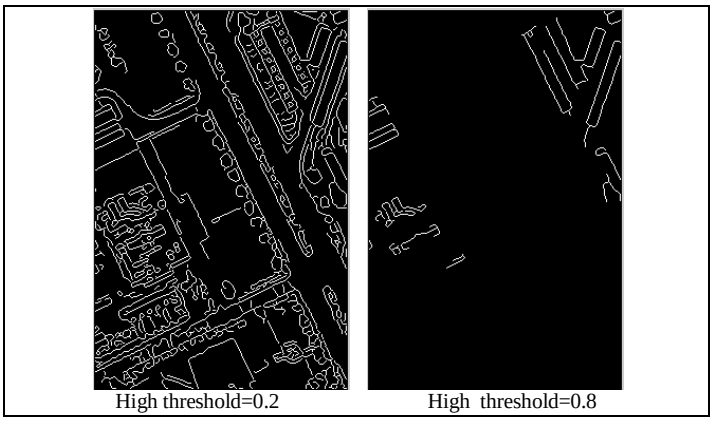


Fig. 10. exemple of contour map for image 1

The principle is to compare the contours of each threshold with the reference contours which are the contours of the SVM classification using the spectral and spatial information (Fig.11.).

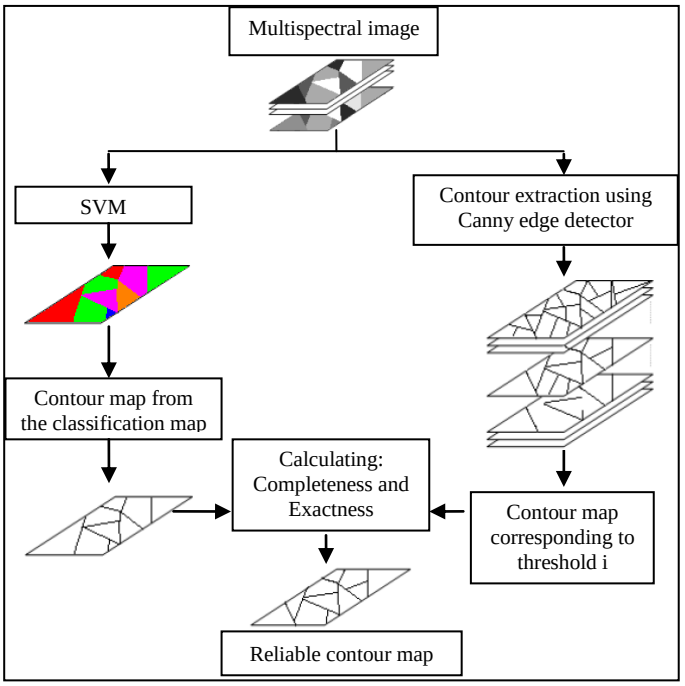


Fig. 11. Selecting reliable contour map

The comparison is made through the calculation of these measures. The constraint is that the selected threshold map is the one in which the extracted contours are the closest to the classification reference contours. The assessment method implemented in our study has a tolerance of a width of three pixels along the edges. The Fig.12. Represents a threshold evaluation for both images. The choice of thresholds of the image contours that gives us a reliable contour map that we have taken the one with a good both Completeness and Exactness, so we have chosen threshold 0.3 for image 1 and 0.4 for image 2 as we can see in Fig.12.

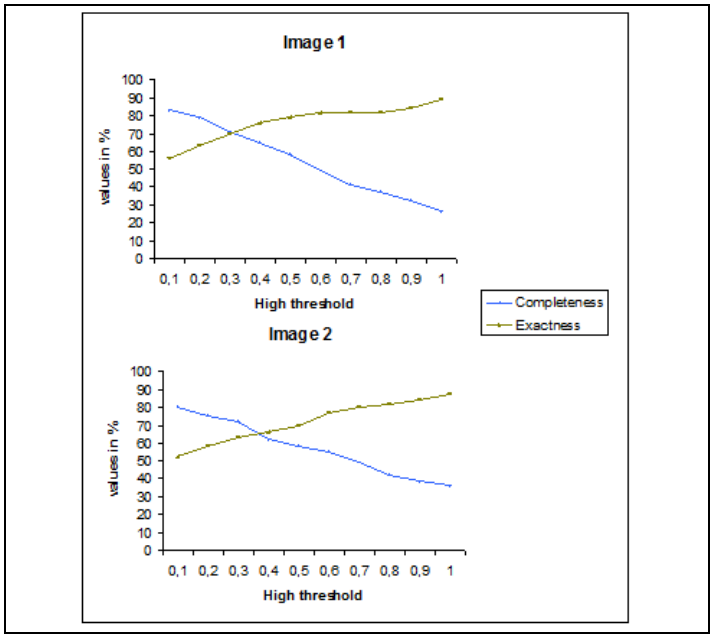


Fig. 12. threshold evalation for the two images

3) Results

To combine spectral, spatial and contour information, we have used the extended weighted summation kernel in (38) that gave the best performance at the first stage of our experiments.

Where the μ_m are varied in the range [0, 1] to satisfy the condition $\sum_{m=1}^{3}\mu_m = 1$. For simplicity and for illustrative purposes, all μ_m were the same for all classes in our experiments. The penalization factor in the SVM was tuned in the range $C = \{10^{-1} \ldots 10^{7}\}$.

In this work, we have computed the participation of contour information in function of spectral and spatial information: $\mu_3 = 1-(\mu_1+\mu_2)$ and we have varied μ_1 *and* μ_2 with a step of 0.1 in the range [0, 1] to satisfy the condition $\sum_{m=1}^{3}\mu_m = 1$.

We have used the Gaussian RBF kernel (28) (with $\sigma = \{10^{-1} \ldots 10^{3}\}$) for all kernels. k_{spect} uses a spectral information, k_{spa} uses Haralick features while k_{cont} uses Fourier descriptors, Hough transform and Zernike moments.

The image (c) in Fig.13. and Fig.14. represent the reliable contour map used to compute contour descriptors' (Hough transform and Zernike moments); while (d) in Fig.13. and Fig.14. represent the classification map resulting by introducing contour (Fourier descriptors, Hough transform and Zernike moments) information with both spectral and spatial information.

A visual analysis of classification maps shows that it is less noisy and the classification performances are increased globally as well as almost all the classes. It matches well with an urban land cover map in terms of smoothness of the classes; and it also represents more connected classes.

Table V lists the accuracy estimates and kappa coefficient of the classification results, we can find different combination of descriptors used to characterize the contour information all models are compared numerically (overall accuracy, kappa coefficient).

Table VI and Table VII present respectively the confusion matrix results for SVM classification using the extended weighted summation kernel (38) based on spectral, spatial and contour information for both images used in experiments.

TABLE V. OVERALL ACCURACY (%) AND KAPPA COEFFICIENT OF CLASSIFIED IMAGES USING THE EXTENDED WEIGHTED SUMMATION KERNEL

Used Descriptors	Image 1		Image 2	
	Overall accuracy	*Kappa coefficient*	*Overall accuracy*	*Kappa coefficient*
Spectral + haralick features	94.48%	0.93	92.55%	0.92
Spectral + haralick features + FD	94.91%	0.93	92.88%	0.92
Spectral + haralick features + ZM	94.59%	0.93	92.68%	0.92
Spectral + haralick features + HT	94.49%	0.93	92.56%	0.92
Spectral + haralick features + FD +HT	95.06%	0.93	93.13%	0.92
Spectral + haralick features + FD + ZM	95.94%	0.94	93.98%	0.93
Spectral + haralick features + HT + ZM	95.81%	0.94	93.72%	0.93
Spectral + haralick features + FD + HT + ZM	96.17%	0.95	94.08%	0.93

The composite kernels offer excellent performance for the classification of multispectral satellite images by simultaneously exploiting both the spatial and spectral information. The weighted summation kernel allows a significant improvement of the classification performances when compared with the two other approaches. So the extended weighted summation kernel has been selected to introduce contour information.

The experimental results indicate a global accuracy value of 93.52%, the addition of contour information, described by the Fourier descriptors, Hough transform and Zernike moments, allows increasing the obtained global accuracy by 1.61% (using all descriptors) which is very promising. Although the Hough transform don't give a remarkable increasing of the overall accuracy, it preserves the edges in the obtained classification map.

TABLE VI. CONFUSION MATRIX RESULTS (%) FOR SVM CLASSIFICATION USING THE EXTENCED WEIGHTED SUMMATION KERNEL WITH ALL DESCRIPTORS FOR IMAGE 1. GLOBAL ACCURACY = 96.17 %

Class name	Asphalt	Green area	Tree	Soil	Building	Shadow
Asphalt	96,52	0,34	1,92	0,00	0,62	0,60
Green area	1,03	96,78	0,00	0,03	0,87	1,29
Tree	0,18	1,36	95,42	0,38	0,00	2,66
Soil	0,00	0,34	0,13	96,94	2,49	0,10
Building	1,94	1,16	0,81	0,08	96,01	0,00
Shadow	0,33	0,02	1,72	2,57	0,01	95,35

TABLE VII. CONFUSION MATRIX RESULTS (%) FOR SVM CLASSIFICATION USING THE EXTENDED WEIGHTED SUMMATION KERNEL WITH ALL DESCRIPTORS FOR IMAGE 2. GLOBAL ACCURACY = 94.08%

Class name	Asphalt	Green area	Tree	Soil	Building	Shadow
Asphalt	93,23	1,00	3,21	0,00	0,64	1,92
Green area	1,04	95,18	0,00	1,08	1,44	1,26
Tree	0,28	1,08	93,91	1,40	0,82	2,51
Soil	3,33	1,26	0,00	93,07	2,34	0,00
Building	1,41	1,06	0,41	2,36	94,76	0,00
Shadow	0,71	0,42	2,47	2,09	0,00	94,31

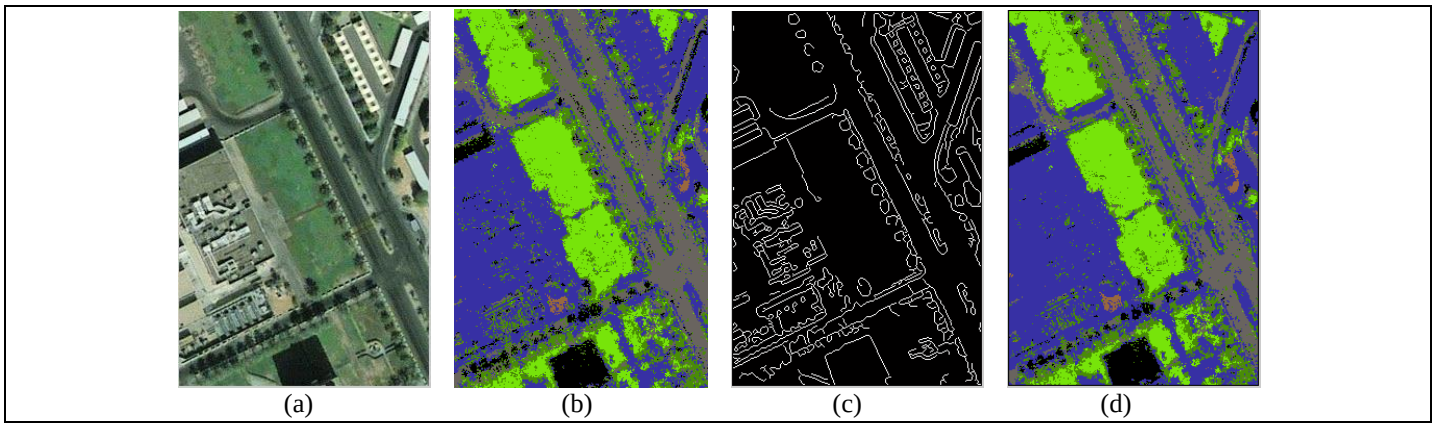


Fig. 13. (a) Original image 1, (b) Classification Map obtained using the weighted summation kernel, (c) A reliable contour map and (d) Classification Map obtained using the extended weighted summation kernel

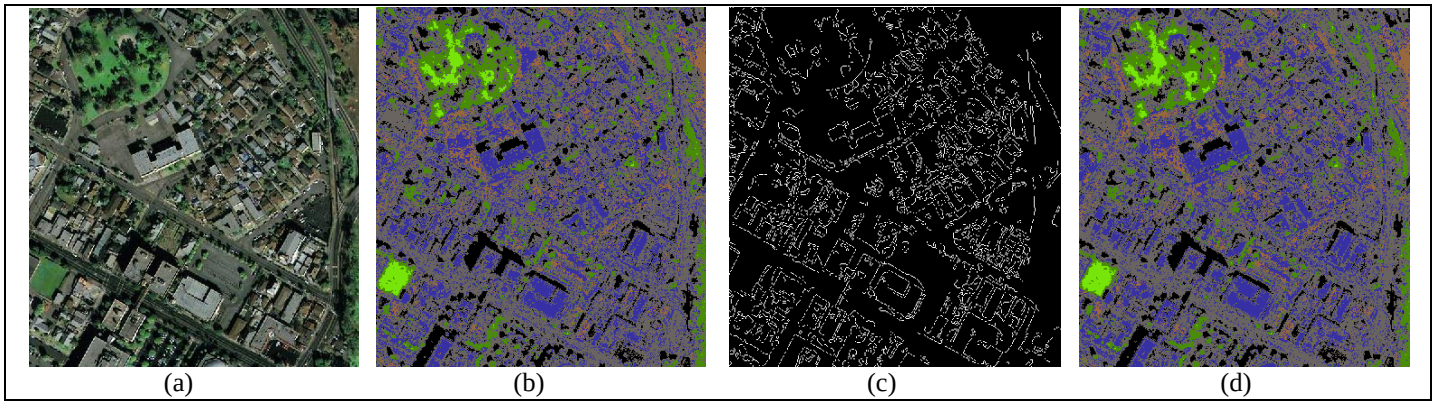


Fig. 14. (a) Original image 2, (b) Classification Map obtained using the weighted summation kernel, (c) the reliable contour map and (d) Classification Map obtained using the extended weighted summation kernel

VI. Conclusion And Future Research Lines

Addressing the classification of high resolution satellite images from urban areas, we have presented three different kernel approaches taking simultaneously the spectral and the spatial information into account (the spectral values and the Haralick features).

The weighted summation kernel allows a significant improvement of the classification performances when compared with the two other approaches. So an extended version of this kernel has been selected to introduce contour information (Fourier descriptors, Hough transform and Zernike moments). This approach exhibits flexibility to balance between the spectral, spatial and contour information as well as computational efficiency.

The proposed method is computationally expensive in comparison with a single kernel-based approach. In order to address this issue, we are planning on exploring the impact of reducing the original data set dimensionality before applying the proposed approach.

We are also planning to explore nonlinear combination methods, and the data-dependent combination methods which assign specific kernel weights for each data instance, to identify local distributions in the data and learn proper kernel combination rules for each region.

Acknowledgment

This work was funded by CNRST Morocco and CNRS France Grant under “Convention CNRST CRNS” program SPI09/12.

References

[1] G. F.Hughes, "On the mean accuracy of statistical pattern recognizers, " IEEE Trans. Inf. Theory, 1968, IT vol. 14 no. 1 pp. 55-63.

[2] D. A.Landgrebe, "Signal Theory Methods in Multispectral Remote Sensing. " New York: Wiley 2003.

[3] C. Samson "Contribution à la classification des images satellitaires par approche variationnelle et équations aux dérivées partielles" : Thesis of doctorate, university of Nice-Sophia Antipolis 2000.

[4] J.R.G.Townshend, "Land cover". International Journal of Remote Sensing, 1992 vol. 13 pp. 1319–1328.

[5] F.G. Hall, J.R. Townshend, E.T. Engman, "Status of remote sensing algorithms for estimation of land surface state parameters." Remote Sensing of Environment,1995 vol. 51 pp. 138–156.

[6] Lu, D.,Weng, Q.,(2007) "A survey of image classification methods and techniques for improving classification performance. " International Journal of Remote Sensing 28:823–870.

[7] C. Huang, L.S. Davis and J.R.G. Townshed, "An assessment of support vector machines for land cover classification. "International Journal of Remote Sensing, 2002 vol. 23 pp. 725–749.

[8] T. Kavzoglu, S. Reis, "Performance analysis of maximum likelihood and artificial neural network classifiers for training sets with mixed pixels." GIScience and Remote Sensing, 2008 vol. 45 pp. 330–342.

[9] M. Pal, and P.M. Mather, "Support vector machines for classification in remote sensing." International Journal of Remote Sensing, 2005 vol. 26 pp. 1007−1011.

[10] G. Zhu, and D.G. Blumberg, "Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel." Remote Sensing of Environment, 2002 vol. 80 pp. 233-240.

[11] B. Scholkopf, K. Sung, C. Burges, F. Girosi, P. Niyogi, T. Poggio, et al. "Comparing support vector machines with gaussian kernels to radial basis function classifiers." IEEE Transactions on Signal Processing, 1997 vol. 45 pp. 2758−2765.

[12] X. Cao, J. Chen, H. Imura, O. Higashi, "A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data", Remote Sensing of Environment, 2009 vol. 113 pp. 2205–2209.

[13] J. Inglada, "Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features", ISPRS Journal of Photogrammetry & Remote Sensing, 2007 vol. 62 pp. 236–248.

[14] A. Bekkari, S. Idbraim, D. Mammass and M. El yassa "Exploiting spectral and space information in classification of high resolution urban satellites images using Haralick features and SVM" IEEE 2ed International Conference on Multimedia Computing and Systems ICMCS’11 , Ouarzazate, Morocco 2011.

[15] M. Fauvel, J.A. Benediktsson, J. Chanussot and J.R. Sveinsson, “Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles” IEEE International Geoscience and Remote Sensing Symposium, IGARSS 07, Barcelona Spain 2007.

[16] L. Lepisto, I. Kunttu, J. Autio, and A. Visa, “Classification method for colored natural textures using gabor filtering,” Image Analysis and Proc., pp. 397–401, Sept. 2003.

[17] L. Lepisto, I. Kunttu, J. Autio, and A. Visa, “Rock image classification using non-homogeneous textures and spectral imaging,” Proc. of the WSCG, 2003.

[18] J.A. Palmason, J.A. Benediktsson, J.R. Sveinsson, and J. Chanussot, “Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis,” IEEE Trans., Int. Geosci. and Rem. Sens., vol. 1, July 2005.

[19] A. Plaza, P. Martinez, R. Perez, and J. Plaza, “Spatial/ spectral endmember extraction by multidimensional morphological operations,” IEEE Trans., Int. Geosci.and Rem. Sens., vol. 40, no. 9, pp. 2025–2041, Sep 2002.

[20] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson,“Classification based marker selection for watershed transform of hyperspectral images,” IEEE Trans., Int. Geosci. and Rem. Sens. Symp., 2009.

[21] G. Roussel, V. Achard, A. Alakian and J.C. Fort ‘benefits of textural characterization for the classification of hyperspectral images’, 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010, pp:1-4.

[22] G. Roussel ‘Développement et évaluation de nouvelles methods de classification spatiale-spectrale d’images hyperspectrales’, Theses Days ISAE Toulouse France, 2010.

[23] M. Sharma, M. Markou and S. Singh ‘evaluation of texture methods for image analysis’ Intelligent Information Systems Conference, The Seventh Australian and New Zealand 2001, pp. 117-121.

[24] W.Y. Chiu, and I. Couloigner, "Evaluation of incorporating texture into wetland mapping from multispectral images" University of Calgary, Department of Geomatics Engineering, Calgary, Canada, EARSeL eProceedings 2004.

[25] R.M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification." IEEE Transactions on Systems Man and Cybernetics, 1973.

[26] J.S. Weszka, C.R. Dyer, and A. Rosenfeld. "A Comparative Study of Texture measures for Terrain Classification." IEEE Transactions on Systems Man and Cybernetics, 1976.

[27] R.W. Conners, and C.A. Harlow, "A Theoretical Comaprison of Texture Algorithms." IEEE Transactions on Pattern Analysis and Machine Intelligence, 1980.

[28] P. Gong, D. J. Marceau and P. J. Howarth ‘A Comparison of Spatial Feature Extraction Algorithms for Land-Use Classification with SPOT HRV Data’, Remote Sensing Environ, 1992. vol. 40 pp. 137-151.

[29] V. Arvis, C. Debain, M. Berducat and A. Benassi, "Generalization of the cooccurrence matrix for colour images: application to colour texture classification" journal Image Analysis and Stereology, 2004 vol. 23 pp. 63-72.

[30] H. Ming-Kuel, "Visual pattern recognition by moment invariants", IRE trans. on Information Theory, 1962 vol. 8 pp. 179-187,.

[31] M. Teague, "Image analysis via the general theory of moments", Journal of the Optical Society of America, 1980 vol. 70 no. 8 pp. 920-930.

[32] R. L. Cosgriff, "Identification of shapes", Ohio State Univ. Res. Foundation, Columbus, 1960 Rep. 820-11, ASTIA AD 254 792.

[33] D. L. Fritzsche, "A systematic method for character recognition", Ohio State Univ. Res. Foundation, Columbus, 1961 Rep. 1222-4, ASTIA AD 268 360.

[34] G. H. Granlund, "Fourier preprocessing of hand printed character recognition", IEEE Trans. Comp., 1972 C-21 pp. 195-201.

[35] C. T. Zhan, R. S. Roskies, "Fourier descriptors for plane closed curves", IEEE Trans. Comp., 1972 C-21 pp. 269-281.

[36] L. Yang, R. Prasad, "Recognition of line-drawing based on generalized Fourier descriptors", International Conference on Image Processing and its Applications, 1992 pp. 286-289.

[37] C. Adamos, W. Faig "Hough Transform in Digital Photogrammetry". In: International Archives of Photogrammetry and Remote Sensing, Washington, USA, 1992 vol. 29, Part B3 pp. 250-254.

[38] D. H. Ballard, "Generalizing the Hough Transform to detect arbitrary shapes". Pattern Recognition, 1981 vol. 13 no. 2 pp. 111- 122.

[39] R. D. Duda, P. E. Hart, "Use of the Hough Transform to detect lines and curves in pictures. " Communication of the ACM, 1972 vol. 15 no. 1 pp. 11-15.

[40] J. D. Klingensmith, R. Shekhar, and D. Geoffrey, "Evaluation of Three-Dimensional Segmentation Algorithms for the Identification of Luminal and Medial–Adventitial Borders in Intravascular Ultrasound Images, " IEEE Transactions on Medical Imaging, 2000 vol.19 pp. 996-1011.

[41] M. S. Nixon, A. S. Aguado, "Feature extraction & image processing, " Academic Press Inc, 2nd revised edition 2007.

[42] R. Roman-Roldan, J. Francisco-Gomez-Lopera, C. Ataellah, J. Martinez-Aroza, P. L. Luque-Escamilla, "A measure of quality for evaluating methods of segmentation and edge detection, Pattern Recognition," 2001 vol. 34 pp. 969-980.

[43] W. C. Siu, "Efficient computation of moments for pattern recognition, " IEEE Pacific Rim Conference on Communications Computers and Signal Processing, 1991 vol. 2 pp. 589-592.

[44] R. Unnikrishnan, C. Pantofaru and M. Hebert, "A measure for objective evaluation of image segmentation algorithms, " IEEE Workshops on Computer Vision and Pattern Recognition, 2005 pp. 34-34.

[45] Y. J. Zhang, "A review of recent evaluation methods for images segmentation, " International Symposium on Signal Processing and its Applications, 2001 vol. 1 pp. 148-151.

[46] L. Chapel, " Maintenir la viabilité ou la résilience d'un système : les machines à vecteurs de support pour rompre la malédiction de la dimensionnalité ? " : Thesis of doctorate, university of Blaise Pascal - Clermont II 2007.

[47] S. Aseervatham " Apprentissage à base de Noyaux Sémantiques pour le traitement de données textuelles " : Thesis of doctorate, university of Paris 13 –Galilée Institut Laboratory of Data processing of Paris Nord 2007.

[48] O. Bousquet, "Introduction au Support Vector Machines (SVM)", Center mathematics applied, polytechnique school of Palaiseau 2001. http://www.math.u-psud.fr/~blanchard/gtsvm/index.html.

[49] M. Fauvel, J. Chanussot and J. A. Benediktsson "A Combined Support Vector Machines Classification Based on Decision Fusion" IEEE International Geoscience and Remote Sensing Symposium, IGARSS 06, Denver, USA 2006.

[50] B. Mak, J. Kwok, and S. Ho, "A study of various composite kernels for kernel eigenvoice speaker adaptation," in IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP04, 2004 vol. 1.

[51] J.-T. Sun, B.-Y Zhang., Z. Chen, Y.-C Lu., C.-Y Shi, and W. Ma, "GE-CKO: A method to optimize composite kernels for web page classification," in IEEE/WIC/ACM International Conference on Web Intelligence, WI04, 2004 vol. 1 pp. 299–305.

[52] SVMlight Version: 6.02 Developed at University of Dortmund, Informatik, AI-Unit Collaborative Research Center on 'Complexity Reduction in Multivariate Data' (SFB475). 2008 http://svmlight.joachims.org/

[53] D. Ziou, and S. Tabbone, "Edge Detection Techniques – An Overview. TR-195", Département de Math et Informatique, Université de Sherbrooke, Québec, Canada, 1997 pp. 1-41.

[54] J. Canny, "A computational approach to edge detection". IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986 volo. 8 no. 6, pp. 679-698.

[55] C. Wiedemann, C. Heipke, H. Mayer, "Empirical evaluation of automatically extracted road axes", In: CVPR Workshop on Empirical Evaluation Methods in Computer Vision, California, , 1998 pp. 172–187.

Performance Evaluation of Two-Hop Wireless Link under Nakagami-m Fading

Afsana Nadia[1], Arifur Rahim Chowdhury[2], Md. Shoayeb Hossain[3] and M. R. Amin[5]
Dept. of Electronics and Communication Engineering
East West University, Dhaka, Bangladesh

Md. Imdadul Islam[4]
Department of Computer Science and Engineering
Jahangirnagar University
Dhaka, Bangladesh

***Abstract*—Now-a-days, intense research is going on two-hop wireless link under different fading conditions with its remedial measures. In this paper work, a two-hop link under three different conditions is considered: (i) MIMO on both hops, (ii) MISO in first hop and SIMO in second hop and finally (iii) SIMO in first hop and MISO in second hop. The three models used here give the flexibility of using STBC (Space Time Block Coding) and combining scheme on any of the source to relay (S-R) and relay to destination (R-D) link. Even incorporation of Transmitting Antenna Selection (TAS) is possible on any link. Here, the variation of SER (Symbol Error Rate) is determined against mean SNR (Signal-to-Noise Ratio) of R-D link for three different modulation schemes: BPSK, 8-PSK and 16-PSK, taking the number of antennas and SNR of S-R link as parameters under Nakagami -m fading condition.**

Keywords—MIMO; MISO; SIMO; TAS; MRC; SER; SNR.

I. INTRODUCTION

Most common protocols used in wireless networks are the decode-and-forward (DF) and amplify-and-forward (AF) methods [1]. AF protocol used the knowledge of the instantaneous channel state information (CSI) of the source to relay (S-R) channel to control the gain by the relay [2]. The CSI assisted relay may also use the knowledge of the CSI of the S-R channel in the gain; we refer to this as CSI-assisted AF (CSIAF). When the direct link between source and destination is neglected, this is referred to a two-hop network. Two-hop relay networks, where the channel from source to destination (S-D) is split into two possibly shorter links using a relay, are attractive when the direct link between the source and destination is in deep fading. The performance analysis of the two-hop relay network has gained a lot of attention [3]. The two-hop relaying techniques are one of the promisingwireless technologies that have been kindling an enormous interest from the wireless community in the last decade [4]. To achieve higher reliability and throughput for wireless networks, half-duplex two-way relaying system has attracted much research interest [5]. Two-hop relaying communication has a number of advantages over direct-link transmission in terms of connectivity, power saving and channel capacity for the high data-rate coverage required for future cellular and ad-hoc networks [6]. Relaying is a convenient solution to satisfy the requirements of the next generation wireless communication systems, such as: high data rates and large coverage areas [7].

Multiple-antenna systems, also known as multiple-input multiple-output (MIMO) radio, can improve the capacity and reliability of radio communication. In this system, the multiple antenna elements are at both the transmitter and the receiver [8]. They were first investigated by computer simulations in the 1980s [9], and later papers explored them analytically. Since that time, interest in MIMO systems has exploded. The multiple antennas in MIMO systems can be exploited in two different ways. One is the creation of a highly effective antenna diversity system; the other is the use of the multiple antennas for the transmission of several parallel data streams to increase the capacity of the system. The multiple antennas also increase the average SNR seen at the combined output [10]. Antenna diversity is used in wireless systems to combat the effects of fading. If multiple independent copies of the same signal are available, we can combine them to a total signal with high quality, even if some of the copies exhibit low quality.

Deploying multiple antennas in wireless relay networks, referred to as MIMO relaying, has been identified as a promising technique to combat fading and increase transmission reliability [11]–[12]. Transmit antenna selection (TAS) with receive maximal-ratio combining (MRC) in MIMO relaying was proposed in [13]. In TAS/MRC relaying, a single antenna is selected at the transmitters and all the antennas at the receivers are MRC combined [14]. In [15], authors presented a framework for the comparative analysis of TAS/MRC and TAS with receive selection combining (TAS/SC) in a two-hop AF relay network. TAS/MRC and TAS/SC are two attractive MIMO protocols.

This paper presents the performance of a two-hop link where the number of antennas and SNR of S-R as parameters is evaluated under Nakagami-m fading environment separately for BPSK, 8-PSK and 16-PSK modulation schemes. The objective of the paper is to observe the relative impact of Nakagami-m fading environment on the two above mentioned modulation schemes under three different conditions; MIMO on the both hop, MISO in the first hop and SIMO in the second hop, SIMO in the first hop and MISO in the second hop. Though Nakagami-m fading affects the two-hop link, hence some additional techniques like: adaptive equalization, combining scheme of MIMO, incorporation of space-time block code (STBC) etc. are recommended to enhance the performance of such links.

The rest of the paper is organized as follows. Section II describes the system model under consideration. In section III, performance analysis and results from the system model are presented. Finally, section IV concludes the paper.

II. System Model

We consider a two-hop wireless network as shown in Fig. 1, where the source node *S*, communicates with the destination node *D*, through the relay node *R* and they equipped with N_S, N_R, N_D antennas respectively. There is no direct link between *S* and *D* and the communication can be performed only through the relay *R*. This introduces fixed gain on the received signal regardless of the amplitude on the first hop, hence in an output signal with variable power, this type of fixed gain relay is cost effective to implement. Here, space diversity technique is used. Space diversity technique employs multiple transmit or receive antennas having some separation between the adjacent antennas [16]. Various techniques are available to combine the signals from multiple diversity branches. As has been mentioned in the introduction that MRC scheme is one of them. MRC represents a theoretically optimal combiner over fading channels as a diversity scheme in a communication system. Theoretically, multiple copies of the same information signal are combined so as to maximize the instantaneous SNR at the output [17]. Here $N_S = i$, $N_R = j$ and $N_D = k$ where $i = 1,2,\ldots n$; $j = 1,2,\ldots n$; and $k = 1,2,\ldots n$.

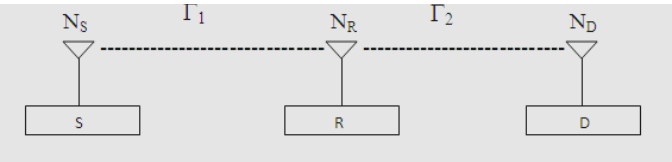


Fig.1. A Two-Hop Wireless Link.

Let Γ_1 and Γ_2 are the random variables representing the SNRs of link S-R and R-D. Then, the equivalent SNR will be

$$\Gamma_{eq} = \frac{\Gamma_1 \Gamma_2}{\Gamma_1 + \Gamma_2 + 1} \quad (1)$$

Let the individual probability density function (PDF) and cumulative distribution function (CDF) of S-R and R-D links are $f_{\Gamma_1}(\gamma_1)$, $F_{\Gamma_1}(\gamma_1)$, $f_{\Gamma_2}(\gamma_2)$, $F_{\Gamma_2}(\gamma_2)$ respectively. Then, the CDF of the equivalent link will be [18]

$$F_{\Gamma_{eq}}(\gamma) = \int_0^\infty Pr\{\gamma_{eq} \le \gamma | \gamma_2\} f_{\Gamma_2}(\gamma_2) d\gamma_2,$$

Which gives, $F_{\Gamma_{eq}}(\gamma) = \int_0^\infty Pr\left\{\frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} \le \gamma | \gamma_2\right\} f_{\Gamma_2}(\gamma_2) d\gamma_2$ (2)

Now, with the condition, $\frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2 + 1} \le \gamma$,

We have, $\gamma_1 \gamma_2 \le \gamma\gamma_1 + \gamma\gamma_2 + \gamma$

Or, $\gamma_1(\gamma_2 - \gamma) \le \gamma(\gamma_2 + 1)$

Or, $\gamma_1 \le \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma}$ (3)

For the condition, when $\gamma_2 > \gamma$, we have

$\gamma_1 \le \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma}$; the range of γ_2 will be $[\gamma, \infty]$ (4)

Again, when $\gamma_2 < \gamma$,

$\gamma_1 \ge \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma}$; the range of γ_2 will be $[o, \gamma]$ (5)

From (1),

$$F_{\Gamma_{eq}}(\gamma) = \int_0^\gamma Pr\{\gamma_1 \ge \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} | \gamma_2\} f_{\Gamma_2}(\gamma_2) d\gamma_2 + \int_\gamma^\infty Pr\{\gamma_1 \le \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} | \gamma_2\} f_{\Gamma_2}(\gamma_2) d\gamma_2$$

Or, $F_{\Gamma_{eq}}(\gamma) = I_1 + I_2$ (6)

Where $I_1 = \int_0^\gamma Pr\{\gamma_1 \ge \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} | \gamma_2\} f_{\Gamma_2}(\gamma_2) d\gamma_2$

and

$$I_2 = \int_\gamma^\infty Pr\{\gamma_1 \le \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} | \gamma_2\} f_{\Gamma_2}(\gamma_2) d\gamma_2$$

We know, the cdf will be

$$F_x(x) = \int_o^x f_x(x)\, dx \quad (7)$$

and, $Pr\{x \le r\} = F_x(r)$ (8)

Then,
$$I_1 = \int_0^\gamma Pr\{\gamma_1 \ge \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} | \gamma_2\} f_{\Gamma_2}(\gamma_2) d\gamma_2$$
$$= \int_0^\gamma \{1 - F_{\Gamma_1}(\frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma})\} f_{\Gamma_2}(\gamma_2) d\gamma_2$$
$$= \int_0^\gamma [1 - \{1 - e^{-\frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} \cdot \frac{1}{\overline{\gamma_1}}}\}] f_{\Gamma_2}(\gamma_2) d\gamma_2$$
$$= \int_0^\gamma e^{-\frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} \cdot \frac{1}{\overline{\gamma_1}}} \frac{1}{\overline{\gamma_2}} e^{-\frac{\gamma_2}{\overline{\gamma_2}}} d\gamma_2 \quad (9)$$

And
$$I_2 = \int_\gamma^\infty Pr\{\gamma_1 \le \frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} | \gamma_2\} f_{\Gamma_2}(\gamma_2) d\gamma_2$$
$$= \int_\gamma^\infty F_{\Gamma_1}\{\frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma}\} f_{\Gamma_2}(\gamma_2) d\gamma_2$$
$$= \int_\gamma^\infty \{1 - e^{-\frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} \cdot \frac{1}{\overline{\gamma_1}}}\} \frac{1}{\overline{\gamma_2}} e^{-\frac{\gamma_2}{\overline{\gamma_2}}} d\gamma_2 \quad (10)$$

Using (6), (9) and (10), we can write

$$F_{\Gamma_{eq}}(\gamma) = \int_0^\gamma e^{-\frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} \cdot \frac{1}{\overline{\gamma_1}}} \cdot \frac{1}{\overline{\gamma_2}} e^{-\frac{\gamma_2}{\overline{\gamma_2}}} d\gamma_2 + \int_\gamma^\infty \{1 - e^{-\frac{\gamma(\gamma_2 + 1)}{\gamma_2 - \gamma} \cdot \frac{1}{\overline{\gamma_1}}}\} \frac{1}{\overline{\gamma_2}} e^{-\frac{\gamma_2}{\overline{\gamma_2}}} d\gamma_2 \quad (11)$$

The symbol error rate (SER) [15] will be

$$P_{SER} = \frac{a}{2} \cdot \sqrt{\frac{b}{\pi}} \int_0^\gamma \frac{e^{-b\gamma}}{\sqrt{\gamma}} F_{\Gamma_{eq}}(\gamma)\, d\gamma \quad (12)$$

Where, *a* and *b* are the constellation-specific constants.

In the following, we have considered three cases depending on the number of antenna of source, relay and destination.

A. Case A: Mimo on Both Hops

In the case A, we consider a distributed wireless network where node S is equipped with a multiple transmitting antennas S_1, S_2 and S_3 and node D is equipped with a multiple receiving antennas D_1, D_2 and D_3 whereas the relay has three antennas R_1, R_2 and R_3 in Fig. 2. These three antennas R_1, R_2 and R_3 play the role of receiving antennas when they receive signals from the transmitting antennas S_1, S_2 and S_3. These three antennas R_1, R_2 and R_3 also play the role of transmitting antennas when they transmit signals to the receiving antennas D_1, D_2 and D_3.

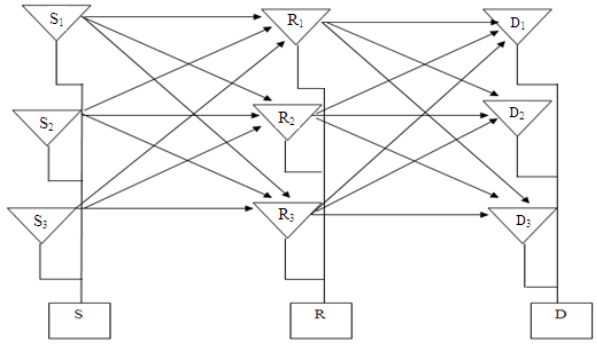


Fig.2. MIMO on Both Hops

B. Case B: MISO in the First Hop and SIMO in the Second Hop

In the case B, node S is equipped with a multiple transmitting antennas S_1, S_2 and S_3 and node D is equipped with a multiple receiving antennas D_1, D_2 and D_3 whereas the relay has a single antenna R_1 in Fig. 3. The antenna R_1 plays the role of receiving antenna when it receives signals from the transmitting antennas S_1, S_2 and S_3. The antenna R_1 also plays the role of transmitting antenna when it transmits signals to the receiving antennas D_1, D_2 and D_3. Therefore, from the path S to R, space diversity is used as there are three transmitters and one receiver. Now, from R to D, the method of MRC is used as there are only one transmitter and three receivers.

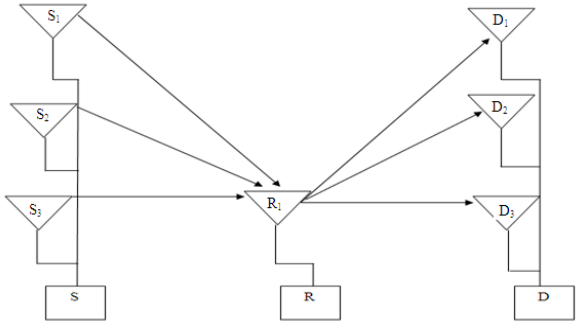


Fig.3. MISO in the First Hop and SIMO in the Second Hop

C. Case C: SIMO in the First Hop and MISO in the Second Hop

In the case C, node S is equipped with a single transmitting antenna S_1 and node D is equipped with a single receiving antenna D_1 whereas the relay has three antennas R_1, R_2 and R_3 in Fig. 4. These three antennas R_1, R_2 and R_3 play the role of receiving antennas when they receive signals from the transmitting antenna S_1. These three antennas R_1, R_2 and R_3 also play the role of transmitting antennas when they transmit signals to the receiving antenna D_1. Therefore, from the path S to R, the method MRC is used as there are only one transmitter and three receivers; from R to D, orthogonal scheme is used as there are three transmitters and one receiver.

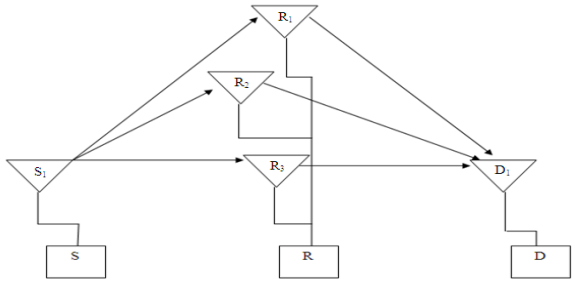


Fig.4. SIMO in the First Hop and MISO in the Second Hop

III. Results

Fig. 5 shows the variation of SER against the instantaneous SNR of R-D link taking $N_S = N_R = N_D = 3$. The following modulation schemes: BPSK, 8- PSK and 16- PSK are considered taking SNR between S-R link as 3dB and 2 dB. Here, we see from Fig. 5 that the SER decreases very rapidly till 5 dB of the mean SNR between R-D. After 5 dB improvement is very slow. For any further improvement, we have to increase the SNR between the source and the relay as visualized form Fig. 5.

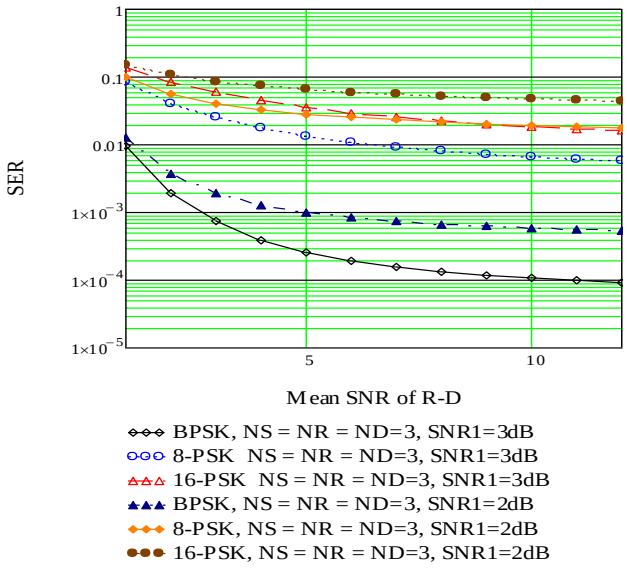


Fig.5. Impact of Modulation Scheme and SNR of link-1

The impact of number of antennas is visualized from Fig. 6. Similar to the previous case like Fig. 5, it is observed that the SER decreases with the mean SNR. Here, the parameter of the curves is the number of antennas of MIMO link. Incorporation of a signal antenna at each step increases the performance of the SER tremendously. The impact of the number of antennas is more prominent of BPSK scheme than the other two modulation schemes.

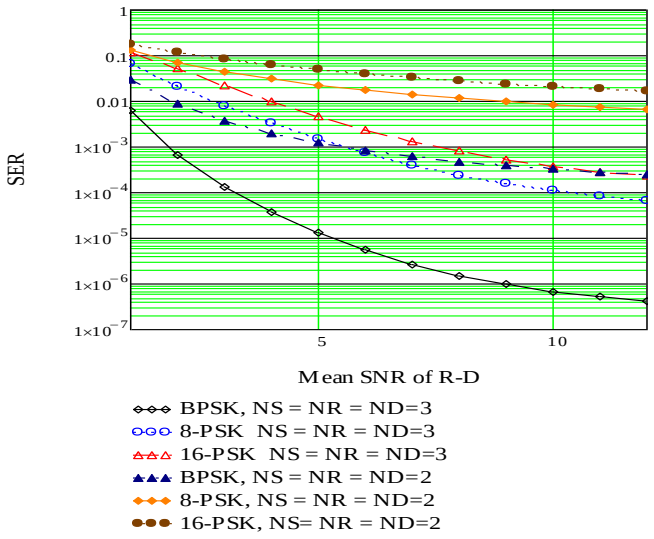


Fig.6. Impact of Modulation Scheme and the number of antennas

Next, we use MISO technique in the first hop and SIMO technique in the second hop so that the space diversity can be applied at the first hop and MRC can be used at the second hop. In this case, performance is heavily improved with the incorporation of one additional antenna at the source and the destination as shown in Fig. 7. But the overall performance of such a scheme is inferior to the MIMO case.

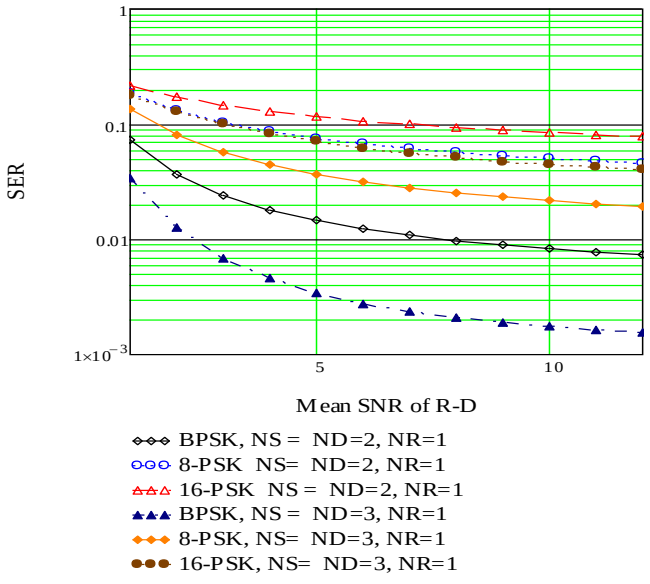


Fig.7. Impact of Modulation Scheme, numbers of transmitting and receiving antennas

Finally, we use single antenna at both the source and the destination but multiple antennas on the relay, hence MRC and orthogonal scheme can also be applied in this case. Fig. 8 shows the performance of the scheme for N_R= 2 and N_R= 4 cases. This scheme gives marginally better performance than that of the case of Fig. 7 for the same number of antennas and modulation schemes.

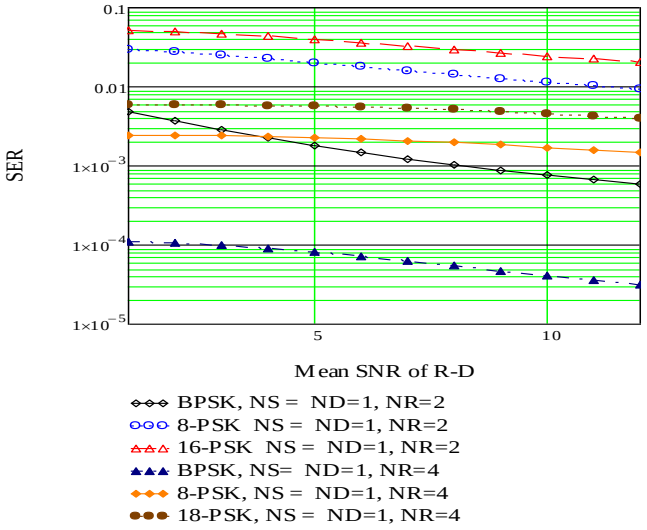


Fig.8. Impact of Modulation Scheme and numbers of relaying antennas

IV. CONCLUTION

The finding of the paper is that, instead of using multiple antennas at the sender and the receiver one can use multiple antennas only on the relay but single antenna at the sender and the receiver. Therefore, SIMO in the first hop and MISO in the second hop is better scheme than MISO in the first hop and SIMO in the second hop case. Although, MIMO on the both hops is better than anyone of above model. But if we use SIMO-MISO scheme, cost could be minimized. So, SIMO-MISO combination is the best in context of SER and cost. We have used this technique only for Nakagami-m fading. In future, the above phenomenon can also be observed for Rayleigh, Rician and K-fading cases. Furthermore, equalizer could be used to improve the overall system performance. This work is under investigation and will be reported soon in future.

REFERENCES

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks efficient protocols and outage behaviour", IEEE Trans. Info. Theory, vol. 50, no. 12, pp. 3062–3080, 2004.

[2] M. O. Hasna and M.-S. Alouini, "A performance study of dual-hop transmissions with fixed gain relays," IEEE Trans. Wireless Commun., vol. 3, no. 6, pp. 1963–1968, Nov. 2004.

[3] Lei Cao, Li Chen, Xin Zhang, and Dacheng Yang, "Asymptotic performance of two-hop amplify-and-forward MIMO relaying withtransmit antenna selection" Global Telecommunications Conference (GLOBECOM 2010), Dec 2010.

[4] Daniel Benevides da Costa, Member, IEEE, and Michel DaoudYacoub, "Outage Performance of Two Hop AF Relaying Systems with Co-Channel Interferers over Nakagami-m Fading" in IEEE communications letters, vol. 15, no. 9, september 2011.

[5] T. Oechtering, E. Jorswieck, R. Wyrembelski, and H. Boche, "On theoptimal transmit strategy for the MIMO bidirectional broadcast

channel," IEEE Trans. Commun., vol. 57, no. 12, pp. 3817–3826, Dec. 2009.

[6] SoumendraNathDatta and SaswatChakrabarti, "Unified Error Analysis of Dual-Hop Relay Link in Nakagami-m Fading Channels," IEEE Commun. Lett. Vol. 14, no. 10, pp. 897-899, Oct. 2010.

[7] Himal A. Suraweera and Geroge K. Karagiannidis, "Closed-Form Error Analysis of the Non-Identical Nakagami-m Relay Fading Channel," IEEE Commun. Lett., vol. 12, no. 4, pp. 259-261, April 2008.

[8] D. Gesbert, M. Shafi, D. Shan Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," IEEE J. Selected Areas Comm., vol. 21, pp. 281–302, 2003.

[9] J. H. Winters, "On the capacity of radio communications systems with diversity in Rayleigh fading environments," IEEE J. Selected Areas Comm., vol. 5, pp. 871–878, June 1987.

[10] Andreas F. Molisch and Moe Z. Win, "MIMO Systems with Antenna Selection – an overview", Mitsubishi Electric Research Labs, TR-2004-014, March 2004.

[11] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," IEEE Trans. Inform. Theory, vol. 51, pp. 29–43, Jan. 2005.

[12] B. Khoshnevis, W. Yu, and R Adve, "Grassmannianbeamforming for MIMO amplify-and-forward relaying," IEEE J. Sel. Areas Commun., vol. 26, pp. 1397–1407, Oct. 2008.

[13] P.L. Yeoh, M. Elkashlan, and I.B. Collings, "Exact and asymptotic SER of distributed TAS/MRC in MIMO relay networks, " IEEE Trans. Wireless Commun., vol. 10, no. 3, pp. 751-756, Mar. 2011.

[14] H.A. Suraweera, G. K. Karagianniadis, Y. Li, H. K. Garg, A. Nallanathan, and B. Vucetic, " Amplify-and-forward relay transmission with end-to-end antenna selection," in Proc. IEEE WCNC, Sydney, Australia, Apr, 2010, pp. 1-6.

[15] M. Elkashlan, P. L. Yeoh, N. Yang, T. Q. Duong, and C. Leung, "A comparison of two MIMO relaying protocols in Nakagami-m fading", IEEE Transactions on Vehicular Technology, vol. 61, no. 3, March 2012.

[16] Himadri S. Saha, Md. Imdadul Islam and M.R.Amin, "Performance evaluation of two-hop wireless link under Rayleigh and Nakagami-m Fading", International Journal of Engineering and Technology, vol. 2, no. 1, Jan. 2012.

[17] Simon Haykin and Michael Moher, "Modern Wireless Communications," Pearson Education, New Delhi, India, 2005.

[18] T. A. Tsiftsis, G. K. Karagiannidis, P. T. Mathiopoulos, and S. A. Kotsopoulos, "Nonregenerative dual-hop cooperative links with selection diversity," EURASIP J. Wireless Commun. Netw., vol. 2006, no. 2, pp. 17862-1–17862-8, Apr. 2006.

AUTHOR PROFILE

Afsana Nadia has completed her B.Sc. Engineering in Electronics and Telecommunication Engineering from Daffodil International University, Dhaka, Bangladesh in 2011 and M.S. Engineering in Telecommunication Engineering from East West University, Dhaka, Bangladesh in 2012. She is the associate member of Bangladesh Computer Society. Her fields of research are computer networks, wireless and mobile communications.

Arifur Rahim Chowdhury has completed his B.Sc. in Electronics and Telecommunication Engineering degree from the People's University of Bangladesh, Dhaka in 2011 and M.S. Engineering in Telecommunication Engineering from East West University, Dhaka, Bangladesh in 2012. His field of study is wireless communications and networks.

Md. Shoayeb Hossain has completed his B.Sc. in Information and Telecommunication Engineering degree from Darul Ihsan University, Dhaka, Bangladesh in 2010 and M.S. Engineering in Telecommunication Engineering from East West University, Dhaka, Bangladesh in 2012. His field of study is wireless and mobile communications.

Md. Imdadul Islam has completed his B.Sc. and M.Sc Engineering in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology, Dhaka, Bangladesh in 1993 and 1998 respectively and has completed his Ph.D degree from the Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh in the field of network traffic engineering in 2010.

He is now working as a Professor at the Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh. Previously, he worked as an Assistant Engineer in Sheba Telecom (Pvt.) LTD (A joint venture company between Bangladesh and Malaysia, for Mobile cellular and WLL), from Sept.1994 to July 1996. Dr Islam has a very good field experience in installation of Radio Base Stations and Switching Centers for WLL. His research field is network traffic, wireless communications, wavelet transform, OFDMA, WCDMA, adaptive filter theory, ANFIS and array antenna systems. He has more than hundred research papers in national and international journals and conference proceedings.

M. R. Amin received his B.S. and M.S. degrees in Physics from Jahangirnagar University, Dhaka, Bangladesh in 1984 and 1986 respectively and his Ph.D. degree in Plasma Physics from the University of St. Andrews, U. K. in 1990. He is currently working as a Professor of Electronics and Communications Engineering at East West University, Dhaka, Bangladesh. He served as a Post-Doctoral Research Associate in Electrical Engineering at the University of Alberta, Canada, during 1991-1993. He was an Alexander von Humboldt Research Fellow at the Max-Planck Institute for Extraterrestrial Physics at Garching/Munich, Germany during 1997-1999.

Dr. Amin was awarded the Commonwealth Postdoctoral Fellowship in 1997. Besides these, he has also received several awards for his research, including the Bangladesh Academy of Science Young Scientist Award for the year 1996 and the University Grants Commission Young Scientist Award for 1996. His current research fields are wireless communications and networks and also nonlinear plasma dynamics. He is a member of the IEEE.

Segmentation on the Dental Periapical X-Ray Images for Osteoporosis Screening

Enny Itje Sela
Department of Informatics
STMIK AKAKOM
Yogyakarta, Indonesia

Sri Hartati[1], Agus Harjoko[1], Retantyo Wardoyo[1],
Munakhir MS[2]
[1]Department of Computer Science and Elecronics
[2]Department of DentomaxillofacialRadiology
UniversitasGadjahMada
Yogyakarta, Indonesia

***Abstract*—Segmentation on the trabecular of dental periapicalX-Ray images is very important for osteoporosis screening. Existing methods do not perform well in segmenting the trabecular of dental periapical in X-Ray images due to the presence of large amount of spurious edges. This paper presents a combination of tophat-bothat filtering, histogram equalization contrasting and local adaptive thresholding approach for automatic segmentation of dental periapical in X-Ray images. The qualitative evaluation is done by a dentist and shows that the proposed segmentation algorithm performed well the porous of trabecular features of dental periapical. The quantitative evaluation used fuzzy classification based on neural network to classify these features. It were found accuracy rate to be 99,96% for training set and around 65% for testing set for a dataset of 60 subjects.**

Keywords—dental periapical X-Ray; osteoporosis; porous trabeculae; segmentation.

I. INTRODUCTION

Computer Aided Diagnosis (CAD) has been showing greater significance for bringing such effective and voluminous number of medical images possible [1]. Computer algorithms play a major part in extracting data from a medical image such as facilitating and automating the delineation of anatomical structures, identifying bone cracks and various other biomedical applications [2]. One major of such techniques lie in the realm of X-Ray image analysis.

Radiograph is the oldest form of non-invasive, painless, economical and easily observable digital imaging technique and it is widely used during various stages of treatment. Bone structure analysis is a technique that requires details that can only be obtained as features from an X-Ray image. A method used to extract such information is called segmentation which involves the grouping or classification of pixels in an image into groups containing common characteristics [2].

Segmentation of bones in medical images is very important for medical applications such as fracture detection. Segmentation of bones in X-ray images is a very difficult and challenging task that is not well understood [3]. In fact, the segmentation alone can determine the eventual success or failure of the analysis at hand. Segmentation involves working on a number of images processing tasks such as noise removal and image enhancement. In an anterior-posterior periapicalX-Ray, the image is too noisy, too blur and too dark. As a consequence, some boundary edges of trabecular and its porous may appear to be connected each other. These difficulties cause general segmentation methods inappropriate.

Combination of the tophat filtering, contrast stretching, and Otsu thresholding methods [16] was fail to segment our periapicalX-Ray images. Other methods such as combination of the tophat filtering, contrast stretching, wiener noise removal, color quantization and spatial segmentation [4], combination of a 3x3 median filtering, binarization, and erosion [5], combination of the Gausian filtering, subtraction the original image with the image of the filtering, then Otsu threshold method [6] were tried to segment periapical X-ray images and result unsuccessfully segmentation.

Another method has been proposed to perform for segmentingperiapical X-ray images. The segmented image consists of the trabecular and it's porous. The porous then extracted as features of osteoporosis disease. These features can be used to predict a subject's femoral/lumbar World Health Organization (WHO) bone mineral density. The WHO classifiers the subjects as follows: osteoporotic, osteopenia, and normal [7].

II. MATERIALS

A. Subjects

Ethical clearance has been obtained from the local ethics committeeof Faculty of Dentistry, UniversitasGadjahMada, Yogyakarta for this study. The informed consent obtained from all subjects. The subjects consisted of 60 consenting females (average age 61.3 years;and ranges 41-80). Each subject was collectedsubject's age, height, weight and BMI at the time of radiographic examination.None of the subjects was known to have endocrine, metabolic, or skeletal disorder. None of the subjects was on hormonal replacement therapy or taking calcitonin, bisphosphonates, or fluorides except of low doses of calcium or vitamin D.

B. Dental Periapical X-Ray

All periapicalX-Ray carried out from the Department of Radiology of Prof. SoedomoDental Hospital, Faculty of Dentistry UniversitasGadjahMada (Fig 1). They were taken by a radiographer using dental X-ray Villa SISTEMI medicaliendos ACP CEI specification 70 kVp, 8 mA, and 3.2s and the image receptor photostimulable phosphor plate (PSP).

Periapical X-Ray processing used digital radiography (DBSWin 4.5, Durr Dental). The size of this bitmap image is 1252 x 1645 pixels.

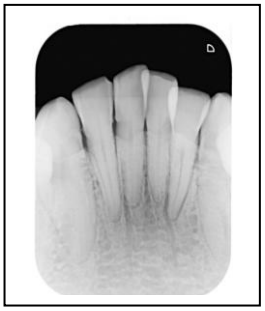


Fig.1. Dental Periapical X-Ray Image

C. BMD

Assessment of Bone Mineral Density(BMD) on femoral neck and lumbar spine carried out from the Department of Radiology Dr. Sardjito Hospital using densitometer Dual energy X-ray absorptiometry (DXA) specifications 76kV, 1.5mA, during 1 min 14s (femoral), and 1 min 27s (spine). The procedure is operated by a radiographer. Subjects were classified into one of three groups contained women who were classified according to the WHO classification.

III. METHODOLOGY AND DESIGN

A. Methodology

In this study, there are some steps to obtain the segmented images. First step is selection ROIs from originally images. Second step is segmentation process. Last step is evaluation process. All of the steps can be shown on Figure 2.

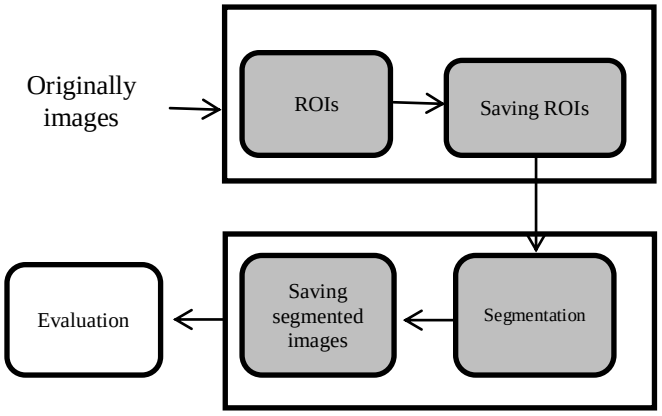


Fig.2. Research Methodology

Selection of ROIs.

All of ROIswere selected around the trabeculararea for each patient, as decided independently by an observer. To obtain the ROIs, a dentist should make a point on the trabecular area. Then the system makes rectangle automatically. The maximum size of all rectangles was 400 x300 pixels, with a pixel size of 0.02 mm. All of ROIs are saved in bitmap format.

Segmentation.

The segmentation process is used to separate the trabecular and porous objects all ROIs resulted from previous step. Trabecular are presented in the white pixel and porous presented in black pixel. Data X-Ray were scanned at 600 dpi, made uniform in overall intensity by blurring the image by applying a tophat bothat filtering. Top hat filtering and bottom hat filtering can be used together to enhance contrast in an image. Contrasting by histogram equalization[8]was used to distribute the intensity of the pixels in the interval [0 .. 255], and last, adaptive threshold [10, 11] with a window of 40x40 pixels. Thisprocess was resulted a binary image.

Segmentation procedureis as follows [8] :

Input: bitmap format images, gray scale, 8 bit (img)
Output: a binary image (outimage)

1) Filtering

Tophat performs morphological top-hat filtering on the grayscale or binary input image using the structuring element. Steps of tophat filtering are:

a. Erosion.

$$I_2 = I_1 \ \Theta \ B \text{.....(1)}$$

I_1 = initial image, I_2 = image erosion
B = structure element disk (r=1)

Steps to perform erosion are:

a) Comparing each pixel with the color of the pixel center I_1 B by superimposing B with I_1 so that the center of B is exactly same with the image pixel positions I_1.

b) If all the pixels in B exactly the same as all the pixels in I_1 (foreground) pixel I_1 then set its value to the foreground pixel value, otherwise the value of the pixel value will be set according to the value of the background.

b. Dilation on I2.

$$I_3 = I_2 \ \oplus \ B \text{ .....(2)}$$

I_3 = dilation result.

Steps to perform dilation are

c) Comparing each pixel with the color of the pixel center I2 B by superimposing B with I2 so that the center of B is exactly same with the image pixel positions I2.

d) If at least one pixel to pixel B is equal to the value of the object (foreground) in I2 then the pixel value is set to the foreground pixel value and if all the pixels that are related to the background pixel value set in I2 are like background pixel value.

c. Subtract.

$$I_4 = I_1 - I_3 \quad \text{.....(3)}$$

Bothat performs morphological bottom-hat filtering on the grayscale or binary input image. This process is the reverse of the tophat. The steps for this process are:
a. Dilation on I_1

$$I_5 = I_1 \oplus B \quad \text{.....(4)}$$

b. The process of erosion on I5

$$I_6 = I_5 \ \Theta \ B \quad \text{.....(5)}$$

c. Subtract I_1 to I_6

$$I_7 = I_1 - I_6 \quad \text{.....(6)}$$

1. Contrasting.
   This step is used to distribute the intensity of the pixels in the interval [0..255]. The equation to perform histogram equalization can be seen below.

$$n(g) = \max(0, round\left[(L-1 * \frac{c(g)}{N}\right] - 1) \quad \text{....(7)}$$

N = number of pixels in the image vectors
g = initial value of gray level value of (L-1)
L = maximum value of gray level
c (g) = number of pixels that have a value equal to g or less

$$c(g) = \sum_{i=1}^{g} h(i) \quad \text{.....(8)}$$

g = 1,2, ……., (L-1)
h (i) = initial histogram.

2. Adaptive segmentation thresholding with the average value[9][10]
   a. Set a constant value of N = 40
   b. Make a subimage window of K1 with size NxN on the image.
   c. Find the average pixel intensity subimage K1
   d. Perform the threshold with the average value of K1

```
if K1 (i, j) <= threshold
        outimage (i, j) = 1;
else
        outimage (i, j) = 0
```

IV. RESULT

A custom computer program was used to obtain morphologic variables from the digitized periapical X-Ray. The same ROIs used for the radiographic jaw density were used for this analysis. A dentist used our system with give a point on the trabecular using mouse (Figure 3). Then, the system made 400x300 pixel of ROI automatically.Figure 4 shows the result of ROI from our periapical dental X-Ray.

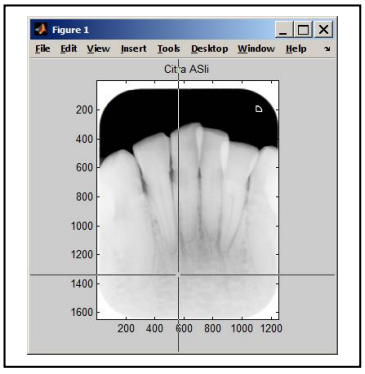


Fig.3. Input a point on the trabecular bone

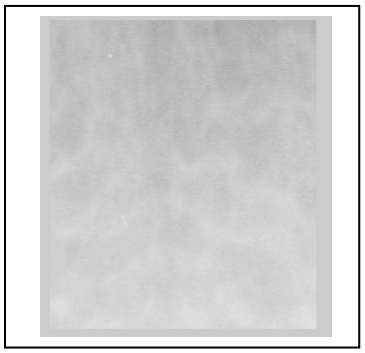


Fig.4. ROI of trabecular bone from digitized X-Ray

Based on the above segmentation procedure, the results can be presented as follows. Fig 5 shows the filtering image using tophat and bothat filtering of ROI. This image uses tophat and bothat filtering with a disk-shaped structuring element disk (r=1) to remove the uneven background illumination from an image and enhance contrast in the image.

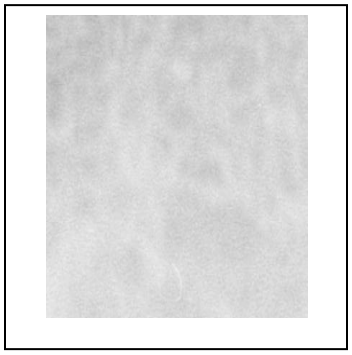


Fig.5. Result of tophat and bothat filtering of ROI

Fig 6 shows the result of adjusting image seen in Fig 5 using histogram equalization. This image can show the trabeculae (white pattern) and porous (black pattern) clearly.

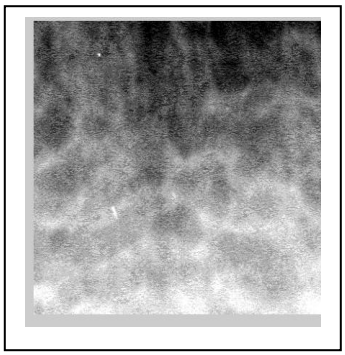


Fig.6. Contrasting image seen in Fig 5

Fig 7 shows the result of adaptive local threshold on adjusted image seen in Fig 6. The size of the resulting binary image is smaller than the original image depending on the size of the window. In this study, the size of the window is 40x40

pixels. This operation result a binary image with 360x260 pixels in size.

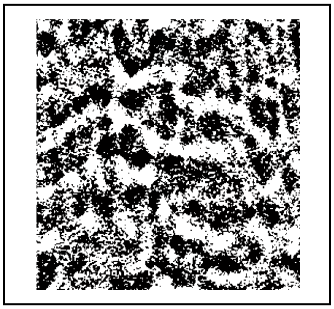


Fig.7. Segmented Image

There is a difficulty to evaluate the segmentation. A dentist then evaluate segmented image by comparing the segmented image with another segmented image resulted in previous research [6][15]. It is said that the binary image resulted from the segmentation is well. In this study, the quantitated evaluation is performed for classification. The binary image is analyzed to measure features including porosity[13],number of vertex of porous [11] and perimeter of porous [11] after morphology operation (clear border, filling, dilation, and erosion). The features are saved in text data format. Then thefuzzypattern classificationusingbackpropagationlearningalgorithm [14] is used to training and testing these features.

Samples of 60 acquired dataset are divided into training set (12 normal, 11 osteopenia, 7 osteoporotic) to generate weight of neural network and testing set (14 normal, 12 osteopenia, 4 osteoporotic) to test the capability of system to get the output class.The result of classification using multilayer backpropagation with 10^{-6} MSE (MSE goal was 6.90 x 10^{-7}), 650 epochs, 0.1 learning rate, and sigmoidal transfer function. The performance of neural network is shown at Fig8.

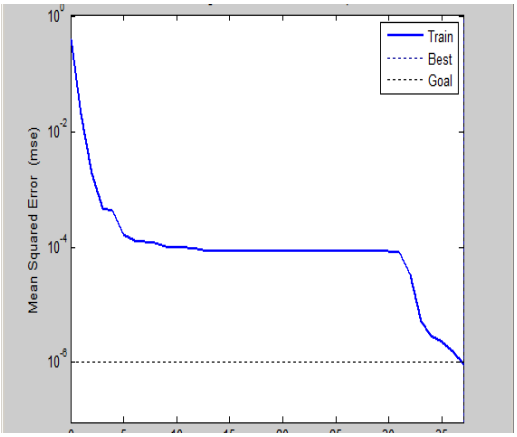


Fig.8. Performance Neural Network

The training results were found accuracy rate to be 99,96% and are shown as Table 1. Test results were found to be around 65% and shown in Table 2.

TABLE I. PERCENTAGE ACCURACY TRAINING RESULT

Class	Femur	Lumbar
Normal	99,96	99,83
Osteopenia	99,83	99,93
Osteoporotic	99,99	99,96

TABLE II. PERCENTAGES ACCURACY TESTING

Class	Femur	Lumbar
Normal	68,56	65,18
Osteopenia	68,25	62,49
Osteoporotic	61,32	60,27

There are some limitations in this study. The qualitative evaluation of segmentation needs another method for comparing. As a consequent, actually it could not be proven that the segmented images are robust. On the quantitated evaluation, since the number of normal data is quite large compared to osteopenia and osteoporotic data, and thenthis uneven distribution could lead to accuracy result.More participant are needed to get more data for the future study, and last;more ROI can be considered in an image.

V. CONCLUSION

The combination of tophat-bothat filtering, histogram equalization contrasting, and local adaptive thresholding segmentation method can be performed to clearly separate porous trabecular bone in periapical dental X-Ray images. Porous trabecular analysis such as porosity, perimeter of porous, and number of vertex of porous is valuable and promising areas in osteoporosis screening. The experimental resultsshows that the features of porous trabecular bone were used for osteoporosis screening with the classification accuracy around 65%. This result suggests evaluating the segmented image using qualitative method. Improving accuracy rate could be done using different methods.

ACKNOWLEDGMENT

We thank to DIKTI, The Ministry of Education and Culture of Republic Indonesia for BPPS scholarship, the Prof. Soedomo Dental Hospital, Faculty of Dentistry, Universitas Gadjah Mada and the Dr. Sardjito Hospital, Yogyakarta.for data supporting.

REFERENCES

[1] M. Park, B. Kang, andS.J. Jin, “Computer aided diagnosis system of medical images using incremental learning method, Intl Journal Expert Systems with Applications”, 3, 36, Pergamon Press, Inc. Tarrytown, NY, 2009.

[2] A.A. Tirodkar, “A multi-stage algorithm for enhance X-Ray image segmentation”, International Journal of Engineering Science Technology (IJESE), vol. 3, no. 9, September 2011.

[3] F. L.Ding, K. Wee, T. Howe, “Automatic segmentation of femur bones in anterior-posterior pelvis X-Ray images”, CAIP 2007: Springer-Verlag, Berlin, 2007.

[4] D. Herumurti, A.Z. Arifin, R. Sulaeman, A. Asano, A. Taguchi, T. Nakamoto, and K. Uchimura, “Weighted fuzzy ARTMAP for osteoporosis detection”, Proceeding 16th Korea-Japan Joint Workshop on Frontiers of Computer Vision, Seoul, 2007.

[5] W.G.M. Geraets, J.G.C. Verheij, P.F. Van der Stelt, K. Horner, C. Lindh, K. Nicopoulou-Karayianni, R. Jacobs, E.J. Marjonovic, J.E. Adams, and H. Devlin, "Selecting regions of interest on intraoral radiographs for predictions of bone mineral density", DentomaxillofacRadiol 37, pp. 375-379, 2008.

[6] B.D. Lee, and S.C. White, "Age and trabecular features of alveolar bone associated with osteoporosis", Oral Surg Oral Pathol Oral RadiolEndod , Elsevier Inc, 2005.

[7] R. Silvakumar, "Identification of early osteoporosis using intensity slicing method", Global Journal of Researches in Engineering General Engineering, vol. 12, issue. 2, version. 1.0, May 2012.

[8] R. Gonzalez, and R.E. Wood, Digital imageprocessing 2nd edition, Prentice-Hall Inc, USA, 2007.

[9] N. Milstein, "Image segmentation by adaptive threshoding", Technion-Israel Institute of Technology, 1998.

[10] J. Sauvola, and M. PietikaKinen, "Adaptive image binarization", Pattern Recognition,no. 33, pp. 225-236, 2000.

[11] H.J. Sommer, "Area, centroid and area moments for polygonal objects",www.me.psu.edu/ **sommer**/me481/notes_07_02.doc, accessed 10 Feb 2013.

[12] S.K. Pal, and S. Mitra, A review of: neuro-fuzzy pattern recognition methods in soft computing, John Wiley & Sons, New York, 1999.

[13] A.M. Badawi, "Osteoporosis classification using fuzzy rule based and neural networks", accessed April 2011.

[14] P. Augat and S. Schorlemmer, "The role of cortical bone and its microstructure in bone strength", Age and Ageing, 35-S2: ii27-ii32, Oxford University Press, 2006.

[15] S.C. White, and D.J. Rudolph, "Alterations of the trabecular pattern of the jaws in patients with osteoporosis", Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 88, pp. 628-635, 1999.

[16] Z. Abidin, andA. Z. Arifin, "Analisis kerapatan trabecular bone berbasis graph berbobot pada citra panorama gigi untuk identifikasi osteoporosis", JUTI, vol. 2, no. 7, pp. 59-64, Juli 2009.

A Hybrid Approach for Co-Channel Speech Segregation based on CASA, HMM Multipitch Tracking, and Medium Frame Harmonic Model

Ashraf M. Mohy Eldin
Department of Computer Engineering
Arab Academy for Science, Technology and Maritime Transport
Cairo, Egypt

Aliaa A. A. Youssif
Department of Computer Science
Faculty of Computers and Information
Helwan University
Cairo, Egypt

***Abstract*—This paper proposes a hybrid approach for co-channel speech segregation. HMM (hidden Markov model) is used to track the pitches of 2 talkers. The resulting pitch tracks are then enriched with the prominent pitch. The enriched tracks are correctly grouped using pitch continuity. Medium frame harmonics are used to extract the second pitch for frames with only one pitch deduced using the previous steps. Finally, the pitch tracks are input to CASA (computational auditory scene analysis) to segregate the mixed speech. The center frequency range of the gamma tone filter banks is maximized to reduce the overlap between the channels filtered for better segregation. Experiments were conducted using this hybrid approach on the speech separation challenge database and compared to the single (non-hybrid) approaches, i.e. signal processing and CASA. Results show that using the hybrid approach outperforms the single approaches.**

***Keywords*—CASA (computational auditory scene analysis); co-channel speech segregation; HMM (hidden Markov model) tracking; hybrid speech segregation approach; medium frame harmonic model; multipitch tracking, prominent pitch.**

I. INTRODUCTION

In everyday life, speech doesn't arrive to our ears in a clean way, but rather corrupted by various types of noise including speech of other competing talkers in what is known as cocktail party effect. Human auditory system is remarkably capable of focusing on the target speech and separating it from noise.

On the contrary, artificial speech processing systems are designed to deal with clean, noise free speech. These systems need a front end component that segregates the target speech from other interferences. Competing speech is the most difficult kind of interference because of the similarity of temporal and spectral characteristics between target and interfering speeches. Work on speech segregation dates back to 70s [1].

The complexity of the speech segregation problem is related to the number of speakers and channels (i.e. microphones) used to record the speeches. The most difficult situation is when only one channel is used, i.e., co-channel speech segregation, as all spatial cues used in segregation are lost, e.g., inter-aural time, phase, and level differences.

Many approaches have been investigated to solve the co-channel speech segregation problem. The earliest approachesare the general signal processing approaches. R. H. Frazier [3] used adaptive comb filter with frequency spacing of pass bands varying with the fundamental frequency of the speech. T. F. Quatieri and R. G. Danisewicz [4] used the sinusoidal model of speech which assumes that speech consists of a sum of sin waves with varying amplitudes, frequencies, and phases over time. They used a minimum mean-squared error estimation combined with the sinusoidal model. D. S. Benincasa and M. I. Savic [5] used a technique to separate the co-channel mixed speech of 2 talkers by using constrained nonlinear optimization to separate overlapping voiced speech.

Another category of approaches is the computational auditory scene analysis (CASA). G. J. Brown and D. L. Wang [7] explained how CASA could be used in speech segregation showing how both monaural and binaural cues could be used for co and multichannels speech segregation. They also explained how to integrate CASA with speech recognition. L. Ottaviani and D. Rocchesso [9] proposed a system with 2 stages, pitch analysis using enhanced summary autocorrelation function (ESACF) and signal re-synthesis using highly zero-padded Fourier transform and its inverse. P. Li *et al.* [18] used objective quality assessment of speech (OQAS) combined with CASA. They used OQAS as a guide to lead CASA grouping. X. Zhang *et al.* [19] introduced the new concept of dynamic harmonic function (DHF) and replaced the conventional autocorrelation function (ACF) with DHF to suppress invalid peaks. Blind source separation (BSS) is a statistical approach that tries to recover a set of original signals from observed mixtures by assuming the linearity of the mixing process. A standard approach of BSS is independent component analysis (ICA) [10], which assumes the statistical independency of all sources. ICA needs 2 conditions to be satisfied to solve the speech segregation problem, namely, there must be a number of observed mixtures equal to or greater than the number of source signals, and all source signals must be perfectly aligned [2].

Obviously, these 2 conditions are not met in co-channel speech segregation problem. To solve this problem, a so called underdetermined BSS was invented [11]. In this technique, a priori knowledge obtained through training must be available. An example of the use of underdetermined BSS can be found in [12], in which case the priori knowledge was a set of time domain basis functions learned in a training phase. A comparison between CASA and BSS can be found in [13]. The

comparison yields that CASA is more suitable to natural situations as it does not need a lot of conditions required by BSS. However, in the presence of these conditions, BSS may outperform CASA. This may suggest the join of the 2 approaches in a hybrid one.

Model based approaches could be used to solve the problem of co-channel speech segregation. These approaches consist of 3 steps. First, training phase is used to obtain patterns of sources. Second, patterns whose combinations model the observed signal are chosen. Third, selected patterns are used to estimate the sources directly or used to build filters to get the sources from the filtered observed signals. A. M. Reddy and B. Raj [14] used a model with minimum mean squared error estimator for co-channel speech segregation. H. A. T. Kristjansson and J. Hershey [15] used the male and female speech fine structure and the source signals strong high frequency resolution model. S. T. Roweis [16] used a simple factorial hidden Markov model (HMM) system which is trained on recordings of single speakers and then uses the co-channel observed signal to separate the mixture by calculating the masking function and re-filtering. The masking function is simply a non-stationary reweighting of the individual speakers' sub bands. D. E. M. J. Reyes-Gomez and N. Jojic [17] broke the mixed speech signal into multiple frequency bands. For each individual band they built separate HMM. Those separate HMMs are coupled together to model the whole mixed speech.

Hybrid approaches try to benefit the advantages of 2 or more different approaches to get better segregation results. J. Ming *et al.* [25] combined a missing feature technique to improve the robustness against crosstalk and noise with: Wiener filter to enhance the speech, hidden Markov model to reconstruct the speech, and speaker dependent/independent modeling to recognize both the speaker and the speech. P. Li *et al.* [26] combined Gaussian mixture models (GMMs) and max vector quantizers (MAXVQ) with CASA to separate co-channel mixed speech. Pitch is considered the most important cue in CASA as it is used for both simultaneous and sequential grouping. Accordingly, multipitch determination algorithms (MPDAs) were the subject of many researches. MPDAs may be either: time-domain, frequency domain, or time-frequency-domain. Time-domain MPDAs depend on the speech waveform temporal characteristics, e.g., autocorrelations. Frequency-domain MPDAs uses the short term spectrum to detect the fundamental frequency. Time-frequency-domain MPDAs obtain the signals using multichannel front end, then band-filter these signals, and finally perform time-domain analysis [2].

An example of time domain MPDAs is presented byA. de Cheveigne[22]. He extended the average magnitude difference function (AMDF) in a two-dimensional way by cancelling one of the 2 speakers and estimating the pitch of the other one. An example of frequency domain MPDAs is the one suggested by F. Sha and L. K. Saul [23]. They proposed an approach with instantaneous frequency estimation as front end and nonnegative matrix factorization as back end. An example of time frequency MPDAs is the one introduced by M. Wu *et al.* [24]. In their approach, periodicity information is extracted across different frequency channels, and the pitch tracks are formed using HMM.

The proposed approach combines a model based MPDA with time and frequency one for multipitch tracking. The enhanced pitch tracks are used as cues for enhanced CASA segregation.

II. Proposed ApproachOverview

The proposed approach consists of 2 stages of enhancements, pitch tracking enhancement and CASA segregation enhancement, Fig. 1. The 1st stage consists of 4 steps, HMM multipitch tracking, prominent pitch enrichment, grouping based on pitch continuity, and extracting the 2nd pitch (for frames with only one pitch deduced) using medium frame harmonics.

The chosen algorithm for multipitch tracking is the one developed by Z. Jin and D. L. Wang [21]. The reason behind choosing this algorithm is the ability to use it for different speech corpuses without the need to be trained on the new ones. This makes it a general algorithm and more preferable than other algorithms. Although it was developed taking reverberation into account, it is usable for normal conditions.

Medium frame harmonic extraction is inspired by the work of Q. Huang and D. Wang [6]. They used both short and long frames for pitch state deduction and pitch calculation. However, in the proposed approach, pitch hypothesis from the MPDA of Z. Jin and D. L. Wang [21] besides the concept of pitch continuity are used to judge the pitch state. So, only one type of frames is used and this type was neither short nor long but medium to enhance the resolution of the Fourier analysis without compromising neither the stationary assumption needed for Fourier transform nor the calculation speed.

For the segregation stage CASA algorithm proposed by G. Hu and D. L. Wang [8] is used with some changes. The center frequency range of the gamma tone filter banks was maximized to reduce the overlap between the channels filtered for better segregation. Also a new mechanism for time frequency (T-F) unit labeling that depends on the pitch tracks of both talkers was used.

III. HMM Multipitch Tracking

The list describes the changes from the base algorithm:

- The base algorithm assumed that the sampling frequency of the mixed speech is 16 KHz. For this paper, the sampling frequency of the test database is 25 KHz. This means that the typical pitch range for both male and female from 80 to 500 Hz or time lags from 2 to 12.5 msnow corresponds to 50 to 313 samples, i.e. pitch state space which consists of the union of 0, 1, and 2 pitch hypothesis, since this algorithm tracks up to 2 pitches simultaneously,is described by (1).

$$S = S_0 Y S_1 Y S_2 \quad (1)$$

Where $S_0 = \{\phi\}$, $S_1 = \{\{\tau_1\} : \tau_1 \in [50,313]\}$, and $S_2 = \{\{\tau_1,\tau_2\} : \tau_1,\tau_2 \in [50,313], \tau_1 \neq \tau_2\}$. ϕ indicates the absence of a pitch, and τ_1 and τ_2 represent the time lags of the 2 pitches in sample points.

- Z. Jin and D. L. Wang mentioned that broadband noise distorts the spectral peaks of the speech and causes

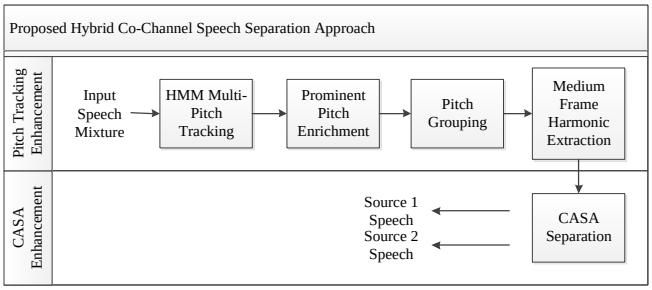


Fig. 1 Proposed Hybrid Approach

HMM search to be biased towards S_2. To overcome this, they performed two independent Viterbi searches. The 1st assumes the presence of one pitch in the frame at maximum. The 2nd performs the search normally trying to find up to 2 pitches in the frame. They used linear discriminant analysis (LDA) to decide which one of the 2 searches is correct for the frame in question.

Since the database used to test the proposed approach is noise free, this step is removed and the search is always performed for up to 2 pitches in the frame.

This enhanced the results, although slightly, however it speeded up the whole process in a good way.

IV. Prominent Pitch Enrichment

HMM multipitch tracks are enriched with the prominent pitch calculated using summary autocorrelation across all channels. For each frame, if zero pitch is deduced by the HMM multipitch tracking, i.e. zero pitch hypothesis, S_0, this hypothesis is unconditionally overridden to one pitch hypothesis, S_1 and the pitch is simply the prominent pitch. If one pitch is deduced by the HMM multipitch tracking, i.e. one pitch hypothesis, S_1, if the prominent pitch is far from the pitch deduced by more than a threshold, 32, the one pitch hypothesis is overridden to two pitch hypothesis, S_2, and the second pitch is simply the prominent pitch.

V. Pitch Grouping

Pitches are assigned to the proper source, 1 and 2, based on pitch continuity as follows:

- The previous frame in the following paragraph is the first frame preceding the current one with more than 0 pitches deduced.
- If the current and previous frames both have 2 pitches deduced, then the pitches of the current frame are assigned to the tracks that achieves the minimum distance between corresponding pitch lags.
- If the current frame has 2 pitches and the previous frame has only one pitch deduced, then the current frame pitch nearest to the previous frame one is put in the same track and the other current frame pitch is put in the opposite track.
- If the current frame has one pitch and the previous frame has 2 pitches deduced, then the current frame pitch is put in the same track of the nearest pitch of the previous frame.
- If the current and previous frame has only one pitch deduced, then if the distance between the 2 pitch lags is within a threshold, 32, then the current frame pitch is put in the same track as the previous frame pitch, otherwise, in the opposite track.

VI. Medium Frame Harmonic Extraction

The following steps are used to get the other source's pitch, for frames with only 1 pitch obtained:

- Previous pitch, F_{prev} for the track that needs to estimate the current pitch is obtained by iterating previous frames the same way done in pitch grouping.
- Also, next pitch, F_{next} is obtained by iterating next frames.
- Fourier transform is obtained for the current frame with medium frame length, 50 ms that is sufficient to get good resolution of the harmonics compared to short frame of 30 ms while maintaining less complexity of calculation compared to long frame of 90 ms.
- All Fourier components after a threshold of 4000 Hz are removed. This is because the main energy of voiced speech is concentrated in the low frequency.
- The remaining Fourier components are divided into bands of 200 Hz. For each band peaks that are not less in magnitude than 1/5 of the highest peak in the band are obtained. They will form a vector of chosen harmonics for both pitches of the 2 sources, F_{vec}.
- From F_{vec}, all the harmonics (multiple integers and also half, quarter, and 1/8) that belongs to the pitch already known for that frame, F, including the pitch itself, are removed. Now, the vector contains only candidates of the other pitch and their harmonics.
- Candidate pitches, F_{cand} are those that exist in Fvec and are not far from F_{prev} or F_{next} by 8 Hz.
- For each pitch in F_{cand}, the harmonic order (number of harmonics that exist in F_{vec}) and the average frequency deviation of those harmonics from the ideal ones (multiple integers of the candidate pitch) is calculated. Only those pitches with harmonic order that is not less than 9/10 of the maximum harmonic order are chosen. The pitch with minimum frequency deviation from the chosen ones is simply the other pitch for that frame.

VII. CASA Segregation

The list describes the changes from the base algorithm:

- Before segregation, pitch tracks need to be refined from pitches suspected to be error. This enhances the segregation as leaving a T-F unit without assigning to a source (to be assigned later based on grouping) is better than assigning it to the wrong source. Pitches in tracks less than 5 contiguous frames are considered suspected and removed.
- In the peripheral analysis, input signal is passed through a bank of 128 gamma tone filters centered

from 21 to 12500 Hz instead of from 80 to 5000 Hz. The frequency range is maximized to reduce the overlapping between gamma-tone filtered channels for better segregation.

- Unit Labeling:

Both pitch tracks are used to label each T-F unit as either belonging to source 1 or 2 instead of just using the target pitch track to label the T-F unit as either belonging to the target source or interference. The following points worth mentioning:

- Error in segregation is directly related to error in pitch tracking.
- When using only the pitch track of the target source, error in separating target source is directly related to error in the pitch track of the target source.
- When using pitch tracks of both users, error in segregation of both sources is directly related to the average error of both pitch tracks.
- Accordingly, if the interest is to only separate one source (target source), if the target pitch track error is less than the average error, then using only target pitch track in unit labeling is better than using both pitch tracks. Otherwise, using both pitch tracks is better.
- However, if the interest is to separate both sources, it is always better to use both pitch tracks.

Features used in unit labeling are correlogramA_H (2) and envelope correlogramA_E(3).

$$A_H(c,m,\tau)=\frac{\sum_n h(c,mT_f-nT_s)h(c,mT_f-nT_s-\tau T_s)}{\sum_n h^2(c,mT_f-nT_s)} \quad (2)$$

$$A_E(c,m,\tau)=\frac{\sum_n h_E(c,mT_f-nT_s)h_E(c,mT_f-nT_s-\tau T_s)}{\sum_n h_E{}^2(c,mT_f-nT_s)} \quad (3)$$

Where c is the channel number, m is the frame number, τ is the time delay at which autocorrelation A is calculated, $n$is the digitized time, T_f is the time shift from one frame to the next (10 ms) and T_s is the sampling time, and h is the channels response transduced by Meddis model.

Unit labeling is done using the following steps:

- If the current frame has 0 pitch deduced, then all channels T-F units are not labeled.
- If the current frame has only 1 pitch deduced, if AH/max(AH)>0.85 if the unit is marked as 1, or AE/max(AE)>0.7 if the unit is marked as 2, the unit is labeled to the same track of the pitch, otherwise, it is labeled to the opposite track.
- If the current frame has 2 pitches deduced, the previous ratios are calculated for both pitches. The unit is labeled to the track of the pitch that satisfies the condition. However, if neither or both pitches satisfy the condition, the unit is not labeled at all.

VIII. Experimental Results

A. Database

A different database was chosen to test the proposed approach rather than using the database used to test the base algorithms.

This is to make sure that these base algorithms (as was mentioned by their authors) are easily generalized to new speech corpuses with no training needs. The speech separation challenge [27] was particularly chosen because it is considered one of the most complex databases used in co-channel speech segregation problem. This is because it uses a small vocabulary set which leads to close similarity of the speeches of the competing talkers which makes the pitch tracking and segregation more difficult. In order to get accurate results, the whole database, a total of 900 mixed speech samples, was used rather than selecting a subset. The test was conducted on 0dB target to masker ratio (TMR) which is considered the most difficult situation in co-channel speech segregation problem as both talkers equally masks each other.

B. Pitch Tracking Enhancements

The pitch tracking enhancement stage will be compared with the base HMM multipitch tracking algorithm [21]. The comparison measures will be the gross error, E_{gs}, and the fine error, E_{fn}. E_{gs} is the percentage of frames where the deduced pitch differs from the ground truth pitch by more than 20%. E_{fn} is the average deviation from the ground truth pitch for the frames with no gross error. Ground truth pitches are calculated using summary autocorrelation of the frames of the original speeches of each source before mixing. This is not error free, some ground truth pitches will not be correct. However, the same error will be added (approximately) to both the base approach and the proposed one and will not affect the comparison. Since this paper is interested in separating the speech of both sources, the sum of gross errors of both pitch tracks and the sum of fine errors using the enhanced proposed pitch tracking stage will be compared with the base algorithm. Table. I shows the pitch tracking results.

C. Whole Segregation Approach Enhancements

The proposed approach will be compared to: the one proposed by G. Hu and D. L. Wang [8], the one proposed by Z. Jin and D. L. Wang [20], and the traditional harmonic selection approach. The following points need to be stated:

- The multi pitch tracker proposed by M. Wu *et al.* [24] and used in [8] is replaced by the base pitch tracker [21] since [24] needs to be trained on the speech separation challenge and the target of this paper is to enhance algorithms that need no training.
- The same suggested unit labeling step that uses both pitch tracks will be used for the base algorithm [8].
- The previous 2 points actually enhances the algorithm [8]. If the comparison of the proposed approach was done with [8] without these 2 points, much better

TABLE 1. PITCH TRACKING RESULTS

		Average E_{gs} %		Average E_{fn} %	
		Base Approach	Proposed Approach	Base Approach	Proposed Approach
Different Gender		59	51.9	4	4.9
Same Gender	Male	52.7	46.7	4.8	5.5
	Female	68.8	59.2	5	5.8
Same Talker	Male	55.6	48.1	5.4	5.6
	Female	69.1	59.1	5.1	5.5

results are obtained. However, this is done to get more accurate comparison results.

- Since the test database contains only speech interference, the speech model of [20] is used only for faster calculations. The binary masks for both talkers are used to label the T-F units as opposed to [20] which only uses the target talker's mask. T-F unit is labeled to talker 1 if mask 1 = 1 and mask 2 = 0. Similarly, T-F unit is labeled to talker 2 if mask 1 = 0 and mask 2 = 1. In the grouping stage, only T-F units labeled were taken into consideration.

- Harmonic selection is included in the comparison to show that CASA approaches outperform the traditional signal processing approaches. The pitch tracks obtained using the base pitch tracker [21] was used to select the harmonics of both talkers (the proposed enhanced pitch tracker stage could also be used for comparison and the same conclusion of the superiority of CASA approaches to the signal processing ones would be reached). If 2 pitches are tracked for a frame, those 2 pitches are used to select the harmonics of both talkers. If only one pitch is tracked, it is used to select the harmonics of the respected talker. The other talker's signal is the difference between the mixed signal and the first talker's one. If no pitch is tracked, the choice is to assign the frame to the first talker.

The comparison measure is the signal to noise ratio (SNR). CASA segregation gives some amplification (though not uniform) to the segregated speech. In order for the SNR to be accurate, the original speech had to be compensated for such amplification. This was accomplished by applying all the stages of CASA to the original speech except using all 1s mask for unit labeling. Original speech was left as it is for SNR calculation of harmonic selection. Since this paper is interested in both talkers' speeches, the average SNR of both talkers will be compared. Table. II shows the segregation results.

IX. ANALYSISAND DISCUSSION

A. Pitch Tracking Enhancements

The proposed pitch tracking enhancement stage shows better results than the base approach. Enhancements range from 6 to 10% decrease in E_{gs} with overall average of about 8%. As expected, the decrease in E_{gs} results in increase in E_{fn} since more pitches now are taken into account while calculating E_{fn} with their deviation from the ground truth pitch added. However, the increase in E_{fn} is in the range of 0.2 to 0.9% with overall average of 0.6%. This is a small increase in E_{fn} which suggests that the pitches obtained from the enhancement stage are very close to the ground truth pitches with small deviation.

The following points are worth mentioning:

- E_{gs}, whether for base and proposed approaches, is very high. This is due to the nature of the speech separation challenge database as explained in sec. VIII-A. It is expected to have better (less) E_{gs} for normal speech conditions with less similarity between competing talkers speeches.

- The base algorithm has a problem in tracking the female pitches. This is suggested from the higher E_{gs} for female same gender and same talker mixtures than corresponding male mixtures. E_{gs} for same gender female mixture is higher than male mixture by 16.1%.For same talker, it is higher by 13.5%. The proposed approach enhanced same gender female pitch tracking by 9.6% as opposed to 6% for male. Also, for same talker, the proposed approach enhanced female tracking by 10% as opposed to 7.5% for male. This means that the proposed approach could deduce some of the female pitches missed by the base approach, mostly in the prominent pitch enrichment step.

B. Whole Segregation Approach Enhancements

The proposed approach shows better segregation results than the compared approaches. Enhancements range from 3.1% for the different gender case to 27% for same talker female case. The following points are worth mentioning:

- Enhancements are better when the case is worse. This is apparent from the fact that less enhancements happened in different gender case whereas best ones happened in same talker case. Also, enhancements for male-male are less than enhancements for female-female cases. The following points will try to explain the reason behind this.

- Different gender case exhibits minor enhancements. This is due to that the proposed approach depends on enhancing the pitch tracking and minimizing the T-F units overlapping between the 2 talkers. For the different gender case, the overlapping was originally small because female T-F units tend to be in the higher frequency channels whereas male T-F units tend to be in the lower frequency channels.

- For same gender and same talker cases, there was a reasonable amount of overlapping between the T-F units of both talkers as their T-F units tend to occupy the same range of frequencies. This means that the step of minimizing the overlap made better enhancements in these cases than the different gender case.

- This also means that the step of pitch tracking enhancements made fewer enhancements than the step of overlap minimization. This suggests that more enhancements in pitch tracking are needed.

TABLE II. SEGREGATION RESULTS

		SNR dB			
		G. Hu and D. L. Wang [8]	Proposed Approach	Harmonic Selection	Z. Jin and D. L. Wang [20]
Different Gender		3.2	3.35	0.83	3.25
Same Gender	Male	2.44	2.83	0.6	2.43
	Female	2.2	2.68	1.02	2.21
Same Talker	Male	1.87	2.31	0.53	1.86
	Female	1.74	2.16	0.78	1.7

- The minor differences in enhancements between same gender and same talker cases and also between male and female in each case are due to the minor better enhancements in pitch tracking for these cases.

X. CONCLUSIONAND FUTURE WORK

This paper proposed enhanced co-channel speech segregation approach. It also proposed enhanced MPDA that could be used on its own for multiple purposes. More future enhancements are suggested in the following lists:

Enhancing the pitch tracking for female talkersby detecting their presence in the mix using their higher pitch range, then using different parameters for the base algorithm tuned for females than those tuned for males.

Using hybrid algorithm for segregation by detecting the frame state. For voiced-voiced or unvoiced-unvoiced, CASA approaches may be used for segregation. For voiced-unvoiced, low pass and high pass filters could be used to get the speech of each talker.

ACKNOWLEDGMENT

The authors would like to thank G. Hu and D. L. Wang [8] and Z. Jin and D. L. Wang [20] [21], for providing their algorithms and codes.

REFERENCES

[1] K. S. Ananthakrishnan and K. Dogancay(2009), "Recent trends and challenges in speech separation systems research - a tutorial review," in TENCON 2009, IEEE Region 10 Conference, Hong Kong, 23-26 Nov., 2009(C1).

[2] Y. Mahgoub, B.Eng., and M.Eng., " Co-channel speech separation using state-space reconstruction and sinusoidal modeling," Ph.D. dissertation, Carleton University, Ottawa, Ont., Canada, Canada, 2010.

[3] R. H. Frazier, "An adaptive filtering approach toward speech enhancement," M.S. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA, June 1975.

[4] T. F. Quatieri and R. G. Danisewicz, "An approach to co-channel talker interference suppression using a sinusoidal model for speech," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 38, pp. 56-69, Jan. 1990.

[5] D. S. Benincasa and M. I. Savic, "Co-channel speaker separation using constrained nonlinear optimization," in ICASSP-1991, Munich, Germany, Apr. 1997, vol. 2, pp. 1195-1198.

[6] Q. Huang and D. Wang, "Single-channel speech separation based on long-short frame associated harmonic model," Digital Signal Processing Journal, vol. 21, issue 4, pp. 497-507, July 2011.

[7] G. J. Brown and D. L. Wang, "Separation of speech by computational auditory scene analysis," in Speech Enhancement (J. Benesty, S. Makino, and J. Chen, eds.), New York: Springer, 2005, ch. 16, pp. 371-402.

[8] G. Hu and D. L. Wang, "An auditory scene analysis approach to monaural speech segregation," in Topics in Acoustic Echo and Noise Control (E. Hansler and G. Schmidt, eds.), New York, NY, USA: Springer, 2006, ch. 12, pp. 485-515.

[9] L. Ottaviani and D. Rocchesso, "Separation of speech signal from complex auditory scenes," in COST G-6 Conference on Digital Audio Effects, Limerick, Ireland, Dec. 2001, pp. 87-90.

[10] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," Neural Networks, vol. 13, pp. 411-430, May-June 2000.

[11] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," Signal Processing 81, 2001, pp. 2353–2362.

[12] G. Jang, T. Lee, and Y. Oh, "Single-channel signal separation using time-domain basis functions," IEEE Signal Processing Letters, vol. 10, pp. 168-171, June 2003.

[13] A. J. W. van der Kouwe, D. Wang, and G. J. Brown, "A comparison of auditory and blind separation techniques for speech segregation," IEEE Transactions on Speech and Audio Processing, vol. 9, pp. 189-195, Mar. 2001.

[14] A. M. Reddy and B. Raj, "A minimum mean squared error estimator for single channel speaker separation," in INTERSPEECH 2004 - ICSLP, Jeju Island, Korea, Oct. 2004, pp. 2445-2448.

[15] H. A. T. Kristjansson and J. Hershey, "Single microphone source separation using high resolution signal reconstruction," in ICASSP-2004, Montreal, QC, Canada, May 2004 , vol. 2, pp. 817-820.

[16] S. T. Roweis, "One microphone source separation," in Advances in Neural Information Processing Systems (NIPS-2001) (T. K. Leen, T. G. Dietterich, and V. Tresp, eds.), Cambridge, MA, USA: MIT Press, Dec. 2001, vol. 13, pp. 793-799.

[17] D. E. M. J. Reyes-Gomez and N. Jojic, "Multiband audio modeling for single channel acoustic source separation," in ICASSP-2004, Montreal, Canada, May 2004, vol. 5, pp. 641-644.

[18] P. Li, Y. Guan, B. Xu, and W. Liu, "Monaural speech separation based on computational auditory scene analysis and objective quality assessment of speech," IEEE Transactions on Audio, Speech, And Language Processing, vol. 14, no. 6, Nov. 2006.

[19] X. Zhang, W. Liu, and B. Xu, "Monaural voiced speech segregation based on dynamic harmonic function," EURASIP Journal on Audio, Speech, and Music Processing vol. 2010, Article ID 252374, 2010.

[20] Z. Jin and D. L. Wang, "Reverberant speech segregation based on multipitch tracking and classification," IEEE Transactions on Audio, Speech, And Language Processing, vol. 19, no. 8, Nov. 2011.

[21] Z. Jin and D. L. Wang, "HMM-based multipitch tracking for noisy and reverberant speech," IEEE Transactions on Audio, Speech, And Language Processing, vol. 19, no. 5, July 2011.

[22] A. de Cheveigne, "A mixed speech F0 estimation algorithm," in EUROSPEECH-1991, Genova, Italy, Sep. 1991, pp. 445-448.

[23] F. Sha and L. K. Saul, "Real-time pitch determination of one or more voices by nonnegative matrix factorization," in Advances in Neural Information Processing Systems (NIPS 2004) (L. K. Saul, Y. Weiss, and L. Bottou, eds.), Cambridge, MA, USA: MIT Press, Dec. 2004, vol. 17, pp. 1233-1240.

[24] M. Wu, D. Wang, and G. Brown, "A multipitch tracking algorithm for noisy speech," IEEE Transactions on Speech and Audio Processing, vol. 11, pp. 229-241, May 2003.

[25] J. Ming, T. J. Hazen, and J. R. Glass, "Combining missing-feature theory, speech enhancement, and speaker-dependent/-independent modeling for speech separation," Computer Speech and Language vol. 24, pp. 67–76, 2010.

[26] P. Li, Y. Guan, S. Wang, B. Xua, and W. Liu, "Monaural speech separation based on MAXVQ and CASA for robust speech recognition," Computer Speech and Language, vol. 24, pp. 30–44, 2010.

[27] M. Cooke, J. R. Hershey, and S. J. Rennie, "Monaural speech separation and recognition challenge," Computer Speech and Language, vol. 24, pp. 1–15, 2010.

An Adaptive Multimodal Biometrics System using PSO

Ola M. Aly
Ministry of Military Production
Egypt, Cairo

Gouda I. Salama
Egyptian Armed Forces
Egypt, Cairo

Tarek A. Mahmoud
Egyptian Armed Forces
Egypt, Cairo

Hoda M. Onsi
Vice dean of post graduate and research, Faculty of Computers and Information Cairo University
Egypt, Cairo

***Abstract*—Multimodal biometric systems which fuse information from a number of biometrics, are gaining more attentions lately because they are able to overcome limitations in unimodal biometric systems. These systems are suited for high security applications. Most of the proposed multibiometric systems offer one level of security. In this paper a new approach for adaptive combination of multiple biometrics has been proposed to ensure multiple levels of security. The score level fusion rule is adapted using (PSO) Particle Swarm Optimization to ensure the desired system performance corresponding to the desired level of security. The experimental results prove that the proposed multimodal biometric system is appropriate for applications that require different levels of security.**

Keywords—multibiometric; match score fusion; PSO; Irsi; Palmprint; Finger_Knuckle

I. INTRODUCTION

The biometric technologies cover a wide range of applications that can be used to verify person identity by measuring human physiological or behavioral characteristics. Biometric characteristics including fingerprint, facial features, iris, voice, signature, and palmprint, finger-knuckle, gait etc. are now widely used in security applications. Unimodal biometric systems perform person recognition based on a single source of biometric information. Such systems are often affected by some problems such as noisy sensor data and non-universality, inter-class similarities, and spoof attacks. Thus, due to these practical problems, the error rates associated with unimodal biometric systems are quite high and consequently it makes them unacceptable for deployment in security critical applications [1].

Some of the problems that affect unimodal biometric systems can be avoided by using multimodal biometric systems. They address the issue of non-universality. It becomes increasingly difficult (if not impossible) for an impostor to spoof multiple biometric traits of an individual. Morever multibiometric systems may also be viewed as fault tolerant systems.

Multibiometric systems which fuse information from multiple biometric sources can be classified into one of six categories [2]: Multi-sensor systems, Multi-algorithm systems, Multi-instance systems, Multi-sample systems, Multimodal systems and Hybrid systems. Depending on the level of information that is fused, the fusion scheme can be classified as sensor level, feature level, score level and decision level fusion. The score level fusion is the most commonly used approach in multibiometric systems.

Most of the multimodal biometric systems proposed in literature have used a fixed combination rule and a fixed threshold to achieve the desired performance. These systems offer a fixed level of security and often have to contend with high false rejection rate if the security level is the highest. Therefore, the performance of these systems is not adaptive to the requirements of the varying level of security [3].

There are wide ranging applications where a biometric system with multiple levels of security is desirable. In this paper, an adaptive multimodal biometric system has been proposed to ensure different levels of security. This system can automatically select the best fusion rule and the optimum decision threshold to achieve the best performance corresponding to the desired security level.

The remainder of this paper is organized as follows: Section (II) describes the related works. Section (III) introduced the proposed multimodal biometric system. Section (IV) introduces the experimental results and discussion. Finally the paper is concluded in section (V).

II. RELATED WORKS

Beginning from 2000, multibiometric recognition systems in score level fusion have gained much attention and several fusion rules have been proposed. Authors in [4] [5] have provided comparisons between fixed and trained rules in combination strategies. It has been shown that the trainable fusion strategies do not necessarily perform better than fixed combination rules.

Sim et al. [6] have proposed an interesting approach to achieve high security using multimodal biometrics. Their approach has involved performing continuous verification using user's passively collected fingerprint and face biometric data. However, this approach requires continued physical presence of the user and therefore is not suitable for certain kind of applications including the popular access control applications.

Frischholz and Deickmann [7] have developed BioID system which offers multiple security levels by employing different decision strategies on the biometric modalities (face, lip motion and voice) being fused. When the required security level is low, it may well be enough to make a decision based on the agreement of two out of three modalities. On the other hand, for high security applications, this system demands agreement of all the three modalities. However, BioID system did not provide a systematic way to vary the level of security. Instead, a system administrator made a decision on the decision strategies to be adopted to achieve the desired performance.

Tronci et al. [8] have recently investigated another aspect of multimodal problem that focused on the dynamic selection of matching scores from all the available matching scores. The best matching score from a set of matching scores was selected based on the likelihood of input user being genuine or impostor. However the utility of this approach was quite limited as the achieved performance was not consistent.

Kanhangad et al. [9] have presented a promising approach to the adaptive management of multimodal biometrics to adaptively ensure the desired performance. The authors have proposed an algorithm based on Particle Swarm Optimization (PSO) to optimally combine the individual biometric sensor decisions. The proposed algorithm selected the fusion rule and sensor operating points that minimize a given cost function.

Kumar et al. [10] have introduced an adaptive combination system of multiple biometrics to ensure the optimal performance for the desired level of security using PSO. They have used different biometric combinations (iris, palmprint), (face , speech) and (fingerprint , hand geometry). The experimental results showed that the proposed score-level approach generated fairly stable performance and required smaller number of iterations to generate better performance as compared to the decision level approach.

Anzar and Sathidevi [11] have proposed an efficient PSO integration weight optimization scheme using d-prime statistics to determine the optimal weight factors for the complementary modalities. They have used fingerprint and voice biometrics in the score level fusion. The proposed method has reduced the False Acceptance Rrate (FAR) under varying noise conditions by estimating the optimal integration weight using stochastic optimization technique and Leave-One-Out Cross Validation techniques.

III. THE PROPOSED MULTIMODAL BIOMETRIC SYSTEM

Fig 1 shows the block diagram of the proposed system for optimized matching scores level fusion using Particle Swarm Optimization (PSO). Given three biometrics iris, finger-knuckle and palmprint. The feature vectors are extracted from each biometric separately. Then the matching score for each biometric sample is calculated according to the corresponding templates.

The proposed work is concerned with the development of multimodal biometric system that can dynamically choose from different fusion rules according to the desired level of security. The required level of security is an external parameter that is supplied to the system.

This level of security according to Bayesian sense is quantified by two parameters[10]: (CFA) the global cost of falsely accepting an imposter and (CFR) the global cost of falsely rejecting a genuine user. The Bayesian cost E to be minimized by the multimodal biometric system is the weighted sum of FAR and FRR as shown in eq. 1:

$$E = C_{FA}\, F_{AR}(\eta) + C_{FR}\, F_{RR}(\eta) \qquad (1)$$

Where

$$C_{FA} + C_{FR} = 2$$

Where

$F_{AR}(\eta)$:false acceptance rate

$F_{RR}(\eta)$:false rejection rate

(η): decision threshold

C_{FA} : [0, 1] and C_{FR} : [0, 1]

The main goal of the proposed multimodal biometric system is to minimize the cost function E by selecting the appropriate score level fusion rule and the decision threshold. This is achieved by the Particle Swarm Optimization (PSO) approach.

A. Unimodal biometric systems

1) Iris Identification System

Among biometric technologies, iris-based authentication systems have more advantages than other biometric technologies do. Iris patterns are believed to be unique due to the complexity of the underlying environmental and genetic processes that influence the generation of iris pattern. Iris offers an excellent recognition performance because the false match and false non-match errors are very small [12].

The iris identification system consists of three stages, the first stage is the iris analysis which involves iris localization and iris normalization. The second stage is the feature extraction and encoding. The last stage is the recognition stage which involves identification and verification.

In this paper Daugman's algorithm is used for performing iris localization which is based on applying an integro-differential operator to find the iris and pupil contours [13]. Only the significant features of the iris must be encoded in order to generate the iris code for the matching process. In the proposed system, log-Gabor filter [14] [15] is used for extracting the features from the iris image. Finally matching is performed using the calculated Hamming distance (HD) which is a measure of the number of different bits between the two iris codes[16].

2) Palmprint Identification System

Palmprint based personal identification has become an increasing active research topic over the years. The Palmprint is rich in information not only has the unique information available as on the fingerprint but has far more amount of details in terms of principal lines, wrinkles and creases.

In the proposed Palmprint identification system a preprocessed image database is used, then log-Gabor filter is

performed for extracting the features from the Palmprint image and Hamming distance is calculated during the matching stage [17] [18].

3) Finger-Knuckle Print Identification System

The usage of finger-knuckle biometric for personal identification has shown promising results and generated a lot of interest in biometrics [19]. finger-knuckles of the human hand are characterized by special creases on them. These creases differ from person to another.

In the proposed finger-knuckle identification system, a preprocessed image database is used then the features are extracted from the finger-knuckle image. Linear Discriminant Analysis (LDA) is performed to extract the only significant features from the finger-knuckle image. In this proposed system, the LDA is used to both reducing the dimensionality of the feature vector and performing the classification algorithm. [20] [21].

applied which transforms scores into a common range [0, 1]. The normalized scores are given by [22]:

$$S_i' = \frac{Si - S_{\min}}{S_{\max} - S_{\min}} \qquad (2)$$

Where

S_i' : the normalized matching scores

S_i : the matching scores,

i= 1,2, ….., n and n: number of matching scores

$S_{\min}$ & $S_{\max}$: the min and max match scores

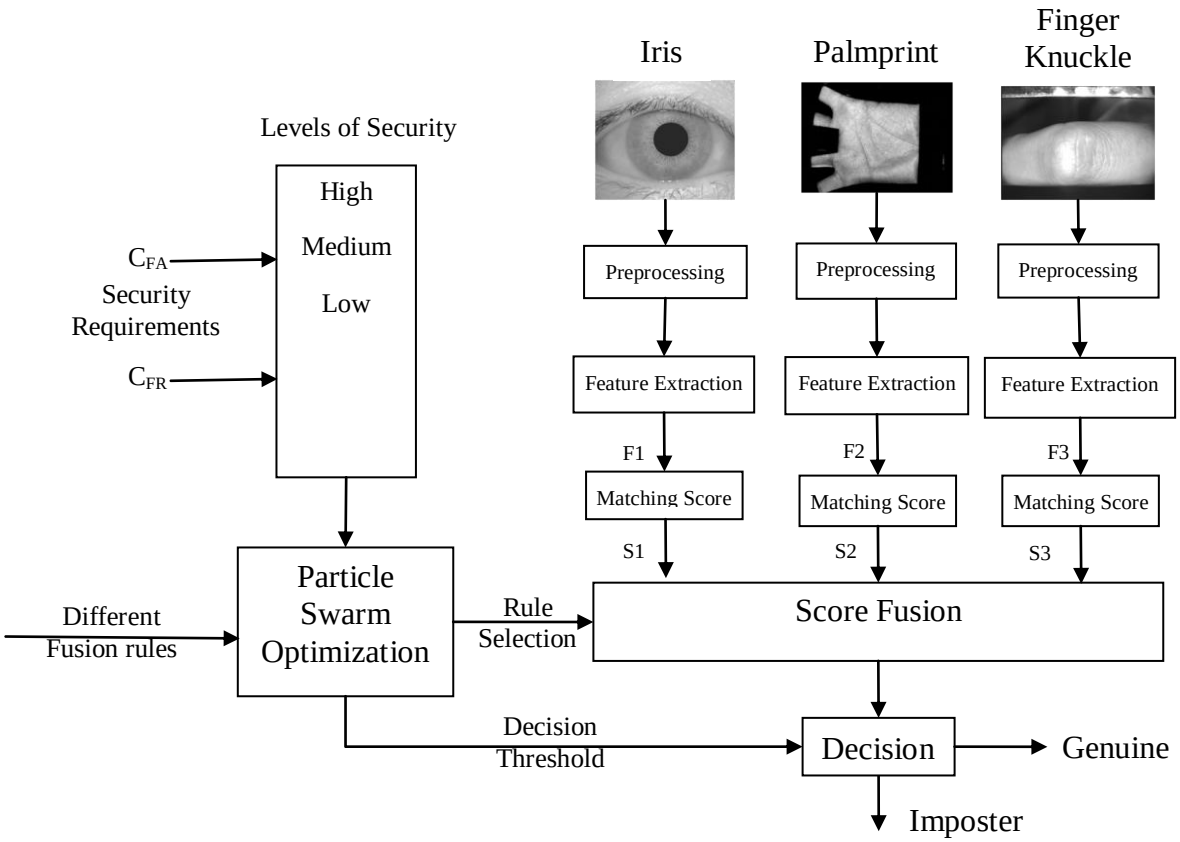


Fig. 1. The block diagram of the proposed system

B. Score Level Fusion

Score level fusion refers to the combination of matching scores provided by the unimodal classifiers in the system. This is the most widely used fusion approach, as evidenced by the experts in the field. But before the fusion step, these matching score should be normalized. In this paper Min-max method is

In order to combine the scores reported by the three matchers, different score level combinations could be applied, such as sum, product, weighted sum rule and min rules:

$$Sum = \sum_{i=1}^{n} S_i \quad (3)$$

$$\Pr oduct = \prod_{i=1}^{n} S_i \quad (4)$$

$$Weighted\ \_Sum = \sum_{i=1}^{n} w_i S_i \quad (5)$$

Where:

N : number of match scores wanted to be fused

S : the matching score

w_i : The weight for each score which calculated as follow

$$W_i = \frac{EER_i}{\sum_{i}^{m} EER_i} \quad (6)$$

Where EER_i is the unimodal biometric error.

m: the number of biometrics.

C. Particle Swarm Optimization (PSO)

PSO is an evolutionary, stochastic, population-based optimization algorithm whose goal is to find a solution to an optimization problem in a search space. The PSO algorithm was developed by Kennedy and Eberhart in 1995 [23]. The main idea of PSO is inspired from the social behavior of organisms, such as birds in a flock. The PSO algorithm imitates the behavior of flying birds and their means of information exchange to solve optimization problems. Each particle (representing a bird in the flock), characterized by its position and velocity, represents the possible solution in search space. Behavior of the particles in the PSO imitates the way in which birds communicate with each other, while flying. During this communication, each bird reviews its new position in the space with respect to the best position it has covered so far. The birds in the flock also identify the bird that has reached the best position/environment. Upon knowing this information, others in the flock update their velocity (that depends on a bird's local best position as well as the position of the best bird in the flock) and fly towards the best bird. The process of regular communication and updating the velocity repeats until reaching a favorable position.

In a similar manner, the particle in the PSO moves to a new position in the multidimensional solution space depending upon the particle's best position (also referred to as local best position (Pak) and global best position (Pgk). The Pak and Pgk are updated after each iteration whenever a suitable solution is located by the particle (lower cost). The velocity vector of each particle represents/determines the forthcoming motion details. The velocity update equation of a particle of the PSO, for instance (t+1), can be represented as follows[24]:

$$v_{ak}(t+1) = \omega\, v_{ak}(t) + c_1\, r_1\big(p_{ak}(t) - x_{ak}(t)\big) + c_2 r_2\left(p_{gk}(t) - x_{ak}(t)\right) \quad (7)$$

Where

ω is the inertia weight between 0-1 and provide a balance between global and local search abilities of the algorithm. The accelerator coefficients c_1 and c_2 are positive constants, and r_1 and r_2 are two random numbers in 0-1 range. The corresponding position vector is updated by:

$$x_{ak}(t+1) = x_{ak}(t) + v_{ak}(t+1) \quad (8)$$

Equation (7) indicates that the new velocity of a particle in each of its dimensions depends on the previous velocity and the distances from the previously observed best solutions (positions of the particle). In the implementation of this paper each particle is characterized by three variables; the fusion rule and the decision threshold and the corresponding FAR and FRR for each threshold.

IV. EXPERIMENTAL RESULTS AND DISCCUSION

In this work three different databases for three biometric modalities (iris, palmprint and finger-knuckle) are used. Firstly the results for each unibiomtric system will be presented, and then the results of fusion of two or three biometrics at match score level using PSO will be introduced.

Generally, the performance of the biometric identification system is measured by False Acceptance Rate (FAR) and False Rejection Rate (FRR) or Genuine Acceptance Rate (GAR). The system should have a high GAR with a corrosponding low FAR,FRR and Total Error Rate (TER) [25] [26].

FRR, FAR, GAR and TER are determined as follow:

$$FAR(\%) = \frac{false\ accep\tan ce\ numbers}{No\ of\ imposter\ test} X100\% \quad (9)$$

$$FRR(\%) = \frac{false\ rejection\ numbers}{No\ of\ client\ test} X100\% \quad (10)$$

$$GAR(\%) = 100 - FRR(\%) \quad (11)$$

$$TER(\%) = FRR(\%) + FAR(\%) \quad (12)$$

A. Unimodal Experimental Results

For iris images, CASIA iris Image Database is used [27], includes 2500 iris images from 250 eyes for each eye. 200 persons have been selected, for each person 6 Iris images are used for training and 4 for testing.

For palmprint images, PolyU palmprint database is used [28], contains 7752 grayscale images corresponding to 386 different palms (10 samples for each hand). 200 persons have been selected, for each person we have 6 palmprint images for training and 4 for testing.

For finger-knuckle images, database images introduced in [29] is used, collected from 165 volunteers (12 samples for each user), including 125 males and 40 females. 200 persons have been selected, for each person 8 finger-knuckle images for training and 4 for testing.

Table 1 shows the results of iris, palmprint and finger-knuckle identification systems. It could be noticed that the TER is too much to be suitable for high security applications.

TABLE I. COMPARISON OF UNIMODAL BIOMETRIC RESULTS

Biometric Type	GAR %	FAR %	FRR %	TER %
Iris	**97**	7.14	3	**10.14**
Palmprint	**96.76**	0.00	3.24	**3.24**
Finger_Knuckle	**85.50**	0.00	14.50	**14.50**

B. Matching score fusion results

The goal of this experiment is to evaluate the system performance when using a unimodal biometric system versus a multibiometric system using match score fusion by the aid of PSO as an optimizer.

In this experiment, three score level combinations are considered including sum, product and weighted sum. The PSO is employed to dynamically select the appropriate decision threshold and the best fusion rule to minimize the Bayesian cost for the corresponding required security level. Each particle in PSO is characterized by three variables; a variable representing one of a different score level fusion rules, decision threshold and the corresponding FAR and FRR for each threshold. The performance of PSO is largely depending upon the parameter chosen.

The parameters of PSO in these experiments were determined as follows:

- Population size is 30
- The inertia weight ω is an important parameter as it controls the effect of the previous velocity vector of the swarm on the new one. It is experimentally found that ω in the range [0.8,1.2] yields a better performance [30]. It is selected and fixed at 0.8.
- The acceleration constants c_1, c_2 are set to 1.
- Velocity limitation V_{max} is set to 1.

Table II shows the result of the classification rate including FAR, FRR, TER and GAR for the proposed multimodal biometric fusion approach by the aid of PSO as an optimizer. It is clear that the performance of the proposed multimodal biometric system outperforms the unimodal systems and strongly reduces the TER. The proposed system achieves significant results with best GAR 98.40% and TER 2.60%.

TABLE II. RESULTS OF THEPROPOSED ADAPTIVE MULTIMODAL BIOMETRIC SYSTEMS USING PSO

Biometric Type	GAR %	FAR %	FRR %	TER %
Iris-knuckle-fusion	**98**	0.00	2	**2**
Iris-Palmprint-fusion	**98.40**	1	1.60	**2.60**
Knuckle-Palmprint-fusion	**97.25**	0	2.75	**2.75**

Figs. 2 and 3 show the average of the minimum weighted error rate and the standard deviation of the minimum error of the proposed score level adaptive combination scheme using iris and finger-knuckle modalities. Fig. 4 shows the adaptive rule selected at score level versus the variation of security levels.

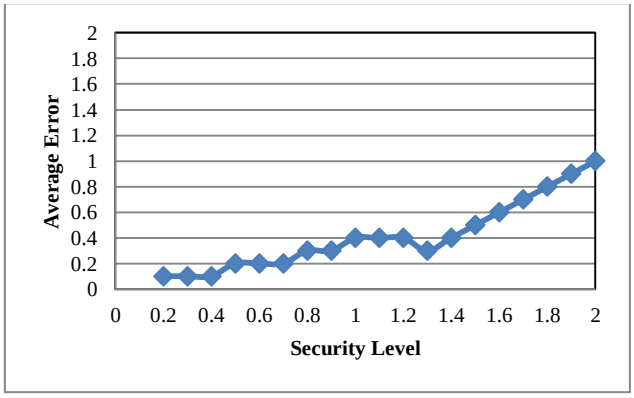


Fig. 2. Average minimum error from the score level approach using the adaptive combination of iris and finger-knuckle modalities

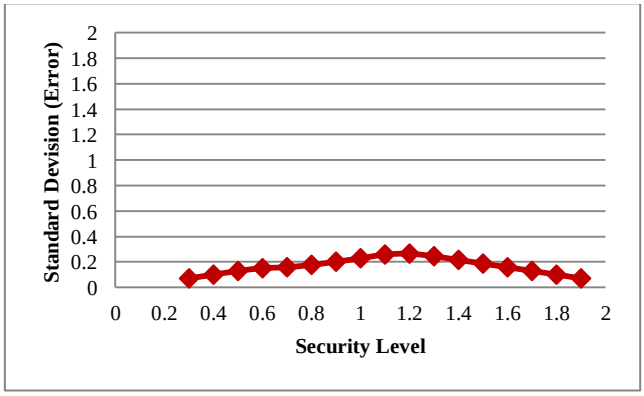


Fig. 3. Standard deviation of the minimum error, from each run, using score level approach for iris and finger-knuckle modalities

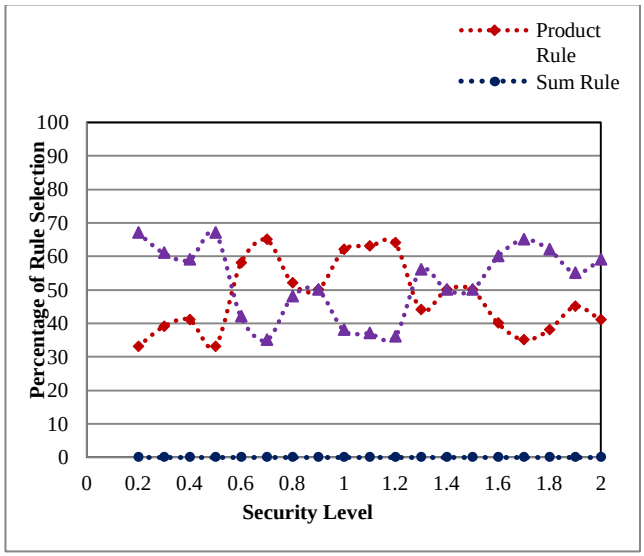


Fig. 4. Adaptive selection of fusion rules using score level combination for iris and finger-knuckle modalities

Figs. 5 and 6 show the average of the minimum weighted error rate and the standard deviation of the minimum error of the proposed score level adaptive combination scheme using iris and palmprint modalities. Fig. 7 shows the adaptive rule selected at score level versus the variation of security levels.

Figs. 8 and 9 shows the average of the minimum weighted error rate and the standard deviation of the minimum error of the proposed score level adaptive combination scheme using finger-knuckle and palmprint modalities. Fig. 10 shows the adaptive rule selected at score level versus the variation of security levels.

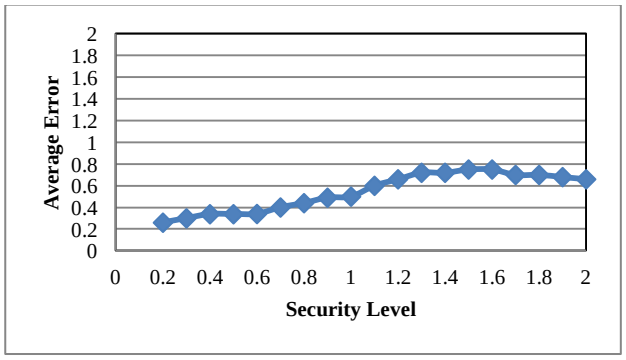


Fig. 5. Average minimum error from the score level approach using the adaptive combination of iris and palmprint modalities

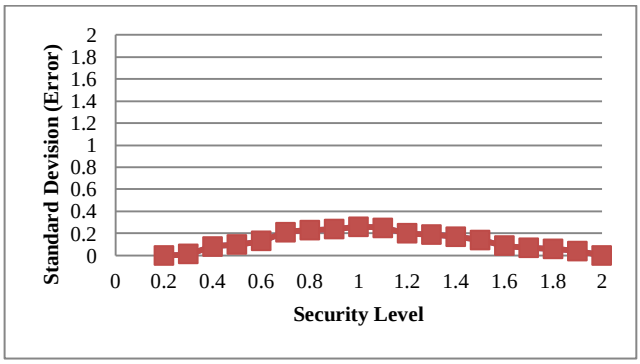


Fig. 6. Standard deviation of the minimum error, from each run, using score level approach for iris and palmprint modalities

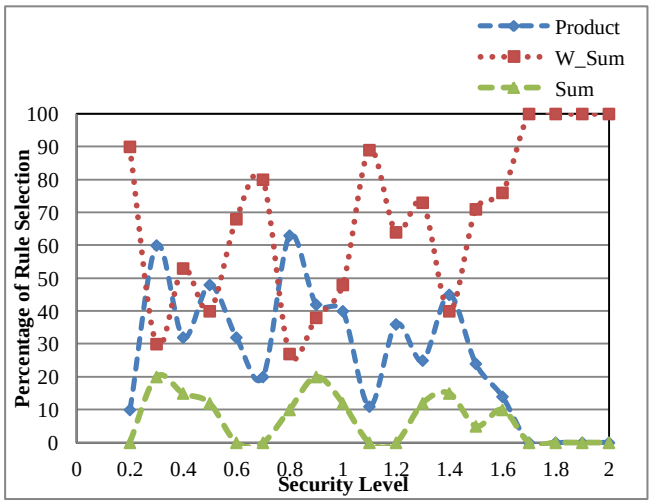


Fig. 7. Adaptive selection of fusion rules using score level combination for iris and palmprint modalities

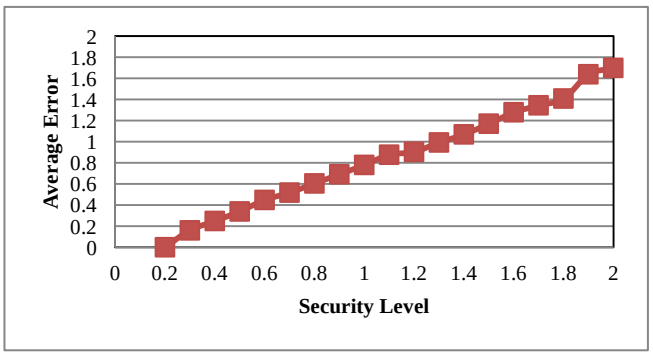


Fig. 8. Average minimum error from the score level approach using the adaptive combination of Knuckle and palmprint modalities

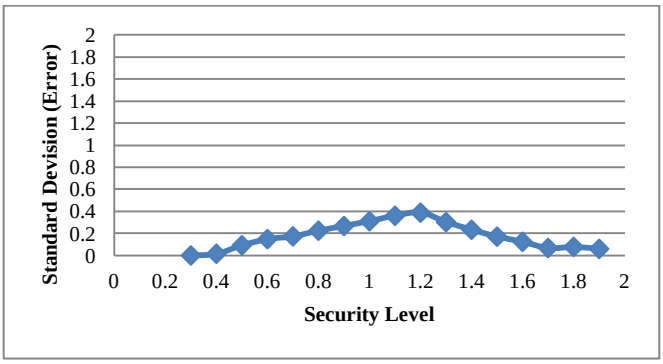


Fig. 9. Standard deviation of the minimum error, from each run, using score level approach for knuckle and palmprint modalities

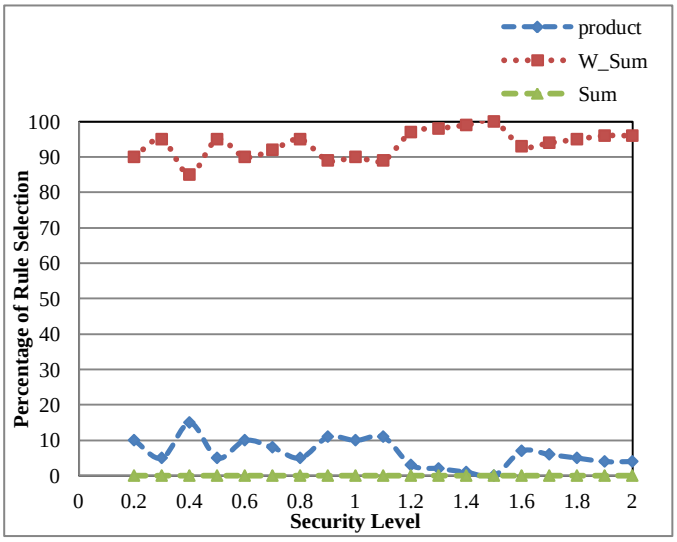


Fig. 10. Adaptive selection of fusion rules using score level combination for knuckle and palmprint modalities

As shown in Figs. 2,5 and 8 the security level is equal to the sum of cost of false acceptance (CFA) and cost of false rejection (CFR). It could be notice that errors increases as the security level is increased from one level to another.

Figs. 3, 6 and 9 shows the standard deviation of the minimum error versus the security level for the score-level fusion approach using different pairs of biometric modalities, It could be observed that the results of the proposed scheme are significantly stable, i.e., have smaller standard deviation, and therefore require significantly smaller number of iterations. From these experiments, it was shown that by using combinations at score-level only ten iterations are adequate to achieve the stable results as compared to the 100 iterations needed in case of decision-level approach [31].

Figs. 4, 7 and 10 show the adaptive rule selected at score level versus the variation of security levels. It can be observed that the sum rule was less choice rul during iterations for any security level. However, the product and weighted sum rules were chosen interchangeably through different levels of security.

V. Conclusion

In this paper, a new multimodal biometric identification system is proposed using three modalities including iris, palmprint and finger-knuckle with fusion at matching score level. The main objective of this work is to develop a reliable approach for the adaptive combination of multiple biometric modalities to ensure desired level of security. The proposed method uses PSO to achieve adaptive combination of multiple biometrics from their matching scores.

The PSO is used to optimize the selection of score level combination, its corresponding parameters, and the decision threshold. In this work only 3 fusion rules have been suggested, there may be several other score level combination approaches which may perform better, i.e., achieve minimum cost E and can be easily incorporated in the proposed framework. The experimental results shown in Figs. 4, 7 and 10 illustrate the dynamic rule selection of these score level combinations to ensure the desired level of security.

The results prove that the proposed multimodal biometric system improves the identification rate and outperforms the unimodal biometric systems using different biometric combinatioins. Morever, the TER is strongly decreases to 2.60% at 98.4% Identification rate.

References

[1] F. Wang and J. Han, “Robust multimodal biometric authentication integrating iris, face and palmprint”, Information Technology and Control, vol.37, no.4, pp. 326-332, 2008.

[2] A. Ross, K. Nandakumar and A. Jain. Handbook of Multibiometrics1st edition. Springer, New York, USA, 2006.

[3] A. Kumar, V. Kanhagad and D. Zhang, “A new framework for adaptive multimodal biometrics management”, IEEE Transactions on Information Forensics and Security, vol. 5, no. 1, March 2010

[4] D. Tax, M. Breukelen, R. Duin, and J. Kittler, “Combining multiple classifiers by averaging or multiplying,” Pattern Recognition, vol. 33, pp. 1475-1485, 2000.

[5] F. Roli, S. Raudys, and G. Marcialis, “An experimental comparison of fixed and trained fusion rules for crisp classifier outputs,” Intl. Workshop on Multiple Classifier Systems, Cagliari, Italy, Jun. 2002.

[6] T. Sim, S. Zhang, R. Janakiraman and S. Kumar, “Continuous verification using multimodal biometrics,” IEEE Trans. Patt. Anal. Machine Intell. , vol. 29, no. 4, pp. 687-700, Apr. 2007.

[7] R. W. Frischholz and U. Deickmann, “BioID: A multimodal biometric identification system,” IEEE Comput., vol. 33, no. 2, Febraury. 2000.

[8] R. Tronci, G. Giacinto and F. Roli, “Dynamic score selection for fusion of multiple biometric matchers”, Proc. 14 IEEE International Conference on Image Analysis and Processing, Modena, Italy, pp. 15-20, 2007.

[9] V. Kanhangad, A. Kumar, and D. Zhang, “Comments on an adaptive multimodal biometric management algorithm,” IEEE Trans. Sys. Man & Cybern., Part-C , vol. 38, no. 5, pp. 438-440, 2008.

[10] A. Kumar, , V. Kanhangad and D. Zhang, “ A new framework for adaptive multimodal biometrics managemen”, IEEE Transactions on Information Forensics and Security, vol. 5, pp. 92-102, Mar. 2010

[11] S. Anzar and P. Sathidevi, “An efficient PSO optimized integration weight estimation using D-prime Statistics for a multibiometric system”, International Journal on Bioinformatics & Biosciences (IJBB) vol.2, no.3, September 2012.

[12] J. Daugman “How iris recognition works”, IEEE Trans. On Circuits and Systems for Video Technology. vol. 14, no. 1, pp. 21-30. 2004.

[13] J. Daugman, “High Confidence Visual Recognition of Persons by a Test of Statistical Independence,” IEEE Trans. Pattern Anal. Machine Intell, vol. 15, pp. 1148-1161, 1993

[14] D.Gabor, " Theory of communication ", J. Inst. Elect. Eng. London, Vol. 93, No. 3, 1946.

[15] A. Kumar and A. Passi, "Comparison and Combination of Iris Matchers for Reliable Personal Authentication," Pattern Recogniion., Vol. 23, No. 3, pp. 1016–1026, March 2010.

[16] L. Ma, T. Tan, Y. Wang and D. Zhang, "Efficient Iris Recognition by Characterizing Key Local Variations", IEEE Transactions on Image Processing, Vol. 13, No. 6, pp739-750, June 2004.

[17] R. Chu, Z. Lei, Y. Han and S. Li," Learning Gabor magnitude features for palmprint recognition", ACCV, pp. 22–31, 2007.

[18] A. Kong, "Palmprint Identification Based on Generalization of IrisCode", PhD Thesis, University of Waterloo, Canada, 2007.

[19] A. Kumar and C. Ravikanth, "Personal authentication using finger knuckle surface", IEEE Trans. Information. Forensics & Security, vol. 4, no. 1, pp. 98-110, March 2009

[20] H. Yu and J. Yang, "A direct LDA algorithm for high dimensional data with application to face recognition," Pattern Recognition, September 2001.

[21] P. Navarrete and J. Ruiz-del-Solar, "Analysis and Comparison of Eigenspace-Based Face Recognition Approaches," International Journal of Pattern Recognition and Artificial Intelligence, vol. 16, no. 7, November 2002.

[22] A. Pour, K. Faez, and R. Amirfattahi "Multimodal biometric system using face, ear and gait biometrics", 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010).

[23] J. Kennedy and R. Eberhart, "Particle swarm optimization". IEEE Int'l Joint Conf. on Neural Networks, Perth, Australia. (1995).

[24] M. Clerc and J. Kennedy, "The Particle Swarm-Explosion, Stability, and Convergence in a Multidimensional Somplex space," IEEE Trans. Evolutionary Comp. , vol. 6, p. 58-73, 2002.

[25] T. Sabareeswari and S. Stuwart," Identification of a Person Using Multimodal Biometric System", International Journal of Computer Applications vol. 3, no.9, 2010.

[26] L. Hong and A. Jain, "Integrating faces and fingerprints for personal identification", IEEE Trans. Pattern Anal. Mach. Intell. , Vol. 20, No. 12, pp. 1295– 1307, 1998.

[27] Chinese Academy of Science Institute of Automation, Database of the Eye Grayscale Images. http://www.sinobiometrics.com

[28] PolyU Palmprint Database, http://www.comp.polyu.edu.hk/~biometrics/

[29] L. Zhang, "Finger-Knuckle-Print: A new Biometric Identifier", ICIP, Cairo, Egypt, 2009.

[30] Y. Shi and R.C. Eberhart, "A modified particle swarm optimizer", proc. IEEE conference on evolutionary computation, 1998.

[31] K. Veeramachaneni, L. Osadciw and P. Varshney, "An Adaptive Multimodal Biometric Management Algorithm," IEEE Trans. Sys. Man & Cyber., vol. 35, no. 3, pp. 344-356, August 2005.

Interaction Protocols in Multi-Agent Systems based on Agent Petri Nets Model

Borhen Marzougui
Emirates College of Technology
Abu Dhabi, UAE, P.O.BOX 41009

Kamel Barkaoui
CNAM
292 Rue Saint-Martin, Paris, France

***Abstract*— this paper deals with the modeling of interaction between agents in Multi Agents System (MAS) based on Agent Petri Nets (APN). Our models are created based on communicating agents. Indeed, an agent initiating a conversation with other can specify the interaction protocol wishes to follow. The combination of APN and FIPA Protocols schemes leads to a set of deployment formal rules for points where model interaction can be successfully implemented. We introduce some models FIPA standard protocols.**

Keywords—Interaction; MAS; APN; Model; formalism; FIFA; Protocols.

I. INTRODUCTION

Multi Agent systems composed of multiple interacting agents. Each Agents have to interact together to achieve their tasks. This Achieving tasks can be done through interaction. So, the aim of interaction is to allow the synchronization of agents and the exchange of messages.

The basic function of an interaction protocol is to provide a way for agents to communicate effectively without having to explicitly plan for each speech act by delimiting the space of possible answers [5]. When using an interaction protocol, we assume that, during analysis, it must be made to ensure that follow the protocol will achieve the goals associated with final states. Protocol is more efficient with less information needs to be transmitted and less time is spent in communication. All Agents attend different interact protocols appropriately between them. For example, responding to message, performing actions in their respective fields, or updating their local states. So, Protocols can be taken as a way to specify the policy that agents will follow in their interactions with others [12]. This Policy will determine the conditions under which a request can be satisfied.

Sometimes, when a problem solved by two or more modules it becomes quite complex. The protocol describes the communication will. This has led researchers to propose protocol engineering, specify properties that a protocol should satisfy and provide multiple modeling formalisms. This is to simplify the representation of communication protocols.

The definition of generic software architecture for interoperability between agents in a dynamic environment seems to be interesting for the operation of these protocols.

Considering this, interaction between agents is the main part of cooperation in MAS. For this reason, it is important that suitable theories, frameworks, methodologies, and tools are provided to support the building of all communicates agent. As a result, identifying, designing, and studying protocols via which agents interact is an important and active research area within the Multi Agent systems field [21]. Indeed, several research groups have developed their own models such as General Magic, KAOS, OMG, ZEUS and FIPA has defined an environment of existence and operation of agents and a platform that describes the agents, their creation, and deletion authentication.

In this context, we involve formal specifications methods to describe some complex properties of engineering protocols. We focus in this work on modeling interaction between agents with Agents Petri Nets (APN) [13], [14], and [15]. Indeed, the complementarily between Multi Agent System (MAS) and APN becomes very advantageous: an approach by agents requires a powerful and expressive formalism that allows him to model the behavior of a set of agents that interact.

This paper is organized as follows: Section 2 describes the interaction protocols. We propose in Section 3 our interaction models based on APN and MAS. In section 4, we present an example of the application of our Models. In the last section we conclude this paper by giving some perspectives.

II. INTERACTION PROTOCOLS

Address the problem of interaction in the field of SMA it's to provide the means to analyze and design the various forms of interaction that agents can use to accomplish their tasks and fulfill their goals. So, the solutions consist to assure an interaction protocols.

Indeed, the interaction, with the organization, is one of the basic concepts of multi-agent systems. According to [4], "for an agent, interact with one another is both the source of his power and the source of his problems." Indeed, it is the cooperation of agents who brings a kind of intelligence or ability to solve problems rather complex, but also because of their many conflicts that arise.

Interactions come from the dynamic linking of several agents through a set of reciprocal actions. There are several types of interactions that depend on three parameters: goals, resources and skills.

Depending on the types of agents involved direct interaction can also take many forms. This can be expressed for using reagents exchange of simple signals (as in the case of the Eco-resolution) and cognitive agents, using language and communication protocols developed. It is inspired by social interaction (communication between humans) and

supports a vision of the interaction and high communication [10]. Thus, researches in MAS consider that communication models are more complex, like the philosophy of language. However, there are other agent interaction models providing an indirect communication.

An interaction protocol is a set of rules that govern the communication between several agents [5]. It allows to describe explicitly conversational sequences when the interaction between agents (who can say what to whom and when). These protocols are used to define a sequence of messages communicated between agents and describe how agents should react to messages received during interactions [9]. For a given state of the protocol, there are a finite number of messages in transmission and reception.

With a conceptual manner, [22] presents a conceptual taxonomy of currently known/available agent interaction models (see Figure 1), trying to define advantages and issues related to them, both from a conceptual and a technical point of view.

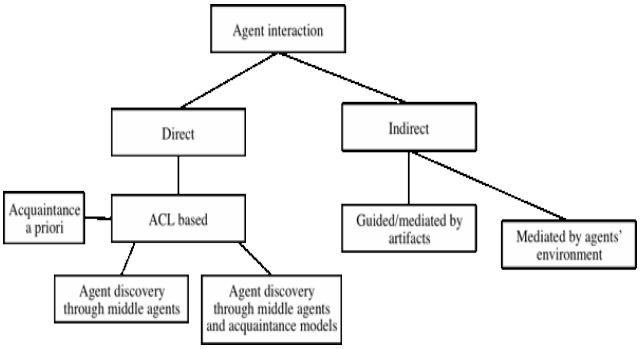


Fig.1. The proposed taxonomy of Agent Interaction Models [22].

If an agent agrees to use a protocol then he agrees to comply with this protocol and to abide by its syntax and semantic rules (on the architecture of the protocol defining the actions that agents must perform when sending and receiving a message).

A. Types of Interaction Protocols

Interaction protocols can be classified according to the types of agents (cooperative, competitive or shared goals) [8]. Among which are:

1) Coordination Protocols

They enable agents to manage (maintain, adjust or abandon) their commitments in cases where the circumstances in which they were developed, evolve. Among coordination protocols include acquaintance networks for distributed task allocation and Contract Network. The major advantage of the latter is that it allows the coordination of tasks between the agents who are ensuring the most possible optimal allocation.

2) Cooperation Protocols

Cooperation protocols consist to decompose tasks into subtasks and distribute them among different agents specifying who does what, with what resources, for what purposes and under what constraints. This strategy aims to reduce the complexity of tasks and optimize resource utilization. There are various mechanisms for allocating tasks such as election where tasks are assigned pursuant to an agreement or a vote.

3) Negotiation Protocols

Negotiation protocols are used in the case where agents have different goals or the use of a resource by agents can prevent another agent to achieve its goal. The protocol followed in the negotiation and decision-making process that determines each agent uses its positions and criteria for agreement [16].

B. FIPA Protocols

FIPA [6] provides the description of a set of protocols for high-level interaction, including the request for action, establishing contract (Contract Net) and several types of auctions .

1) Basic protocols

These protocols are often used and implicitly. They are listed in [6] and specified in the ACL. They allow an agent to simply ask another to perform an action (request protocol) to request information (query protocol), etc.. In the following we mention some of them:

The FIPA Request Protocol: This protocol allows an agent to request another agent to perform a certain action as shown in Figure Fig.2. The officer receiving the request shall, upon receipt thereof, indicate whether it accepts or rejects the request. The agent accepts the request must also notify the applicant when the action concerned by the request is made.

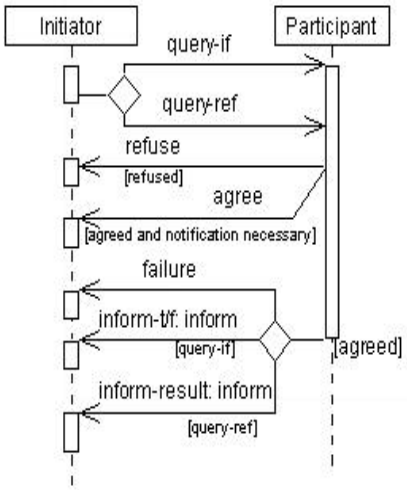


Fig.2. AUML Representation of protocol FIPA query [6]

Conditional query protocol FIPA: This protocol allows an agent to request another agent to perform an action when a certain condition is met. The agent accepts the request must wait until the condition is met to perform the requested action. It must then inform the initiator of the request that the action was performed.

The FIPA request protocol: It allows an agent to make an inquiry. The officer receiving the request can then accept or

refuse to provide information. It must of course give the requested information if it accepts the request.

2) Network protocols contractual FIPA

This protocol specifies how to use the sharing protocol tasks Contract Net [20] using FIPA-ACL as a language of communication. This protocol allows an agent (the manager or originator in Figure Fig.3.) To make a bid for performed job, agents who wish to carry out the task in question (or participants) must provide their services. Depending on the offers received, the manager decides to whom he attributes the accomplishment of the task. In fact, it determines which agent is awarded the contract for completion of the task. Finally, the agent who gets the contract must inform the manager when the task is completed.

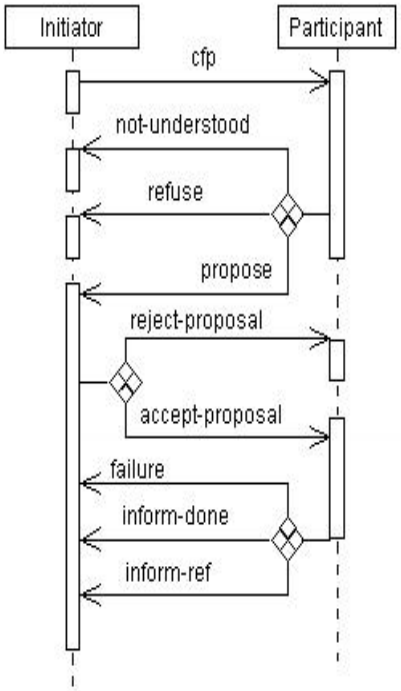


Fig.3. AUML Representation of protocol FIPA Contract Net [6]

3) Protocols FIPA auction

Protocols of this family are widely used in the field of electronic commerce. It generally refers to two different versions of auction protocols which are English and Dutch.

Protocol FIPA Dutch Auction: In a Dutch auction, the seller sets a starting price that is far beyond the actual value of the property that is for sale. Then the price is reduced until a buyer accepts announces that the proposed price. The property is then sold to the purchaser.

The English auction protocol FIPA: This protocol (Figure Fig.4.) allows an agent to use an auction to sell type English property. The seller sets a starting price that is lower than the desired selling price. Buyers who wish to purchase the property are encouraged to build on the property offering a higher amount than the current implementation of the auction. The auction ends when no one wants to raise the bet and the property is granted to the best buyer.

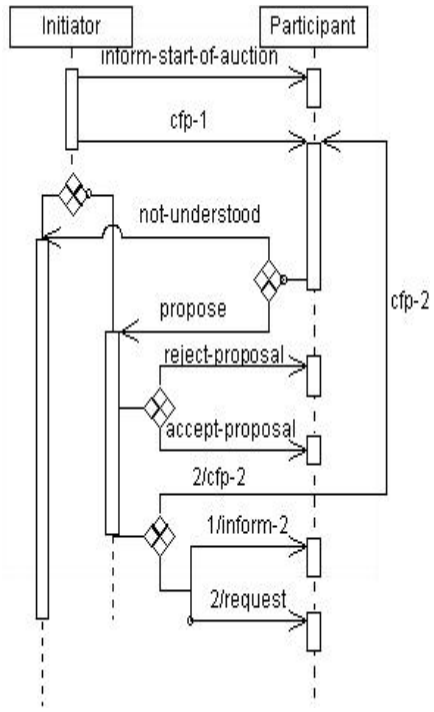


Fig.4. FIPA English Auction Protocol [6]

III. MODELING INTERACTION PROTOCOLS

A. Engineering protocols in MAS

Several studies in the literature [18], [4], [12] and [17] were interested in the proposal of an interaction protocol engineering that will guide the designer from specification to validation. Some of those researches have correspondence with the communication protocols in distributed systems as [18] and [12]. Others researchers have developed a process for development of interaction protocol specific to multi-agent systems.

B. Properties of an interaction protocol in a MAS

When looking the model of protocol for dialogue agents, there are two essential aspects which are: specification and flexibility [3].

Flexibility: It is very important to design a flexible protocol insofar as it is possible to achieve the desired goal without affecting the autonomy of agents [14]. According to [7], it is interesting to consider interaction protocols of small a designed as micro-protocols, and combine them to form more complexes and more specific task. This is particularly useful since most commonly found similar dialogue structures in different interaction protocols. Allowing the composition of micro-protocols then promotes reuse [2] and contributes to the specification of flexible protocols and its extension become easier.

Specification: It is important to use a formalism that allows specifying interaction protocols with a high level of abstraction. According to [7], a protocol must be independent of the domain and architecture of agents who will use it. [7] Indicates that it is also important to adopt a declarative

approach to explicitly state the rules of protocols. In fact, the formalism should allow specifying interaction protocols as clearly as possible while having a good power of expression. In addition, a protocol must be specified as to allow possible to verify properties such as deadlock, termination, etc..

C. APN Model for Interaction Protocol

Precisely, the underlying idea of Agent Petri Nets is that they can properly represent the agent and its autonomy in communication with other agents in its environment or other environments, while maintaining a fairly simple and understandable graphical representation. APN model transitions correspond to actions that can be performed, places are the variables of the states containing tokens corresponding to agents and arcs, according to their orientation, determines the activation conditions of a transition and its effect on the state.

Formalism is considered reliable if it ensures some important properties such as synchronization, competition but also reusability. It is therefore natural to compose models of complex protocols from simple protocols or a set of protocol elements by connecting their bows and places synchronization.

This is why we try to give our early model open connection between two agents A1 and A2, the protocol is said elementary and aims to establish a connection between two agents and can be reused in other models. Indeed, A1 sends a connection request (call) to A2 by sending request message (F (A1, A2)). After receipt of the request, A2 can accept the call by sending a message (m') to agree or to refuse. So we can define m as (m' = {agree, refuse}).

The protocol ends with the receipt of the agreement From A2. A2 is connecting to new environment. So, A1 and A2 cross their transition (Success: T7). If A2 refuses the request of A1 then A1 receiving a message refuses (Failed: T8) as shown in Figure Fig 5.

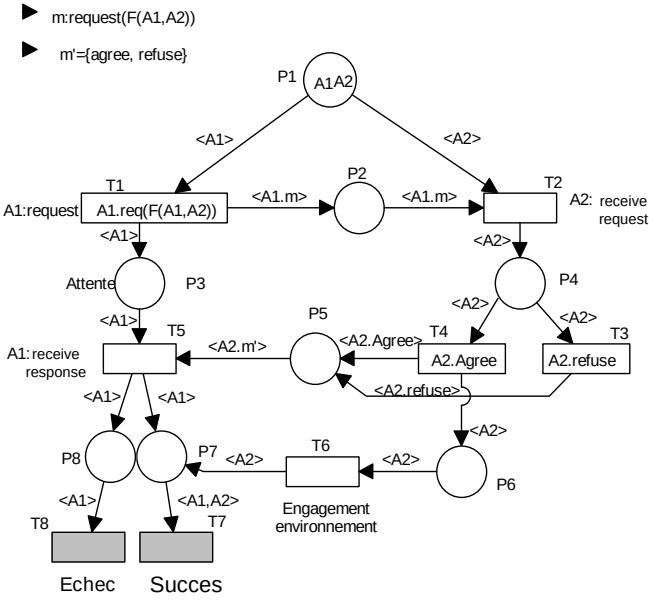


Fig.5. Open_connection APN model

In the case of connection failure due to a refusal on the part of A2, we can consider checkpoints in our model APN. Indeed, A2 sends <A2.refuse> message and can return to its original state. A1, after receiving this message, returns to its original state and another instance of the protocol can be triggered.

We opt to FIPA standards that are more used to present the interaction between agents process. In the remainder of this session, we modeling two simple protocols defined in FIPA which are "Inform" and "Request". We present a variant of "FIPA-Contract Net" protocol involving more than two agents.

1) FIPA-Inform

It is a simple communicative act to pass information from one agent to another. There are two agents interact: A1 sends a message inform (T1) to A2. A2 receives this message and processes (T2). The conversation ends when both agents cross their transitions (T3) and (T4) as shown in the figure below:

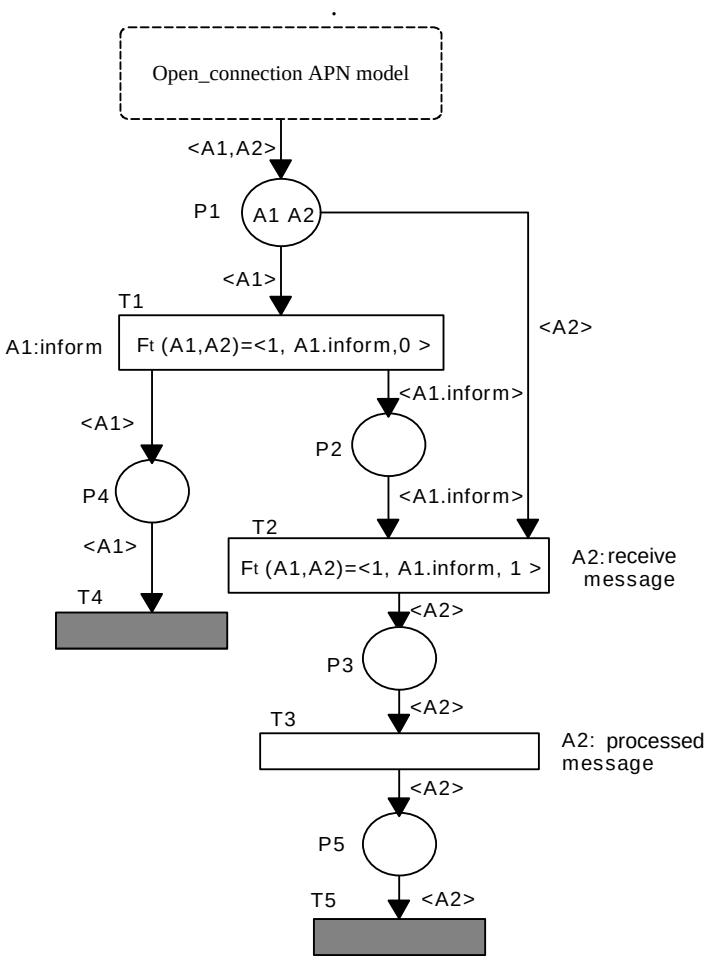


Fig.6. APN model of FIPA-Inform protocol

It was assumed in the protocol that the two agents are already in communication (connection opening). Basicly, we using basic "Open_connection APN model". It is a reusability method.

A1 send the message "inform» using the function F_t (A1, A2) = <1, A1.inform, 0>, indicate that the two agents in communication are A1 and A2. A1 is the transmitter, the receiver is A2 and the message sent by A1 is "inform". Receiving the message is validated by the value 1 in the third field of the Ft1 function (A1, A2) upon receipt.

With F_t () function we can model sending message to inform several agents always keeping the same syntax: the recipients are in brackets and the transmitter is A1, for example, inform A2, A3 and A4 is presented by: F_t (A1, A2, A3, A4) = <1, A1.inform, 0>.

2) FIPA-Request

The idea is to present a communication protocol between two agents A1 and A2. An agent A1 sends a request to another agent A2 to perform an action P. The receiver may grant or refuse to perform the action. In case of refusal; the receiving agent is obliged to disclose the reason for the rejection. This is one of the FIPA-Request protocol as shown in the diagram below.

Fig.7 describes the same protocol using Petri Nets. Each agent executes a Petri net whose places correspond to its state or the condition of the conversation and transitions correspond to sending and receiving messages.

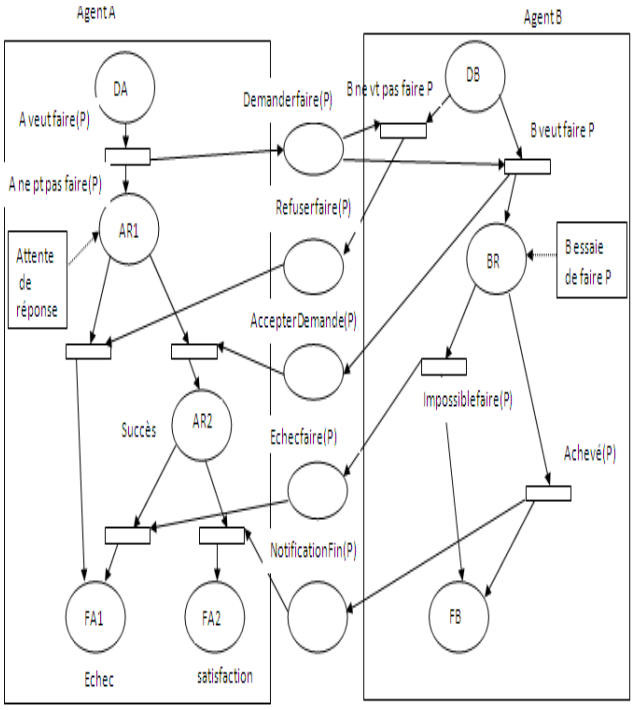


Fig.7. Petri Nets model for FIPA-Request Protocol [8]

Interpretation:

Despite the simplicity of protocol, several places, transitions and arcs were used to model the state of the conversation and agents throughout their communication.

In the conventional model, the designer has to model two cases each time. For example, B want execute P and the inverse case. The number of places used tokens are not distinguished and are increase.

So, the goal is to create a valid model for the two agents in which the location of the officer's decision must be explicit, this is possible with the use of tokens as agents identified by their Names.

We try to model this same protocol by APN and we refine our model by integrating primitives of ACL language.

A1 sends a connection request to A2 with the primitive request. A2 may accept the application, it responds him in this case with a message <A2.agree>, and <A2.refuse> if he refuses demand. In case of non understanding, A2 sends <A2.not-understood>.

In the case of acceptance of A2, it tries to Run P: send a message <A2.inform-done>. In the case of failure the message sended is <A2.failure>. However, this failure leaves the possibility to redo the task A2. To do this control, we must add checkpoints in our model.

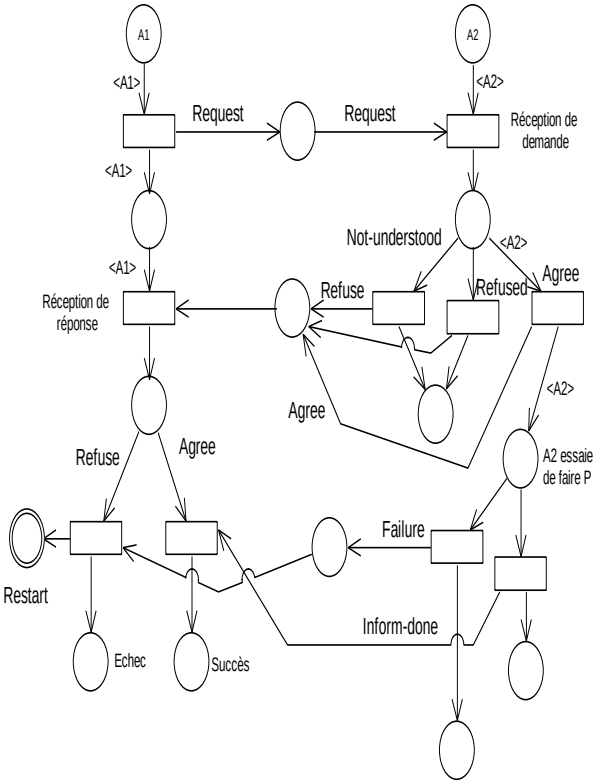


Fig.8. APN model for FIPA-Request

The Petri net of figure Fig.8 model the protocol stated above that describes the statements relating to the interaction between the two agents. We distinguish three possible situations: success, failure due to rejection of the application and failure in achieving the task.

Formally, the models specify how the interaction between these two agents occurs and what performative are used at each step of the conversation.

The following figure illustrates the APN model for FIPA-Request protocol detailed with messages exchanged between agents and the functions used.

Note that in this model, it is always possible to capture the current state of the conversation or the agent through current location of tokens (agents).

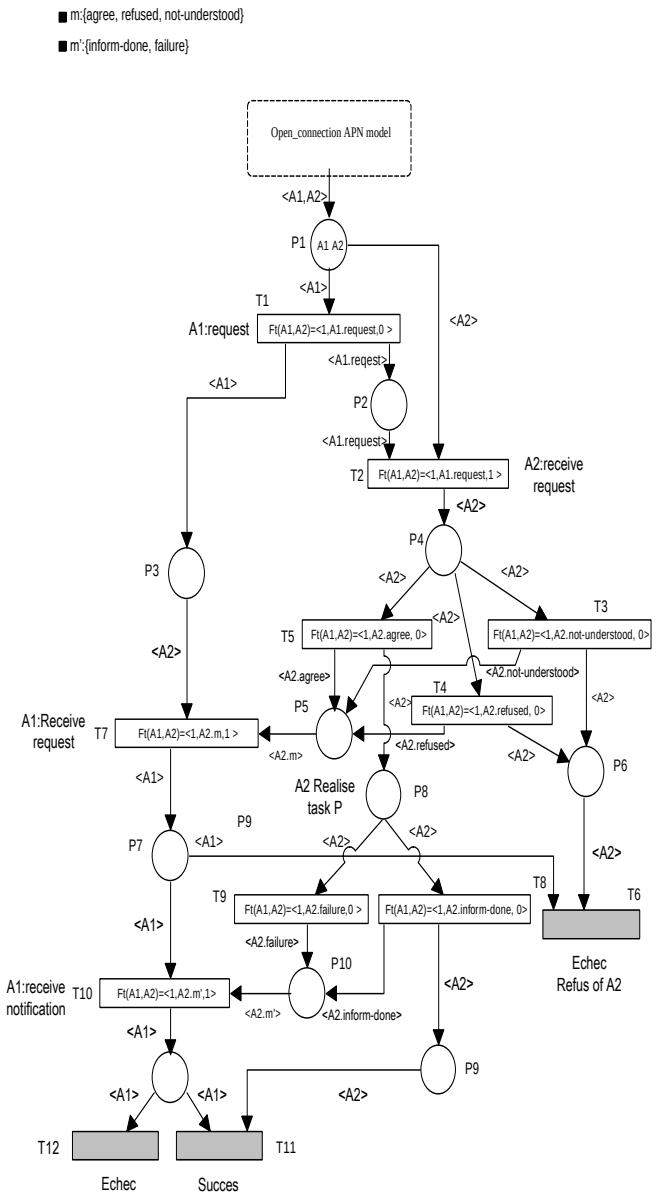


Fig.9. Detailed APN Model for FIPA-Request

Interpretation:

In this model it was supposed that the two agents are engaged in the same environment of communication. First, the connection is created by Open_connection APN model between A1 and A2.

In addition, we detailed our APN model FIPA-Request specifying the different exchanged between the two agents. We propose the structure of each message in our model with the function Ft (Ai, Aj): Ai is the transmitter and A2 is the receiver.

In both cases of failure, a new instance of the protocol can be triggered and checkpoints or host states can be added. Indeed, A2 must specify the reasons for refusal. This refusal can be either because it does not have the skills to do the job or because he does not want this job. In this second case A1 can throw a new conversation.

The first case is due to the refusal of A2: the two agents will cross the end transition T6 and can return to the initial state by adding an arc from T6 to P1.The second case of failure is due to a problem in the realization: A2 may decide to repeat the task, then add a arcc from T9 to P8.

Note that the agents in question are cognitive agents having the ability to make decisions and act autonomously while following the rules of protocol. An agent can get stuck in a state of waiting for an answer.

However, in order to more improve our models based on APN, you can insert a timing mechanism that uses a delay () function and a maximum R beyond which the agent leaves the wait state. This solution allows us to avoid an agent stuck wait a long time.

3) FIPA-Contract Net

In the following section, we will try to show the power of formalism APN in modeling protocols involving multiple agents such as FIPA-Contract Net Protocol. In this protocol, a moderator agent chooses an agent that he does not know his skills to perform a task by broadcasting a request message to perform a task P.

Our goal is not modeling the local behavior of the agent, for it was assigned to the moderator agent to choose the first positive response and refuse all that come after. This agent can cancel the negotiation during the conversation as shown in Figure Fig.10.

There are several possible scenarios:

- All agents do not accept the offer of the moderator: failure.
- There is a positive response. In this case, three cases are possible:
- If the moderator accepts the offer (accept-proposal) :success
- If the moderator cancels negotiation (cancel) : failure
- the moderator refused all offers except the first positive (reject-proposal)

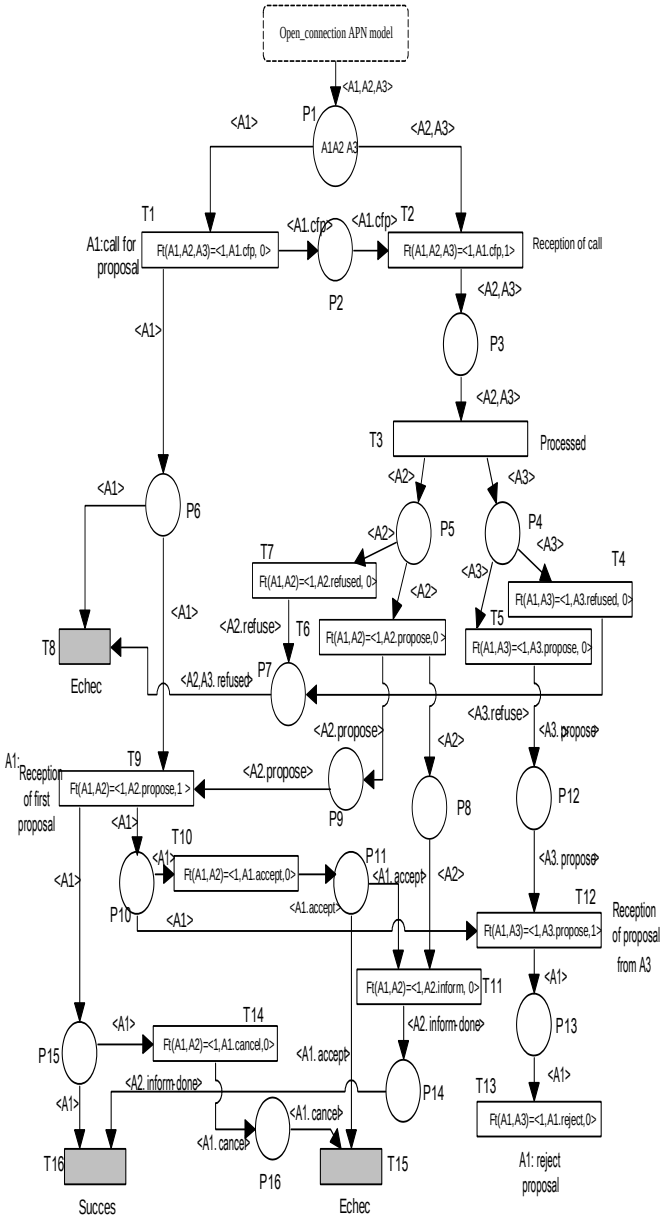


Fig.10. APN model for FIPA-Contract Net Protocol

IV. Example: Using Interaction Model To Ensure The Coordination Between Agents.

This section intends to stress the versatility of multi-agent systems and its implementation with the interaction schemes. We propose an APN model for the process of coordination in multi-agent systems based on interaction APN model (Figure. 11). The choice of modeling the coordination is done considering the importance of preparation for communication between agents. The resulting model helps agents to successful communication (cooperation or competition).

APN model for coordination between Agents based on APN Interaction present a set of rules or the process using to ensure that the common goal should be controlled by implementing of correct protocols.

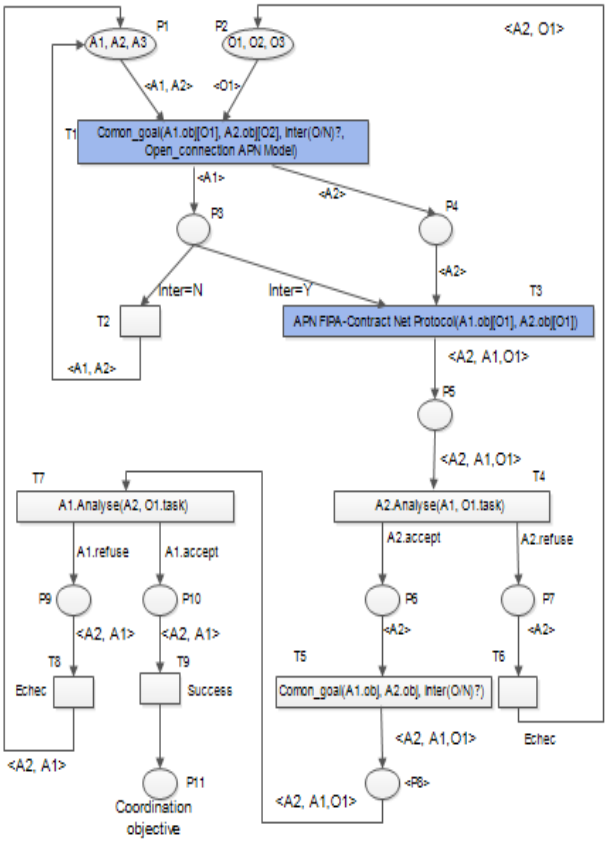


Fig.11. APN model for Coordination

List of places:

P1: Set of Agents in MAS.

P2: Set of objectives which may be subject to coordination between agents.

P3: The state indicating the compatibility of the objectives between two agents in coordination.

P4: The state indicating the presence of agent involved in the coordination to know the result of his study of the similarity with the objective that initiating by the second agent.

P5: Reception of the comparison result for coordination with the initiator agent by the participant agent. This comparison is a logical type (yes / no). If the comparison gives "yes" we can complete the process of comparison otherwise we cannot complete because the objectives of the two agents are not compatible.

P6: Acceptance of the participant agent to coordinate with the initiator about the objective studied.

P7: The refusal of the participant agent to coordinate with the initiator about the objective studied.

P8: The reception by the initiator of the agreement of the participant agent for the treatment about coordination.

P9: The cancellation of the coordination by the initiator agent because of the evolution in environment.

P10: The end of coordination with success between the two agents.

List of transitions:

T1: The study of goals compatibility between the two connected agents. This study was made by R1. Which may lead to good result and both agents, after comparing objectives, cannot complete their coordination if we move to a new coordination between other agents. The connection is created by Open_connection APN model between A1 and A2 to assure the objective O1.

In addition, we detailed our APN model FIPA-Request specifying the different exchanged between the two agents

T2: The release of coordination about the goal led by A1 with initiator agent A2.

T3: A1 informs A2 by the possibility of coordination about a task given to the compatibility of their goals. We apply APN FIPA- Contract Net Protocol.

T4: The treatment of the coordination demand about a specific task received from A1. This request can be accepted or rejected by A2.

T5: It is the acceptance of the request coordination proposed by A1.

T6: The refusal of the request coordination proposed by A1. T7: The treatment of the ability to continue coordination with A2 on the task indicated.

T8: It is unable to complete the coordination about the task chosen by A1.

T9: Is the achievement of the objective of coordinating set by A1.

V. CONCLUSION

In this paper, we introduce a model for specifying Interaction protocols in MAS based on APN. Our framework combines a FIPA Protocols approach to specify interaction protocols of agents with coordination. So, our model is meant to compete with existing FIAPA protocols for it applies to mediated interaction scenario rather than direct. It is undeniable that the use of interaction protocols for conversations greatly facilitates the development of systems based on communicating agents. We believe that the limitations inherent in other formalisms described necessitate the use of a formalism supporting competition and factorization for modeling such complex and competing interactions.

The major contribution of APN model is the power expressing based on agents. This formal method can verify correctly the interaction between them by specifying the messages exchanged during the conversation and during interaction. Some issues remain open for future developments, such as parameterization of protocols. For example, during an auction, how long an agent is permitted to wait before performing task? Otherwise, we can extend our model by incorporating a timeout mechanism and exception handling to avoid blocking during conversations.

REFERENCES

[1] Augeraud M., Collé F. and Sarramia D., "Design center interaction: Application to the design of interactive simulation", in International Conference on INFORSID'06, Hammamet, Tunisia, Juin 2006.

[2] Burkhard H.D., "Liveness and Fairness Properties in Multi-Agent Systems", International Joint Conference on Artificial Intelligence IJCAI'93, Chambery, France, 1993.

[3] Chaib-draa B., Jarras I. and Moulin B., "Multi-Agent Systems: General principles and applications", in J. P. Briot et Y. Demazeau "Agent and Multi Agents System", Hermes, 2001.

[4] Charif Y. and Sabouret N., "Interaction protocol for service composition in the room", JFSMA'06, Annecy, France, pages 253-266, 2006.

[5] Chicoisne G., "Dialogue between natural agents and artificial agents: An application to virtual communities", PhD thesis, National Institut of Polytechnique of Grenoble, pages 71-74, 2004.

[6] Foundation for Intelligent Physical Agents. FIPA 2001, "Specification : Agent Communication Language", Available at http://www.fipa.org.

[7] Greaves M., Holmback H. and Bradshaw J., "What Is a Conversation Policy", In LNCS, Issues on Agent communication, Vol. 1916, 2000.

[8] Hernandez I.R., "Modeling, formal specification and verification of interaction protocols: a role-based approach", PhD thesis, National Institut of Polytechnique of Grenoble, pages 66-86, 2004.

[9] Koji M., Jin-Hua S., and Yasuhiro O., "Study on Common Coordinate System by using Relative Position of Other Autonomous Robot"., SICE Annual Conference, Japan, August 20-23, 2012,

[10] Koning J.L. and Pesty S., "Communication patterns in principles and architectures of multi-agent systems ", IC2 Collection, Hermes Science Publications, Paris, 2001.

[11] Koning J.L. and Hernández I.V., "Generating Machine Processable Representations of Textual Representations of AUML", in Third International Workshop, AOSE'02, Bologna, Italy, LN CS, Vol. 2585, Springer, 2003.

[12] Mazouzi H., "Engineering of interaction protocols: distributed multi-agent systems to systems", Ph.D. Thesis, University of Paris IX-Dauphine, France, pages 83-85, 2001.

[13] Marzougui. B., Hassine. K., Barkaoui. K., "A New Formalism for Modeling a Multi Agent Systems: Agent Petri Nets", Journal of Software Engineering and Applications (JSEA), vol. 3(12), pp. 1118-1124, 2010.

[14] Marzougui. B., Hassine. K., Barkaoui. K., "Method for Verification of a Multi Agents System", Intelligent Systems, Modelling and Simulation (ISMS), Page(s): 62 – 65, 2011.

[15] Marzougui. B., Hassine. K., Barkaoui. K., "Modeling Migration of Mobile Agents", Business Process Management Workshops, volume 132 of Lecture Notes in Business Information Processing, page 530-540, Springer, 2012.

[16] Pauchet A., "Cognitive modeling of human interactions in a collaborative planning framework", PhD Thesis, University of Paris IX-Dauphine, France, page 117-118, 2004.

[17] [Sib02] Sibertin-Blanc C.,"A layered model for the engineering of interaction protocols", Formalizing Competing Activities, FAC'02, Toulouse, France, March 2002.

[18] [Seg00] El Fallah Seghrouchni A., "Coordination of agents: models and algorithms protocols", HDR research, University of Paris 13, pages 64-66, Janvier 2000.

[19] Singh S., Kearns M. and Littman M., "Graphical models for game theory", In Proc. of the 17th Conference in Uncertainty in Artificial Intelligence, pages 253-260, Morgan Kaufmann Publishers, 2001.

[20] Smith R. G., "The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver", IEEE Transactions on Computers, Vol. C-29, No. 12, 1980.

[21] Stephen C., Martin P., Mariusz N., Peter H, Ontologies for Agents: Theory and Experiences:Ontologies for Interaction Protocols, Whitestein Series in Software Agent Technologies 2005, pp 1-17

[22] Stefania B., Sara M. and Giuseppe V., Agent Based Modeling and Simulation: An Informatics Perspective Journal of Artificial Societies and Social Simulation **12** (4) 4, 2009.

Analysis of Child Computer Interaction in Edutainment and Simulation Games Application on Android Platform in Indonesia

Setia Wirawan
Faculty of Computer Science and Information Technology
Gunadarma University
Depok, Indonesia

Faris Fajar Muhammad, Lutfi Dwi Saifudin, Mustafa Ibrahim
Faculty of Computer Science and Information Technology
Gunadarma University
Depok, Indonesia

Dewi Agushinta R.
Faculty of Computer Science and Information Technology
Gunadarma University
Depok, Indonesia

***Abstract*—Child Computer Interaction (CCI) has become a challenge in utilizing the technology as education media. The increasing number of children, who use advanced gadgets in Indonesia such as smartphones and tablet PCs, provides a new space for developing interactive educational game application for kids. Indonesia is a country with the biggest number of Android-based game application downloaders in the world of service providers. Modeling serious game that has been chosen to deliver the concept is Edutainment and Simulation Games. This paper will analyze and review the application of the two concepts of the game, using data on the ranking from one of the top application service providers in analytic applications. The game application developers are expected to understand CCI and develop applications concepts that suit the needs of children in Indonesia. This will create market opportunities in the Indonesian game industry in the future.**

Keywords—Child Computer Interaction; Edutainment; Game application; Simulation Games

I. INTRODUCTION

In the recent years, the use of information and communication technologies used by the children is increasing. Especially with the development of advanced gadget with a variety of forms. Ranging from smartphones, to tablet PCs, they dominate the market with a lot of communication tools. Business people in information technology domain have also begun to take special care in the kids market.

This change provides an opportunity for the developers to start developing applications that fit the diverse needs of all players. Both in terms of usability, and in terms of user satisfaction. This has brought about the development application to the generation of more interactive and allows the interaction of children and computers to be studied. This interaction is called the Child Computer Interaction (CCI). The interaction process combines the input and perspectives of. several scientific disciplines that support research and industrial aspects related to interactive systems for children. CCI is defined as" the study of Activities, Behaviours, Concerns and Abilities of Children as they interact with computer technologies, often with the intervention of others (mainly adults) in situations that they partially (but generally do not fully) control and Regulate" [1, 2].

Child Computer Interaction (CCI) is a sub-science of the development of Human Computer Interaction (HCI), with an average age of interest in study of children ranging between the ages of five to twelve years [1]. The discipline of psychology is also emphasized to establish the mindset and behavior of childrens to use the system architecture designed. Contribution of science and language arts also affects how the display allows children to explore the information and instructions to be carried out.

CCI research used many approaches to the art of design, ergonomic characteristics, as well as the interactions that occur. However, many implementations of this approach is not suitable for children. It is based on the different abilities and expectations that each user get [3]. There are many designers created the application design that are not oriented for child user, experience in playing an application. If in the first time a child cannot use the application, the developer is deemed to have failed in developing the application [4].

Basically, there are three main categories that support the development of children. Those are Cognitive Development; Physical Development and Social/ Emotion Development [3]. Cognitive Development related to the processing of information, intellectual, and mental health of a child. Physical Development related to the coordination of movements that occur during the interaction between the child with the system, particularly with the existing interface. While Social/ Emotional Development describes how the socialization and emotional processes engaged by the user in response to the interaction. Whether positive or negative, the influence of these three will determine how the design of the application is made. Although it is not dominant and balanced in its implementation in the application for the child. These three aspects should be contained in it. This is what creates the characteristics of each application used.

One of the most common methods used to approach the CCI is a game application. Especially revealed that play is very important element for healthy child development [5]. The implementation has been realized by many applications that are developed in android platform. The emergences of many models are applied to make the demand for utilization of instructional media used in the application. Edutainment and

Simulation games become the main aim for developer as children like to play in. This article analyzes how the children in Indonesia can play both the current concept of the application.

II. CCI IN GAME APLICATION ON ANDROID PLATFORM

The development of games application on android platform, cannot be separated from the development of mobile industry since the first era of mobile phone was popularized in the end of the 20th century. At that time, mobile application only consists of a few simple games and applications to support mobility such as calendar, alarm, calculator, etc [6]. Fast development in the mobile industry creates an evolution of the mobile content and mobile application. Until now mobile phone has its own operating system. Platform as a special term has change the standard environment for mobile phone to a more powerful gadget that support human activity. This change also give a drastic impact to the development of game application.

Today, the ease of the touchscreen also gives an influence to the application developer. Design interface, game concept, or either the variant of feature have been added to the application to reach all the user, especially children. In the development, the design assumption of the application which support CCI for the children, only choose an interface that suitable in adults. By adding some animation, and brightful color make automatically suit to the children [7]. But, actually it is not as easy as we just figure out. The demand of using game application media has changed, not only educative, but also has some aspect that children need.

Serious games are considered as interest thing to talk to be developed in learning CCI because it designed to give effect to the user beside the pure aspect, which is entertainment [8]. Study process also has to be covered well. Edutainment and Simulation games are two type of serious game concept which is the most often developed. Both in terms of usability, and in terms of the density of existing knowledge.

A. Edutainment Games

Conventional education process used to deliver the knowledge. Limited human resources and time have become a constraint to deliver information to children. The new alternative way is needed to do study activity by entertaining and giving happiness without any assistant. Edutainment application becomes the new concept from it paradigm.

Edutainment refers to any kind of education that also entertains even though it is usually associated with video games with educational aims. The primary target group was preschool and young children [9]. Indonesian children from 5 years old until 7 years old become their users target. In general, edutainment represents the basic knowledge which can be used in life in the society with focus on reading, mathematics, and science [9]. Edutainment applications are more support to cognitive development of a child, by using interesting and easy to understand picture. This is because of edutainment only developing low level thinking.

The phenomenon of Edutainment signed by the increase of relation between needed for education and expectation of the quality of entertain given [10]. There are many ways to present Edutainment application. Edutainment representation can be divided from its purpose and its content: for improving user life control (informal education), which usually presented by a discussion or a narration. And the second one is giving user an experience (skills education), which is usually presented with such experience simulation [11]. “Alphabet Find” are one of application who represented edutainment application.

Fig.1. Application “Alphabet Find” [12]

The information learning process to a child in edutainment application will be created as easy as possible and simple as well. So, the children will be able to use the application easily. As we can see in the figure 1, the use of letter and number plus some colorful color effect help children to understand what the information is about. When the child understand what is the purpose of the application they played, then the game will be interesting to play from the point of player view. However, edutainment software will fail since it resulted in what has been described as “boring games and drill-and-kill learning” [13].

B. Simulation games

If viewed from certain perspective, the reaction for making decision based on condition will be different for each user. User experience and abilities will take a big part in how the user will react when facing the issue. To make studying those issue more effective, simulation model was created. Literally, simulation is popularly defined as a partial representation of reality which selects crucial characteristics of a real situation and makes a replica of them, within an environment or place which is basically free of risk [14].

With the potential exists inside gaming industry, many professionals and tutor attracted to simulation games. Firstly, there is a big change in the field of conventional teaching. It is the development from “learning by hearing” process to

"learning by doing" [15]. Moreover, Simon (1996) note "how we learn already change from memorizing information to find and use the information player found". An interactive simulation will give children chance to face the issues directly and solve it with their own way. The second reason is some prove from empiric study claimed that game can become a tools to improve learning a complex material [15]. And the last one that makes this concept interesting is simulation game is more effective than the other learning method because at the same time it trains affective and cognitive process [16]. An example for game application based on simulation game is social reality, driving simulation, operation simulation, etc.

Fig.2. "Paradise Island" application [17]

In the figure 2, game aplication "Paradise Island" described a condition close to reality used for training media. There is an island player can manage and organize their resource. By managing those two things, user can learn how to implement those condition in reality. Player can learn how to think logically.

III. Model Analysis and Assumption

This article uses data from a leading analytic applications. From these data we are sorting it based on initial concepts that will be the first presumption by the children. This is because we see that some kinds of applications is very difficult to find in some literatures. We use the top ranked 50 on a game application that we expect to represent users in Indonesia. The initial concepts used as a starting point to see how the presentation of the concept of Edutainment applications and applications that are presented with the concept of simulation games will be shown on table 1.

As one form of analysis, presentation of the results of the classification form Edutainment application shown in Fig 3.

Based on the literature edutainment game applications, these applications are based on the design of a lightweight and straight to the main of the desired concept.

TABLE I. LITERATURE CONCEPT

Edutainment Games Applications Presentation	Simulation Games Application Presentation
Implementation in the form of elementary instruction for children.	Implementation in the form of modelingreal events.
Children are required to follow rules and instructions.	Children are required to explore the modeling that has been provided.
Game application presents a counting process in mathematics, reading and science.	Game application presents thef process of creativity and problem resolution in accordance with the wishes of the player.
Children are required to know directly what the application presented	Children are required to see what the problem represented

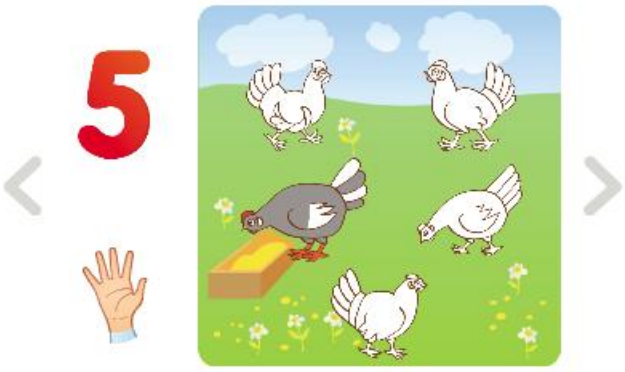


Fig.3. Example of edutainment game application [18]

Introducing numbers with the use of animal images that have been provided for imaging a portrait of five numbers. This application is also intended to stimulate children to use fingers to count. When viewed closely, there is a navigation to continue or return to the other to study the numbers. Based on this, the form of presentation of the application is classified as edutainment applications.

Then, the presentation of simulation games are seen on pieces of game play showed in Figure 4. One ice cream shop modeling simulation was chosen to be used by this application. The children will be invited to see a store with a lot of dessert food.

When kids play this game, then freedom in managing resources will depend on what will the player does. Existing learning theories in this game put emphasis indirectly on the development of a child's creativity in dealing with a case/ matter by solving the problems in the shop. Presumptions about the shop's children are expected to learn how to take care of the store with exploring added to the store.

Fig.4. example simulation games application presented [19]

IV. RESULT

Based on the guidelines presented in Table 1, the results shown in Fig 5.

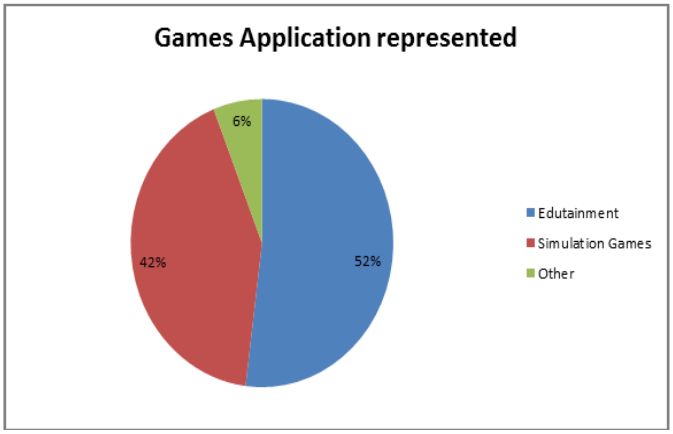


Fig.5. Games Application Represented Diagram [20]

The data who were calculated ,got from Distimo Analytics for SamsungApps for free application in Indonesia, may 23 2013. As shown in Fig 5, the presentation of applications nearly balanced. 52% more game applications are presented in the form of edutainment, while 42% more applications are presented in the form of simulation games. From a total of 50 applications for children who play more often downloaded, half of the entire application is presented to convey learning numbers and letters.

Broadly, the one with most quantity is an application that use the concept of edutainment. Then, speculation about the development of children in Indonesia can be seen. If viewed from how the information processed, such as numbers and letters. Then the conventional learning basic science can not be assumed to provide maximum results. Thought the parents of the importance of basic science studies, making this application becomes one of the needed media. The use of advanced gadget and edutainment applications as a replacement media learning outside of school, seen as an opportunity. This become a chance to learn CCI. Particularly, children have been introduced to the computer environment through the application of edutainment games. So that children can adapt to advanced gadgets faster than they should.

Seeing almost the number between edutainment applications and simulation games nearly equal, it would seem the trend in Indonesian children have been taught how to make decisions early on. Creativity, interests and talents can help children to explore and imagine by using the knowledge they just get. This should have the same goal with the approach to model the children mindset. CCI in Indonesia has entered into a process in which how to create designs provide convenience to the child to solve problems and make decisions. Simulation games have been seen as a bridge linking the transition from basic science to the science management. Thus, the evolution of children's development in Indonesia is better than the previous generation.

Although the form of the presentation of each applications different from each other, the second presentation of the applications contained in the data used is almost the same. Especially the design complexity and the interaction that occurs between children of the application used. We found that the offered level of complexity are depending on what they see (What you see is what you get). From many simulation games applications observed, developers find it is difficult to develop a range of knowledge to be conveyed. So to present simulation games and the experience in it, they insert edutainment concept in order to ease the adaptation.

If viewed from the age of the players that play applications, it can be assumed that the application of simulation games will be played after edutainment applications and quality of interaction built when using these advanced gadgets. Industrial applications will consider the type of game players and the number of players who will be used. So first they will consider developing edutainment applications compared to rather than simulation games. So, that the number of applications is dominated by widely played edutainment applications.

V. CONCLUSION

From the result showed, it is considered as a new step to see how the process of interaction between children and computers occurs in sophisticated equipment in Indonesia.

From the above analysis, it concludes that the process of Child Computer Interaction can be seen from the number of games applications presented in the form of edutainment and simulation games. Edutainment games and Simulation Games have different forms of interaction, but they have the same goal, namely compressing the knowledge that the developers want to convey, especially in the number of players who use their applications. From many assumptions seen from the number of both applications being played. One form is how part of the application is able to provide new opportunities for studying the CCI for children in Indonesia. By developing one form of the application, mobile application industry opportunities is to prepare interactions that can build the quality of Indonesian children mindset, where this could be done by developing a game edutainment applications and simulation games. Given that the two types of market for game developers increasingly widespread applications. Not mentioned the development of interfaces in the developed application, too.

CCI's development itself in Indonesia, as seen on the chart showed that edutainment is still more popular than simulation games although simulation games still lost by edutainment. The difference between simulation and edutainment games are not too much different. This indicates children in Indonesia began wanting to play simulation games. And show that CCI in Indonesia keep growing to fit the international market.

REFERENCES

[1] J.C Read and P. Markopoulos, "Child-Computer Interaction," in:International Journal of Child-Computer Interaction, 2012

[2] J.C. Read and M.M. Bekker, "The Nature of Child Computer Interaction," in: Paper Presented at the HCI 2011, Newcastle, UK, 2011.

[3] Sonia Chiasson and Carl Gutwin , "Design Principles for Children's Technology," 2005.

[4] Amy Bruckman and Alisa Bandlow, "HCI for Kids," Published in: The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, 2002.

[5] K.R. Ginsburg, "The importance of play in promoting healthy child development and maintaining strong parent-child bonds," Published in: Pediatrics 119 ,pp 182-191, 2007.

[6] John F. Clark, "History of Mobile Application," Unpublished

[7] A. Druin, Children as Our Technology Design Partners, in The Design of Children's Technology, Allison Druin ed., Morgan Kaufmann, 1999.

[8] Francesco Bellotti et.al , "Assessment in and of Serious Games: An Overview,"2013.

[9] Jiří Němec, Josef Trna, "Edutainment or Entertainment.Education Possibilities of Didactic Games in Science Education," in: The Evolution Of Children Play - 24. ICCP Word Play Conference, pp. 55-64, 2007.

[10] Tarja Susi, Mikael Johannesson and Per Backlund," Serious Games – An Overview", 2007.

[11] Sari Walldén and Anne Soronen, "Edutainment From Television and Computers to Digital Television." Tampere:University of Tampere ,2004.

[12] Hadron Solutions India Pvt. Ltd., "Alphabet Find", http://img.samsungapps.com/product/2013/0410/000000579673/IconImage_20130410124226086_NEW_WEB_SHOT2_HALF.png,2013.

[13] R. Van Eck, "Digital game based learning: It's not just the digital natives who are restless". EDUCAUSE review, march/april, 2006, pp. 16-30.

[14] P. Hodson, M. Connoly and D. Saunders, "Can computer-bases learning support adult learners?," Journal of Further and Higher Education, 2001, pp. 325-335

[15] Rosemary Garris, Robert Ahlers, James E. Driskell, Games, Motivation, and Learning: A Research and Practice Model. 2002.

[16] R. D. Tennyson and R. L. Jorczak, "A conceptual framework for the empirical study of instructional games," in H. F. O'Neil & R. S. Perez (Eds.), *Computer games and team and individual learning*, 2008, pp. 39-54). Oxford, UK: Elsevier.

[17] GIGL, "Paradise Island", https://play.google.com/store/apps/details?id=com.seventeenbullets.android.island&hl=en,2013

[18] Oleg Alexandrovich Demyanov , "Math and Number of Kids Lite", http://apps.samsung.com/mercury/topApps/topAppsDetail.as?COUNTRY_CODE=IDN&productId=000000567698,2013

[19] Quanshuang Qin, "Ice Cream Shop", http://apps.samsung.com/mercury/topApps/topAppsDetail.as?COUNTRY_CODE=IDN&productId=000000586654, 2013.

[20] Distimo Analytics ,http://www.distimo.com/leaderboards/samsung-apps/indonesia/kids/free for statistic 23 may 2013.

Security Concerns in E-payment and the Law in Jordan

Mohammad Atwah Al-ma'aitah
Mis Department
Al-balqa Applied University
Amman - Jordan

***Abstract*—Recently communications and information technology became widely used in various aspects of life. The internet becomes the main network for information support. Using of internet enabled public and private organizations to develop its business and expand its activities. Private organizations applied the principles of e-commerce to improve the quality of services which provided to customers. While public sector organizations started to apply the principles of e-government in an effort to increase efficiency and effectiveness and achieve maximum equality among citizens. One of the major challenges raised by widespread use of e-government and e-commerce application is security issues especially e-payment. This paper discusses the present law in Kingdom of Jordan which deal with the problem of frauds and violation of consumers' rights and privacy when they making e-payment. In addition this paper tries to make comprehensive study on e-payments and the law to decide if there more legislation is needed.**

***Keywords*—E-payment systems; Cyber crime; Web security; Law.**

I. Introduction

E-payment systems (EPS) have become a most important factor in the growth of electronic commerce and e-government application. In addition EPS system may determine success and fail of these applications. An electronic payment system is an essential part in new business-to-consumer and business-to-business e-commerce [1] [2] [3]. The development of e-business contributed considerably to the development of electronic payment systems so as to meet the needs of e-procurement processes and to facilitate the completion of the transactions [4]. Tsiakis and Sthephanides argued that "the security and trust issues that are essential for every electronic payment mechanism in order to be accepted and established as a common medium of financial transactions" [5]. As a result of this development in this area, the diseased soul's owners were increased to ambitions to commit cyber crime and penetrate privacy of others especially that the numbers of workers in the field of information were increasing. Then more legislation is required to deal with these crimes.

II. Literature Review

A. Evolution of Electronic Payment

Organizations were used to deal with financial dealings in the traditional way such as paper work. But with beginning of communication and internet technology most of financial procedures dealt with it electronically. On the other hand the appearance of internet and the raise of electronic communications technology impact considerably the growth of E-payment [6].

B. Electronic payment systems (EPS)

The desired e-payment service appears after e-commerce shaped [7]. E-Payment is defined as "the transfer of an electronic value of payment from a payer to a payee through an e-payment mechanism" [8]. Prior researcher such as [7] [9] [10] categorize the instruments used for e-payment as shown in Fig (1).

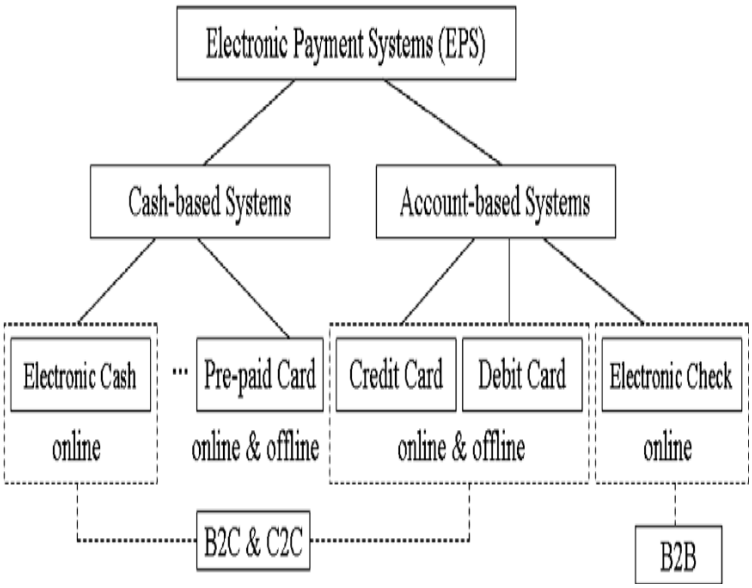


Fig.1. Classification of electronic payment systems: [7] [9] [10].

In addition [7] argued that " Electronic-cash, pre-paid cards, credit cards, and debit cards are widely used in B2C and C2C EC". According [11] he defined e-money as "Broadly, e-money is defined subject to exceptions as monetary value represented by a claim on the issuer that is (i) electronically stored; (ii) issued on receipt of funds for the purpose of making payment transactions; and (iii) accepted as payment by someone other than the issuer". New mechanisms discussed by [12] in deeply which called Web ATMs. Via this new technology the payment of transaction can obtain and send right in several seconds. It is a really suitable way of online payment tool enabling clients to assign or transfer funds, effect payments, and make account inquiries 24 hours a day, all year

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

round. In addition they argued that "With Web ATMs and Web 2.0, we now have an opportunity to make amends. Modern electronic financial services will be more flexible and modular, allowing capabilities to be added as and when required on Web. Such safe payment services will not only extend their sales opportunities but also benefit to the whole economic benefits, cost-effective of banking industry. Most importantly, this does mean that there is a big breakthrough on the innovative payment instrument of money flow of e-commence".

Some Suggested steps by [13] that have to be taken by organizations to protect personal information in the particular security risks within organizations. In determining appropriate security measures, organizations should:

- Identify the security risks to the personal information that is being held.
- Build up policies and procedures to reduce those identified risks.
- Apply suitable IT security settings governing system access; and
- - Monitor and measure performance against relevant Australian and International standards

Ma'aitah and shtat [14] agreed that Importance of Authorization and Importance of encryption influence the perceived security of E-finance transactions; these features can contribute toward enhancing the perceptions of the users that the web and online transactions including E-Finance transaction are secure, and encourage them to use the online system and do financial transactions online.

Tsiakis and Sthephanide [5] discussed the concept of security and trust in electronic payments. Although they discussed some of requirement and properties necessary to build successful electronic payment systems as listed follow:

1) Integrity: confidence that information has not been changed after the data was signed.

2) Authentication: The process by which one person verifies that another entity is who they claim to be.

3) Fraud prevention and tolerance: prevention of parties from fraud and from economic losses in the case the system crashes or the network fails.

4) Privacy: information must not be revealed to not permit people.

5) Divisibility: option of numerous denominations.

6) Transferability: spending of token without the need to contact the issuer.

7) Payment anonymity: the payee will know only pseudonym of the payer

C. Web Security

Web security means the capability of the web to maintain and protect the individual sensitive information from any changing, waste, disclosure, destruction or in use by unauthorized persons such as Internet intruders andhackers. The web security system must avoid unauthorized users to use the computer system and manage access to the network from inside and outside the organization [15], [16].

Security is the life of E-Commerce and it has grown to be the most significant concern for its expansion [17]. However the two most vital areas preventing the successful implementation of E-Commerce worldwide are the Internet methodologies and E-Commerce dealings security

III. E-PAYMENT AND SECURITY THREATS

The lack of security measures still plays as a major risk factor in electronic payment process. Hence; the privacy of personal data on the level of the individual or institutions is not attained. This will lead to loss of individual data or spy on financial and administrative processes. Nevertheless, barriers could arise in term of buying and selling through electronic means such as the abuse of the privacy of individuals and the lack of electronic laws structure that are available in many countries to detect them from such a great danger as piracy and spy ware.

In a result of that people started to be afraid of electronic payment even though it is equipped with a lot of advantages been offered. This was confirmed by a study that was conducted by the researcher

A survey questioner was distributed with a random sample of 300 respondent in which different levels of instruction was applied to the questions to focus on the reason behind the lack of use of electronic payment (buying and selling) to answer whether a lack of security, privacy and laws are necessary to protect the operations of electronic payment. The following results show that there is a significant relationship between the lack of security, privacy and laws are necessary to protect the operations of electronic payment see Table below I:

TABLE I. Response answer

1. 62%	answered	yes
2. 29%	answered	no
3. 9%	answered	do not know

All findings obtained refer that respondent's fear of using the electronic payment concept because of the security and privacy of data, nevertheless; the increasing number of hacker's and data penetration possibility which transferred through the Internet.

Nowadays; technological innovation has played a major role in the development of business. One major innovative tool is the Websites. Websites is a great technological innovation that changed the way we do business [18]. Because of The increase use of Websites, millions of virtual stores are already available and the number is increasing dramatically. Unfortunately, parallel numbers of cyber crimes done by hackers are also accruing. Website security is still an open issue since we lack comprehensive international laws against cyber criminals. If laws do exist against hackers, the implementation is quite hard considering the complexity of the cyber world

Hackers refer to any computer fanatic who breaks into computer systems for the sake of gaining information or

conducting playful mischief and stealing or corrupting data [19]. In relation to the previous definition hackers are everyone feels that the security forces are not yet able to provide reasonable protection for internet users. Hence, the aim of this research focuses on the legal role in the protection of data privacy and security through electronic payment.

IV. Law regarding Security of e-payments in Jordan

In the case of Jordan, like many developing countries, a lack of legal legislation for the protection of any form of electronic payment. At the same time taken in consideration cyber crime, increase with the information technology revolution and the use of networks for data transmission between individuals and institutions that form a real threat to all users. Therefore a special law has to be issued to protect the privacy of citizens and institutions and protection of their moral and financial.

A. Cyber crime can be divided into two main types as follows:

- Crime committed through the use of electronic media as a tool to commit the crime as it is threatened by electronic means or slander and libel and others.

- The crime that targeted electronic means or their contents as the destruction of an information system or steal information from a site or network or electronic information system or infringing on privacy and confidentiality.

B. Jordan Information Systems and Cyber Crime Law:

On 16/9/2010 issued by the Government of Jordan Information Systems Crimes Act No. 30 of 2010 meaningful issuing this law limit the crimes that occur through the use of the Internet, which use spread due to the large and rapid development in the field of Communications and Information Technology.

The Information Systems Crimes Act on maintaining the rights and privacy of personal and financial rights as well as all that would affect the security and stability of the country by selecting the sanctions to curb violations and / or the excesses of the Internet users.

Noting here that the enactment of such a law it is absolutely imperative that the Jordanian Penal Code, as amended text in Article III that "no offense, but the text does not spend any penalty or measure did not provide for the law to them while committing the crime, and is considered the crime completely if has actions implemented without of time to get the result, "meaning no crime or punishment except the text so it has assigned men of law the validity of acts considered crimes inventory and determine the punishment through the specific provisions into law

C. Crimes addressed by Jordan Information Systems and Cyber Crime Law:

Law deals mainly some electronic crimes that have evolved and became her legal terms set it apart from traditional crimes legislation does not criminalize traditional, and of those crimes:

1) *Penetrate the e-mail site.*
2) *Piracy.*
3) *Denial of Service.*
4) *Illegal access to information system or network computers.*
5) *Send viruses.*
6) *Sabotage devices and systems remotely, or using electronic systems.*
7) *Destruction of information that is available to the public affecting national security.*
8) *Steal information stored by electronic means.*
9) *Impersonate capacity.*
10) *Change the information stored or transmitted by electronic means.*

The law has been divided into 17 articles were allocated legal articles 5, 6, 7 electronic crimes and states:

1) *Article 5:*

Any person who intentionally captures or intercept or eavesdrop on what is sent through the Internet or any information system shall be punished by imprisonment for not less than one month nor more than one year or a fine of not less than (200) two hundred dinars and not more than (1000) thousand dinars, or both penalties.

2) *Article 6:*

A. All of intentionally got without a permit through the Internet or any information system, data or information related to credit card data or information to be used in the implementation of financial transactions or electronic banking is punishable by imprisonment for a term not less than three months and not exceeding two years or a fine not less than (500) five hundred dinars and not more than (2000) thousand dinars, or both penalties.

B. All of the used through the Internet or any information system intentionally without a legitimate reason data or information regarding the credit card or the data or information to be used in the implementation of financial transactions or electronic banking for himself or to other data, information or money or services belonging to others punished imprisonment for a term not less than one year and a fine of not less than (1000) thousand dinars and not more than (5000) five thousand dinars.?

3) *Article 7:*

Doubled the punishment for the crimes stipulated in Articles (3) to (6) of this law, the right of every person who committed any of them while doing his job or his work or exploitation of any of them.

V. The Official Security Agencies that Deal with Cyber Crime

There is in the General Security Directorate Jordanian administration called CID and U where a special section offenses electronic named Department computer crimes and their website is (http://www.cdd.psd.gov.jo), and through correspondence, official and field visits to this section got excellent results I did not expect that we in Jordan are dealing security with this volume of electronic crimes were as shown in table(II):

TABLE II. The number of cases from 1999 to 2013.

No.	Year	Cases No.	Samples No.	Detected ratio
1.	1999	5	13	71.4%
2.	2000	28	452	63.6%
3.	2001	29	179	77.7%
4.	2002	36	7756	73.5%
5.	2003	43	145	71.4%
6.	2004	56	125	63.6%
7.	2005	66	723	71.4%
8.	2006	74	813	62.9%
9.	2007	84	85	65.6%
10.	2008	96	975	83.3%
11.	2009	105	753	71.4%
12.	2010	111	604	75%
13.	2011	113	347	78.5%
14.	2012	167	262	60%
15.	2013 To date	84	160	56.3%
Total		1097	13392	-

Through the above table which we have collected from the Department of Electronic Crimes and found that it received from several quarters, namely, (CID, preventive security, counter-narcotics, protection of the family, the tourist police, legal affairs, crime scenes, the courts, the General Intelligence Department and Customs).

It was one of the most important issues that deal with the following section:

The issue of ATM card theft by devices have been added to an ATM machine and revealed through the crime scene downtown Amman.

The issue of counterfeiting of currency: and this case had been received from the Drug Enforcement Administration include computers and flash memory belonging to suspects in counterfeiting, forgery and using a program (Encase) confirms that the suspected counterfeiting banknotes.

The issue of electronic articles and comments insulting: Use Program (Email Tracker Pro) your track websites and in cooperation with the Telecommunications Regulatory Authority (TRA) is determined by a person who worked on the publication of an article or comment the abuse of others.

Fraud case in the Forex market Stock Exchange firms: where he worked as a suspect in this case to design Web pages and fake sites similar speculation the company's global Forex and using a program (Encase) devices suspects were examined by a court order.

Email issues and sexual exploitation of adults and children: have been discovered using the program (Email Tracker Pro) and the FTK program.

Indecent assault as a result of issues dating through social networking sites: was discovered using a program (Encase)

The issue of recognition of registration of murder: were discovered using a program (CSL) which used to matching voice to samples.

VI. Discussion And Conclusion

The results of this research demonstrate that there are a number of reasons for not using information and communication Technologies (ICT) such as electronic payment; despite of it is convenience, and saving time and efforts of users. First, one of the most important reasons this study confirms, is that the use of such technologies threats the security and privacy of information of users. Earlier to 2001, Jordan did not have any law to protect deal-mail. However, on December 11, 2001, the Jordanian legislator enacted the provisional law of Electronic Transactions No. 85. This was the first law within the package of legislation of information technology in Jordan, and the second Arabic law after Tunisia that concerns with electronic commerce. While this progress is considered important in this domain, it has been remained in the form of a temporary law.

Second, the issuance of this law was not enough to convince citizens to accept the use of electronic payment. That is, the practice of Jordanian customers was only limited to the use of ATM at banks and Visa cards to complete their purchases in different shops. In contrast, people in a number of developed countries such as US, UK, and Franc use electronic payment and do not use cash to compete their transactions. This indeed refers to the presence of contemporary laws that aim to maintain the security and integrity of the information and the privacy of individuals in all information and communication technologies (ICT) such as electronic payment.

Third, the law in its current content and application still unable to address all forms of threat that users of information and communication technologies face. This is due to the absence of clear and comprehensive texts that express all expected forms of electronic crimes, in addition to the lack of judges specialized in electronic crimes and lawyers in this area.

Fourth, some banks and companies in Jordan and particularly in the private sector point out that recently piracy became very common. For instance, there are some hackers who get into customers' accounts through the confidential numbers of their credit cards and transfer money from their accounts to other accounts.

This led Jordanian banks to maintain the reputation action settlement directly with customers whose accounts penetrated by hackers. Most banks which its accounts under threaten

suggest their customers, by sending them short messages to their Mobiles, to not give any figures or confidential information about their accounts to any party whatsoever.

Finally, security and privacy are considered two critical elements in dealing with electronic transactions in general and each other up. The lack of appropriate security means that privacy will be the victim and this leads to the unwillingness of citizens to accept the idea of dealing with the electronic commerce in general, and electronic payment in specific. Although the Jordanian law of 2010 addresses some of the threats related to the security and privacy of information through electronic transactions, it is still insufficient. That is, courts face difficulties to give its judgment on relevant issues within appropriate time framework, as some cases take years to get a definitive legal judgment that saves monetary and non-monetary rights of the victims.

With the above discussion in mind, this research based on monitoring the progress of some issues in the courts, recommends the following:

1) Prepare judges to be specialists in both fields: law and computer. This can be happened by adopting the Ministry of Justice to the idea of sending them to obtain their higher education in law with the main focus on studying contemporary issues related to the use of computer, networks, internet, information security, and databases. The total course hours should be designed in a way that enables these students to be familiar with the nature and different forms of cyber crime exist in the real life. This helps them to deal with different cyber crime issues in the courts.

2) Amend the current laws to cope with the development of using information and communication technologies (ICT) in all areas. This leads to enhance the confidence of institutions and individuals to use electronic payment, that became crucial in distinguishing between developed and developing countries.

3) Protect the rights of individuals and organizations by exerting all the important and required efforts and providing legal security. This also leads to increase the number of people using (ICT) such as electronic payment.

4) Used IP version 6 and web2.0 technology.

5) Enhance culture about computer crimes during public and private sectors

REFRENCES

[1] Katsaros, P. "A roadmap to electronic payment transaction guarantees and a Colored Petri Net model checking approach", Information and Software Technology, Vol. 51, (2009), pp: 235–257.

[2] Kousaridas, A., Parissis, G., and Apostolopoulos, T. "An open financial services Architecture based on the use of intelligent mobile devices", Electronic Commerce Research and Applications, Vol.7, (2008), pp: 232–246.

[3] Cotteleer, M. J., Cotteleer, C. A., and Prochnow, A. "Cutting checks: challenges and choices in B2B e-payments", Communications of the ACM, Vol.50, Issue 6, 2007, pp: 56–61.

[4] Simon, N., and Sutter, G. "Electronic Payments — the Smart Card", Computer Law & Security Report Vol. 18, No.4, pp: 2002.

[5] Tsiakis, T., and Sthephanides, G. "The concept of security and trust in electronic payments", Computers & Security, Vol.24, pp: 10-15.

[6] Shamim, F. "The ICT environment, financial sector and economic growth: a cross-country analysis", Journal of Economic Studies, Vol. 34, No. 4, 2007, pp. 352-370.

[7] Kim., C., Tao., Wang , and Shin, N. "An empirical study of customers' perceptions of security and trust in e-payment systems", Electronic Commerce Research and Applications, Vol. 9, (2010), pp: 84–95.

[8] Weir, C. S., Anderson, J. N., and Jack, M. A. "On the role of metaphor and language in design of third party payments in E-Banking: usability and quality", International Journal of Human-Computer Studies, Vol.64, No.8, (2006), pp: 70–784.

[9] Guan, S., and Hua, F. "A multi-agent architecture for electronic payment", International Journal of Information Technology and Decision Making, Vol.2, No.3, (2003), pp: 497–522.

[10] Dai, X., and Grundy, J. "Net Pay: an off-line, decentralized micro-payment system for thin-client applications", Electronic Commerce Research and Applications, Vol.6, (2007),Pp: 91–101.

[11] Kemp., R. "Mobile payments: Current and emerging regulatory and Contracting issues", computer law & security review Vol.2, No.9, (2013), pp: 175:179.

[12] Tsai., W, Huang. B, Tsaur. T, and Lin., S. "The application of Web ATMs in e-payment industry: A case study", Expert Systems with Applications, Vol.37, (2010), pp: 587–597.

[13] Kennedy., G. "Asia-Pacific news", computer law & security review, Vol.27, (2011), pp: 563:570.

[14] AL-ma'aitah, M. and Shatat, A. "Empirical Study in the Security of Electronic Payment Systems ", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No (2011).

[15] Hopwood, W. "Security in a Web-Based Environment, Managerial Finance", Vol.26, (2001), pp.42-52.

[16] Efraim, T. Michael, C.H. and Jae, L.K. "Electronic Commerce: A Managerial Perspective", Prentice Hall, 4th Ed, 2006.

[17] Jun., S. and Punit, A. "The more secure the better? A study of information security readiness", Industrial Management & Data Systems, Vol.111 Issue: 4, (2011), pp.570-588.

[18] Wells, J. Valacich, J. and Hess, T. "What signal are you sending? how website quality influences perceptions of product quality and purchase intentions", Journal MIS Quarterly, Vol.35, (2011), pp:373:396.

[19] Mookerjee, V., Mookerjee, R., and Bensoussan, A. "When Hackers Talk: Managing Information Security Under Variable Attack Rates and Knowledge Dissemination", Journal of Information Systems Research, Vol.22, Issue 3, (2011),pp: 606-623 .

Identification–Oriented Control Designs with Application to a Wind Turbine Benchmark

Silvio Simani
Department of Engineering
University of Ferrara
44122 Ferrara (FE), Italy

Paolo Castaldi
Aerospace Engineering Faculty
University of Bologna
47100 Forlì (FC), Italy

***Abstract*—Wind turbines are complex dynamic systems forced by stochastic wind disturbances, gravitational, centrifugal, and gyroscopic loads. Since their aerodynamics are nonlinear, wind turbine modelling is thus challenging. Therefore, the design of control algorithms for wind turbines must account for these complexities, but without being too complex and unwieldy. Therefore, the main contribution of this study consists of providing two examples of robust and viable control designs with application to a wind turbine simulator. Due to the description of the considered process, extensive simulations of this test case and Monte–Carlo analysis are the tools for assessing experimentally the achieved features of the suggested control schemes, in terms of reliability, robustness, and stability, in the presence of modelling and measurement errors. These developed control methods are finally compared with different approaches designed for the same benchmark, in order to evaluate the properties of the considered control techniques.**

Keywords—control algorithms; fuzzy modelling and control; recursive estimation; adaptive PI controllers; wind turbine model.

I. INTRODUCTION

Wind turbines are complex nonlinear dynamic systems forced by gravity, and stochastic wind disturbance, which are affected by gravitational, centrifugal, and gyroscopic loads. Their aerodynamics are nonlinear, and unsteady, whilst their rotors are subject to complicated turbulent wind inflow fields driving fatigue loading. Therefore, wind turbine modelling and control are challenging tasks [1], [2]. Accurate models should contain many degrees of freedom in order to capture the most important dynamic effects. Moreover, the rotation of the turbine adds further complexity to the dynamics modelling. In general, off–the–shelf commercial software usually is not adequate for wind turbine dynamics modelling, but special dynamic simulation codes are required. It is clear that the design of control algorithms for wind turbines has to take into account these complexities. On the other hand, control algorithms must capture the most important turbine dynamics, without being too complex [1], [2].

Today's wind turbines employ different control actuation and strategies to achieve the required goals and performances. Some turbines perform the regulation action through passive control methods, such as in fixed–pitch, stall control machines. In these machines, the blades are designed so that the power is limited above rated wind speed through the blade stall. Thus, no pitch mechanism is needed. Rotors with adjustable pitch are often used in constant–speed machines, in order to provide turbine power control better than the one achievable with blade stall. Therefore, blade pitching can be regulated to provide constant power above rated wind speed, in order to provide good power regulation in the presence of gusts and turbulence. Large commercial wind turbines can employ also yaw control to orient the machine into the wind. A yaw error signal from a nacelle–mounted wind direction sensor is used to calculate a control error. The control signal is usually just a command to yaw the turbine at a slow constant rate in one direction or the other. The yaw motor is switched on when the yaw error exceeds a certain amount and is switched off when the yaw error is less than some prescribed amount. Some recent control studies were addressed *e.g.* in [3], [4], [5], [6]. It is worth noting that the main disadvantage of these approaches consists of the need of an accurate model of the process under investigation, followed by the control design strategies, which usually require advanced mathematical methodologies.

On the other hand, this work describes the application of two control methods, which are quite direct and straightforward, as well as their testing through extensive simulations for a wind turbine prototype, which is freely available for the Matlab® and Simulink® environments [2].

In particular, the first proposed strategy consists of a scheme relying on a fuzzy identification approach to model–based control design. In contrast to pure nonlinear identification methods, fuzzy systems are capable of deriving nonlinear models directly from measured input–output data without detailed system assumptions, with arbitrary degree of accuracy. In particular, Takagi–Sugeno (TS) fuzzy prototypes are exploited, whose parameters are obtained by identification procedures from the data of the monitored process. The suggested fuzzy approach is motivated also by previous works by one of the same authors [7]. It is worth noting that the works by one of the same author [8], [9] presented a totally different solution to the design of the fuzzy regulators. In fact, even if the papers [8], [9] and the present study share the common fuzzy clustering methodology, this contribution focuses on the direct fuzzy *regulator identification*, whilst [8], [9] were based on fuzzy PI controllers, whose parameters were computed from the identified fuzzy prototypes.

With reference to the second control method proposed in this work, the application of an on–line identification mechanism in connection with a model–based adaptive control design is considered. This control scheme belongs to the field of adaptive control. On–line parametric model identification

schemes represent an alternative for developing experimental models for complex systems, such as wind turbine systems. Therefore, this paper suggests the implementation of controllers based on adaptive identification schemes, used for the on–line estimation of the controlled process, which can be affected by uncertainty and errors. The recursive Frisch extended to the adaptive case making use of exponential forgetting is considered here [10]. It also overcomes potential numerical difficulties with the existing recursive scheme. The ability of the adaptive scheme to track changes in the system parameters is exploited here in connection with the on–line computation of time–varying controller parameters, in order to maintain the required control performances. The use of this identification procedure is motivated by its easy integration into the Simulink® toolbox for the design of on–line controllers [11].

The effectiveness of the proposed control strategies has been assessed on data sequences acquired from the considered benchmark. Several simulation results show the achieved performances with respect also to different control methods specifically developed very recently for the same wind turbine benchmark [12]. In particular, three alternative control schemes are considered in this work, which are based on Unknown Input Observers (UIOs) [6], virtual sensors/actuators (VAS) [5], and LMI–based LPV controllers [13]. Since it is necessary to evaluate the impact on the designed control systems of modelling uncertainties, disturbance, and measurement errors, the overall scheme verification uses extensive Monte–Carlo simulations for the analysis and the assessment of the robustness, the stability, and their final performance evaluation. In fact, as shown in the following, the wind turbine system may contain elements that cannot be described by any analytical model obtained via first principles.

Finally, the paper has the following structure. Section II provides an overview of the wind turbine system considered in this work. Section III recalls the strategy exploited for the identification of the fuzzy controller. On the other hand, the second parameter–varying controller design is described in Section IV-A. The achieved results are summarised in Section V. Section VI ends the paper by highlighting the main achievements of the work.

II. Wind Turbine Benchmark Description

The three–blade horizontal axis turbine considered in this paper works according to the principle that the wind is acting on the blades, and thereby moving the rotor shaft. In order to up–scale the rotational speed to the needed one at the generator, a gearbox is introduced.

Note that a controller for a wind turbine operates in principle in four zones. Zone 1 is start–up of the turbine, zone 2 is the so-called power optimisation, zone 3 corresponds to constant power production, and zone 4 with high wind speed. Since the benchmark model works in normal operating conditions, only zone 2 and zone 3 are considered here, as described e.g. in [14]. In zone 2, which will be denoted as region 1, the turbine is controlled to obtain optimal power production. This working condition is known as partial load condition. On the other hand, zone 3 corresponds to region 2, i.e. the so-called full load working situation.

The wind turbine model in the continuous–time domain is briefly recalled in the following. The aerodynamic model is defined as in (1):

$$\tau_{aero}(t) = \rho\, A\, C_p\big(\beta_r(t), \lambda(t)\big)\, v^3(t)/2\, \omega_r(t) \qquad (1)$$

Where ρ is the density of the air, A is the area covered by the turbine blades in its rotation, $\beta_r(t)$ is the pitch angle of the blades, v(t) the wind speed, whilst $\lambda(t)$ is the tip–speed ratio of the blade [12]. $C_p(.)$ represents the power coefficient, here described by means of a two–dimensional map (look–up table) [12]. Equation (1) is used to compute $\tau_{aero}(t)$ based on an assumed estimated wind speed v(t), the measured $\beta_r(t)$ and the rotor speed $\omega_r(t)$. Due to the uncertainty of the wind speed v(t), the estimate of $\tau_{aero}(t)$ is considered affected by an unknown measurement error, which motivates the approaches proposed in this study. A simple one–body model is used to represent the drive train, whilst the hydraulic pitch model is described as a closed–loop transfer function of the hydraulic pitch system modelled as a second order transfer function [14]. The converter dynamics are modelled by a first–order transfer function, and the measurement sensors are modelled by adding the actual variable values with stochastic noise processes. These noise signals are described as Gaussian processes with fixed mean and standard deviations values, depending on the considered measurement sensors. A more accurate description of the benchmark model can be found in [12].

With these assumptions, the complete continuous–time description of the system under diagnosis has the form of (2):

$$\dot{x}_c(t) = f_c\big(x_c(t), u(t)\big)$$
$$y(t) = x_c(t) \qquad (2)$$

where $u(t) = [\beta_r(t), \tau_g(t)]^T$ and $y(t) = x_c(t) = [P_g(t), \omega_g(t)]^T$ are the control inputs and the monitored output measurements, respectively. $f_c(.)$ represents the continuous–time nonlinear function that describes the complete behaviour of the wind turbine process. Regarding the input and output signals, $\omega_g(t)$ is the generator speed measurement, $P_g(t)$ the generator power measurement, and $\tau_g(t)$ generator torque measurement. Finally, the model parameters, and the map $C_p(\beta_r, \lambda)$ are chosen in order to represent a realistic turbine, which is used as benchmark system in this study [12].

Finally, the next Section III will recall the scheme for obtaining the fuzzy description of the wind turbine controller.

III. Fuzzy Identification for Controller Design

This section recalls the approach exploited for obtaining the fuzzy description of the wind turbine controller, whilst the proposed controller model estimation is shown in Section III-A, which represents one of the main contributions of the paper.

The approach suggested in this section employs fuzzy clustering techniques to partition the available data into subsets characterised by linear behaviours. Relationships between clusters and linear regression are exploited, thus allowing for the combination of fuzzy logic techniques with system identification tools. In addition, an implementation in the Matlab® Toolbox of the Fuzzy Modelling and IDentification (FMID) technique presented in the following is available [15].

In this study, TS fuzzy models are exploited [16], as they are able to provide the mathematical description of the nonlinear system. The switching and the scheduling between the submodels is achieved through a smooth function of the system state, the behaviour of which is defined using fuzzy set theory.

In more detail, the fuzzy modelling and identification algorithm is based on a two–step procedure, in which at first, the operating regions are determined using the data clustering technique, and in particular, the Gustafson–Kessel (GK) fuzzy clustering, since already available in [15]. Then, in the second stage, the estimation of the controller parameters is achieved using the identification algorithm already proposed by one of the same authors in [7], which can be seen as a generalisation of classical least–squares.

The TS fuzzy models have the form of:

$$y(k+1) = \Sigma^{M}_{i=1}\mu_i(\mathbf{x}(k))\, y_i / \Sigma^{M}_{i=1}\mu_i(\mathbf{x}(k)) \quad (3)$$

where $y_i = \mathbf{a}^{T}_{i} + b_i$, with $\mathbf{a}_i$ the parameter vector (regressand), and b_i is the scalar offset. $\mathbf{x} = \mathbf{x}(k)$ represents the regressor vector, which can contain delayed samples of u(k) and y(k).

The *antecedent* fuzzy sets $\mu_i(.)$ are extracted from the fuzzy partition matrix [16]. The *consequent parameters* $\mathbf{a}_i$ and b_i are estimated from the data using the procedure presented in [7]. This identification scheme exploited for the estimation of the TS model parameters has been integrated into the FMID toolbox for Matlab® by one of the same authors. This approach developed by one of the same authors is usually preferred when the TS model should serve as predictor, as it computes the consequent parameters via the Frisch scheme, developed for the Errors–In–Variables (EIV) descriptions [10].

A. Fuzzy Controller Estimation

This section addresses the design of the nonlinear fuzzy controller of the wind turbine process, which relies on the *model inverse control principle*, and solved within the fuzzy identification framework.

Therefore, the fuzzy identification scheme will be used for both predicting the wind turbine behaviour and *estimating the inverse model controller structure*. An optimal control strategy is thus obtained by minimising a cost function, which includes the difference between the desired and controller outputs, and a penalty on the system stability. Constraints on the complete system stability are thus included as a part of the optimisation problem. Generally, a non–convex optimisation problem must be solved at each control sample, which hampers the direct and practical application of the approach. However, to solve this problem, the optimisation scheme described in [7], which is based on a parameterised search technique, is applied at a higher level to formulate the control objectives and constraints.

The proposed method is implemented for model with a unitary delay $n_k = 1$. For systems with larger delays, the same method can be applied after performing $n_k - 1$ steps of prediction with the fuzzy model. When $n_k = 1$, the general rule–based model (3) corresponds to the following regression system:

$$y(k+1)=f(\mathbf{x}(k), u(k)) \quad (4)$$

The inputs of the model are the current state $\mathbf{x}(t) = [y(k), \ldots, y(k–n+1), \ldots, u(k-1), \ldots, u(k–n)]^T$ and the current input u(k). The output is a prediction of the system's output at the next sample y(k+1). The objective of the control algorithm is to compute the control input u(k), such that the system output at the next sampling instant is equal to the desired (reference) output r(k+1). In principle, this can be achieved by inverting the model of the process. Given the current state $\mathbf{x}(k)$ and the reference r(t+1), the control input is given by:

$$u(k) = f^{-1}(\mathbf{x}(k), r(k+1)) \quad (5)$$

where the reference r(k+1) is replaced by y(k+1). Generally, it is difficult to find the analytical inverse function $f^{-1}(.)$. Therefore, the method exploited here makes use of the identified fuzzy TS of the process under investigation for providing the particular state $\mathbf{x}(k)$ at each time step k. From this mapping, the inverse mapping (5) is easily *identified* as a model in the form of (3), provided the controlled system is stable.

Therefore, the series connection of the controller and the identified inverse model, should give an identity mapping (perfect control):

$$y(k+1) = f(\mathbf{x}(k), u(k)) = f(\mathbf{x}(k), f^{-1}(\mathbf{x}(k), r(k+1))) = r(k+1) \quad (6)$$

when u(k) exists such that $r(k+1) = f(\mathbf{x}(k), u(k))$. However, due to modelling errors, noise, and disturbance, by means of the fuzzy identification procedure, the difference $|r(k+1)–f(\mathbf{x}(k),u(k))|$ is made arbitrarily small by an appropriate choice of the identification parameters, *i.e.* the fuzzy membership functions, the number of clusters, and the regressand. The *process fuzzy model* is used for the recursive prediction of the state vector $\mathbf{x}(k)$. Apart from the computation of the membership degrees, both the process model and the controller are estimated using standard matrix operations and linear interpolations, which makes the algorithm suitable for *real–time implementation*.

Note however that, with the fuzzy control strategy proposed here, disturbances acting on the process, measurement noise and model–plant mismatch can cause differences in the behaviour of the process and of the model. A mechanism to compensate this error can be exploited *e.g.* via *on–line adaptation* of the process model. On–line adaptation can be applied to cope with the mismatch between the plant and the fuzzy model. In many cases, a mismatch occurs as a consequence of (temporary) changes of process parameters. Therefore, Section IV motivates the adaptive strategy based on linear models, whose parameters are adapted on–line and exploited for the controller parameter estimation.

IV. Recursive Identification for Adaptive Control

This section describes the recursive approach exploited for obtaining the mathematical description of the wind turbine system, which is used for the design of the second control strategy. A modification of the Frisch scheme algorithm is proposed here to identify dynamical Errors–In–Variables (EIV) models [10, 17]. For the update of the estimated model parameters, a recursive bias–compensating strategy is also

implemented. Thus, a recursive Frisch scheme identification approach is extended to enhance its on–line applicability. It is shown that by incorporating adaptation via the introduction of exponential forgetting, the algorithm is able to compensate for the systematic errors, which arise in the original scheme [10]. Therefore, this adaptive recursive Frisch scheme is able to deal with linear time–varying systems, and it is used in connection with the design of an adaptive control scheme, shown in Section IV-A.

Thus, the recalled scheme is used for the on–line identification of the process modelled by the following transfer function G(z):

$$G(z) = B(z^{-1})/A(z^{-1}) =$$

$$= (b_1 z^{-1} + \ldots + b_{nb} z^{-nb})/(1 + a_1 z^{-1} + \ldots + a_{na} z^{-na}) \quad (7)$$

where a_i, b_i, n_a and n_b represent the unknown parameters and the structure of the model, defining the polynomials $A(z^{-1})$ and $B(z^{-1})$ whilst z is the discrete–time complex variable.

The parameter vector describing the linear relationship is given by:

$$\theta = [a_1 \ldots a_{na}\, b_1 \ldots b_{nb}]^T \quad (8)$$

and its extended version is denoted with:

$$\vartheta = [1\ \theta^T]^T \quad (9)$$

Hence, an alternative expression for the considered difference equation is given by:

$$\Psi^T(k)\, \vartheta = 0 \quad (10)$$

Where:

$$\Psi(k) = [-y(k), \ldots, -y(k-na+1), \ldots, u(k-1), \ldots, u(k-nb)]^T \quad (11)$$

Is the extended regressor vector.

The Frisch scheme provides estimates for the measurement errors affecting the input and output signals u(k) and y(k), *i.e.* σ_u, σ_y, and θ for a linear time–invariant dynamical system. Moreover, it can also be exploited to determine the polynomial orders n_a and n_b, as shown *e.g.* in [7].

However, since this work is oriented to the design of an adaptive controller, the polynomial orders are assumed to be fixed in advance. In this adaptive control application, it is essential to obtain on–line estimates of the model parameters $\theta(k)$ in (7), while the process generating the data is running. In fact, this application study is focusing on adaptive control, where the control action at time step k relies on a current estimate of the plant model, which is estimated using data up to the sample k. Therefore, the Frisch scheme relying on batch expressions has to be modified in a recursive algorithm, as described in [10]. The on–line identification method described here was implemented by the author in the Matlab® and Simulink® environments, and integrated in the Simulink® toolbox [11].

Finally, once the time–varying parameters $\theta(k)$ of the discrete–time linear model approximating the nonlinear process (2) have been computed at each time step k, the adaptive controller is designed as described in Section IV-A.

A. Adaptive Controller Design

With reference to the particular benchmark under diagnosis, adaptive controllers for processes of second order ($n_a = n_b = n = 2$) are exploited. Moreover, the considered adaptive controllers are based on the trapezoidal method of discretization.

Therefore, with reference to (7), with $n_a = n_b = n = 2$, the transfer function parameters estimated on–line are:

$$\theta = [a_1\ a_2\ b_1\ b_2]^T \quad (12)$$

Note that the subscript k for model and controller parameters is dropped in order to simplify equations and formulas.

The control law corresponding to discrete–time PI adaptive controller has the form:

$$u(k) = K_p \left[e(k) + \tfrac{1}{2} T_s/T_I \left(e(k) - e(k-1)\right)\right] + u(k-1) \quad (13)$$

where e(k) is the tracking error, *i.e.* e(k) = r(k) – y(k), with r(k) the set–point or reference signal, T_s the sampling time. The (time–varying) controller variables K_p and T_I are now computed from the time–varying model parameters $\theta(k)$ [11].

In particular, the controller parameters K_p and T_I are computed using the Ziegler-Nichols relations depending on the (time–varying) critical gain and the critical period of oscillations [11]. Also these variables are functions of the time–varying model parameters $\theta(k)$.

Finally, the next Section V will show the results achieved by using the control design schemes described above.

V. Simulation Results

Regarding the fuzzy modelling method, the GK clustering algorithm with M = 3 clusters and a number of shifts n = 2 was applied to the estimation and validation sampled data sets $\{P_g(k), \omega_g(k), \beta_r(k)\}$, with k = 1, 2, …, N and $N = 440 \times 10^3$. On the other hand, a number of clusters M = 3 and n = 2 was considered for achieving a suitable clustering of the sampled data sets $\{P_g(k), \omega_g(k), \tau_g(k)\}$. After clustering, the TS model parameters for each output $P_g(k)$ and $\omega_g(k)$ were estimated.

Therefore, the two outputs y(t) of the wind turbine continuous–time model (2) are approximated by two TS fuzzy prototypes (3). The relative mean square errors of the output estimations are 0.0254 for the first output, and 0.0125 for the second one. The fuzzy model estimation procedure was implemented in order to guarantee the identification of stable TS prototypes via the parameters n, μ_i, a_i, and b_i. The fitting capabilities of the estimated fuzzy models can be expressed also in terms of the so–called Variance Accounted For (VAF) index [16]. In particular, the VAF value for first output was bigger than 90%, whilst bigger than 99% for the second one. Hence, the fuzzy multiple models seem to approximate the process under investigation quite accurately.

Regarding the fuzzy controllers, the experimental set–up employs 2 (Multiple–Input Single–Output) MISO fuzzy regulators used for the control of the blade pitch angles $\beta_r(t)$ and the generator control torque $\tau_g(t)$, respectively, that were

identified according to the fuzzy inverse model scheme suggested in Section III-A. Also in this case, the GK clustering algorithm was applied again for the estimation of the two fuzzy inverse model regulators. A number of M = 3 clusters and a number of shifts n = 3 were applied to the estimation and validation sampled data sets $\{\beta_r(k), P_g(k), \omega_g(k)\}$. On the other hand, a number of clusters M = 3 and n = 3 were considered again for achieving a description of the second fuzzy inverse model regulator via the clustering of the data $\{\tau_g(k), P_g(k), \omega_g(k)\}$.

The controller capabilities were assessed in simulation by considering different data sequences. In Table II the per–cent Normalised Sum of Squared tracking Error (NSSE) values defined as:

$$NSSE^2\% = 100\ \Sigma^N_{k=1}(r(k) - y(k))^2 / \Sigma^N_{k=1} r^2(k) \qquad (14)$$

are computed for the designed fuzzy controllers. Table II refers to the full–load operation, where the performance depends on the generator speed, ω_g, with respect to the nominal one, ω_{nom}.

With reference to the second adaptive control design, the two outputs $P_g(t)$ and $w_g(t)$ of the wind turbine continuous–time nonlinear model (2) were approximated by 2 time–varying MISO discrete–time second order prototypes of the type (7) with 2 inputs. The approach described in Section IV for SISO models can be easily extended to the MISO case. Using these two on–line identified prototypes, the model–based approach for determining the adaptive controllers shown in Section IV-A was exploited. Thus, the parameters of the adaptive controllers were computed on–line. In particular, the adaptive regulator parameters in (13) were computed analytically at each time step k. Simulations were performed in the same conditions of the fuzzy controllers, and 2 adaptive regulators were used. As an example, the initial values for the parameters of the on–line estimation algorithm are listed in Table I.

In order to analyse the performance of the proposed adaptive strategy, Table II reports the NSSE values computed also for these controllers. In full load operation the performance depends on the generator speed ω_g with respect to the nominal one, ω_{nom} = 162 rad/s.

According to these simulation results, good tracking capabilities of the suggested controllers seem to be reached, and the adaptive solution seems better than the fuzzy one.

TABLE I. INITIALISATION PARAMETERS OF THE ADAPTIVE ALGORITHM.

Parameter	Value
$\theta(0)$	$[0.1, 0.20, 0.30, 0.4]^T$

TABLE II. FUZZY AND ADAPTIVE CONTROLLER NSSE% VALUES.

Data Set	Fuzzy	Adaptive
Data set #1	16.57%	12.95%
Data set #2	17.85%	13.67%

A. Robustness Evaluation

In this section, further experimental results were reported. They regard the performance evaluation of the developed control schemes with respect to modelling errors and measurement uncertainty. In particular, the simulation of different data sequences was performed by exploiting the wind turbine benchmark simulator, and a Matlab® Monte–Carlo analysis. In fact, the Monte–Carlo tool is useful at this stage as the control strategy performances depend on the error magnitude due to the model approximation and uncertainty, as well as on input–output measurement errors.

In particular, the nonlinear wind turbine simulator originally developed in the Simulink® environment [14] was modified by one of the authors in order to vary the statistical properties of the signals used for modelling possible process parameter uncertainty, and measurement errors. Under this assumption, Table III reports the nominal values of the considered wind turbine model parameters with respect to their simulated but realistic uncertainty.

The Monte–Carlo analysis describes these variables as Gaussian stochastic processes, with zero–mean and standard deviations corresponding to the maximal error values in Table III.

Therefore, for performance evaluation of the control schemes, the best, average, and worst values of the NSSE% index were computed, and experimentally evaluated with 500 Monte–Carlo runs, as shown in Table IV.

TABLE III. REALISTIC WIND TURBINE UNCERTAINTY.

System	Error	Physical Effect
$\beta_r(t)$	11%	Pitch position measurement accuracy
$\omega_g(t)$	18%	Generator speed measurement accuracy
$\tau_g(t)$	21%	Generator torque measurement accuracy
$P_g(t)$	20%	Electrical power measurement accuracy
Pitch system	49%	Hydraulic system pressure change
Drive–train	5%	Drive train dynamics change
Converter dynamics	50%	Offset in Converter torque control

TABLE IV. NSSE% VALUES FOR THE MONTE-CARLO ANALYSIS.

Test case	Fuzzy	Adaptive
Best case	15.57%	11.05%
Average case	17.94%	13.72%
Worst case	19.94%	15.04%

In particular, Table IV summarises the values of the considered performance index NSSE% according to the best, worst and average cases, with reference to the possible combinations of the parameters described in Table III. Table IV shows that the proposed control schemes, and in particular the adaptive solution, allow for good control performances

even in the presence of considerable error and uncertainty effects.

B. Comparative Simulations

This section compares the control methods suggested in this paper with respect to alternative control approaches.

The first control scheme proposed for comparison purposes relies on UIOs [6], which were used for estimating measurements used by the control system. The UIO–based scheme is chosen since it enables the possibility to include robustness towards the uncertainty of the wind speed, which is difficult to measure.

On the other hand, the second control approach uses the idea of virtual sensors/actuators (VAS) [5]. An estimation of the uncertainty acting on the process is provided on the basis of a batch least squares approach. The use of on–line disturbance estimation is essential for all compensation approaches.

The third approach relies on an LMI–based method for designing and synthesising the LPV controller, which is based on the LPV controller design method presented in [13]. In fact, the considered wind turbine model has varying parameters caused by nonlinearities in the aerodynamic model along the nominal operating trajectory and due to the model uncertainty.

In order to provide a brief but clear insight into the above-mentioned techniques, the comparison was performed in the same previous working conditions, and based on the NSSE% index suggested at the beginning of Section V. Table V summarises the results obtained by comparing the three control techniques recalled above with the ones proposed in this study.

Table V shows that the scheme using the VAS and the LPV strategies allow to achieve better performances in terms of tracking error. However, the LPV controller can increase the computational time considerably with respect to the other solution, without any gain scheduling, whilst the LPV and UIO control methods can require larger computational effort at the design stage.

Few further comments can be drawn here. When the modelling of the dynamic system can be perfectly obtained, in general model–based control strategies are preferred. On the other hand, when modelling errors and uncertainty are present, alternative control schemes relying on adaptation mechanisms, or passive robust control methods, showed interesting robustness properties. The fuzzy logic–based scheme relies on the learning accumulated from off–line simulations, but the on–line estimation stage could be computationally heavy. Finally, regarding the proposed methods using LPV or fuzzy tools, they seem rather simple and straightforward, even if optimisation stages can be required.

TABLE V. MONTE-CARLO ANALYSIS AND NSSE VALUES WITH THE COMPARED CONTROLLERS.

Test case	UIO	VAS	LPV
Best case	26.55%	19.85%	10.95%
Average case	27.72%	17.94%	13.02%
Worst case	28.44%	16.33%	14.94%

C. Experimental Results

Also the stability properties of the overall control strategies were checked by means of a Monte–Carlo campaign based on the wind turbine benchmark. In fact, as pointed out above, the Monte–Carlo analysis represents the only method for estimating the efficacy of the developed control schemes when applied to the monitored process.

It is worth noting that the work [18] provided an analytical demonstration of the stability of an adaptive control scheme for wind turbines. However, model parameter variations, recursive Frisch scheme adaptive methods, or complete wind models were not taken into account there.

All simulations were performed by considering noise signals modelled as Gaussian processes, according to the standard deviations reported in Table III. The wind turbine benchmark simulator generated different wind sequences. Moreover, the initial conditions of the dynamic models and recalled in Section II (*i.e.* the drive–train, the generator/converter, and the pitch system) were changed randomly. Therefore, the random wind signal v(t), the parameters of Table III, and the dynamic model initial conditions allowed to obtain different sequences of the wind turbine signals $\beta_r(t)$, $\tau_g(t)$, $\lambda(t)$, $\omega_g(t)$, and $P_g(t)$ for each Monte–Carlo simulation.

As an example of a single Monte–Carlo run, Fig. 1 highlights that the main wind turbine model variables, such as the generator torque $\tau_g(t)$, the tip-speed–ratio $\lambda(t)$, and the generator power $P_g(t)$ remain bounded around the reference values, proving the overall system stability in simulation, even in the presence of disturbance and uncertainty. These results refer to the case of partial load operation with the fuzzy controllers.

Fig. 1 shows also that in the first part of the simulation the output power P_g becomes larger than the theoretical one, $P_{g,max}$, as the kinetic energy from the rotor shaft is converted into electrical energy produced by the generator. On the other hand, $P_{g,max}$ can be above the generated power, since the inertia of the rotor is accelerated before $P_{g,max}$ can be matched.

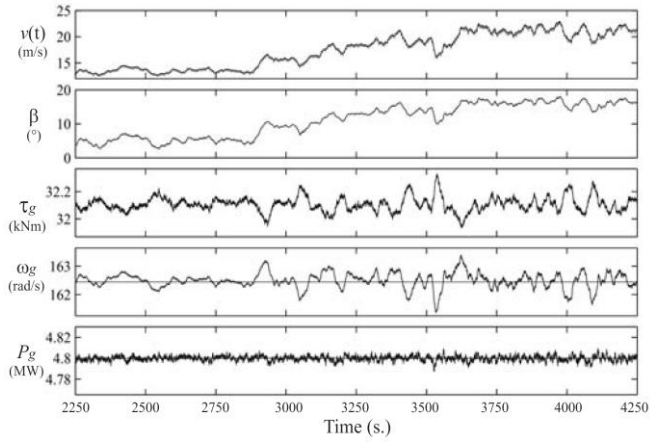


Fig. 1. Simulations of the wind turbine benchmark with the fuzzy controllers.

TABLE VI. HIL RESULTS WITH RESPECT TO THE NSSE% INDEX.

Controller	Partial load	Full load
Fuzzy	38.72%	18.94%
Adaptive	28.81%	14.83%

In order to evaluate the potential of utilising the proposed control algorithms also in real applications and investigate their capability to on–board implementation, the remainder of this section presents the results of the Hardware In the Loop (HIL) tests. These experimental results serve to validate definitely the designed control algorithms considering almost real conditions that the wind turbine may experiment with during its working situations. For this purpose, HIL test–bed already described in [9] was exploited, in order to provide the capabilities to validate the developed control algorithms in an almost real–time condition. The results achieved from one test are summarised in Table VI for the proposed identified fuzzy and adaptive controller solutions.

Table VI illustrates that there are some deviations between the achieved results, but consistent with the ones from the Monte–Carlo analysis. Although there are some deviations between the simulation and the experimental results, these deviations are not critical and the results obtained are accurate enough for future wind turbine real applications.

VI. FUTURE WORKS

The increasing dimensions of wind turbines lead to the increase in the loads on wind turbine structures. Because of increasing rotor size and spatially varying loads along the blade, individual blade pitch control can reduce the negative effects of sub-rotor-sized turbulent structures. Additional pitch control loops can be used to damp the tower motion or additional structural vibrations in the full load working condition [13].

Given the complexity of the wind turbine system, the stability of the complete plant plus control system cannot be proven. The multiple control loops interact, as do the multiple degrees of freedom of the turbine, especially as wind turbines become larger and have lower natural frequencies. A unified Multiple-Input Multiple-Output (MIMO) framework for individual blade pitch control can achieve significant load reduction for floating offshore wind turbines with strong coupling across degrees of freedom [13].

Because wind turbine control is often achieved using two distinct control loops for the working regions between partial load and full load conditions, the transition between these regions can be problematic. For some turbines, the maximum structural damage occurs due to extreme and fatigue loads during this transition. Often, the act of switching between these region controllers contributes to the problem.

Advanced control strategies, for example, uses an additional control region between partial and full load conditions to facilitate switching between these two conditions. The primary objective of this control strategy, described in [13], [1] is to connect these switching controllers linearly in the generator torque versus generator speed plane. Unfortunately, this linear connection does not result in smooth transitions, and the discontinuous slopes in the torque control curve can contribute to excessive loading on the turbine.

Wind turbines can also be damaged when they are stopped as a result of supervisory control action due to high winds or fault conditions [13], [1], [2], [3]. However, little or no active control is performed when the turbine is stopped, although the yaw angle can be changed to accommodate changes in wind direction, which can prevent some damage.

In addition to the possibility of improving control when the turbine is stopped, advanced fault detection and turbine protection schemes are of interest to the wind industry [2], [3]. Stopping the turbine in the case of emergency, which might entail pitching the blades to a predetermined stop position at maximum pitch rate and setting the mechanical brakes with which the rotor is equipped, can also cause damage to the machine and must be done only when a turbine failure is suspected.

Finally, controller performance depends on modelling accuracy. For instance, as shown in Section V, realistic modelling error in the optimal tip-speed ratio can cause an energy loss of around a few per cent in the partial load condition [13], which can be a significant loss in this industry. Even disregarding model errors, the dynamical behaviour of a wind turbine changes over time due to wear, debris build-up on the blades, and environmental conditions. As such, adaptive methods shown in this work can be used to tune controllers to improve performance compared to time-invariant methods [8], [9].

While wind turbine dynamics can be modelled using first principles, this study shown that efficient methods for obtaining models from measurements also exist, including the development of closed-loop identification methods for deter mining linear parameter-varying models [7]. These models can be used for robust control. Modelling of wind turbines and wind farms is further discussed in [1], [13].

VII. CONCLUSION

The paper is focused on two examples of control designs for a nonlinear wind turbine prototype. The proposed control designs represent viable and easy-to-use methods for the straightforward derivation of proper controller models, as data-driven and system identification from data approaches are exploited. Tests on the considered benchmark process and Monte–Carlo analysis were the tools for assessing experimentally the properties of the proposed control schemes, in the presence of modelling and measurement errors. The developed control methods were also compared with different approaches, in order to evaluate the considered techniques. These comparisons highlight that the proposed design methodologies can constitute reliable and robust approaches for application to real wind turbine processes.

REFERENCES

[1] P. F. Odgaard and K. Johnson, “Wind Turbine Fault Diagnosis and Fault Tolerant Control – an Enhanced Benchmark Challenge,” in Proc. of the 2013 American Control Conference – ACC, (Washington DC, USA), pp. 1–6, IEEE Control Systems Society & American Automatic Control Council, June 17 – 19 2013.

[2] P. F. Odgaard, J. Stoustrup, and M. Kinnaert, "Fault–Tolerant Control of Wind Turbines: A Benchmark Model," IEEE Transactions on Control Systems Technology, vol. PP, no. 2013, pp. 1–15, 2013. Available Online. DOI: 10.1109/TCST.2013.2259235.

[3] M. Sami and R. J. Patton, "An FTC Approach to Wind Turbine Power Maximisation Via T–S Fuzzy Modelling and Control," in Proceedings of the 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS 2012 (C. Verde, C. M. Astorga Zaragoza, and A. Molina, eds.), vol. 8, (National Autonomous University of Mexico, Mexico City, Mexico), pp. 349–354, August 2012. DOI: 10.3182/20120829–3–MX–2028.00131.

[4] P. F. Odgaard and J. Stoustrup, "Fault Tolerant Wind Speed Estimator used in Wind Turbine Controllers," in Proceedings of the 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS 2012 (C. Verde, C. M. Astorga Zaragoza, and A. Molina, eds.), vol. 8, (National Autonomous University of Mexico, Mexico City, Mexico), pp. 331–337, August 2012. DOI: 10.3182/20120829-3-MX-2028.00009.

[5] D. Rotondo, F. Nejjari, V. Puig, and J. Blesa, "Fault Tolerant Control of the Wind Turbine Benchmark using Virtual Sensors/Actuators," in Proceedings of the 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS 2012 (C. Verde, C. M. Astorga Zaragoza, and A. Molina, eds.), vol. 8, (National Autonomous University of Mexico, Mexico City, Mexico), pp. 114–119, August 2012. DOI: 10.3182/20120829-3-MX-2028.00185.

[6] P. F. Odgaard and J. Stoustrup, "Fault Tolerant Control of Wind Turbines using Unknown Input Observers," in Proceedings of the 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS 2012 (C. Verde, C. M. Astorga Zaragoza, and A. Molina, eds.), vol. 8, (National Autonomous University of Mexico, Mexico City, Mexico), pp. 313–319, August 2012. DOI: 10.3182/20120829-3-MX-2028.00010.

[7] S. Simani, C. Fantuzzi, R. Rovatti, and S. Beghelli, "Parameter identification for piecewise linear fuzzy models in noisy environment," International Journal of Approximate Reasoning, vol. 1, pp. 149–167, September 1999. Publisher: Elsevier.

[8] S. Simani, "Data–Driven Design of a PI Fuzzy Controller for a Wind Turbine Simulated Model," in Proceedings of the IFAC Conference on Advances in PID Control – PID'12 (R. Vilanova and A. Visioli, eds.), vol. 2, (University of Brescia, Brescia, Italy), pp. 667–672, DII, Faculty of Engineering, University of Brescia, Italy, IFAC, March 28–30 2012. Invited paper. ISBN: 978-3-902823-18-2. DOI: 10–3182/20120328–3–IT–3014.00113.

[9] S. Simani, "Application of a Data–Driven Fuzzy Control Design to a Wind Turbine Benchmark Model," Advances in Fuzzy Systems, vol. 2012, pp. 1–12, November 2nd 2012. Invited paper for the special issue: Fuzzy Logic Applications in Control Theory and Systems Biology (FLACE) . ISSN: 1687–7101, e–ISSN: 1687-711X. DOI: 10.1155/2012/504368.

[10] J. G. Jens G. Linden, T. Larkowski, and K. J. Burnham, "Algorithms for recursive/semi–recursive bias-compensating least squares system identification within the errors–in–variables framework," International Journal of Control, vol. 85, pp. 1625–1643, June 2012. DOI:10.1080/00207179.2012.696145.

[11] V. Bobál, J. Böhm, J. Fessl, and J. Machácek, Digital Self–Tuning Controllers: Algorithms, Implementation and Applications. Advanced Textbooks in Control and Signal Processing, Springer, 1st ed., 2005.

[12] P. F. Odgaard, J. Stoustrup, and M. Kinnaert, "Fault Tolerant Control of Wind Turbines – a Benchmark Model," in Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, vol. 1, (Barcelona, Spain), pp. 155–160, June 30 – July 3 2009. DOI: 10.3182/20090630–4–ES–2003.0090.

[13] F. D. Bianchi, H. D. Battista, and R. J. Mantz, Wind Turbine Control Systems: Principles, Modelling and Gain Scheduling Design. Advances in Industrial Control, Springer, 1st ed., July 2007. ISBN: 1–84628–492–9.

[14] P. F. Odgaard and J. Stoustrup, "Unknown Input Observer Based Scheme for Detecting Faults in a Wind Turbine Converter," in Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, vol. 1, (Barcelona, Spain), pp. 161–166, IFAC – Elsevier, June 30 – July 3 2009. DOI: 10.3182/20090630–4–ES–2003.0048.

[15] R. Babuška, Fuzzy Modelling and Identification Toolbox. Control Engineering Laboratory, Faculty of Information Technology and Systems, Delft University of Technology, Delft, The Netherlands, version 3.1 ed., 2000. (Available at http://lcewww.et.tudelft.nl/˜babuska).

[16] R. Babuška, Fuzzy Modeling for Control. Kluwer Academic Publishers, 1998.

[17] S. Van Huffel and P. Lemmerling, eds., Total Least Squares and Errors–in–Variables Modeling: Analysis, Algorithms and Applications. Springer–Verlag, 1st ed., February 2002. ISBN: 1402004761.

[18] K. E. Johnson, L. Y. Pao, M. J. Balas, V. Kulkarni, and L. J. Fingersh, "Stability analysis of an adaptive torque controller for variable speed wind turbines," in Proceedings of the 43rd IEEE Conference on Decision and Control – CDC'04, vol. 4, (Paradise Island, Bahamas), pp. 4087–4094, December 14–17 2004.

Multimodal Biometric Technology System Framework and E-Commerce in Emerging Markets

Chike Obed-Emeribe (PhD)
Head Research,
Association for Promoting Interest in Mathematics and Sciences (APRIMATHS)
Abuja, Nigeria

***Abstract*— It is self-evident that the game changer of our modern world – the "internet" has endowed the twenty first century man with enormous potentials and possibilities. Ranging from enhanced capabilities in business (e-business), governance (e-governance), politics, social interaction and information exchange. The internet has indeed shrinked the global distance that once posed a great barrier and limited man's endeavours in the preceding centuries.**

Amidst the great advantages derivable from the use of internet for various purposes lie inherent security threats. To a large extent, these security hindrances have been addressed in advanced nations of the world, as a result, internet phenomenon has pervaded all aspects of the advanced nation's economy. This is evident in different electronic platforms that are available for the delivery of various products and services. On the contrary, the application of internet in various aspects of commerce has been hampered by the challenges of security limitations due to identity issues in the developing/emerging economies. Due to these security threats, business owners and the general public in less-developed world demonstrate great sense of apathy in the use of available electronic options for the purpose of commerce.

Against the backdrop of the above, and the poor infrastructure basis of the developing nations, this research paper analyses and proposes the implementation of multimodal biometric technology frameworks with novel server architecture to tackle the security threats inherent with e-commerce in the developing world.

Keywords—Biometrics; Multimodal; Frameworks; e-commerce; Emerging Markets.

I. INTRODUCTION

Considering the enormity and sensitivity of transactions expected to be carried out via internet platforms in the wake of the accelerated internet access of the modern age, adequate means of identification and verification (authentication) should be evolved in other to give business owners, consumers and the general public an assurance of safety.

Quite a number of means of ensuring safety of transactions in e-commerce platforms have been developed in the recent past, but the inadequacies of such methods have precipitated into incessant cases of internet fraud and online theft, a situation that calls for urgent action to guarantee internet security particularly in less advanced nations. Quite recently, the issue of cyber crime and cyber security has been in the front burner.

As a solution to this internet security threat, biometric technology was evolved to ensure the security of all e-commerce transactions. The origin of biometrics can be traced back to the primordial Greek society [1] The technology of biometrics entails the use of intrinsic physical, behavioural and psychological features of individuals as a means of identification and verification (ie Authentication) . The most commonly used biometric features for the purpose of identification and identity management include: facial features, hand geometry, vascular pattern, fingerprints, retina, iris, keystroke, handwriting, gait and voice. These features have either being used singly or in combination in different security applications with the attendant advantages of robustness, universality, permanence and accessibility [3].

Due to the peculiar nature of developing world, in terms of infrastructure development and public awareness, a suitable framework for the implementation of biometric technology in e-commerce is highly imperative. This framework when fully implemented will enable business owners to adopt e-commerce as well as encourage consumers to engage in e-commerce transactions.

In this research paper, the potential prototype framework proposed is both suitable and safe to be implemented in any developing economy with such peculiarity. Under this framework, backend server architecture is delivered, in such a way as to give the multimodal mix of the biometric design. One of the bases for this research endeavour is the fact that about 79% of world population lives in the developing countries. This statistics shows that there are huge economic potentials in these countries in terms of e-commerce which is yet to be tapped owning to the barriers of poor infrastructure, legal issues, socio-cultural bottlenecks as well as lack of trust.

Given the fact that e-commerce has been proven to improve market efficiency, operational effectiveness, access to markets and linkages, establishing a highly secure and robust e-commerce system in the emerging market economies is truly a well come development.

II. ANALYSIS AND DESIGN OF MULTIMODAL BIOMETRIC TECHNOLOGY

It has been advocated at different quarters that data and system security is the next frontier of information technology in the coming centuries. As more people access the internet infrastructure, more businesses go online, and most traditional operations become internet based, reliable means of user

identification and verification become of high essence. The only means of attaining this height of online internet security is via biometric technology.

Basically, a complete biometric system majorly is characterized by three elements namely;

- Enrollment sub-system
- Template representation
- Matching process subsystem.

These three main elements are depicted in the figure below:

Fig.1. The general block diagram representation of a typical Biometric System.

A. THE ENROLLMENT STAGE:

At this stage, data samples are collected from the enrollee. Mostly devices such as scanners and readers are employed for this purpose. This stage is usually crucial as any mistake will lead to identity misrepresentation.

B. THE TEMPLATE REPRESENTATION STAGE:

At this stage of biometric operation, data samples obtained at the enrollment stage are gathered and stored for future referencing. This operation is usually carried out by some specific software tools.

C. MATCHING PROCESS SUBSYSTEM:

Here, input data is compared with the already store data template within the system for the purpose of identification and verification.

III. BIOMETRIC PROCESSES AND CLASSIFICATIONS

As earlier stated, the biometric process entails capturing the unique biological, behavioural or psychological features of a particular individual with a view to identification and verification. This process can be done basically in two modes, viz: unimodal and multimodal forms. Each of these modes of biometric authentication has its computational requirements as well as inherent advantages and demerits. The adoption of any particular mode depends on the expected outcome of that application.

IV. THE UNIMODAL BIOMETRIC SYSTEM

As the name suggests, this is a type of biometric process that uses only single biometric feature such as fingerprint, iris, retina, vascular pattern etc for the purpose of authentication of individuals either in e-commerce platforms, e-governance etc. The unimodal biometric system is the most commonly system used in e-commerce due to its simplicity and affordability.

Although no particular biometric system can be said to be 100% efficient, but the loophole of the unimodal biometric systems currently employed in most e-commerce raises various questions on the degree of the security of certain sensitive transactions being done in e-commerce today. Among these weaknesses are: limited degree of freedom for users, noisy input data during enrollment and use, inter-class variations, distinctiveness, non-universality and spoof-attacks. All these weakness found to exist in the unimodal biometric system gives room for a more robust system devoid of such loopholes to evolve. Hence the multimodal biometric system came into being.

V. THE MULTIMODAL BIOMETRIC SYSTEM

This combines two or more unique biometric features of the individuals for authentication. The combination of these features could be in terms of multiple snapshots of a single fingerprint, faces or palm or any combination of choice. However, "the biometric characteristics of an individual is normally a biological feature which can either be genetically implied possibly environmentally altered or feature acquired or learned over time that can be used to recognize or identify the individual"[4]. How secured a system is, usually is based on the amount of time it will take an impostor to decipher the ciphered biometric digital data streams. Multimodal biometric system, due to its high computational requirements is more robust and less prone to attack by an impostor.

"A recent study undertaken within the context of a fingerprint biometric system indicated that a blend of multiple enrollment templates or multiple fingers of a specific user can improve a fingerprint verification system with greater accuracy" [2]. While great attention has been given to the front end processes of this biometric system mode, the backend biometric server architectural framework has received limited attention especially as it concerns emerging economies.

Thus, the next section of this paper lays down a framework for proper implementation of a multimodal biometric server authentication system based on the combination of fingerprints and palm biometric features.

VI. FRAMEWORK FOR IMPLEMENTATION

Under this framework, the fundamental prototype architecture at the backend consists of the e-commerce database server, a Transaction Process Monitor Server, and a multimodal biometric database server. See fig1.5. Similarly, the frontend stage (client/user system) comprises the hardware input device (e.g. a pad with a scanner for both fingerprint and palm made to be used for enrollment of intended e-commerce customers.

However, here, multimodal server architecture was selected due to the fact that the number of potential e-commerce users is expected to grow with time (this is a feature that is commonly found in emerging markets with high population growth rate). The superiority of the multimodal biometric system considered in this paper is in the fact that the palm geometry is most suitable for verification whereas the fingerprint which reveals a lot about an individual's identity is used for identification [2]. This is most suitable and secured for the developing economy where trust barely exists. This framework is such that the actual authentication takes place in

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

the multimodal biometric server, whereas, the Transaction Monitor Server will be responsible for the encryption.

A typical transaction flow for the proposed framework is as shown below: Also, see Fig2.

- User initiates a transaction from his/her client system using the e-commerce website.
- Transaction process monitor server comes up with both private and public key using the RSA algorithm. While the public key is sent to the client computer, the private key is retained in the transaction process monitor.
- At the front end, the Rijindael crypto method will be used to generate the key for encrypting the fingerprint and palm of the user.
- Further encryption of the user fingerprint and palm is done using RSA algorithm with the public key earlier received from the transaction process monitoring server.
- Fully encrypted data is then transmitted to the e-commerce server, and then to the Transaction Monitoring Server.
- Concurrency control is then ensured by the Transaction process Monitoring Server after which the encrypted data is transferred to the multimodal biometric server which decrypts the data and carries out comparison with the biometric data earlier stored.
- If a match is found, information will be transferred to the Transaction Monitoring Server which in turn sends a signal to the e-commerce server to allow the user's transaction or otherwise.

The cryptographic techniques adopted in this framework as well as the physical layout of the servers involved make it most adaptable and suitable in emerging market environment.

VII. The Architectural Framework

In almost all developing economies, infrastructure is lacking. In this case, adequate network and telecommunication infrastructure upon which e-commerce depends are epileptic to the extent that e-commerce cannot perverse the length and breadth of the country in question. Against this backdrop, a suitable backend and frontend architectural framework is highly required. See Fig1.5.

Basically, in engineering network architecture, the physical layout of network entities vis-a-vis server setup can be either a centralized architecture or decentralized architecture [2]. A typical example of a decentralized architecture is the peer-to-peer- (P2P) networks in which case there is no central controlling system.

Contrarily, the centralized architecture has a central controlling system or server. This is usually a case where application and operating system controlling the main operations or the whole setup is domiciled in the server (i.e. client/server framework). Under this arrangement, the client initiates a request for the required resources from the central server. That is, the client does not have an intelligence of its own. The distributed system based on centralized architectural framework has the advantage of adequate access control, but the issues of concurrency and failure recovery have been a major setback, especially in an environment with poor network infrastructure.

As a matter of fact, due to e-commerce requirement that the business owners determine what transaction is required or otherwise on their platform, the centralized architecture (i.e. client/server model) is still most suitable for e-commerce processes and operations.

In view of this fact, e-commerce operations in a developing economy with poor network infrastructure can best be based on a multi-tier architectural framework so as to tackle the issues of concurrency and failure recovery. By concurrency it implies a situation whereby numerous users are accessing a server at the same time. When this happens, if the network architecture is not based on a robust framework, failure rate will be high leading to poor quality of service (QoS) delivery to the consumers. Thus, in an emerging economy, a multi-tier client/server architectural framework is most suitable so as to take care of the issue of concurrency and failure recovery, especially in view of the huge potentials of the markets in the e-commerce sector occasioned by high population of the emerging market economies.

Under this architectural framework, the Transaction Processing (TP) Server will be configured to restrict processes from running concurrently in order to avoid the contention for resources between the e-commerce database server, and the multimodal biometric server. A typical multi-tier client/server architecture that can operate under this framework is the 3-tier architecture as depicted in fig1.5 below, in this, case the e-commerce server houses the database for the e-commerce business, the middleware which serves as the Transaction Processing Monitoring Server and also the multimodal biometric database server for the template capture from the enrollees.

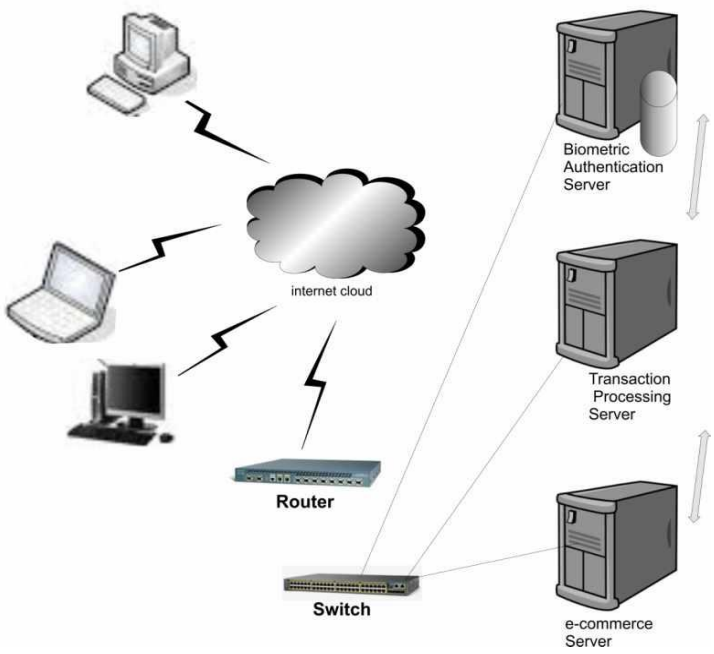


Fig1.5 : Backend Multimodal Biometric Server Architectural Layout

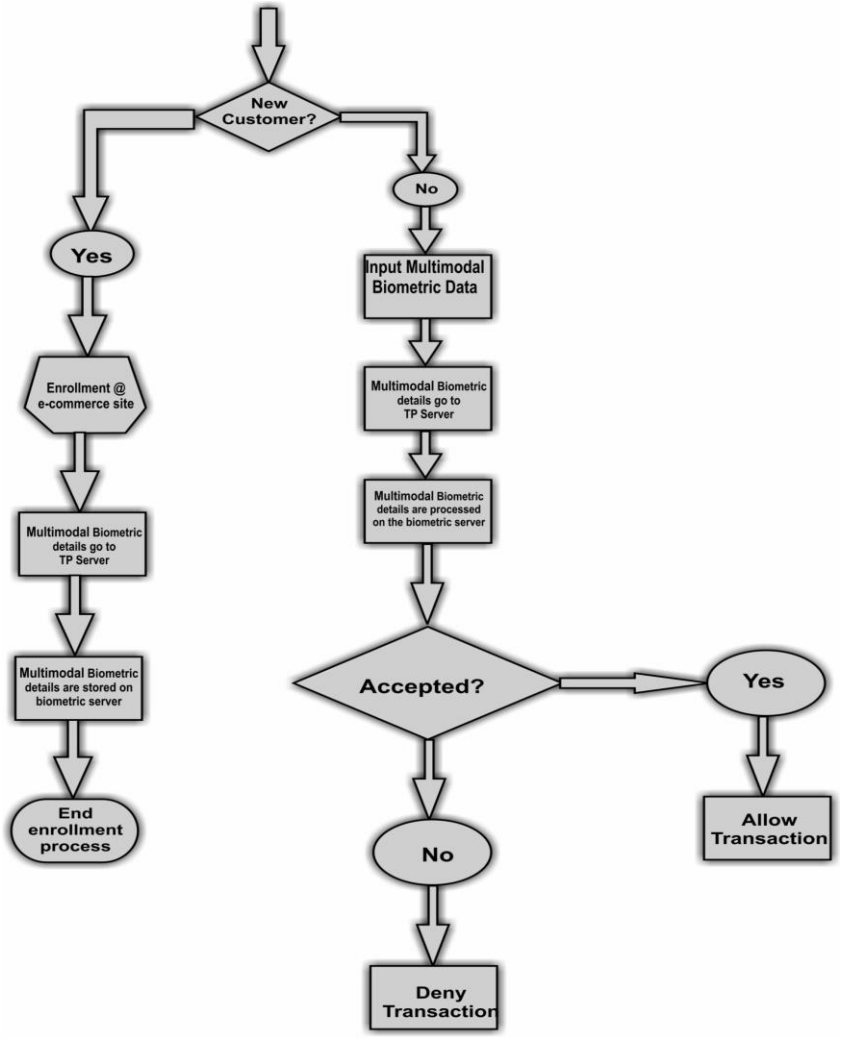


Fig 2: Process Flow of the Multimodal Biometric Architectural System Framework

VIII. CONCLUSION

The paper delved into the analysis of multimodal biometric technology framework and its application to e-commerce. This application was directed to emerging market economies where telecommunication and other network infrastructure are grossly inadequate, yet due to its growing population holds enormous potentials for e-commerce business. The framework for the implementation of the identified multimodal biometric system for the developing economies where trust and identity theft is still a challenge was laid. Information gathered from secondary sources revealed that one of the main reasons business owners, consumers, and general public show high apathy in e-commerce participation is apart from poor infrastructure, lack of security of their transactions.

Hence, in view of this fact, implementation of multimodal biometric system using the laid down framework in e-commerce for emerging market economies will boost public confidence as well as ensure increased participation in e-commerce business transactions.

For the purposes of future research in this area, it is hereby suggested that more work is required in area of data encryption techniques to enhance the security of biometric database. With a more robust data encryption method, more time will be required by any impostor to carry out identity theft there increasing the security of the entire biometric system.

REFERENCES

[1] Mguire, M. (2009), "The birth of biometric security", Anthropology Today, 25,2,pp.9-14, EBSCOhost, [Online]. Available at: http://ehis.ebscohost.com.ezproxy.liv.ac.uk/eds

[2] Kunle Adetunmbi, (2013), "Biometrics in e-commerce", MSc thesis submitted to the University of Liverpool.

[3] Dantcheva, A. Velardo, C.D'Angelo, A.Dugelay (2011), "Bag of soft biometrics for person identification", Multimedia tools & applications, EBSCOhost, [Online]. Available at: http://ehis.ebscohost.com.ezproxy.liv.ac.uk/eds

[4] Schatten, M.Baca, M & Cubrilo M. (2010), "Towards a general definition of biometric systems", International Journal of Computer Science issues (IJCSI), 7,4, pp. 1-7, Computers & Applied Sciences complete, EBSCOhost, [Online]. Available at: http://ehis.ebscohost.com.ezproxy.liv.ac.uk/eds

[5] Al-Dala'in, T. Summons, P. & Suhai, (2009), "A prototype design for enhancing Customer trust in online payments", A journal of computer Science, [Online}. Available at: http://ehis.ebscohost.com.ezproxy.liv.ac.uk/eds

[6] Anderson, R. (2008), "Security Engineering: A guide to building dependable distributed systems". 2nd edition, Indianapolis, Wiley Publishing Inc.

[7] Ariyaeeinia, A. (2003), "Biometrics on the internet", IEE proceedings, vision, image & signal processing [Online]. Available at: http://ehis.ebscohost.com.ezproxy.liv.ac.uk/eds

[8] Asha, S. & Chellappan, C. (2012), "Biometrics: an overview of the technology, issues and applications", International Journal of computer applications [Online]. Available at: http://research.ijcaonline.org/volume39/number.

[9] Basha, A. Palanisamy, V. & Purusothaman, T. (2011), "Efficient Multimodal Biometrics Authentication using Fast Fingerprint verification and enhanced Iris features", Journal of Computer Science, Computers and applied science complete, EBSCOhost, [Online]. Available at: http://ehis.ebscohost.com.ezproxy.liv.ac.uk/eds.

[10] Boukhari, A. Chitroub, S. & Bouraoui, I. (2011), "Biometric signature of private key by reliable Iris recognition based on flexible-ICA Algorithm", International Journal of communications network & system sciences, computers and applied science complete, EBSCOhost [Online]. Available at: http://ehis.ebscohost.com.ezproxy.liv.ac.uk/eds

[11] Corbitt B., Thanasankit, T. & Yi, H., (2003), "Trust and e-commerce: A study of consumer perceptions", Electronic commerce Research & Applications, Business source premier, EBSCOhost, [Online]. Available at: http://www.sciencedirect.com.ezproxy.liv.ac.uk/eds

[12] Dihir, V. Singh, A. Kumar, R & G Singh, (2010), ' Biometric Recognition: A modern era security', International Journal of Engineering & Technology, vol2, no.8, [Online]. Available at: http://www.ijest.info/docs/IJEST

[13] Driscoll, E.C & Fowler, R.C, (1989), 'A comparison of centralized versus distributed architectures in Biometric access control systems', security technology, 1989. Proceedings, 1989 International Camahan Conference, [Online]. Available at: http://ieeexplore.ieee.org/stamp/stamp.jsp

[14] Elbirt, A.J (2005), "Who are you? How to protect against identity theft", Technology & Society Magazine, IEEE, Vol.24, no 2, pp. 5-8, summer 2005, [Online]. Available at: http://ieeexplore.ieee.org/stamp/stamp.jsp

[15] Jones, P. Williams, P. Hillier, Comfort .D, (2007), "Biometrics in retailing", International Journal of Retail & Distribution Management, Vol.35 Iss 3, pp.217-222, [Online]. Available at: http://linkspringer.com.ezproxy.liv.ac.uk/content/pdf

Analysis of MIMO Systems used in planning a 4G-WiMAX Network in Ghana

E.T. Tchao[1], K. Diawuo[2], W.K. Ofosu[3], E. Affum[4]
Department of Electrical/Electronic Engineering, Kwame Nkrumah Univ. of Science and Tech, Kumasi, Ghana[1, 2, 4]
Department of Electrical Engineering Technology, Penn State Wilkes-Barre, USA[2].

***Abstract*—with the increasing demand for mobile data services, Broadband Wireless Access (BWA) is emerging as one of the fastest growing areas within mobile communications. Innovative wireless communication systems, such as WiMAX, are expected to offer highly reliable broadband radio access in order to meet the increasing demands of emerging high speed data and multimedia services. In Ghana, deployment of WiMAX technology has recently begun. Planning these high capacity networks in the presence of multiple interferences in order to achieve the aim of enabling users enjoy cheap and reliable internet services is a critical design issue. This paper has used a deterministic approach for simulating the Bit-Error-Rate (BER) of initial MIMO antenna configurations which were considered in deploying a high capacity 4G-WiMAX network in Ghana. The radiation pattern of the antenna used in the deploying the network has been simulated with Genex-Unet and NEC and results presented. An adaptive 4x4 MIMO antenna configuration with optimally suppressed sidelobes has been suggested for future network deployment since the adaptive 2x2 MIMO antenna configuration, which was used in the initial network deployment provides poor estimates for average BER performance as compared to 4x4 antenna configuration which seem less affected in the presence of multiple interferers.**

Keywords—WiMAX; Performance; BER; MIMO; Interference

I. INTRODUCTION

Broadband internet is quite expensive in Ghana and in many Sub-Saharan African countries and as such, subscribers in many African countries with high data demand have no option than to rely on network operators who offer poor Digital Subscriber Lines (DSL) access and long customer connection times. Broadband internet subscribers were relieved and looked forward to enjoying lasting high data rate service at cheaper cost when deployment of WiMAX begun in Ghana.

It is a fact that planning a new network with a limited number of subscribers is not the real problem. The difficulty is to plan a network that allows future growth and expansion to meet the high demand for mobile data services. In 2nd generation systems, coverage planning was the most important and also the sufficient issue for operating the network [1].

Coverage prediction and capacity estimation were mostly well separable. In 3G and 4G networks, users operate on the same frequency carrier and as such the number of simultaneous connections directly influences the system capacity. Because of the need for 3G and 4G networks to continue offering higher data rate service as the network grows, planning these networks cannot be separated into coverage and capacity estimation as there is the need to increase capacity to serve multiple services like speech, internet and high data rate interactive services [2].

Even though 4G networks need to increase capacity in order to serve the increasing user demand, no breakthroughs in coding or modulation schemes are to be expected and additional spectrum resources are scarce [1]. This development has stimulated research into interference cancellation techniques [3], multicell processing [4], and cognitive radio [5] to improve the spectral efficiency of today's 4G wireless networks. However, none of these techniques is likely to carry the expected increase in mobile data traffic alone and further Base Station densification seems necessary. Base Station densification results in interferences since most of these 4G technology reuses frequency [6]. The resulting interferences as Base Station densification increase in wireless communication systems has motivated studies into Multi-Input Multi-Output (MIMO) systems with interferences in [7-12].

It implies that, in order to plan the deployed WiMAX network in Ghana to continue providing its subscribers higher data rate services as subscribers grow exponentially on the network, more Base Stations will have to be deployed. These Base Stations densification will in turn lead to high level of interference in the network. These expected high levels of interference will undoubtedly reduce network performance of the WiMAX network which will be disappointing to subscribers in Ghana who were eagerly looking forward to a long term solution to the expensive and unreliable broadband services in the country.

In order to help maintain the present high network performance, BER analysis of the current MIMO system in the presence of multiple interference sources is very important. BER performance analysis plays an important role in helping broadband wireless networks achieve maximum capacity while maintaining an acceptable grade of service and good speech quality [12].

This paper studies the potential sources of interference in the network and presents analysis into the BER performance of the MIMO antenna configuration which was used in the initial network deployment in the presence of multiple interferers. The analytical expressions are validated against Monte Carlo numerical methods and results presented accordingly.

II. Initial Wimax Network

The network under study covers 55 sq km in the central parts of Accra and Tema. This network was designed to maximize coverage and as such Base Stations have been optimally placed to augment coverage problems. Initial planning met the main requirements of achieving the simulated coverage area and low Co channel and Adjacent channel interference levels as predicted for maintaining good quality of service[13].

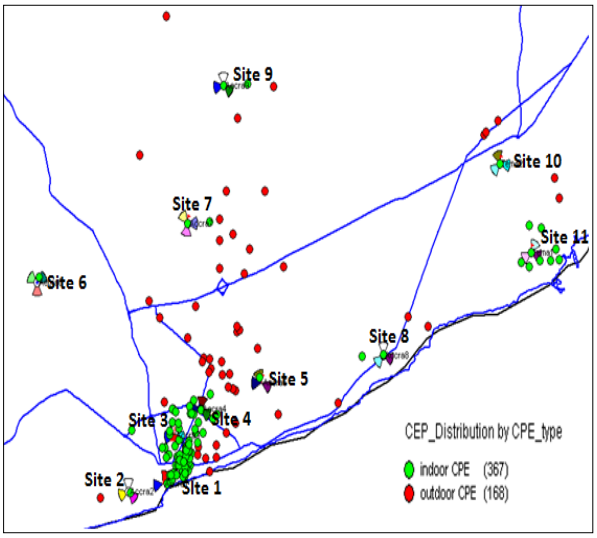


Fig.1. Distribution of CPE in the Deployed Network.

Currently, the number of Mobile Stations (MS) in the network has doubled and in order to serve the increasing demand by subscribers, more Base Stations will have to be deployed. Because of the short inter site distances, as seen in Table 1, deploying more Base Stations in the network will result in smaller cell sizes being created.

TABLE I. Inter-Site Distance For The Deployed Network [14]

Site A	Site B	Inter- site distances (km)
Site 1	Site 2	1.938
Site 1	Site 3	1.072
Site 1	Site 4	2.423
Site 4	Site 5	1.563
Site 5	Site 6	6.20
Site 5	Site 8	7.542
Site 5	Site 7	8.590
Site 7	Site 9	5.342
Site 6	Site 7	5.734
Site 10	Site 11	4.494
Site 10	Site 8	23.33

Since a frequency reuse scheme of PUSC 1x3x3 was used in deploying the network in Figure 1 under a dense urban Sub-Saharan African environment, it is likely that as more Base Stations are added and cell sizes become smaller, there will an exponential increase in interferences and therefore the need for proper evaluation of the performance of the initial MIMO antenna configuration in the presence of multiple interferers.

III. MIMO Channel Model

One important practical implementation problem is that MIMO systems require knowledge of the channel conditions. Therefore, we consider a narrowband block fading channel with n_R receiving antennas, n_T transmit antennas from an interfering source, possibly representing several different users characterized by [15]:

$$\boldsymbol{y} = \varsigma\boldsymbol{\omega} + \varsigma\boldsymbol{I}\,\boldsymbol{\omega}\boldsymbol{I} + z$$

Here, $\boldsymbol{y} \in \mathbb{C}^{n_R}$ is the received signal vector. $\boldsymbol{\omega} \in \mathbb{C}^{n_T}$ is the transmitted complex Gaussian distributed signal vector with zero mean and covariance $\beth = \mathrm{E}[\boldsymbol{\omega}\boldsymbol{\omega}^{\varsigma}], \boldsymbol{\omega}_I \in \mathbb{C}^{n_I}$ is the interfering complex Gaussian distributed signal vector with zero mean and covariance $\beth_{\boldsymbol{I}} = \mathrm{E}\left[\boldsymbol{\omega}_I\boldsymbol{\omega}_I^{\varsigma}\right]$, and $\boldsymbol{z} \in \mathbb{C}^{n_R}$ is additive zero-mean white noise with entries $z_a \sim \mathcal{N}_c(\boldsymbol{0},\boldsymbol{1})$. The channel matrices $\varsigma \in \mathbb{C}^{n_R \times n_T}$ and $\varsigma_I \in \mathbb{C}^{n_R \times n_I}$ has been model separately for correlated Rician fading. Thus, they can be written as

$$\begin{cases} \varsigma = \bar{\varsigma} + \boldsymbol{\rho}^{\boldsymbol{1}/2}\varsigma_w\boldsymbol{\Psi}^{1/2} \\ \varsigma = \bar{\varsigma}_I + \boldsymbol{\rho}_I^{\,1/2}\varsigma_{w,I}\boldsymbol{\Psi}_I^{\,1/2} \end{cases} \qquad (1)$$

Where $\bar{\varsigma}$ and $\bar{\varsigma}_I$ represent the mean values of ς and ς_I, respectively, and are related to the presence of LOS components, $(\varsigma_w)_{ab}, (\varsigma_{w,I})_{ab} \sim \mathcal{N}_c(\boldsymbol{0},\boldsymbol{1})$, and the positive semidefinite matrices $\boldsymbol{\Psi}(\boldsymbol{\Psi}_I)$ and $\boldsymbol{\rho}(\boldsymbol{\rho}_I)$ are the transmit signal (interference) and receive signal (interference) spatial correlation matrices respectively. The covariance between the different entries of ς and ς_I satisfies the identities

$$\begin{cases} \mathrm{cov}\big((\varsigma)_{ij}(\varsigma)_{i'j'}\big) = (\boldsymbol{\rho})_{ii'}(\boldsymbol{\Psi})^*_{jj'} \\ \mathrm{cov}\big((\varsigma_{\boldsymbol{I}})_{ij}(\varsigma_{\boldsymbol{I}})_{i'j'}\big) = (\boldsymbol{\rho}_{\boldsymbol{I}})_{ii'}(\boldsymbol{\Psi}_{\boldsymbol{I}})^*_{jj'}. \end{cases}$$

Extending these to multiple interfering transmitters, as in the case in network under discussion, (1) can be applied to model multiple interfering transmitter under the assumption of a common receive correlation matrix for each interfering source.

Indeed, assume that we have N_I interfering users, so that the channel of the interferer is $i\ (i = 1, \dots, N_I)$ of the form $\varsigma_I^i = \bar{\varsigma}^i + \boldsymbol{\rho}^{1/2}\boldsymbol{W}^{\boldsymbol{i}}\boldsymbol{\Psi}_{\boldsymbol{I}}^{\boldsymbol{i}\,\boldsymbol{1}/\boldsymbol{2}}$ and

$\boldsymbol{y} = \varsigma\boldsymbol{\omega} + \sum_{i=1}^{N_I}\varsigma_I^i\boldsymbol{\omega}_I^i + \boldsymbol{z}$. Then we set

$$\boldsymbol{\omega} \triangleq \left[{\boldsymbol{\omega}_I}^{1^{\mathrm{T}}}, \dots, {\boldsymbol{\omega}_I}^{N_{\mathrm{I}}{}^{\mathrm{T}}}\right]^{\mathrm{T}}, \quad \boldsymbol{\Psi} \triangleq \oplus_{i=1}^{N_I}\boldsymbol{\Psi}_I^i,$$

$$\boldsymbol{W} \triangleq [\boldsymbol{W}^1, \dots, \boldsymbol{W}^{N_1}], \text{ and } \bar{\varsigma} \triangleq \left[\bar{\varsigma}^1, \dots, \bar{\varsigma}^{N_1}\right].$$

Different receive correlation matrices for the interfering transmitter can be modeled by introducing virtual delays in combination with a wideband channel model, as proposed in [15][16][17]. Following the Approach in [18] we define

$$\begin{cases} \grave{\varsigma} \triangleq \bar{\varsigma}\beth^{1/2}\grave{\boldsymbol{\Psi}} = \boldsymbol{\Psi}^{1/2}\beth\boldsymbol{\Psi}^{1/2} \\ \grave{\varsigma} \triangleq \grave{\varsigma}_I\beth_I^{\,1/2}\grave{\boldsymbol{\Psi}}_I = \boldsymbol{\Psi}_I^{\,1/2}\beth_I\boldsymbol{\Psi}_I^{\,1/2} \end{cases} \quad (2)$$

Then, the transmitted signal and interference covariance matrices are implicitly accounted for into $\grave{\varsigma}, \grave{\varsigma}_I, \grave{\boldsymbol{\Psi}}$ and $\grave{\boldsymbol{\Psi}}_I$.

According to these definitions, the total received power is given by [18]

$$\mathrm{E}[\|\boldsymbol{y}\|^2] = \|\acute{\varsigma}\|^2 + \mathrm{tr}(\boldsymbol{\rho})\mathrm{tr}(\grave{\boldsymbol{\Psi}}) + \|\acute{\varsigma}_I\|^2 + \mathrm{tr}(\boldsymbol{\rho_I})\mathrm{tr}(\grave{\boldsymbol{\Psi}}_I) + n_R$$

Splitting the total received power components into direct and diffuse parts we obtain Rician factors

$$K = \frac{\|\acute{\varsigma}\|^2}{\mathrm{tr}(\boldsymbol{\rho})\mathrm{tr}(\grave{\boldsymbol{\Psi}})} K_I = \frac{\|\acute{\varsigma}_I\|^2}{\mathrm{tr}(\boldsymbol{\rho_I})\mathrm{tr}(\grave{\boldsymbol{\Psi}}_I)} \quad (3)$$

Then signal-to-noise and interference-to-noise power ratios can be defined as

$$\begin{cases} SNR = \frac{(K+1)\mathrm{tr}(\boldsymbol{\rho})\mathrm{tr}(\grave{\boldsymbol{\Psi}})}{n_R} \\ INR = \frac{(K_I+1)\mathrm{tr}(\boldsymbol{\rho})\mathrm{tr}(\grave{\boldsymbol{\Psi}}_I)}{n_R} \end{cases} \quad (4)$$

It can be noticed that the definitions of the Rician factors and of the SNRs in (3) and (4) depend on the full transmit covariance matrices $\beth$ and $\beth_I$ and not only on their traces, unless they are proportional to the identity matrix. This will be an issue when the channel capacity of the MIMO channel in the presence of multiple interferers is evaluated against the SNR.

IV. WiMAX ANTENNA FEATURES

An Adaptive Antenna Systems (AAS) was used in deploying the WiMAX network under discussion. The advantage in using adaptive antennas is that, it utilizes beam-forming techniques for focusing and directing the wireless signal between the base station and the receiver station. This reduces interference from other external sources and noises, as the beam is focused directly between two points.

The use of these techniques provides advantages on the basis of coverage, self-installation, power consumption, frequency re-use and bandwidth efficiency. Use of beam form techniques and MIMO under WiMAX reduces interference while increasing throughput and efficiency [19].

Antennas used in wireless communications have unwanted upper sidelobes. Sidelobes are known sources of interference in wireless networks. When the antenna has no downtilt, the upper sidelobes travel upwards and there is little chance of interference.

However, when downtilt is applied to the antenna as in the case of the deployed network, these unwanted upper sidelobes are now directed towards neighboring cells as shown in Figure 2. This has the potential of causing interference in the network since frequency is being reused.

Antenna pattern for the coverage simulations has been summarized in table 2 below. According to [20], all BS antenna elements have the beam pattern defined by 3GPP2. It is given by the formula.

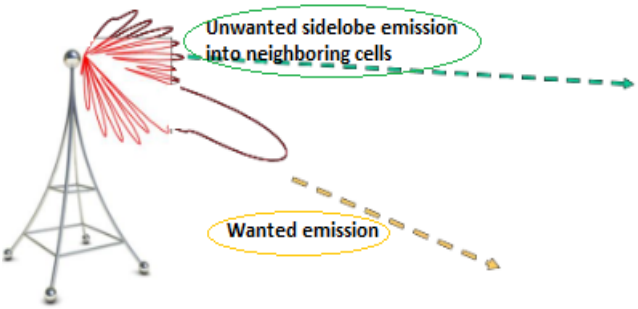


Fig.2. Antenna Radiating Pattern scenarios

$$G(\theta) = G_{max} + \max[-12\left(\frac{\theta}{\theta_{3dB}}\right)^2, -G_{FB}]$$

Gmax in (dBi) is the maximum antenna gain in the boresight direction. θ is the angle of arrival relative to boresight, θ_{3dB} is the 3dB beam width. G_{FB} is the front-to-back ratio in dBi

In order to understand the effect multiple Base Station deployments will cause to the overall network performance of the MIMO systems, the radiating pattern of the antenna has been modeled in the next section and discussed

V. SIMULATION RESULTS

Simulation of the radiation pattern of the WiMAX antenna has been done with Genex-Unet and NEC and shown in Figures 3, 4 and 5. The simulation parameters have been summarized in Table 2

From the non-optimal unsuppressed sidelobes as seen in the antenna's vertical radiation pattern in Figure 5, each BS in the network is a potential source of interference.

Performing a bit-error-rate simulation can be a lengthy process. In order to make sure that our results are statistically significant and allows for the different MIMO configurations to be compared in a fair manner, a Monte Carlo approach was used to model the BER performance of a BS in the presence of 10 interfering Base Stations. The result of the simulation has been shown in Figure 6.

TABLE II. RADIATING PATTERN SIMULATION PARAMETERS

Frequency Range	2300 MHz ~2700MHz
VSWR	≤1.5
Input Impedance	50Ω
Gain	18 dBi±0.5dBi
Polarization	±45°
Horizontal Beamwidth (3dB)	60±5°
Vertical Beamwidth (3dB)	7°±0.5°
Electrical Downtilt	2°
Isolation Between Ports	≥30dB
Front to Back Ratio	≥30dB±2 dB
Cross Polarization Ratio	≥18dB
Upper Sidelobe Suppression	≥18dB
Null-Fill	≤18dB
Max, power	250W

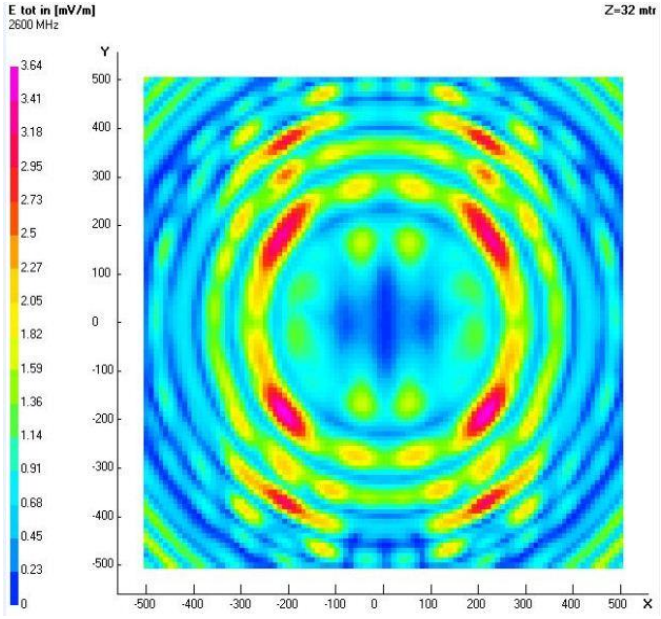


Fig.3. Antenna radiating pattern

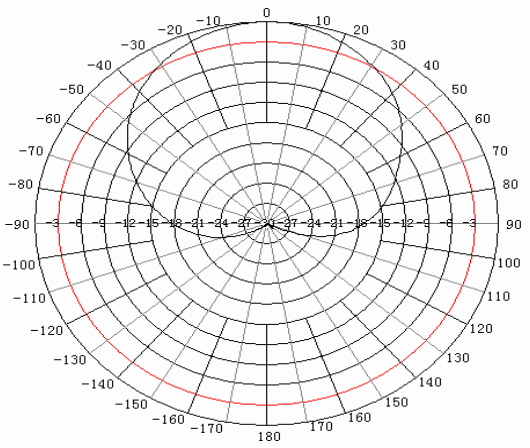


Fig.4. Horizontal Radiating Pattern

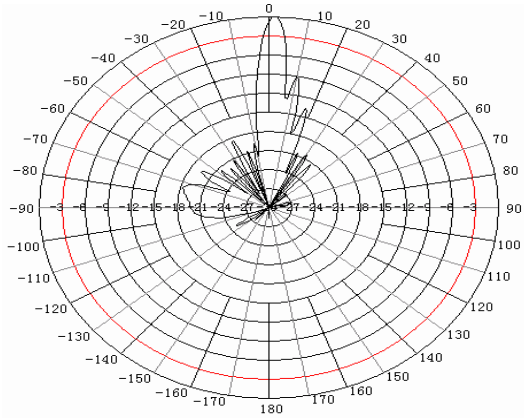
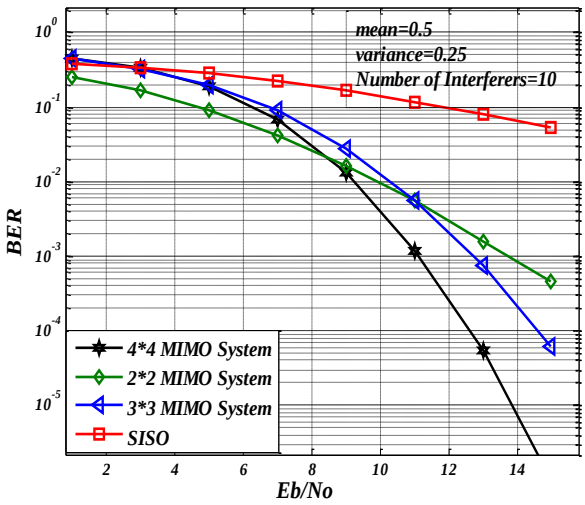


Fig.5. Vertical Radiating Pattern



.

Fig.6. Performance of MIMO System in the presence of 10 interferers.

A bounding technique of specifying the mean and variance was applied to ensure that the simulated BER estimates are fairly accurate.

VI. Discussion Of Results

Interference reduces the integrity of signals travelling between transmitters and receivers. In cellular networks, this can often be caused by frequency reuse, therefore, stray radiation from one cell can disrupt transmission in other nearby cells using the same frequency. As the deployed WiMAX network uses a PUSC 1*3*3 frequency reuse scheme, the presence of unsuppressed sidelobes as seen from the antenna radiating pattern simulation in Figure 5 can significantly contribute to interference.

The BER simulation result using the Monte Carlo approach shows that the number of transmit (n_T) and receive (n_R) antennas affect the BER performance of the System. This complements the earlier results by Erwin *et all* on dealing with asymptotic statistics of the mutual information for correlated Racian Fading MIMO channels with interference. The results seems to be important for large number of antennas as the performance of the 4x4 MIMO system seemed the least affected by the variance and the number of interferers.

With the comparatively low performance of the 2x2 MIMO configurations, data transmitted on the network could be corrupted in the presence of multiple interferers. This may result in data having to be re-transmitted, and this reduces network performance. With Studies predicting a 1000X increase in data traffic from 2010-2020 [21], BER performance in the presence of interference is a critical design issue so as not to put severe strain on deployed 4G networks.

With the 4X4 MIMO configurations giving better performance in the presence of interference, it is recommended that future deployment of Base Stations in the network should use 4x4 MIMO antennas with optimized sidelobes suppression techniques.

VII. Conclusion

Broadband wireless networks provide significant and promising growth marketplace for the telecom companies in Ghana to deliver a variety of applications and services to subscribers. In order for any network to sustain its initial high network performance after deployment, parameter analysis is necessary especially as the network grows and the number of potential interference to the network increases.

In the context of system performance, it can be concluded that 4x4 MIMO configuration provides satisfactory result than the initial 2x2 MIMO configuration in the presence of multiple interferers. A 4x4 MIMO antenna configuration with upper sidelobes optimally suppressed is recommended to be used in subsequent deployments especially in the presence of significant unsuppressed sidelobe emissions of the current antenna.

Evaluation of the effects of using antennas with upper sidelobes optimally suppressed against using the current antennas with typical upper sidelobe suppression technique will be presented when enough data is available from the field trial measurements.

References

[1] W. Webb, Wireless Communications: The Future. John Wiley & Sons Ltd, 2007.

[2] M. Webb, M. Beach, and A. Nix, “Capacity limits of MIMO channels with co-channel interference," in Proc. IEEE Veh. Technol. Conf. Spring, May 2004, pp. 703-707, Milan, Italy.

[3] J. Andrews, “Interference Cancellation for Cellular Systems: A Contemporary Overview,” IEEE Wireless Commun. Mag., vol. 12, no. 2,pp. 19—29, Apr. 2005.

[4] D. Gesbert, S. V. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu,“Multi-cell MIMO Cooperative Networks: A New Look at Interference,”IEEE J. Sel. Areas Commun., vol. 28, no. 9, Dec. 2010.

[5] R. Tandra, A. Sahai, and S. Mishra, “What is a Spectrum Hole and What Does it Take to Recognize One?” Proceedings of the IEEE, vol. 97, no. 5, pp. 824 –848, May 2009.

[6] J. Hoydis and M. Debbah, “Green, Cost-effective, Flexible, Small Cell Networks,” IEEE ComSoc MMTC E-Letter”, vol. 5, no. 5, Sep. 2010.

[7] A. Shah, A. M. Haimovich, M. K. Simon, and M. S. Alouini, “Exact bit-error probability for optimum combining with a Rayleigh fading Gaussian cochannel interferer," IEEE Trans. Commun., vol. 48, no. 6, pp. 908-912, June 2000.

[8] R. S. Blum, “MIMO capacity with interference," IEEE J. Sel. Areas Commun., vol. 21, no. 5, pp. 793-801, June 2003.

[9] Y. Song and S. D. Blostein, “MIMO channel capacity in co-channel interference," in Proc. 21st Biennial Sym. Commun., Jan. 2002, pp. 220-224, Kingston, Canada.

[10] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, “MIMO capacity throughput correlated channels in the presence of correlated interference and noise: a (not so) large N analysis," IEEE Trans. Inf. Theory, vol. 49, no. 10, pp. 2545-2561, Nov. 2003.

[11] M. Kang, L. Yang, and M. S. Alouini, “Capacity of MIMO channels in the presence of co-channel Interference," Wireless Commun. Mob. Comput., no. 7, pp. 113-125, Mar. 2006.

[12] M. Chiani, M. Z. Win, and H. Shin, “MIMO networks: the effects of interference," IEEE Trans. Inf. Theory, vol. 56, no. 1, pp. 336-349, Jan. 2010.

[13] E.T. Tchao, W.K. Ofosu, K. Diawuo: ” Radio Planning and field trial measurement for a 4G WiMAX network in a Sub- Saharan Africa Environment”, Conference proceedings, IEEE Wireless Telecommunications symposium, Phoenix, Arizona, 17-19th April 2013

[14] E.T. Tchao, W.K. Ofosu, K. Diawuo, E. Affum, K. Agyekum; “Interference Simulation and Measurements for a Deployed 4G-WiMAX Network in an Urban Sub-Saharan African Environment”, International Journal of Computer Applications Volume 71-No. 14, June 2013. Published by Foundation of Computer Science, New York, USA

[15] Erwin, R and Giorgio, T., “Asymptotic Statistics of the Mutual Information for Spatially Correlated Racian Fading MIMO Channel with Interference” IEEE Transaction on Information Theory, Vol. 75, No. 7, April 2010

[16] Riegler E., Taricco G., “Asymptotic ergodic capacity region and rate optimization of a multiple access OFDM MIMO channel with separately correlated Rician fading,” in Proc. IEEE GLOBECOM 2008, New Orleans, LA, Nov. 30-Dec. 4 2008

[17] Riegler, E. and Taricco, G., ”Asymptotic ergodic capacity of wideband MIMO channels with separately correlated Rician fading, ” in Proc. IEEE GLOBECOM 2008, New Orleans, LA, Nov. 30-Dec. 4 2008

[18] Taricco, G., “Asymptotic mutual information statistics of separately correlated Rician fading MIMO channels,” IEEE Trans. Inf. Theory, vol.54, no. 8, pp. 3490-3504

[19] WiMAX Forum White paper: “TDD-TDD Interference Analysis involving Synchronized WiMAX systems” 18 September, 2009

[20] Krishna Ramadas and Raj Jain: WiMAX System Evaluation Methodology", WiMAX Forum White Paper, Jan 2008 [online: www.cse.wustl.edu/ on 14th December, 2012]

[21] Gee Rittenhouse: ”A 1000X of data-Technology Directions in the Wireless market”, Conference proceedings, IEEE Wireless Telecommunications symposium, Phoenix, Arizona, 17-19th April 2013

A Hybrid Model for Secure Data Transfer in Audio Signals using HCNN and DD DWT

B. Geetha vani [1]
Research scholar, Dept. of CSE,
JNTU Kakinada. AP, India

Prof. E. V. Prasad [2]
Professor, Dept. of CSE & Rector,
JNTU Kakinada, AP, India

***Abstract*—In today's world, there are a number of cryptographic and steganography techniques used in order to have secured data transfer between a sender and a receiver. In this paper a new hybrid approach that integrates the merits of cryptography and audio steganography is presented. First, the original message is encrypted using chaotic neural network and the resultant cipher text is embedded into a cover audio using Double Density Discrete Wavelet Transform (DD DWT). The resultant stego audio is transmitted to the receiver and the reverse process is done in order to get back the original plain text. The proposed method presents a Steganography scheme along with the cryptography scheme which enhances the security of the algorithm.**

Keywords—Cryptography; Hopfield Chaotic Neural Network; Audio Steganography; Double Density Discrete Wavelet Transform.

I. INTRODUCTION

The increasing internet usage stems from the growing availability of global communication technology that has led to electronically included information gathering and distribution. However, the challenge it presents in terms of information security is enormous. The need to secure information within the global network is of paramount importance so that the user information is preserved until it reaches its destination undisclosed. Providing a secure framework that conceals information content and sender/receiver identity is a matter of prime interest.

The two popular approaches to information confidentiality are Cryptography and Steganography [1,2,3]. Cryptography is the study and practice of protecting information by data encoding and transformation techniques. Steganography, a concealed writing, is the art and science of hiding the fact that communication is taking place.

Steganography techniques, based on the cover file, can be categorized as Image steganography, Text Steganography, Audio steganography and Video steganography. In digital audio steganography system, secret message is embedded in audio file. The binary sequence of audio file is slightly changed by adding secret message to it. This modification

Should not be made available to the human ear. Embedding secret messages in audio file is more difficult than embedding information in digital image. In order to hide secret messages, various methods for embedding information in digital audio have been introduced. These methods range from simple techniques that insert information in the form of noise in audio signal to more powerful methodologies using signal processing techniques. Many audio steganography methods use Least Significant Bit (LSB) insertion technique [4,6] to hide the secret message. But techniques have been developed to detect secret message which is present at LSB position [12]. Hence, an improvement over this is use of robust audio steganography techniques using wavelets [11, 16, 17].

In the proposed system, audio steganography method uses double density discrete wavelet transforms. In order to provide better security chaotic neural network is used for encrypting the secret message. The rest of the paper is organized as follows. In Section 2, Literature review is presented. In Section 3, the proposed system is described. In Section 4, experimental results of proposed approach are shown. Concluding remarks are provided in section 5.

II. LITERATURE REVIEW

A. Audio Steganography Techniques

Audio Steganography techniques can be adopted either in temporal domain or transform domain.

1) Temporal domain Techniques

a) LSB

LSB is one of the earliest, simplest and commonly used methods, for hiding information in audio steganography. In LSB method, as shown in Fig.1, the least significant bits of the cover media/original audio is altered to include the secret message. Even though this is a simple method, an attacker can easily extract the secret message from the stego object.

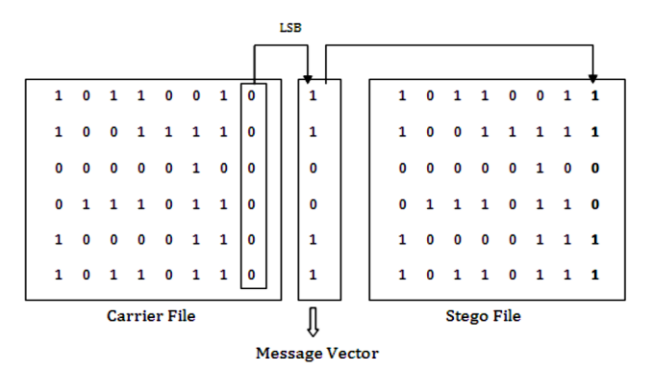


Fig. 1. LSB Insertion

b) Parity coding

Parity coding technique operates on a group of samples instead of individual sample. Here individual samples are grouped and parity of each group is calculated. For inserting message bit one by one, check the parity bit of a group of samples. If the parity bit and message bit matches then no operation need to be performed else change the LSB's of any one of the individual samples in that group to make the parity bit equal to the message bit.

c) Echo hiding

In echo hiding method data is embedded in the echo part of the host audio signal. The echo is a resonance added to the host signal and hence the problem with the additive noise is avoided here. While using echo hiding, to avoid echo audibility, three important parameters to be considered are initial amplitude, offset (delay), and decay rate. The method suffers from lenient detection and low detection ratio. Due to its low embedding rate and low security use of this technique is not interesting among researchers.

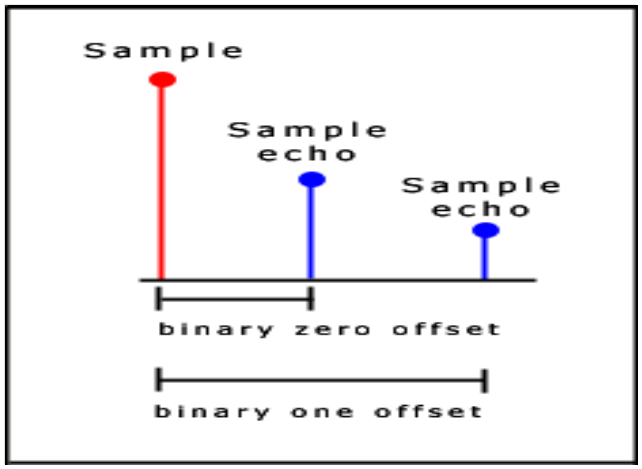


Fig. 2. Echo Hiding

B. Transform domain Techniques

Frequency domain techniques and wavelet domain technique comes under transform domain. The main techniques that can usually adopted in frequency domain include tone insertion, phase coding and spread spectrum technique.

1) Tone insertion

Frequency masking property is exploited in tone insertion method. A weak pure tone is masked in the presence of a stronger tone. This property of inaudibility is used in different ways to embed information.

2) Phase coding

Phase coding method is based on the fact that the phase components are not audible to human as noise components. This method as shown in Fig. 3, embeds the secret message bits as phase shift in the phase spectrum of the original audio signal. The method tolerates better signal distortion, better robustness, but it does not survive low pass filtering.

Here the secret message is inserted only at the phase vector of the first signal segment.

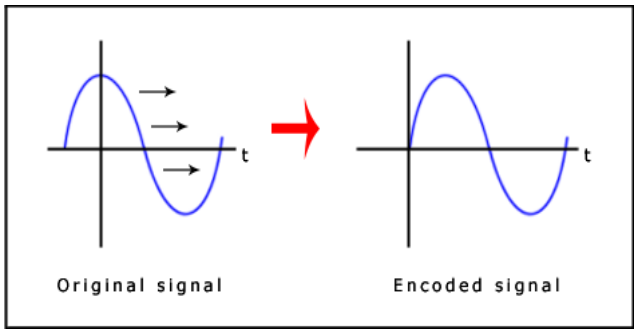


Fig. 3. Phase Coding

3) Spread spectrum technique

This technique, explained in Fig. 4, takes the advantage of masking property of Human Audio System (HAS). A masking threshold is calculated using a psycho-acoustic model. The spread signal now lies below the masking threshold. Apart from phase shifting, here the secret message is distributed along with the host signal. Here the final signal occupies a bandwidth which is more than what is actually required for transmission.

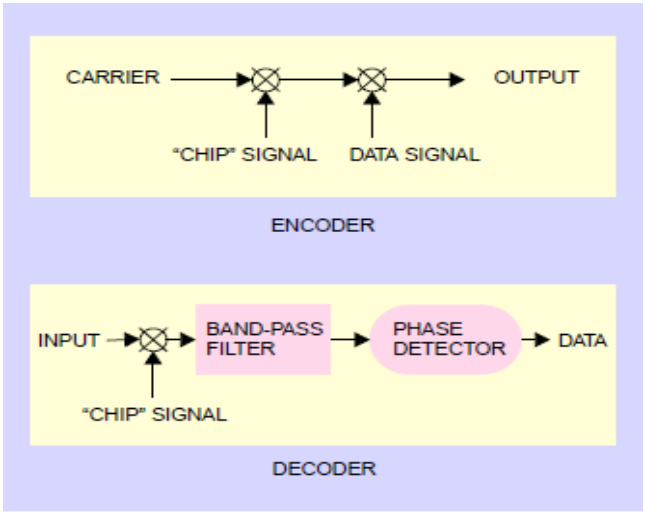


Fig. 4. Spread Spectrum

4) Wavelet domain

Wavelet domain technique, that uses wavelet coefficients, is suitable for frequency analysis because of its multi-resolution property and provides access to both most significant parts and details of spectrum. With the use of inverse transform, the stego signal can be reconstructed.

III. Proposed System

In the Proposed hybrid model, Steganography is combined with Cryptography to transmit message in a highly secured manner and makes the system theoretically and practically unbreakable. For embedding the information, the steps involved are as follows,

1) Get the Plain text which is to be sent to the recipient.

2) Transform the plaintext in to cipher text by applying an encryption process using chaotic neural network.

3) Embed the cipher text inside the cover audio file using double density discrete wavelet transform.

4) The resulted stego audio file is communicated through any communication channel to the receiver.

For extraction of information, the steps involved at the receiver side are as follows,

1) On receiver side Median noise filter is applied to remove noise from the stego audio.

2) By applying inverse DD DWT the embedded message is extracted from the Stego audio file.

3) Since the obtained message is in the scrambled form, decryption is performed.

4) Finally, the receiver is able to read the actual secret message sent at the sender's end.

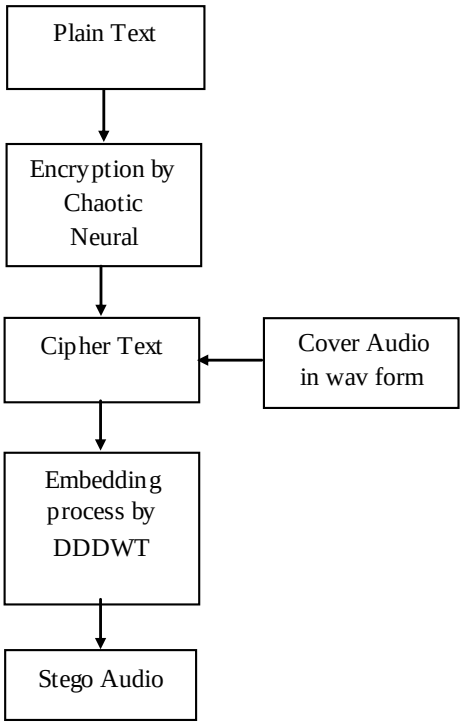


Fig. 5. Embedding process

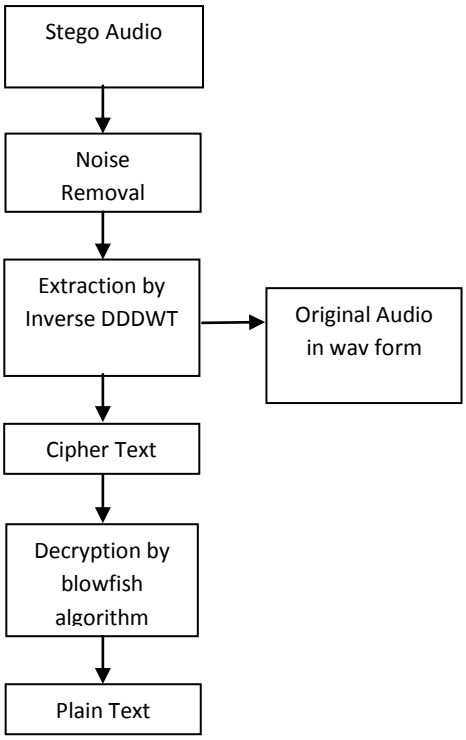


Fig. 6. Extraction process

A. Cryptographic Approach

1) Hopfield Chaotic Neural Network Based Encryption

The encryption methodology adopted for encrypting text characters plays a vital role in deciding the embedding capacity and the level of robustness and security of the entire Steganography system. Hopfield Chaotic Neural network is a suitable environment for cryptography because of some interesting properties like ergodicity, sensitive dependence of initial conditions and control parameters and high speed of information transmission. Yu et al [7] designed a delayed chaotic neural network based cryptosystem, which makes use of the chaotic trajectories of two neurons to generate basic binary sequences for encrypting plaintext. In Chaotic Neural Network, the weights and biases are determined by a chaotic sequence, a binary random deterministic sequence, and are used to mask or to scramble the original information. The encryption algorithm [8] is used for obtaining the cipher text. The Chaotic neural network consumes less computational power and the sequence generated using this is unpredictable leading to highly secured and efficient in terms of power.

2) Double-Density Discrete Wavelet Transform

The double-density DWT is an improved critically sampled DWT with following additional properties:

a) It employs one scaling function and two distinct wavelets, which are designed to be offset from one another by one half.

b) The double-density DWT is over complete by a factor of two and

c) It is nearly shift-invariant.

In two dimensions, this transform outperforms the standard DWT in terms of denoising. However, there is room for improvement because not all of the wavelets are directional. That is, although the double-density DWT utilizes more wavelets, some lack a dominant spatial orientation, which prevents them from being able to isolate those directions.

B. Implementation of DD-DWT

To implement the double-density DWT, it is necessary to first select an appropriate filter bank structure. The filter bank proposed in Fig. 7 illustrates the basic design of the double-density DWT.

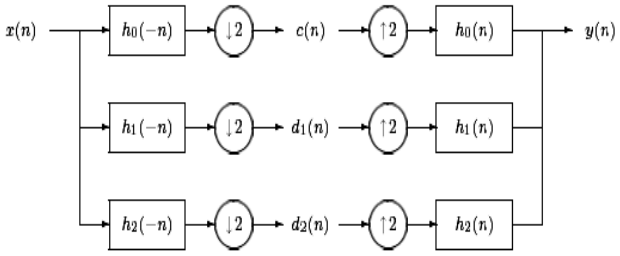


Fig. 7. A 3-Channel Perfect Reconstruction Filter Bank.

The analysis filter bank consists of three analysis filters - one low - pass filter denoted by $h_0(-n)$ and two distinct high-pass filters denoted by $h_1(-n)$ and $h_2(-n)$. As the input signal x(n) travels through the system, the analysis filter bank decomposes it into three sub-bands, each of which is then down-sampled by 2. From this process the signals c (n), $d_1(n)$, and $d_2(n)$, which represent the low frequency (or coarse) sub-band, and the two high frequency (or detail) sub-bands are obtained.

The synthesis filter bank consists of three synthesis filters - one low-pass filter denoted by $h_0(n)$ and two distinct high-pass filters denoted by $h_1(n)$ and $h_2(n)$ - which are essentially the inverse of the analysis filters. As the three sub-band signals travel through the system, they are up-sampled by two, filtered, and then combined to form the output signal y (n).

C. Significance of the Hybrid Model

The proposed method integrates two different techniques for the secured data transmission. They are

- Enciphering & Deciphering phase with the Cryptography
- Embedding & Extraction of data with the Steganography

1) It's quite hard for the eavesdroppers to realize the chaotic neural network encryption hence probability of attack is less when compared with the normal encryption algorithms.

2) Embedding process is done using DD DWT, such that the resulting stego audio is similar to original audio and provides robustness.

3) In this model on the receiver side median filter is used to remove noises in stego audio.

IV. Results And Analysis

Experiments have been conducted to prove the efficiency of the proposed algorithm. The Quantitative performance of the proposed algorithm is evaluated based on Peak signal to noise ratio (PSNR) and Mean Square Error (MSE)which are given in equations 1 & 2 respectively.

The peak signal to noise ratio (PSNR) was utilized to evaluate the stego audio quality. PSNR [15] is often expressed on a logarithmic scale in decibels (dB), it is defined as:

$$PSNR = 10 * \log_{10}(255^2/MSE)\ (dB) \qquad (1)$$

Where MSE (15) is the mean square error between the cover and stego audio. For a cover audio whose size is defined in terms of W and H, MSE is defined as:

$$MSE = \frac{1}{W*H}\left(\sum_{i=1}^{W}\sum_{j=1}^{H}(A_{ij}\text{-}A^{1}_{ij})^2\right) \qquad (2)$$

Where W is the amplitude of the signal, H is the frequency of the audio signal and A_{ij} and A^{1}_{ij} are pitch values of cover and Stego audio.

An audio file with ".wav" extension has been selected as cover file. Modification of bits should not degrade sound quality. Figure 8 shows graph of original audio which is used as host file. Figure 9 shows graph of audio after embedding and figure 10 shows graph of recovered audio after extraction. Graph of original audio, embedded audio and recovered audio is nearly same. These graphs are plotted Sample Numbers versus amplitude. The simulation was carried out in MATLAB R2010b software.

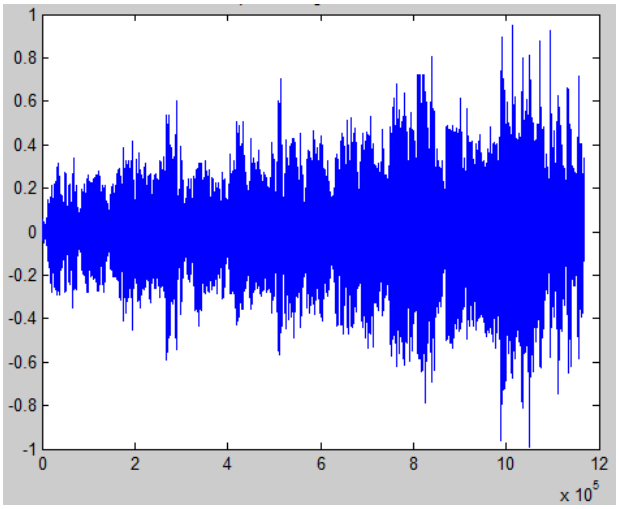


Fig. 8. Original Audio.

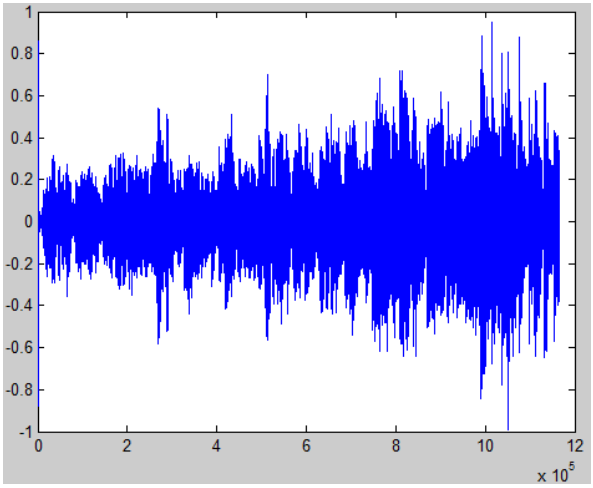


Fig. 9. Stego Audio

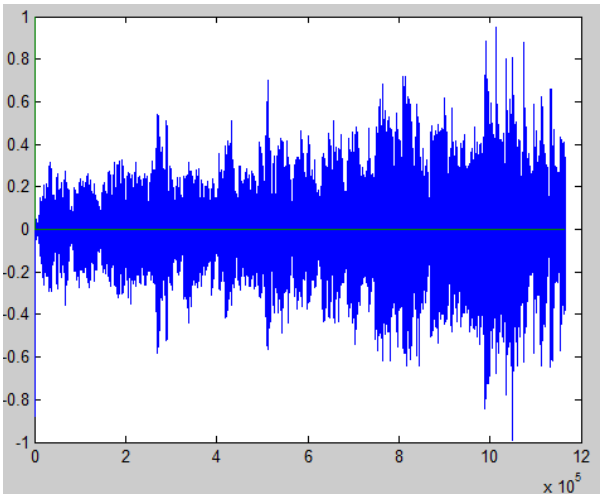


Fig. 10. Recovered Audio

Fig. 11 is Original Message and Fig. 12 is Recovered Message. These two messages are 100% similar.

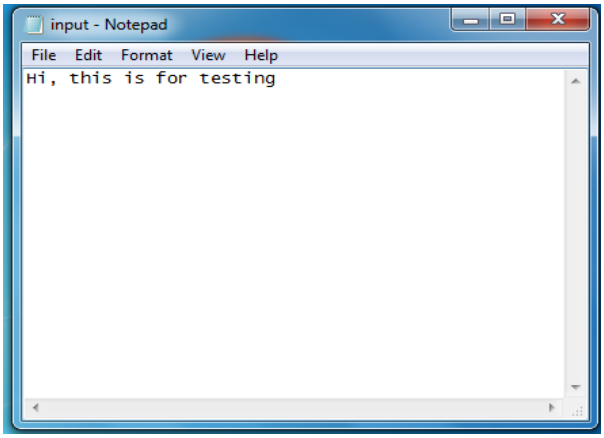


Fig. 11. Original text message

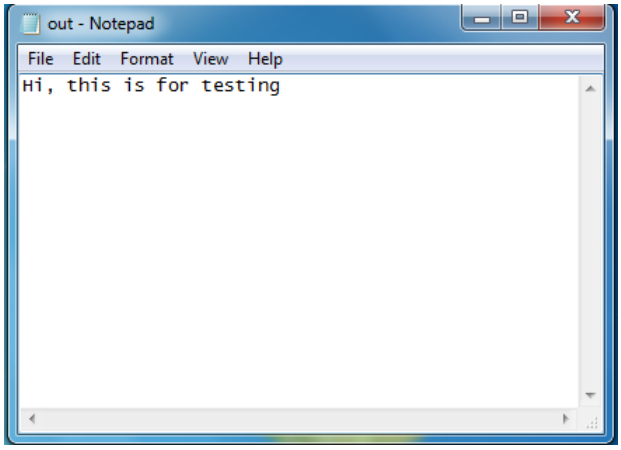


Fig. 12. Recovered message

TABLE I. MSE values of different audio files for different message sizes

Message size in bytes	Male	Female	Male song	Female song
10	$2.24*10^{-5}$	$1.75*10^{-5}$	$5.01*10^{-6}$	$6.8*10^{-6}$
20	$5.05*10^{-5}$	$3.94*10^{-5}$	$1.12*10^{-5}$	$1.54*10^{-5}$
33	$8.97*10^{-5}$	$7.00*10^{-5}$	$2.00*10^{-5}$	$2.74*10^{-5}$
57	$1.40*10^{-4}$	$1.09*10^{-4}$	$3.13*10^{-5}$	$4.28*10^{-5}$

TABLE II. PSNR values of different audio files for different message sizes

Message size in bytes	Male	Female	Male song	Female song
10	70.55	71.63	77.05	75.70
20	67.03	68.10	73.53	72.18
33	64.53	65.61	71.03	69.68
57	62.59	63.67	69.10	67.74

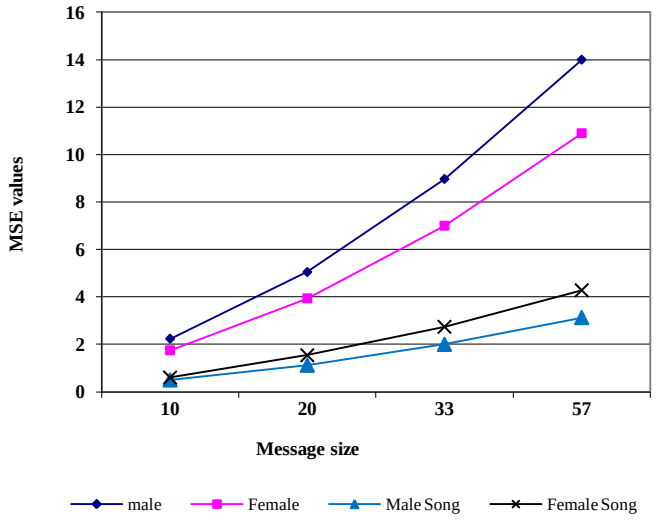


Fig. 13. Graphical representation of MSE values for Different Songs

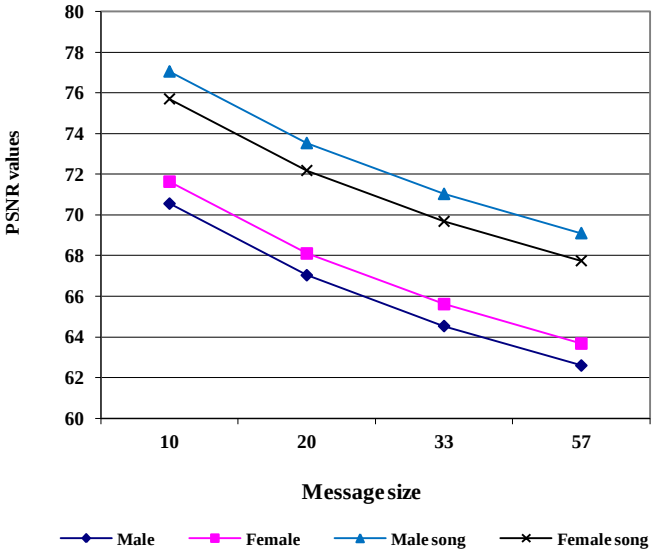


Fig. 14. Graphical representation of PSNR values for different songs

TABLE III. MSE values of different categories for Audio file with same text content

Audio file	MSE
Hip-hop	$4.7*10^{-6}$
Jazz	$4.4*10^{-6}$
Pop	$4.6*10^{-6}$
Rock	$4.5*10^{-6}$

TABLE IV. PSNR values for different categories of audio file with same text content

Audio file	PSNR
Hip-hop	77.31
Jazz	77.56
Pop	77.36
Rock	77.51

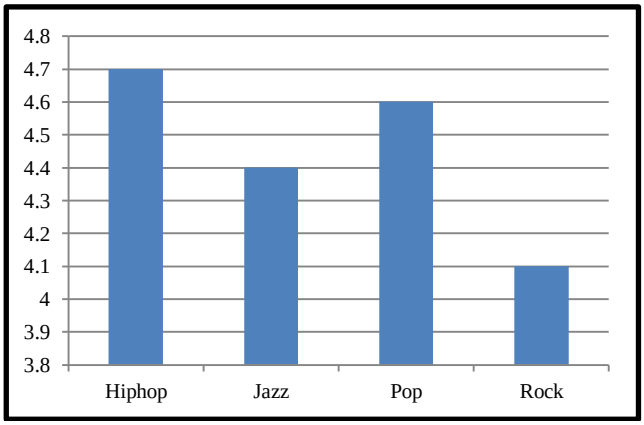


Fig. 15. Graphical representation of MSE values for Different music files

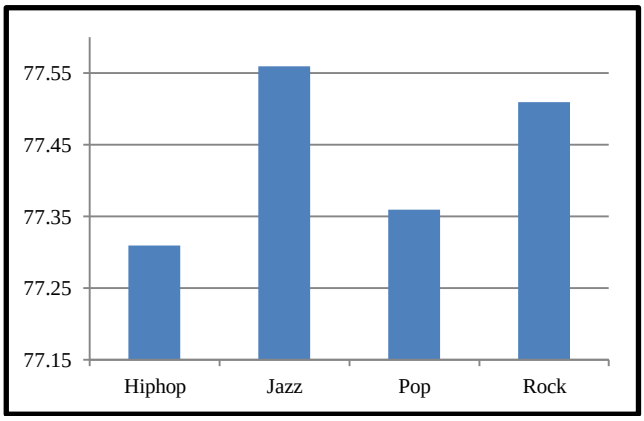


Fig. 16. Graphical representation of PSNR values for Different Music files

TABLE V. CPU time taken for encryption by CNN and with Blowfish algorithm for text files of different sizes

Text Size (in bytes)	CNN algorithm (in seconds)	Blowfish algorithm (in seconds)
50	0.0156	2.0156
100	0.0313	2.0756
150	0.0469	2.1044
200	0.0625	2.1406

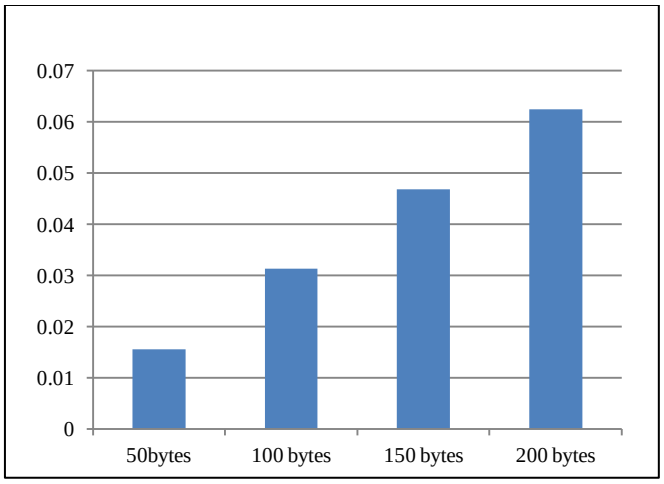


Fig. 17. Graphical representation of CPU time with CNN encryption algorithm

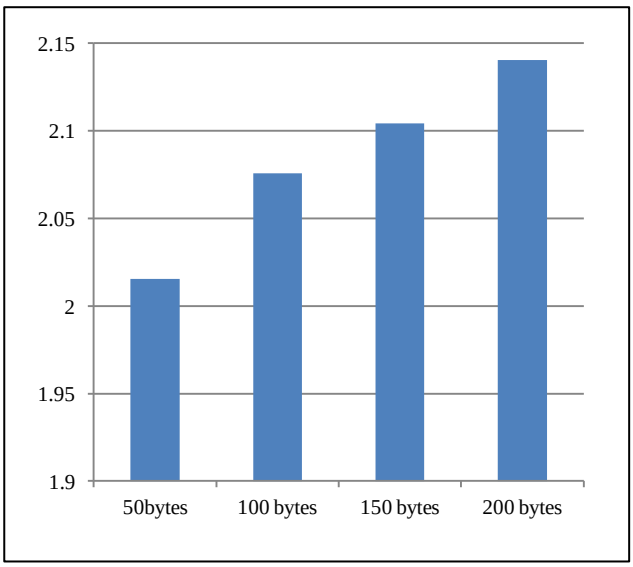


Fig. 18. Graphical representation of CPU time with Blowfish encryption algorithm

V. CONCLUSION

In this Paper, a novel method of Audio Steganography algorithm that uses Double Density Discrete Wavelet Transforms is presented. For providing better security, chaotic neural network encryption scheme is included. The qualitative performance of the proposed system is analyzed. Various sample audio files and music files are considered in .wav format and the MSE and PSNR values obtained after embedding the data have been recorded. The PSNR, MSE varies depending on the amount of data embedded in the audio file and the size of the audio file and better PSNR and low MSE values are obtained with the proposed algorithm. Also it can be observed that CNN encryption algorithm takes less time for the encryption process and is secure than the Blowfish algorithm. The Proposed system shows better performance in terms of both capacity and security. In future, this work can be extended to video data.

REFERENCES

[1] Petitcolas F.A., Anderson R.J., Kuhn M.G, "Information Hiding – A Survey" IEEE. Special Issued on Protection of Multimedia content, 1062-1078 July, 1999.

[2] Katzenbeisser.S, Peticoats.F. "Information hiding Techniques for Steganography and Digital watermarking", Artech House Inc. 2000.

[3] L.Driskell, "Wavelet-based Steganography," Crypto - logia, vol. 28, no. 2, pp.157-174, 2004.

[4] Tian.H, "A Covert Communication Model Based on Least Significant Bits Steganography in voice over IP", Proceedings of 9th International Conference for young computer scientists, IEEE computer Society, pp. 647-652, 2008.

[5] Deng.K, Tian.Y, Yu.X, Yang.Y "A unified block and stream cipher based file encryption" Journal of Global Research in Computer Science, vol.2, No.7, pp.53-57, 2011.

[6] M.Wakiyma, Y.Hidaka, K.Nozaki "A Steganography by low bit coding method with wave files", Sixth International Conference on Intelligent Infomation Hiding and multimedia signal processing, pp.530-533, Oct 2010.

[7] Yu W, Cao J. "Cryptography based on delayed neural networks". PhysLetter A ; 356:333–8. 2006

[8] B.Geethavani, E.V.Prasad. "High Secure Image Steganography Based on Chaotic Neural Network". IJCSNS, Volume13, No.3, pp1-6, March 2013

[9] A.Delforouzi , M.Pooyan, "Adaptive digital audio Steganography based on integer wavelet Transform" Circuits, Systems & Signal Processing, Vol.27, pp.247-259, Mar 2008.

[10] Atoum.M.S, Rababah.O.A , Al.Athili, New technique for hiding data in audio file". Journal of Computer Science, pp.173-177, Apr 2011.

[11] Jisna Antony, Sobin c. Sherly A. P "Audio Steganography in Wavelet Domain – A Survey" International Journal of Computer Applications, Vol.52, No.13, Aug. 2012.

[12] Andrew.D.Ker, "Steganalysis of Embedding in two least significant bits" IEEE Transactions on Information forensics and Security, Vol 2, No.1, Mar 2007.

[13] Mohammad Saleem, Mamoum Suleman, Subariah Ibrahim,"A Steganography method based on hidng secret data in MPEG Audio layer III", International Journal of Computer Science and Network Security, Vol.2, No.5, pp. 184-188, May 2011.

[14] Siwar Rekik, Driss Guerchi, Sid-Ahmed Selouani , Habib Hamam "Speech steganography using wavelet and Fourier Transforms" Rekik et al. EURASIP Journal on Audio, Speech, and Music Processing 2012.

[15] Marghny H Mohamed, Naziha M AL-Aidroos, Mohamed A Bamatraf "A combined image Steganography technique based on edge concept and dynamic LSB", International journal of Engineering Research and Technology,Vol.1 No.8,Oct 2012.

[16] Waffaa S. Ahmed , Loay E. George" Audio Hiding Using Wavelet Transform with Amplitude Modulation" Journal of Al-NahrainUniversity,Vol.16,pp.183-188 March 2013.

[17] Mansour Sheikhan, Kazem Asadollahi , Reza Shahnazi "Improvement of Embedding Capacity and Quality of DWT-Based Audio Steganography Systems" World Applied Sciences Journal, Vol.13, pp.507-516, 2011.

Physical Activity Identification using Supervised Machine Learning and based on Pulse Rate

Mobyen Uddin Ahmed
Centre for Applied Autonomous Sensor Systems
Örebro University
Örebro, Sweden

Amy Loutfi
Centre for Applied Autonomous Sensor Systems
Örebro University
Örebro, Sweden

***Abstract*— Physical activity is one of the key components for elderly in order to be actively ageing. Pulse rate is a convenient physiological parameter to identify elderly's physical activity since it increases with activity and decreases with rest. However, analysis and classification of pulse rate is often difficult due to personal variation during activity. This paper proposed a Case-Based Reasoning (CBR) approach to identify physical activity of elderly based on pulse rate. The proposed CBR approach has been compared with the two popular classification techniques, i.e. Support Vector Machine (SVM) and Neural Network (NN). The comparison has been conducted through an empirical experimental study where three experiments with 192 pulse rate measurement data are used. The experiment result shows that the proposed CBR approach outperforms the other two methods. Finally, the CBR approach identifies physical activity of elderly 84% accurately based on pulse rate.**

Keywords—Physical activity; Elderly; Pulse rate; Case-based Reasoning (CBR); Support Vector Machine (SVM) and Neural Network (NN)

I. INTRODUCTION

Physical activity or moderate exercise is one of the key components for elderly in order to stay active and maintain a longer life. Research has shown that exercise brings a greater benefit for physical capacity. That means, the elderly who are physically active considering a moderate level of exercise promote improved health than over the elderly who do not exercise but are "physically active" (in motion) throughout the day [15]. Similarly, by doing physical exercise one can control or manage the chronic diseases such as heart disease, stroke, diabetes [16], [17] and also for mental health [18]. Today, most common way of measuring the physical activity is by using accelerometer technology i.e. motion sensors are mounted on the users' wrists, waist, and ankles [19], [20], [21]. However, using accelerometers, physical activity can be measured while persons are performing regular household work but composite activities (i.e. exercise) such as running or playing tennis is still a challenge and the exact placements of sensor is also an issue while considering its potential sensitivity [19]. Similarly, concurrent and overlapping activities are also not easy to measure based on accelerometer signals, such as brushing teeth and walking [22]. In addition, the measurements using an accelerometer do not provide indications of an individual's biomedical signals such as pulse rate or heart rate. Pulse rate has an effect with physical activity i.e. pulse rate increases with activity and vice versa [23] [24] [40] [41]. However, analysis and classification of pulse rate for a specific person is often difficult due to large individual variation while doing exercise. For example, pulse rate can be fluctuated between 65 and 90 for one person and for other it can be between 90 and 110. Again, pulse rate can be increased and decreased sharply for one person and for other person it can be changed steadily. So, it is not so easy to classify and identify physical activity using only on a simple threshold or any general rules. Thus, the application domain requires a machine learning algorithm that can identify personalized elderly's physical activity based on pulse rate data.

Today, supervised machine learning algorithm is a hot topic in Artificial Intelligence (AI) and commonly applied to classify physiological sensor signals data. The main goal of the supervised machine learning algorithm is to build a model based on a set of training samples paired with the corresponding labels of those samples. This model is then used to assign class labels on a set of testing samples where the class label is unknown. [11][12]. However, to identification of appropriate algorithm for a particular classification problem is one of the main challenges. Selection of the appropriate algorithm can be done through an analysis of the application domain, but there is always an unsolved question while considering empirical experiment. Recent research also shows that many researcher have been conducted some empirical experiment on different supervised machine learning algorithms before they select and propose the best algorithm which fit well in their particular application domain [13][14][39].

This paper presents an application of a supervised machine learning algorithm in order to classify physical activity of elderly. To handle the complex data processing, first a feature extraction and selection are performed on the raw data. The raw data is labelled according to the control of the measurements. The extracted features together with their class labels are then applied in three supervised machine learning algorithm to classify physical activity. The algorithms are: 1) Case-based Reasoning (CBR), 2) Support Vector Machine (SVM) and 3) Neural Network (NN). The classification accuracy has been evaluated considering the three experimental data sets with 192 controlled pulse rate measurements data. According to the experiment result, CBR approach has been selected to identify physical activity of elderly based on pulse rate as it outperforms the popular SVM and NN techniques. Here, the obtained sensitivity, specificity and overall accuracy for the CBR approach were 87%, 85% and 86% respectively. Finally, 12 measurements with

unknown class are evaluated by applying the proposed CBR approach to identify physical activity of elderly based on pulse rate. According to the experimental work, the CBR system was able to identify the physical activity 84% accurately.

The rest of the paper is organized as follows: Section 2 presents the materials that are the data collection procedure and the feature extraction from the sensor signal. Section 3, describes the approach and methods, and namely presents three supervised machine learning algorithms. The experimental work is presented in section 4, here; the comparison between the three supervised machine learning algorithms is presented. Section 5 discusses the evaluation result of the identification of physical activity using CBR approach. Finally, Section 6 ends with the summary and concluding the remarks.

II. MATERIALS

To classify physical activity using supervised machine learning algorithm elderlies' pulse rate has been used as a physiological measurement. Several features have been extracted from each pulse rate sensor signal before the classification scheme. Each signal has been labelled according to the procedure of the data collection.

A. Data Collection

Pulse rate measurement data were collected using a wearable sensor called $Wristox_2$. $Wristox_2$ is simple and can easily be integrated on the body which offers a continuous data through Bluetooth communication. In total, 192 pulse rate measurements were collected from 24 elderlies where a 'three phases' procedure is applied in order to collect data and the procedure is adapted from previous study [25].

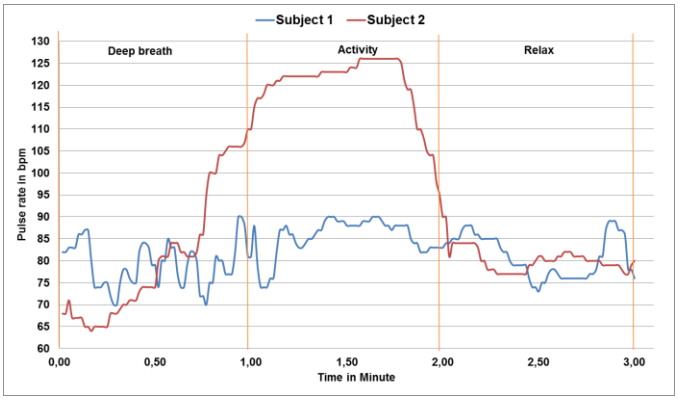


Fig. 1. Example of pulse rate changes between activity and relax state.

Each subject is asked to place the sensor (i.e. WristOx2 sensor) in their finger and follows the phases. The phases are: 1) *deep breath*: inhaling with nose and exhaling with mouth, 2) *physical activity*: walk briskly or running and 3) *relax*: sit down and try to be relaxed. Thus the 192 measurements of pulse rate are controlled and labelled with the names of phases. Finally, each class (i.e. *deep breath, physical activity* and *relax*) contains 64 pulse rate measurements. An example of pulse rate measurement collected from two subjects is presented in Fig 1. As can be seen from Fig 1, the pulse rate is changing between resting and doing exercise. It can be also observed that the changes are very personal, subject 1 (blue colour) has steady changes and subject 2 (red colour) has sharp changes while they are doing exercise. Moreover, the pulse rate of subject one lies between 70 and 90 bpm whereas, for subject 2 it lies between 65 and 126 bpm.

B. Features Extraction

The feature extraction from the sensor signal and feature selection has been conducted with close collaboration of domain expert. Here, time, frequency and time-frequency domains have considered to extracted features. In the *time domain*, statistical features namely maximum value, arithmetic mean and standard deviation of data were considered. To calculate *frequency domain* features, first the power spectral density was calculated from squared amplitude of Discrete Fourier Transform value of data using Fast Fourier Transform algorithm and scaling it to sampling frequency range which is 1 Hz in this case. Zero padding of data was done so that number of data samples became power of two for applying Fast Fourier Transform algorithm. From the power spectral density Low frequency power, High frequency power, Low frequency power to High frequency power ratio, were calculated [26]. The frequency between 0.04 Hz and 0.15 Hz was considered as Low frequency and frequency between 0.15 Hz and 0.4 Hz was considered as High frequency [26]. The power in High and Low frequency region was calculated by numerical integration of Power Spectral Density of the corresponding frequency range. The unit of power spectrum density and power for the pulse rate were BPM (beats per minute) Hz-1 and BPM2 respectively. Similarly, in *time-frequency domain* features, a discrete wavelet transform (DWT) is performed since it can keep the information of both time and frequency. Statistical features of the maximum, arithmetic mean and standard deviations were calculated from the approximation coefficient of wavelet decomposition of level 1 [27]. The function 'Daubechies 2' was used as the mother wavelet. The continuous wavelet transform linked to mother wavelet $\psi(t)$ can be defined by the equation 1.

$$W(a,b) = \int_{-\infty}^{\infty} y(t)\psi_{ab}(t)dt \qquad (1)$$

where $y(t)$ is any square integral function and a, b are scaling and translation parameters respectively. Evaluating the continuous wavelet at dyadic interval the signal can be expressed by the equation 2.

$$y(t) = \sum_{k=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} d_j(k)2^{j/2}\psi(2^j t - k) \qquad (2)$$

Where d_j is the discrete wavelet coefficient of the signal $y(t)$.

Symmetric padding was used to make the data samples power of two to implement discrete wavelet transform [27]. Thus, 9 features from the pulse rate sensor signal are calculated considering the three domains and each measurement is classified according to the condition of the data collection procedure that is the measurements are labelled as *deep breath, physical activity* and *relax.*

III. Approach and Methods

To identify physical activity of elderly, three supervised machine learning algorithms have been used to classify elderlies' pulse rate measurements. The approach of the classification scheme is presented in Fig 2. Here, pulse rate measurements come from a Wristox2 sensor and the measurements have been applied as an input of the classification scheme.

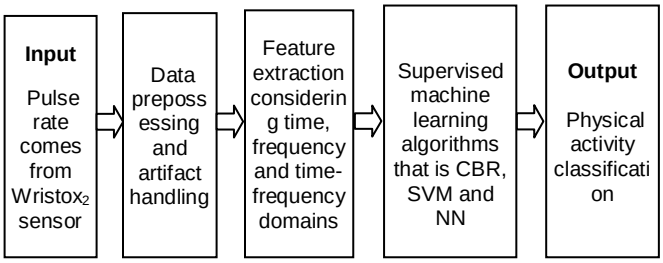


Fig. 2. General approch of classification scheme of pulse rate.

Each measurement is then prepossessed and handled artifact. Here, erroneous data values caused by a loose collection are identified and replaced by the previous samples. The clean measurements are then used to extract a set of various features considering time, frequency and time-frequency domains discussed earlier. The extracted features are sent into three machine learning algorithms where each of them classifies separately. Finally, the classification as output is used for experimental work.

A. Case-based Reasoning (CBR)

Learning from past experience and solve new problems by adapting similar previously solved cases is a cognitive model based on how humans often solve a large group of problems. A requirement is that the similarity of the case also indicates how easy the solution can be adapted to the current situation and reused.

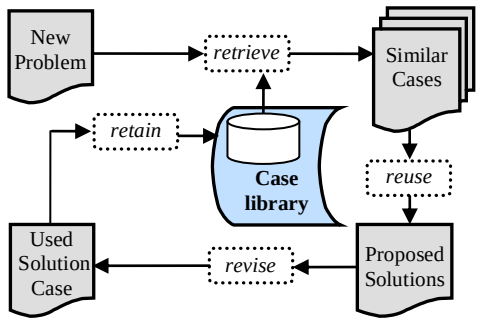


Fig. 3. CBR cycle. Adapted from [28]

A case-based reasoning (CBR) [28],[29] approach can work in a way close to human reasoning e.g. solves a new problem applying previous experiences, which is more common for doctors, clinicians or engineers. CBR has been applied successfully when the domain theory is not clear enough or even incomplete. It is getting increasing attention from the medical domain since it is a reasoning process that also is medically accepted [30] [31] [32] [33] [34] [35] [36] [37][42][43][44]. For example, a clinician/doctor may start his/her practice with some initial experience (solved cases), then try to utilize this past experience to solve a new problem and simultaneously increases his/her case base. So, this method is getting increasing attention from the medical domain since it is a reasoning process that also is medically accepted. Aamodt and Plaza has introduced a life cycle of CBR [28] with four main steps as shown in Fig. 3. Retrieve, Reuse, Revise and Retain present key tasks to implement such kind of cognitive model. In the retrieval step, for any new problem the system tries to retrieve the most similar case(s) by matching previous cases from a case base. If it finds any suitable case that is close to a current problem then the solution is reused (after some adaptation and revision if necessary).

Similarity of a feature value between two cases (i.e. a target case and one case from library) was measured using the normalized Manhattan distance between the feature values of two cases. An example of case base is show in Table I, where a new case is being matched with Cases 1 and 2 from the case base by using a function illustrated in equation 3.

$$sim(T,S) = \frac{\sum_{i}^{n} w_i \times sim(T_i, S_i)}{\sum_{i}^{n} w_i} \quad (3)$$

TABLE I. An Example of Case Representation

Cases	Problem description (input features)									Solution description (output class)
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	
Case 1	104	90.85	5.75	222.32	65.20	3.409	146.27	128.47	8.42	deep breath
Case 2	87	74.53	5.962	199.37	51.53	3.868	119.85	105.62	8.019	relax
..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..
Case 192	107	102.85	3.182	252.19	74.77	3.333	151.25	145.89	4.545	physical activity

Similarity between two cases is then measured using the weighted average of all the features that are to be considered. The function for calculating similarity between two cases T and S with n features is presented in equation 4, where w_i is the weight of the feature i defined by expert of the domain. Note that, in the weight vector w_i is also considered the weight of three domains (i.e. time, frequency and time-frequency features).

Non numeric features such as gender is converted to numeric value by substituting the contextual value with a numeric one (1 for male, 0 for female). The Manhattan distance function is used to calculate the similarity of a feature between two cases. The function is illustrated in equation 4 where T_i and S_i are the *i-th* feature values of target and source case respectively.

$$\mathrm{sim}(T_i, S_i) = 1 - \frac{|T_i - S_i|}{\max\{T_i, Max(i)\} - \min\{T_i, Min(i)\}} \quad (4)$$

Here, *Max (i)* and *Min(i)* represents the *Maximum* and *Minimum* value of the feature *i* obtained from whole case library. Then "max" and "min" functions compare the values between the new case feature T_i and *Maximum* and *Minimum*

values obtained from the case library. The function returns 1 if the values are the same and returns 0 if the values are dissimilar. This is known as a local similarity function.

B. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning that analyses data and identifies pattern and commonly used for classification and regression analysis [1][2][3][4]. Traditionally, SVM algorithm is designed for binary or two-class classification. Providing a training data set with output class where each sample belongs to one of two classes, the SVM training algorithm builds a model that classifies a new example into one class or the other.

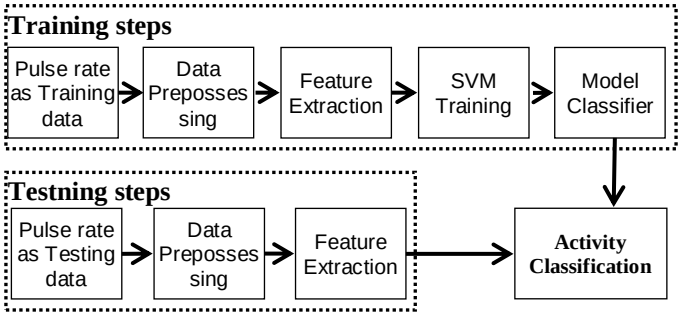


Fig. 4. Training and testing steps of activity classification using SVM.

In linearly separable data, the SVM works based on a distance value between the hyperplane and the two data classes. However, to handle non-linearly separable data, a kernel function [5] could apply where it is used to map non-linearly separable data onto a feature space for classification. Some popular kernel functions that commonly used are linear, polynomial, RBF and sigmoid [6]. However, when there are more than 2 classes (i.e. in our study it is 3 classes), one of the common approach is to used one-versus-all classifiers (also known as "one-versus-rest"), where the target class is determined by choosing the class that is selected by the largest classifiers, that is, let $y \in 1,2,3,\ldots,k$ with $\Theta^{(1)}, \Theta^{(2)}, \Theta^{(3)}, \ldots\ldots, \Theta^{(k)}$ and class *i* will be selected with the largest $\left(\Theta^{(1)}\right)^T x$. Fig 4 presented the training and testing steps. The pulse rate data was pre-processed and features are extracted both for training and testing data. But training data has been used to build a classifier model based on SVM training. Finally, the testing data is used to measure the performance of the model.

C. Neural Network (NN)

Neural Network (NN), is another supervised machine learning inspired by biological neural network, is widely used to model complex relationships between inputs and outputs [38] NN is a network system with many simple processors where the processing elements are referred to as units, nodes, or neurons. These neurons are interconnected and it receives process and transmits numeric data via the connections [7][8]. A NN works based on creating a model by training its network where it is used a set of examples samples that contains the input and its corresponding known target output. Same time, it learns by comparing the network output and target output and makes adjustments on the weights (connections between neurons) in order to move the network outputs closer to the targets. Here, to identify physical activity, this paper applied a feed-forward neural network where the NN is trained by the back propagation algorithm. The NN model is illustrated in Fig 5.

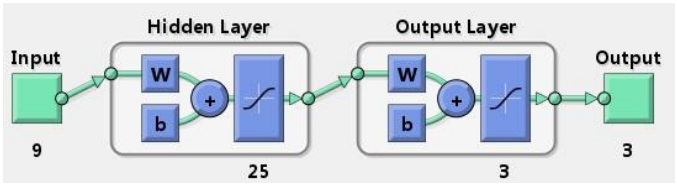


Fig. 5. NN model to identify physical activity based on elderie's pulse rate.

As can be seen in Fig 5, there are 9 features considered as input, 25 hidden neurons, and 3 output neuron with 3 final outputs (i.e. *deep breath*, *physical activity* and *relax* classes).

IV. Experimental Work

The main objective of this experimental work is to find best classification algorithm among the three supervised machine learning methods i.e. CBR, SVM and NN. This experimental work also presented a comparison on this three supervised machine learning methods and finally selects the best method to identify physical activity. For this experimental work, there are 192 cases with three classes (i.e. *Deep breath*, *Activity*, *Relax*) have been used where each case is consist of nine features. These 192 cases are then divided into two sets randomly; the training data set contain 162 cases and the test set contains 30 cases. Thus the experimental work has been conducted with three experiment data sets and they are named as LibraryA and TestA, LibraryB and TestB, LibraryC and TestC. The cases are selected as randomly, however, test sets are containing different cases that mean if a case belongs to one test set then the case is not considered again for another test set.

A. Using Case-based Reasoning (CBR)

The classification accuracy of the CBR retrieval classification scheme has been evaluated by developing a prototypical system where the main goal of the experiment is to see how accurate the CBR approach can classify with the extracted features from the signals. The experimental work has been conducted in two phases. In *first phase*, the training data sets i.e. LibraryA, LibraryB and LibraryC are used to train the CBR classification scheme where weight of the features are adjusted manually to achieve maximum accuracy. Here, for the retrieval, a "leave-one-out" retrieval technique is used i.e. one case is taken from the case library (i.e. 162 cases) as a query case and then the system retrieves the most similar cases. Among the retrieved cases, top similar case is considered, if the top case's class is matched with the query case class then we count the correctly classification as 1. In *second phase,* same procedure is used but the testing data sets are used as in query. This means TestA is evaluated while considering LibraryA; TestB is evaluated while considering LibraryB and TestC is evaluated while considering LibraryC.

The percentage of the correctly classification are presented in Table II.

TABLE II. PERCENTAGE OF CORRECTLY CLASSIFICATION USING CBR

Experiments	Training		Testing	
	Datasets	*Accuracy*	*Datasets*	*Accuracy*
Expriment 1	LibraryA	≈ 85%	TestA	≈ 80%
Expriment 2	LibraryB	≈ 90%	TestB	≈ 87%
Expriment 3	LibraryC	≈ 86%	TestC	≈ 90%
Avarage	In 3 Libraries	≈ 87%	In 3 Test sets	≈ 86%

It can be seen from table II, the accuracy of the CBR approach has been achieved for LibraryA, LibraryB and LibraryC are 85%, 90% and 86% respectively. Similarly, on testing data sets, the accuracy of the CBR approach has been achieved for TestA, TestB and TestC are 80%, 87% and 90% respectively. Moreover, the average accuracy on training data sets is 87% and for testing data sets is 86%. The confusion matrix of the testing data sets i.e. TestA, TestB and TestC are presented in Table III, Table IV and Table V.

TABLE III. CONFUSION MATRIX BASED ON TESTA USING CBR CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	8 (80%)	0	2 (20%)
Activity	2 (20%)	8 (80%)	0
Relax	1 (10%)	1 (10%)	8 (80%)

TABLE IV. CONFUSION MATRIX BASED ON TESTB USING CBR CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	9 (90%)	1 (10%)	0
Activity	1 (10%)	9 (90%)	0
Relax	2 (20%)	1 (10%)	8 (80%)

TABLE V. CONFUSION MATRIX BASED ON TESTC USING CBR CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	9 (90%)	1 (10%)	0
Activity	0	9 (90%)	1 (10%)
Relax	1 (10%)	0	9 (90%)

Fig 6 illustrated the performance of the CBR approach while classifying the elderlies' pulse rate data into three classes i.e. Deep breath, Activity and Relax. Here, the accuracy value shows that the CBR approach can achieve the classification accuracy between 80 and 90 percentages for all the test data sets.

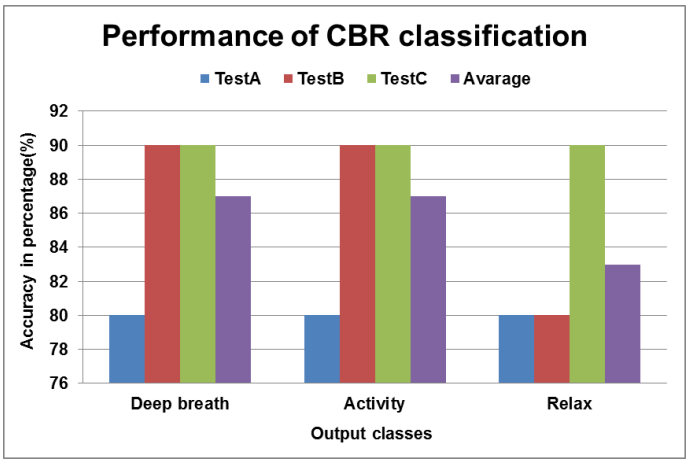


Fig. 6. Performance of CBR classification based on three test data sets.

B. *Using Support Vector Machine (SVM)*

Support Vector Machine (SVM) is applied on the same training (LibraryA, LibraryB and LibraryC) and testing (TestA, TestB and TestC) data sets. The classification accuracy of the SVM classification scheme has been evaluated where the main goal of the experiment is to see how accurate the SVM approach can classify with the extracted features from the signals. Here, the training session has been performed base on the training data set using the LibSVM tool developed by [9] on MATLAB. For the kernel, the RBF kernel function is applied as it gives better accuracy [10]. The SVM model best parameters were chosen after applying 7-fold cross validation.

TABLE VI. PERCENTAGE OF CORRECTLY CLASSIFICATION USING SVM

Experiments	Training		Testing	
	Datasets	*Accuracy*	*Datasets*	*Accuracy*
Expriment 1	LibraryA	≈ 89%	TestA	≈ 63%
Expriment 2	LibraryB	≈ 73%	TestB	≈ 53%
Expriment 3	LibraryC	≈ 74%	TestC	≈ 70%
Avarage	In 3 Libraries	≈ 79%	In 3 Test sets	≈ 62%

As soon as the model is ready, the test data sets i.e. TestA, TestB and TestC are then evaluated and the accuracy values both for training and testing data sets are presented in Table VI. Here, the percentage of the correctly classification are presented as accuracy value. It can be seen from table VI, the accuracy of the SVM approach has been achieved for LibraryA, LibraryB and LibraryC are 89%, 73% and 74% respectively. Similarly, on testing data sets, the accuracy of the SVM approach has been achieved for TestA, TestB and TestC are 63%, 53% and 70% respectively.

Moreover, the average accuracy on training data sets is 79% and for testing data sets is 62%. The confusion matrix of the testing data sets i.e. TestA, TestB and TestC are presented in Table VII, Table VIII and Table IX.

TABLE VII. CONFUSION MATRIX BASED ON TESTA USING SVM CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	8 (80%)	0	2 (20%)
Activity	2 (20%)	5 (50%)	3 (30%)
Relax	4 (40%)	0	6 (60%)

TABLE VIII. CONFUSION MATRIX BASED ON TESTB USING SVM CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	5 (50%)	2 (20%)	3 (30%)
Activity	4 (40%)	5 (50%)	1 (10%)
Relax	3 (30%)	1 (10%)	6 (60%)

TABLE IX. CONFUSION MATRIX BASED ON TESTC USING SVM CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	9 (90%)	0	1 (10%)
Activity	2 (20%)	6 (60%)	2 (20%)
Relax	3 (30%)	1 (10%)	6 (60%)

Fig 7 presents the performance of the SVM approach to classify elderlies pulse rate sensor signals. The accuracy values show that most of the time (i.e. in *Activity* and *Relax* classes) SVM is achieved its classification accuracy between 50% and 60%; however, it also achieved 90% accuracy while considering Deep breath class.

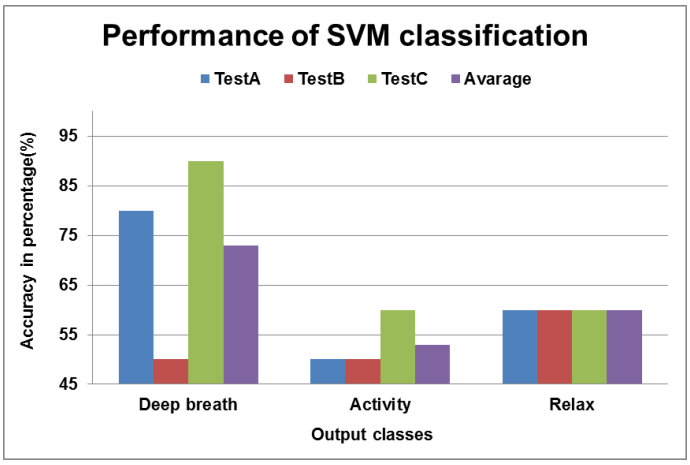


Fig. 7. Performance of SVM classification based on three test data sets.

C. Using Neural Network (NN)

Similarly, Neural Network (NN) is applied using MATLAB on the training sets LibraryA, LibraryB and LibraryC. Here, a feed-forward network with the default tan-sigmoid transfer function is used in the hidden layer and linear transfer function is used in the output layer. The number of hidden layer was fixed as 25; however, several thousand iterations have been performed to achieve a better accuracy. The minimum accuracy for the training data sets have been considered as 80% and to achieve this value, LibraryA and LibraryB are used 12800 times iteration and LibraryC is used 51200 times iteration. Thus three NN models are created based on training data sets and the test data sets are classified and evaluated. The percentage of the correctly classification in terms of accuracy value are presented in Table X.

TABLE X. PERCENTAGE OF CORRECTLY CLASSIFICATION USING NN

Experiments	**Training**		**Testing**	
	Datasets	***Accuracy***	***Datasets***	***Accuracy***
Expriment 1	LibraryA	≈ 80%	TestA	≈ 57%
Expriment 2	LibraryB	≈ 83%	TestB	≈ 60%
Expriment 3	LibraryC	≈ 84%	TestC	≈ 60%
Avarage	In 3 Libraries	≈ 82%	In 3 Test sets	≈ 59%

It can be seen from table X, the accuracy of the NN approach has been achieved for LibraryA, LibraryB and LibraryC are 80%, 83% and 84% respectively. Similarly, on testing data sets, the accuracy of the NN approach has been achieved for TestA, TestB and TestC are 57%, 60% and 60% respectively. Moreover, the average accuracy on training data sets is 82% and for testing data sets is 59%. Considering NN the confusion matrix of the testing data sets i.e. TestA, TestB and TestC are presented in Table XI, Table XII and Table XIII.

TABLE XI. CONFUSION MATRIX BASED ON TESTA USING NN CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	6 (60%)	2 (20%)	2 (20%)
Activity	1 (10%)	7 (70%)	2 (20%)
Relax	5 (50%)	1 (10%)	4 (40%)

TABLE XII. CONFUSION MATRIX BASED ON TESTB USING NN CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	7 (70%)	2 (20%)	1 (10%)
Activity	1 (10%)	8 (80%)	1 (10%)
Relax	4 (40%)	3 (30%)	3 (30%)

TABLE XIII. CONFUSION MATRIX BASED ON TESTC USING NN CLASSIFICATION

	Deep breath	*Activity*	*Relax*
Deep breath	8 (80%)	1 (10%)	1 (10%)
Activity	2 (20%)	7 (70%)	1 (10%)
Relax	4 (40%)	3 (30%)	3 (30%)

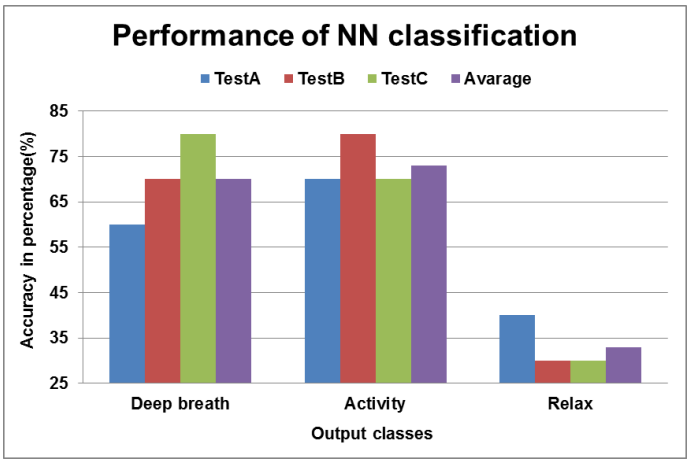


Fig. 8. Performance of NN classification based on three test data sets.

Same as CBR and SVM, Fig 8 presents the performance of the NN approch to classify the eldeires pulse rate signls. Here, the accurecy vlues for Deep breath and Activity lies between 60% and 80%, however, the cases in Relax class has been classified very poorly. According to the Fig 8, the accuracy of the Relax class was between 30% and 40%.

D. Comparision on CBR vs SVM vs NN

As one of the contribution of this paper is to select proper and best supervised machine learning algorithm which can be used to identify physical activity of elderly based on their pulse rate sensor signal. So, in this section, a comparison on CBR, SVM and NN is presented, here only test data sets are considered.

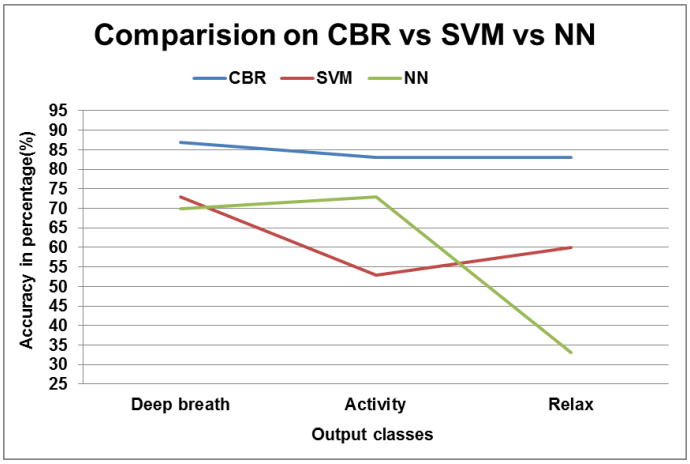


Fig. 9. Classification performance based on CBR, SVM and NN.

The three test data sets (i.e. TestA, TestB and TestC) contains 90 cases totally, among of them each class (i.e. *deep breath*, *Activity*, and *Relax*) consists of 30 cases. Percentage of correctly classification has been calculated for all three classes where CBR, SVM and NN algorithms are applied. The results of percentage of correctly classification in terms of accuracy are presented in Fig 9. As can be seen from Fig 9, CBR can perform well (i.e. > 83%) in correctly classification compare to other two machine learning algorithms (i.e. SVM and NN).

Similar comparison has been conducted considering the sensitivity and specificity analysis. Here, the class *Deep breath* and *Relax* are grouped together and identified as *Normal* group and rest of the class is identified as *Activity* group. Thus, *Normal* group contains 60 samples and *Activity* group contains 30 samples of pulse rate. The sensitivity and specificity analysis is presented in Table XIV.

TABLE XIV. STATISTICAL ANALYSIS OF THE CLASSIFICATIONS

Criteria/Indices	***Values using CBR***	***Values using SVM***	***Values using NN***
Total cases	90	90	90
Cases belong to *Activity* group (P)	30	30	30
Cases belong to *Normal* group (N)	60	60	60
True positive (TP):	26	16	22
False positive (FP):	9	20	29
True negative (TN):	51	40	31
False negative (FN):	4	14	8
Sensitivity = TP / (TP + FN)	≈ 0.87	≈ 0.53	≈ 0.73
Specificity = TN / (FP + TN)	≈ 0.85	≈ 0.67	≈ 0.52
Accuracy = (TP+TN)/(P+N)	≈ 0.86	≈ 0.62	≈ 0.59

It can be seen from Table XIV, using CBR the sensitivity, specificity and overall accuracy are 87%, 85% and 86% respectively; using SVM the sensitivity, specificity and overall accuracy are 53%, 67% and 62% respectively; using NN the sensitivity, specificity and overall accuracy are 73%, 52% and 59% respectively. The comparison on CBR, SVM and NN considering sensitivity, specificity and overall accuracy are presented in Fig 10.

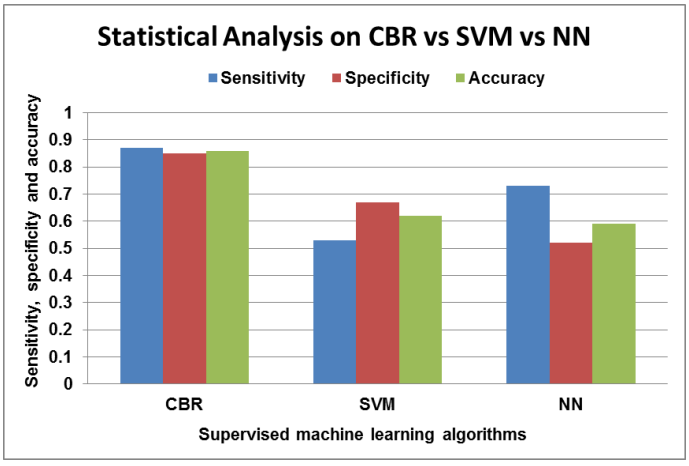


Fig. 10. Classification performance based on CBR, SVM and NN.

According to Fig 10, using CBR the activity classification shows to be superior to SVM and NN considering sensitivity, specificity and overall accuracy which is above 85%. It can be observed that the SVM and NN have achieved their sensitivity, specificity and overall accuracy between 50% and 70%.

V. IDENTIFICATION OF PHYSICAL ACTIVITY USING CBR

Based on the experimental work, the CBR approach is selected to identify physical activity of elderly based on pulse rate. In order to evaluate the physical activity classification by the CBR system, 12 measurements have been collected from 12 subjects. Each measurement is ten minutes long in length and the subjects are asked for walk at least once i.e. one out of ten minutes. Thus, each of the 12 measurements contains physical activity between 0 to 6 times in ten minutes length data. Each case is divided in 10 windows and each of them is 1 minute long (since each case is 10 minutes long). The main objective of this work is to see whether the CBR approach can identify activity. Here, as a case library, the CBR used 192 measurement cases which is categories in 3 classes (i.e. *Deep breath*, *Activity* and *Relax*) discussed earlier. For the retrieval in CBR approach, top similar case is considered to calculate the classification accuracy and the results are presented in Table XV.

TABLE XV. IDENTIFICATION OF PHYSICAL ACTIVITY ON 12 PULSE RATE MEASUREMENTS

Test_Case_id	Number of Activity	Correctly Classification	Missed Classification
Test_Case_1	4	2 (≈ 50.0%)	2 (≈ 50.0%)
Test_Case_2	6	5 (≈ 83.3%)	1 (≈ 16.6%)
Test_Case_3	0	0	0
Test_Case_4	0	0	1
Test_Case_5	1	1 (100%)	0
Test_Case_6	3	3 (100%)	0
Test_Case_7	2	2 (100%)	0
Test_Case_8	3	2 (≈ 67.0%)	1 (≈ 33.0%)
Test_Case_9	4	4 (100%)	0
Test_Case_10	2	2 (100%)	0
Test_Case_11	5	5 (100%)	0
Test_Case_12	1	0	1
Total activity in 12 test cases	31	26 (≈ 84.0%)	5 (≈ 16.0%)

As can be seen from Table XV, around 84% (i.e. 26 out of 31 physical activities) is correctly classified and around 16% (i.e. 5 out of 31 physical activities) is misclassified. However, considering test_case_4, the CBR approach classified one window as activity whereas there was no activity contains in the case.

VI. SUMMARY AND CONCLUSION

This paper presents an application of supervised machine learning algorithm to identify physical activity of elderly based on pulse rate. The pulse rate is used as a physiological parameter since it has an effect with activity that is pulse rate can be increased while performing exercise and decreased while resting. Moreover, the pulse rate sensor is very simple and can easily be integrated on the body than other physiological sensors for example ECG. The contribution of the paper is in two folds: 1) selection of a supervised machine learning algorithm which fits well in this domain 2) identification physical activity of elderly using selected machine learning algorithm and based on pulse rate. To select a supervised machine learning algorithm, this work studied the implementation of three popular classification techniques, i.e. case-based reasoning, support vector machine and neural network. The study was conducted through an empirical evaluation where three experimental libraries of data sets have been used. Each library is containing 192 pulse rate signals and there are 9 features are extracted from each signal. The feature extraction has been done by considering time, frequency and time-frequency domains. The signals are labelled in 3 classes (i.e. *Deep breath*, *activity* and *relax*) according to the control of data collection procedure. Considering the experimental work, a comparison has been done and presented among the three implemented machine learning algorithms, i.e. CBR, SVM and NN. The comparison between these techniques shows that the CBR model yields better results, i.e. the sensitivity, specificity and overall accuracy was above 85%. After selecting the machine learning algorithm, the CBR approach is applied in 12 unknown pulse rate measurements. According to the evaluation the CBR approach was able to identify physical activity 84% accurately, that is 26 out of 31 activities are correctly classified. Thus, the case-based retrieval classification scheme shows the possibility of the identification of physical activities of elderly. However, a comparison considering accelerometer signal with the pulse rate is needed which is now under study. In future, we would like to evaluate the proposed approach considering larger samples and want to calculate energy consumption based on pulse rate. Nevertheless, considering the evaluation result it might be worth and more reliable to use pulse rate measurements besides the accelerometer signal in order to classify physical activity of elderly.

ACKNOWLEDGMENT

The authors would like to acknowledge the projects named “Remote” and “SAAPHO (project reference aal-2010-3-035)” for their support of this research. Also, like to thank the participating subjects for their contributions to this work.

REFERENCES

[1] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995.

[2] C. Burges, A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery 2 (2) (1998) 1–47.

[3] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, Cambridge, 2000.

[4] C. C. Chuang, Z.-J. Lee, Hybrid robust support vector machines for regression withoutliers, Applied Soft Computing 11 (January (1)) (2011) 64–72.

[5] B. Schölkopf, A.J. Smola, Learning with Kernels, MIT Press, 2002.

[6] M. Genton, Classes of kernels for machine learning: a statistics perspective, Journal of Machine Learning Research 2 (2001) 299–312.

[7] S. Simon, Haykin, Neural Networks: A Comprehensive Foundation, Macmillan, New York, 1994.

[8] G. Zhang, Neural networks for classification: a survey, IEEE Transactions on Systems, Man, and Cybernetics Part C 30 (4) (2000) 451–462.

[9] C. Chang, C. Lin, LIBSVM: a library for support vector machine, ACM Transactions on Intelligent Systems and Technology 2 (2011) 27:1–27:27.

[10] C.W. Hsu, C.C. Chang, C.J. Lin, A Practical Guide to Support Vector Classification, Technical Report, Department of Computer Science and Information Engineering, University of National Taiwan, Taipei, 2003, pp. 1–12.

[11] S. Kotsiantis, Supervised Machine Learning: A Review of Classification Techniques. *Informatica* 2007, **31:**249-268.

[12] I. Witten, R. Frank, Data mining: Practical Machine Learning Tools and Techniques. Second edition. San Francisco: Elsevier, Inc; 2005.

[13] Y. J. Kumar, N. Salimb, B. Razac, Cross-document structural relationship identification using supervised machine learning, Applied Soft Computing, Volume 12, Issue 10, October 2012, Pages 3124–3131

[14] S. Yeran , D. Yunyan, Comparison of CBR and SVM method used in the prediction of land use change in Pearl River Delta, China, International Conference on Artificial Intelligence and Computational Intelligence, 2010, PP- 8 – 12

[15] J. S. Brach, E. M. Simonsick, S. Kritchevsky, K. Yaffe, A. B. Newman, and for the Health, Aging and Body Composition Study Research Group, The Association Between Physical Function and Lifestyle Activity and Exercise in the Health, Aging and Body Composition Study. *Journal of the American Geriatrics Society*, **52** (2004), 502–509

[16] H.C. Barry, S.W. Eathorne, Exercise and aging: issues for the practitioner. *Medical Clinics of North Americais*, **78** (1994), 357-376.

[17] F.B. Hu, M.J. Stampfer, G.A. Colditz, A. Ascherio, K.M. Rexrode, W.C. Willett, J.E. Manson, Physical activity and risk of stroke in women. *The journal of the American Medical Association*, **283** (2000), 2961-2967.

[18] A. Salguero, R. Martinez-Garcia, O. Molinero, S. Marquez, Physical activity, quality of life and symptoms of depression in community-dwelling and institutionalized older adults. *Arch Gerontol Geriatr*, **53**, (2011), 152–7.

[19] C. Lin, Y. C. Yang, J. Wang, and Y. Yang, A Wearable Sensor Module With a Neural-Network-Based Activity Classification Algorithm for Daily Energy Expenditure Estimation, IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, **16**, (2012)

[20] K. Taraldsena, S. F.M. Chastinb, I. I. Riphagenc, B. Vereijkend, J. L. Helbostada, Physical activity monitoring by use of accelerometer-based body-worn sensors in older adults: A systematic literature review of current knowledge and applications *Maturitas*, **71** (2012), 13-19.

[21] A. Mannini and A. M. Sabatini, Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers, *Sensors*, **10** (2010) 1154-1175.

[22] O. D. Lara and M. A. Labrador, A Survey on Human Activity Recognition using Wearable Sensors,

[23] I. Åstrand, P. Åstrand, and K. Rodahl, Maximal heart rate during work in older men, *Journal of Applied Physiology*, **14**(1959), 562-566

[24] Z. Jin, J. Oresko, S. Huang, and A. C. Cheng, HeartToGo: A personalized medicien technology for cardiovascular disease prevention and detection, in *Proc. IEEE/NIH LiSSA*, (2009), 80–83.

[25] M.U. Ahmed, A.M. Islam, A. Loutfi, A case-based patient identification system using pulse oximeter and a personalized health profile, *workshop on CBR in the Health Sciences at 20th International Conference on Case-Based Reasoning*, ICCBR12, Springer, Lyon, France , Editor(s):Isabelle B, Stefania M and Cindy M, September, 2012

[26] K. Dingli, T. Assimakopoulos, P.K. Wraith, I. Fietze, C. Witt, N.J. Douglas, Spectral oscillations of RR intervals in sleep apnoea/hypopnoea, *European Respiratory*, (2003), 4

[27] Cvetkovic D., Übeyli E. D., Cosic I., Wavelet transform feature extraction from human PPG, ECG, and EEG signal responses to ELF PEMF exposures: A pilot study, Digital Signal Processing, 2008.

[28] A. Aamodt, E. Plaza, Case-based reasoning: Foundational issues, methodological variations, and system approaches, *AI Communications*, (1994) 39-59.

[29] I. Watson, Applying Case-Based Reasoning: Techniques for Enterprise systems, (1997).

[30] M.U. Ahmed, S. Begum, P. Funk, N. Xiong, B von. Schéele, A Multi-Module Case Based Biofeedback System for Stress Treatment, *international journal of Artificial Intelligence in Medicine*, 2010.

[31] M.U. Ahmed, S. Begum, P. Funk, N. Xiong, B von. Schéele, Case-based Reasoning for Diagnosis of Stress using Enhanced Cosine and Fuzzy Similarity, *Transactions on Case-Based Reasoning on Multimedia Data*, **1(** 2008), 3-19.

[32] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, M. Folke, Case-Based Reasoning Systems in the Health Sciences: A Survey on Recent Trends and Developments, *IEEE Transactions on Systems, Man, and Cybernetics--Part C: Applications and Reviews*, 2010.

[33] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, B Von Schéele, A Case-Based Decision Support System for Individual Stress Diagnosis Using Fuzzy Similarity Matching. *Computational Intelligence (CI)*, **25**, (2009) 180-195.

[34] S. Begum, M.U. Ahmed, P Funk, Case-Based Systems in the Health Sciences-A Case Study in the Field of Stress Management, *WSEAS Transactions on Systems*, **8** (1109-2777), pp 344-354

[35] M.U. Ahmed, S. Begum, P. Funk, A hybrid case-based system in clinical diagnosis and treatment, *International Conference on Biomedical and Health Informatics (BHI)*, 2012 IEEE-EMBS, pp 699-704.

[36] M.U. Ahmed, P. Funk, A Computer Aided System for Post-operative Pain Treatment Combining Knowledge Discovery and Case-Based Reasoning, *20th International Conference on Case-Based Reasoning*, ICCBR'12.

[37] M.U. Ahmed, P. Funk, A Case-Based Retrieval System for Post-operative Pain Treatment, *In the proceeding of International Workshop on Case-Based Reasoning CBR 2011*, IBaI, Germany , New York/ USA, Ed(s):Petra Perner and Georg Rub, September, 2011, pp 30-41.

[38] Z. Jin, J. Oresko, S. Huang, and A. C. Cheng, HeartToGo: A personalized medicien technology for cardiovascular disease prevention and detection, in *Proc. IEEE/NIH LiSSA*, (2009), 80–83.

[39] K.J Kim, Financial time series forecasting using support vector machines,Neurocomputing, Volume 55, Issues 1–2, September 2003, Pages 307–319

[40] M.U. Ahmed, A. Loutfi, Physical Activity Classification for Elderly based on Pulse Rate, 10th International Conference on Wearable Micro and Nano Technologies for Personalized Health Tallinn, June 26 - 28, 2013

[41] M. U. Ahmed, H. Banaee, and A. Loutfi, Health Monitoring for Elderly: An Application Using Case-Based Reasoning and Cluster Analysis, ISRN Artificial Intelligence, vol. 2013, Article ID 380239, 11 pages, 2013

[42] M.U. Ahmed, S. Begum, P. Funk, N. Xiong, B. Schéele von. A Three Phase Computer Assisted Biofeedback Training System Using Case-Based Reasoning, In 9th European Conference on Case-based Reasoning workshop proceedings. Trier, Germany. 2008

[43] M.U. Ahmed, S. Begum, P. Funk, A hybrid case-based system in clinical diagnosis and treatment, *Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on* , vol., no., pp.699,704, 5-7 Jan. 2012

[44] M.U. Ahmed, P. Funk, A Computer Aided System for Post-operative Pain Treatment Combining Knowledge Discovery and Case-Based Reasoning, 20t International Conference on Case-Based Reasoning, ICCBR'12.

Exact Output Rate of Generalized Peres Algorithm for Generating Random Bits from Loaded Dice

Sung-il Pae
Department of Computer Engineering
Hongik University
Seoul, Korea
Email: pae@hongik.ac.kr

***Abstract*—We report a computation of the exact output rate of recently-discovered generalization of Peres algorithm for generating random bits from loaded dice. Instead of resorting to brute-force computation for all possible inputs, which becomes quickly impractical as the input size increases, we compute the total output length on equiprobable sets of inputs by dynamic programming using a recursive formula.**

***Keywords*—Random number generation, Peres algorithm, exact output rate, random bits, loaded dice.**

I. Introduction

Peres algorithm recursively produces unbiased coin flips from biased coin flips, with von Neumann's method as its base [1]. Because it is defined by a simple recursion, Peres algorithm is easy to implement and yet runs fast.

The output *rate* of a procedure that converts a biased Bernoulli source with n-valued distribution $\mathbf{p} = (p_1, \ldots, p_n)$ to unbiased random bits is the average number of output bits per input, and it is known to bounded by Shannon entropy $H(\mathbf{p}) = -(\log_2 p_1 + \cdots + \log_2 p_n)$ [2], [3], [4]. Since Peres algorithm is such a procedure, consequently, its rate is bounded by the entropy bound, $h(p) = -(p\log_2 p + (1-p)\log_2(1-p))$. Interestingly, the rates of Peres algorithm approaches to the entropy bound as the input length tends to infinity [1], and we call such algorithms *asymptotically optimal*.

The exact output rate of Peres algorithm was reported [5] and compared with another asymptotically optimal method by Elias [2]. Recently, a generalization of Peres algorithm was found for generating unbiased random bits from loaded dice, that is, many-valued Bernoulli source [6]. We report, here, a computation of the exact output rate of the three-face case, thus the simplest, among the generalizations of Peres algorithm given in [6].

A. 3-Face Peres Function

Assume our die has three faces with values 0, 1, and 2 with probabilities p, q, and r, respectively, so that $p + q + r = 1$. A sequence in $\{0,1,2\}^N$ is considered to be taken from a source of Bernoulli(p, q, r). Denote by $S_{(n_0,n_1,n_2)}$ the subset of $\{0,1,2\}^N$ that consists of strings with n_0 0's, n_1 1's, and n_2 2's. Then

$$\{0,1,2\}^N = \bigcup_{n_0+n_1+n_2=N} S_{(n_0,n_1,n_2)},$$

and each $S_{(n_0,n_1,n_2)}$ is an equiprobable subset of elements whose probability of occurrence is $p^{n_0}q^{n_1}r^{n_2}$.

Consider the functions on $\{0,1,2\}^2$ defined as follows:

x	$\Pr(x)$	$\Psi_1(x)$	$u(x)$	$v(x)$	$w(x)$
00	p^2	λ	0	0	λ
01	pq	0	1	λ	1
02	pr	0	1	λ	2
10	pq	1	1	λ	1
11	q^2	λ	0	1	λ
12	qr	0	1	λ	0
20	pr	1	1	λ	2
21	qr	1	1	λ	0
22	r^2	λ	0	2	λ

TABLE I. Functions for three-face Peres method

The second column of the table shows the probabilities $\Pr(x)$ for $x \in \{0,1,2\}^2$. Note that

$$\Pr(\Psi_1(x) = 0) = pq + qr + rp = \Pr(\Psi_1(x) = 1).$$

Therefore, the output of Ψ_1 can be regarded as a fair coin flip. Extend the three functions Ψ_1, u, and v to $\{0,1,2\}^*$: for an empty string,

$$\Psi_1(\lambda) = u(\lambda) = v(\lambda) = \lambda,$$

for a nonempty even-length input, define (and the same for u and v)

$$\Psi_1(x_1x_2\ldots x_{2n}) = \Psi_1(x_1x_2) * \cdots * \Psi_1(x_{2n-1}x_{2n}),$$

where $*$ is concatenation, and for an odd-length input, drop the last bit and take the remaining even-length bits.

Define

$$\Psi(x) = \Psi_1(x) * \Psi(u(x)) * \Psi(v(x)) * \Psi(w(x)).$$

This function Ψ, recursively defined, with Ψ_1 as its base, is shown to produce unbiased coin flips and is also asymptotically optimal [6].

In order to compute the exact output rate, consider a decomposition of equiprobable set of inputs $S_{(n_0,n_1,n_2)}$: Fix (n_0, n_1, n_2) such that $N = 2n = n_0 + n_1 + n_2$. Let $C(l_1, l_2, l_3; m_0, m_1, m_2)$ be the subset of $S_{(n_0,n_1,n_2)}$ whose elements are the strings that are combination of l_1 01's, l_2 02's, l_3 12's, m_0 00's, m_1 11's, m_2 22's, where each pair is allowed to be transposed. For example,

$$x = 01|00|10|12|00|20|00|10|02|01|00|20|02|10|20|20|11$$

is in $C(5,6,1;4,1,0)$, and

$$\begin{aligned}\Psi_1(x) &= 010110010111,\\ u(x) &= 10110101110111110,\\ v(x) &= 00001,\\ w(x) &= 110212122122.\end{aligned}$$

So,

$$\begin{aligned}\Psi(x) &= 010110010111 * \Psi(10110101110111110)\\ &\quad * \Psi(00001) * \Psi(110212122122)\\ &= 010110010111\\ &\quad * (1000 * \Psi(10110100) * \Psi(1111) * \Psi(1111))\\ &\quad * (\lambda * \Psi(00) * \Psi(00) * \Psi(\lambda))\\ &\quad * (0001 * \Psi(011110) * \Psi(12) * \Psi(2000))\\ &= 010110010111 * (1000 * 10111 * \lambda * \lambda)\\ &\quad * \lambda * (0001 * 011 * 0 * 11)\\ &= 010110010111100010111000101 1011.\end{aligned}$$

Let $l = l_1 + l_2 + l_3$ and $m = m_0 + m_1 + m_2$ so that $n = m + l$. Then we have a decomposition

$$S_{(n_0,n_1,n_2)} = \bigcup_{\substack{n_0=l_1+l_2+2m_0\\ n_1=l_1+l_3+2m_1\\ n_2=l_2+l_3+2m_2}} C(l_1,l_2,l_3;m_0,m_1,m_2). \quad (1)$$

Call the set $C(l_1,l_2,l_3;m_0,m_1,m_2)$ the (n_0,n_1,n_2)-class of type $(l_1,l_2,l_3;m_0,m_1,m_2)$. If $C = C(l_1,l_2,l_3;m_0,m_1,m_2)$, then

$$\begin{aligned}\Psi_1(C) &= \{0,1\}^l\\ u(C) &= S_{(m,l,0)}\\ v(C) &= S_{(m_0,m_1,m_2)}\\ w(C) &= S_{(l_1,l_2,l_3)}.\end{aligned}$$

Lemma 1 (Structure of (n_0,n_1,n_2)-class [6])**.** *In equiprobable set $S_{(n_0,n_1,n_2)}$, the mapping*

$$x \mapsto \Phi(x) = (\Psi_1(x), u(x), v(x), w(x))$$

is one-to-one correspondence between the (n_0,n_1,n_2)-class of type $(l_1,l_2,l_3;m_0,m_1,m_2)$ and $\{0,1\}^l \times S_{(m,l,0)} \times S_{(m_0,m_1,m_2)} \times S_{(l_1,l_2,l_3)}$.

B. Asymptotic Optimality

Now, consider the truncated versions of Peres function, whose recursion depth is bounded by ν, defined as follows:

$$\Psi_\nu(x) = \Psi_1(x) * \Psi_{\nu-1}(u(x)) * \Psi_{\nu-1}(v(x)) * \Psi_{\nu-1}(w(x)),$$

where $\Psi_0(x) = \lambda$. Since x is from Bernoulli(p,q,r), $u(x)$, $v(x)$ and $w(x)$ are of distributions

$$\begin{aligned}U(p,q,r) &= \left(p^2+q^2+r^2, 2(pq+qr+rp), 0\right),\\ V(p,q,r) &= \left(\frac{p^2}{p^2+q^2+r^2}, \frac{q^2}{p^2+q^2+r^2}, \frac{r^2}{p^2+q^2+r^2}\right),\\ W(p,q,r) &= \left(\frac{qr}{pq+qr+rp}, \frac{pq}{pq+qr+rp}, \frac{rp}{pq+qr+rp}\right),\end{aligned}$$

respectively. The average output length per input of $u(x)$, $v(x)$, and $w(x)$ are $\frac{1}{2}$, $\frac{1}{2}(p^2+q^2+r^2)$, and $(pq+qr+rp)$, respectively. So, the rate ρ_ν of Ψ_ν is

$$\begin{aligned}\rho_\nu(p,q,r) =& (pq+qr+rp) + \frac{1}{2}\rho_{\nu-1}(U(p,q,r))\\ &+ \frac{1}{2}(p^2+q^2+r^2)\rho_{\nu-1}(V(p,q,r))\\ &+ (pq+qr+rp)\rho_{\nu-1}(W(p,q,r)),\end{aligned} \quad (2)$$

and, of course, $\rho_1(p,q,r) = pq+qr+rp$ and $\rho_0(p,q,r) = 0$.

Expanding this formula, we obtain, for example,

$$\begin{aligned}\rho_2(p,q,r) =& (pq+qr+rp) + (pq+qr+rp)(p^2+q^2+r^2)\\ &+ \frac{pqr(p+q+r)}{pq+qr+rp} + \frac{p^2q^2+q^2r^2+r^2p^2}{2(p^2+q^2+r^2)},\end{aligned}$$

and the formula for ρ_ν becomes complicated very fast as ν increases, as we can expect from (2).

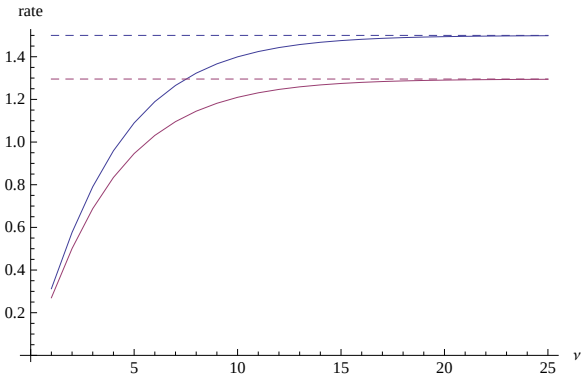


Fig. 1. Rates $\rho_\nu(p,q,r)$ for $\nu = 1,\ldots,25$ and $(p,q,r) = (0.25, 0.25, 0.5)$ and $(p,q,r) = (0.1, 0.3, 0.6)$. The dashed lines indicate the entropy bounds for each value of (p,q,r).

The original (2-face) Peres function [1] was defined as a truncated version and its rate is equal to $\rho_\nu(p,q,0)$. Hence, the (truncated) rate function ρ_ν also generalizes the 2-face case. Since the 2-face Peres function is asymptotically optimal, the corresponding truncated version converges to the Shannon entropy $H(p) = -(p\log_2 p + q\log_2 q)$. We can expect the rate ρ_ν of our 3-face truncated Peres function also converges to $H(p,q,r) = -(p\log_2 p + q\log_2 q + r\log_2 r)$. Fig.1 shows the plot of $\rho_\nu(p,q,r)$ for $\nu = 1,\ldots,25$ and $(p,q,r) = (0.25, 0.25, 0.5)$ and $(p,q,r) = (0.1, 0.3, 0.6)$. Indeed, the rates seem to converge to the corresponding entropy bounds, $H(0.25, 0.25, 0.5) = 1.5$ and $H(0.1, 0.3, 0.6) \approx 1.295$, respectively.

The following theorem, whose proof is given in [6], implies that Ψ is asymptotically optimal.

Theorem 2 ([6])**.**

$$\lim_{\nu\to\infty} \rho_\nu(p,q,r) = H(p,q,r).$$

II. Exact Output Rate

A. Total Output Length on Equiprobable Set $S_{(n_0,n_1,n_2)}$

Define $P(n_0, n_1, n_2)$ to be the total number of output bits over $S_{(n_0,n_1,n_2)}$, that is,

$$P(n_0, n_1, n_2) = \sum_{x \in S_{(n_0,n_1,n_2)}} |\Psi(x)|.$$

Then the rate of Ψ is

$$\begin{aligned}\rho(N) &= \frac{1}{N} \sum_{x \in \{0,1,2\}^N} |\Psi(x)| \Pr(x) \\ &= \frac{1}{N} \sum_{N=n_0+n_1+n_2} P(n_0, n_1, n_2) p^{n_0} q^{n_1} r^{n_2}.\end{aligned}$$

Note that $P(n_0, n_1, n_2)$ is independent on the probability distribution (p, q, r). So, once we compute an appropriate table of values of $P(n_0, n_1, n_2)$, which is computationally the most demanding part, the exact rate $\rho(p, q, r)$ can be easily computed for each (p, q, r). In the following, we give a recursive formula for $P(n_0, n_1, n_2)$ so that its values can be computed, for example, by dynamic programming.

With a bit of abuse of notation, for a class $C(l_1, l_2, l_3; m_0, m_1, m_2)$, use the same symbol P and let

$$P(l_1, l_2, l_3; m_0, m_1, m_2) = \sum_{x \in C(l_1,l_2,l_3;m_0,m_1,m_2)} |\Psi(x)|.$$

Then, by the decomposition (1) we have

$$P(n_0, n_1, n_2) = \sum_{\substack{n_0=l_1+l_2+2m_0 \\ n_1=l_1+l_3+2m_1 \\ n_2=l_2+l_3+2m_2}} P(l_1, l_2, l_3; m_0, m_1, m_2). \quad (3)$$

Now, by the structure lemma, for $C = C(l_1, l_2, l_3; m_0, m_1, m_2)$, the image by Ψ_1 over C is

$$\binom{n}{m, l} \binom{m}{m_0, m_1, m_2} \binom{l}{l_1, l_2, l_3}$$

copies of $\{0, 1\}^l$. So, we have

$$\sum_{x \in C} |\Psi_1(x)| = 2^l \binom{n}{m, l} \binom{m}{m_0, m_1, m_2} \binom{l}{l_1, l_2, l_3} \cdot l.$$

Similarly, we have

$$\begin{aligned}\sum_{x \in C} |\Psi(u(x))| &= 2^l \binom{m}{m_0, m_1, m_2} \binom{l}{l_1, l_2, l_3} P(m, l, 0), \\ \sum_{x \in C} |\Psi(v(x))| &= 2^l \binom{n}{m, l} \binom{l}{l_1, l_2, l_3} P(m_0, m_1, m_2), \\ \sum_{x \in C} |\Psi(w(x))| &= 2^l \binom{n}{m, l} \binom{m}{m_0, m_1, m_2} P(l_1, l_2, l_3).\end{aligned}$$

Since $|\Psi(x)| = |\Psi_1(x)| + |\Psi(u(x))| + |\Psi(v(x))| + |\Psi(w(x))|$, we have

$$\begin{aligned}P(l_1, l_2, l_3; m_0, m_1, m_2) = \\ 2^l \Bigg[&\binom{n}{m, l} \binom{m}{m_0, m_1, m_2} \binom{l}{l_1, l_2, l_3} \cdot l \\ &+ \binom{m}{m_0, m_1, m_2} \binom{l}{l_1, l_2, l_3} P(m, l, 0) \\ &+ \binom{n}{m, l} \binom{l}{l_1, l_2, l_3} P(m_0, m_1, m_2) \\ &+ \binom{n}{m, l} \binom{m}{m_0, m_1, m_2} P(l_1, l_2, l_3) \Bigg].\end{aligned} \quad (4)$$

1) Exploiting Symmetry: If (n_0', n_1', n_2') is a permutation of (n_0, n_1, n_2), then $P(n_0', n_1', n_2') = P(n_0, n_1, n_2)$. Therefore, we need to compute only $P(n_0, n_1, n_2)$ for $n_0 \geq n_1 \geq n_2$.

$$\begin{aligned}P(n_0', n_1', n_2') = P(n_0, n_1, n_2), \quad n_0 \geq n_1 \geq n_2, \\ (n_0', n_1', n_2') \text{ is a permutation of } (n_0, n_1, n_2)\end{aligned} \quad (5)$$

Symmetry in $P(l_1, l_2, l_3; m_0, m_1, m_2)$ is taken care of at this stage.

2) Odd-length Input: In the right-hand side of (4), a recursive call to $P(n_0, n_1, n_2)$ can be made for an odd value of $n_0+n_1+n_2$. In that case, we need to reduce it to even-length, for $n_i > 0$, $i = 0, 1, 2$,

$$\begin{aligned}P(n_0, n_1, n_2) = P(n_0 - 1, n_1, n_2) + P(n_0, n_1 - 1, n_2) \\ + P(n_0, n_1, n_2 - 1), \quad \text{if } n_0 + n_1 + n_2 \text{ is odd,}\end{aligned}$$

and for n_0 and n_1 are positive and $n_2 = 0$,

$$\begin{aligned}P(n_0, n_1, 0) = P(n_0 - 1, n_1, 0) + P(n_0, n_1 - 1, 0), \\ \text{if } n_0 + n_1 \text{ is odd.}\end{aligned}$$

3) Initial Conditions: Clearly, for $n \geq 0$,

$$P(n, 0, 0) = 0. \quad (6)$$

B. Linear Diophantine Equations

Two linear Diophantine equations are involved in the computation. First, given N, we need to find all the nonnegative solutions (n_0, n_1, n_2) such that $N = n_0 + n_1 + n_2$. This is a partition into 3 parts and the solutions can be efficiently generated by methods given in, for example, [7] or [8].

Now, as for the second equation, for a given (n_0, n_1, n_2), we need to generate all the nonnegative integer solutions $(l_1, l_2, l_3; m_0, m_1, m_2)$ of the equations

$$\begin{aligned}n_0 &= l_1 + l_2 + 2m_0, \\ n_1 &= l_1 + l_3 + 2m_1, \\ n_2 &= l_2 + l_3 + 2m_2.\end{aligned} \quad (7)$$

The solutions can be generated efficiently as follows: Since the coefficients for m_0, m_1, and m_2 are dominant, we first list all the possible candidates for (m_0, m_1, m_2) as

$$\begin{aligned}M = \{(m_0, m_1, m_2) \mid m_i \text{ is nonnegative integer s.t.} \\ 0 \leq m_i \leq n_i/2,\ i = 0, 1, 2\}.\end{aligned}$$

For example, given $(n_0, n_1, n_2) = (7, 5, 2)$, the corresponding M is

$$\{(0,0,0),(0,0,1),(0,1,0),(0,1,1),(0,2,0),(0,2,1),$$
$$(1,0,0),(1,0,1),(1,1,0),(1,1,1),(1,2,0),(1,2,1),$$
$$(2,0,0),(2,0,1),(2,1,0),(2,1,1),(2,2,0),(2,2,1),$$
$$(3,0,0),(3,0,1),(3,1,0),(3,1,1),(3,2,0),(3,2,1)\}.$$

Then, for each of these triples (m_0, m_1, m_2), we solve for nonnegative solutions of the equations

$$l_1 + l_2 = n_0 - 2m_0,$$
$$l_1 + l_3 = n_1 - 2m_1,$$
$$l_2 + l_3 = n_2 - 2m_2.$$

This system is non-singular and has a unique *real* solution (l_1, l_2, l_3), and if these l_i's are nonnegative integers, then we take $(l_1, l_2, l_3, m_0, m_1, m_2)$ as a solution. Therefore the number of solutions is bounded by $|M| = (\lfloor n_0/2 \rfloor + 1)(\lfloor n_1/2 \rfloor + 1)(\lfloor n_2/2 \rfloor + 1) \le (N/3 + 1)^3$.

For the case $(n_0, n_1, n_2) = (7, 5, 2)$ given above, the corresponding solutions are now

$$\{(5,2,0,0,0,0),(4,1,1,1,0,0),(5,0,0,1,0,1),$$
$$(3,2,0,1,1,0),(3,0,2,2,0,0),(2,1,1,2,1,0),$$
$$(3,0,0,2,1,1),(1,2,0,2,2,0),(1,0,2,3,1,0),$$
$$(0,1,1,3,2,0),(1,0,0,3,2,1)\}.$$

C. Maximum Output Rate

For any given procedure that converts a length-n input of biased Bernoulli source to unbiased random bits, the maximum average output rate can be obtained, like the rate of Peres algorithm discussed above, by computing the total output lengths on equiprobable sets [4]. In fact, the maximum output rate is obtained by Elias method. For example, for three-face case, let $E_n^3 : \{0,1,2\}^n \to \{0,1\}^*$ be the function corresponding to the Elias method. Then, for (n_0, n_1, n_2) such that $n = n_0 + n_1 + n_2$, the total output length over $S_{(n_0,n_1,n_2)}$

$$Q(n_0, n_1, n_2) = \sum_{x \in S_{(n_0,n_1,n_2)}} |E_n^3(x)|$$

can be computed, from the definition of Elias method, as follows [4]: let the standard binary expansion of $|S_{(n_0,n_1,n_2)}|$ be $\sum_i a_i 2^i$, where a_i is either zero or one. Then

$$Q(n_0, n_1, n_2) = \sum_i i \cdot a_i \cdot 2^i,$$

and the n-maximal output rate for input length n can be computed.

III. Computation Results

Using the recursive definitions for $P(n_0, n_1, n_2)$ and $P(l_1, l_2, l_3; m_0, m_1, m_2)$ given above, we can compute the values of them efficiently using dynamic programming. For example,

$$P(10, 20, 30) = 19\ 38905\ 30631\ 82778\ 17752\ 73600.$$

In comparison,

$$Q(10, 20, 30) = 28\ 46922\ 13778\ 64604\ 61389\ 79776.$$

Fig. 2 shows plots of the exact rates of three-face Peres algorithm $\rho(p, q, r)$, for $(p, q, r) = (0.25, 0.25, 0.5)$ and $(p, q, r) = (0.1, 0.3, 0.6)$, where the input lengths range from 2 to 160. Also shown is the maximal rates of Elias methods in comparison, for the same distributions (p, q, r) and for the same input lenghts.

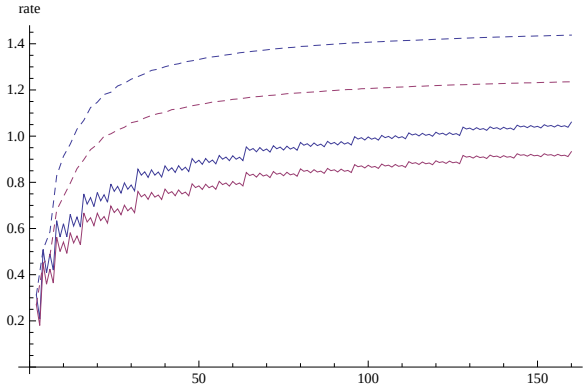


Fig. 2. Exact rates $\rho(p, q, r)$ for $(p, q, r) = (0.25, 0.25, 0.5)$ (shown in blue) and $(p, q, r) = (0.1, 0.3, 0.6)$ (in red), for input lengths $n = 2, \ldots, 160$. Dashed lines are the maximal rates (Elias) for the respective distbutions (again, shown in blue and red).

IV. Remarks

Although the method described here is much more efficient than brute-force calculation over all possible inputs, it still takes a considerable time. For example, it took several hours to obtain the data for Fig. 2 with a decently fast personal computer from the standard of the time when this paper was written.

Acknowledgement

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by Korean government (No. 2009-0077288).

References

[1] Y. Peres, "Iterating von Neumann's procedure for extracting random bits," *Annals of Statistics*, vol. 20, no. 1, pp. 590–597, 1992.

[2] P. Elias, "The efficient construction of an unbiased random sequence," *The Annals of Mathematical Statistics*, vol. 43, no. 3, pp. 865–870, 1972.

[3] S. Pae and M. C. Loui, "Optimal random number generation from a biased coin," in *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, January 2005, pp. 1079–1088.

[4] ——, "Randomizing functions: Simulation of discrete probability distribution using a source of unknown distribution," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4965–4976, November 2006.

[5] S. Pae, "Exact output rate of Peres's algorithm for random number generation," *Inf. Process. Lett.*, vol. 113, no. 5-6, pp. 160–164, 2013.

[6] ——, "A generalization of Peres's algorithm for generating random bits from loaded dice," 2013, submitted.

[7] D. E. Knuth, *The Art of Computer Programming, Combinatorial Algorithms, Part 1*. Addison-Wesley, 2011, vol. 4A.

[8] A. Nijenhuis and H. S. Wilf, *Combinatorial Algorithms: For Computers and Calculators*. Academic Press, 1978.