IJACSA

WHERE WISDOM SHARES

International Journal of Advanced Computer Science and Applications

Volume 3 Issue 4

April 2012

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

SAI

www.ijacsa.thesai.org

IJACSA

WHERE WISDOM SHARES

INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS

SAI

THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org

Editorial Preface

From the Desk of Managing Editor…

IJACSA seems to have a cult following and was a humungous success during 2011. We at The Science and Information Organization are pleased to present the April 2012 Issue of IJACSA.

While it took the radio 38 years and the television a short 13 years, it took the World Wide Web only 4 years to reach 50 million users. This shows the richness of the pace at which the computer science moves. As 2012 progresses, we seem to be set for the rapid and intricate ramifications of new technology advancements.

With this issue we wish to reach out to a much larger number with an expectation that more and more researchers get interested in our mission of sharing wisdom. The Organization is committed to introduce to the research audience exactly what they are looking for and that is unique and novel. Guided by this mission, we continuously look for ways to collaborate with other educational institutions worldwide.

Well, as Steve Jobs once said, Innovation has nothing to do with how many R&D dollars you have, it's about the people you have. 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.

We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJACSA provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

We regularly conduct surveys and receive extensive feedback which we take very seriously. We beseech valuable suggestions of all our readers for improving our publication.

Thank you for Sharing Wisdom!

Managing Editor
IJACSA
Volume 3 Issue 4 April 2012
ISSN 2156-5570 (Online)
ISSN 2158-107X (Print)

Editorial Board

Reviewer Board Members

CONTENTS

A New Application Programming Interface and a Fortran-like Modeling Language for Evaluating Functions and Specifying Optimization Problems at Runtime

Fuchun Huang
School of Engineering and Science, Victoria University
Melbourne, Australia

***Abstract*—A new application programming interface for evaluating functions and specifying optimization problems at runtime has been developed. The new interface, named FEFAR, uses a simple language named LEFAR. Compared with other modeling languages such as AMPL or OSil, LEFAR is Fortran-like hence easy to learn and use, in particular for Fortran programmers. FEFAR itself is a Fortran subroutine hence easy to be linked to user's main programs in Fortran language. With FEFAR a developer of optimization solver can provide pre-compiled, self-executable and directly usable software products. FEFAR and LEFAR are already used in some optimization solvers and should be a useful addition to the toolbox of programmers who develop solvers of optimization problems of any type including constrained/unconstrained, linear/nonlinear, smooth/nonsmooth optimization.**

Keywords-programming language; Fortran computing language; Fortran subroutine; Application programming interface; Runtime function evaluation; Mathematical programming; Optimization problem; Optimization modeling language.

I. INTRODUCTION

In this paper we introduce to readers a new application programming interface for evaluating functions and specifying optimization problems at runtime. FEFAR is a Fortran subroutine For Evaluating Functions At Runtime. It can be linked to a programmer's main program to provide a way to evaluate a function or solve an optimization problem at runtime. The functions or optimization problems must be written in a new language called LEFAR I developed together with FEFAR, for evaluating functions or specifying optimization problems at runtime. LEFAR is similar to but much simpler than Fortran, and that is a big merit as it is easy to learn and use and yet powerful enough and complex enough to be able to specify any functions and optimization problems at runtime. FEFAR and LEFAR are already used in USsolver [1] and UNsolver [2], two binary machine code programs for solving unconstrained smooth or nonsmooth optimization problems.

This paper is organized as following. Section II gives the background and related works, section III explains the FEFAR interfaces and parameters, and give some examples, section IV explains the rules of the LEFAR language with comparison to Fortran language, and section V gives some other resources and a near future work scope.

II. BACKGROUND AND RELATED WORKS

Advanced computing languages such as Fortran [3] and C are compiled language [4]. Unlike interpreted languages [5] such as the S language [6] in SPLUS and R, and the MATLAB language [7], compiled languages must first compile the main program and all other subroutines and functions into a binary machine code program. Programmers of such compiled languages often want to be able to evaluate functions at runtime after the source code of the program has been compiled into a binary machine code program. They may also want to be able to specify/describe optimization problems at runtime by using a modeling language similar to, or, ideally the same, advanced computing language they use, such as Fortran. They may want to keep the source codes of their programs to themselves for commercial reasons but still want others to be able to use or test their software products by giving them a self-executable, directly usable binary machine code program. On the user's end, the binary machine code programs are self-executable and usable immediately; hence the users, in particular the ordinary users but not programmers, are eased from the troubles of finding or purchasing a compiler and compiling the source code programs into binary machine code programs. There are some available modeling languages such as AMPL [8] and OSiL [9] for modeling and specifying optimization problems at runtime though, it is still better to have another modeling language and application programming interface that meet the abovementioned needs better, and that is why I have developed FEFAR and LEFAR, a new application programming interface and modeling language for evaluating functions and specifying optimization problems at runtime, in particular for Fortran programmers. FEFAR is a Fortran subroutine For Evaluating Functions At Runtime. It can be linked to a programmer's main program to provide a way to evaluate a function or solve an optimization problem at runtime. The functions or optimization problems must be written in a new language called LEFAR I developed together with FEFAR, for evaluating functions or specifying optimization problems at runtime. LEFAR is similar to but much simpler than Fortran, and that is a big merit as it is easy to learn and use. Another reason I made it simple is to shorten the processing and running time of LEFAR codes at runtime, subject to yet being powerful enough and complex enough to be able to specify any functions and optimization problems.

III. FEFAR INTERFACES

In the following we use gfortran [12] to illustrate how to use FEFAR, but other compilers work as well. There are several interfaces of FEFAR: FEFAR1.obj, FEFAR2.obj, FEFAR3.obj, etc. The following is a simple test program of FEFAR1.obj linked to the main program at compiling and linking stage.

```
>type FEFARtest.f90
program FEFARtest
real*8 :: f
real*8,dimension(1000) :: b
integer*4 :: k
do;
call FEFAR1(1,b,k,f);
call FEFAR1(2,b,k,f);
call FEFAR1(3,b,k,f);
end do;
end program
>
>gfortran FEFARtest.f90 FEFAR.obj
>a.exe
Input the file name of the function:
rosenbrock.far
    f=          24.1999999999999957
   x1=          -1.2000000000000000
   x2=           1.0000000000000000
>
```

The file “rosenbrock.far” is written in LEFAR language to evaluate the following Rosenbrock function [13] at given x values:

$$F(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2,$$
$$\overline{x}_1 = -1.2, \qquad \overline{x}_2 = 1.0.$$

LEFAR code (that is, the content of the file rosenbrock.f95) is the following:

```
function: Rosenbrock
real                  :: f
real, dimension(2) :: b
integer               :: if123

if(if123==1)
   b(1)=-1.2
   b(2)=1.0
end if

if(if123<=2)
f=100.0*(b(2)-b(1)**2)**2+(1.0-b(1))**2
end if

if(if123==3)
   print,”f=”,f
   print,”x1=”,b(1)
   print,”x2=”,b(2)
end if
end function
```

FEFAR1 is a Fortran subroutine of the following structure:

```
subroutine FEFAR1(if123,b,k,f)
integer*4 :: if123
real*8, dimension(1000) :: b
integer*4 :: k
real*8 :: f
…
end subroutine FEFAR1
```

FEFAR2 is a Fortran subroutine of the following structure:

```
subroutine FEFAR1(if123,b,k,f,g)
integer*4 :: if123
real*8, dimension(1000) :: b,g
integer*4 :: k
real*8 :: f
…
end subroutine FEFAR1
```

FEFAR3 is a Fortran subroutine of the following structure:

```
subroutine FEFAR1(if123,b,k,f,kg,g)
integer*4 :: if123
real*8, dimension(1000) :: b,g
integer*4 :: k
real*8 :: f
integer*4 :: kg
…
end subroutine FEFAR1
```

In calling each FEFAR, a file name is prompted to be inputted at runtime. The value of the integer variable “if123” will be passed to the first integer variable in the runtime file to control which statements in the runtime to be executed or not. For example, initial value assignment statements only need to be executed once, and most other statements need to be executed each time the subroutine is called. Another example is “if123==3” can be used for only displaying values without executing many other statements. Of course, the integer variable “if123” can take any integer value for more complex controls. The argument “b” is a real data type array of dimension 1 for getting initial values of the function from the source code at runtime or setting from the main program the next step *x* values of the function for solving an optimization problem . The argument “f” is a real variable of the value of the function. In FEFAR2 and FEFAR3 argument “g” is a real data type array of dimension 1 for getting array values of dimension 1 returned from the runtime function, with the only difference being that FEFAR3 returns an integer “kg” for the actual number of array values calculated and returned from the runtime function file which are to be used by the main program while FEFAR2 assumes implicitly kg=k. Argument “g” can be used to return gradients of a differentiable function, for example. It can also be used to return values of constraints.

IV. Rules of LEFAR language

LEFAR is a very simple language similar to but far simpler than Fortran language, so in many cases below we just give the Fortran equivalence of most LEFAR statements.

A. Data types

There are only three data types: real, integer, and logical, and they are equivalent to Fortran’s real(len=8) (or equivalently real*8), integer(len=4) (or equivalently integer*4) and logical(len=2) (or equivalently logical*2).

B. Constants

Real constants are specified like 2.0, 2.1, -2.0, where the decimal symbol ‘.’ is necessary, so is the ‘0’ following the ‘.’ even if there are no other decimal digits. Scientific notations such as 2.0D-1 and 2.0E-1 are not used in LEFAR. The following are not allowed: 2., .1, -2., -.3, 2.0D-1. Correct ways are: 2.0, 0.1, -2.0, -0.3, 0.2. Logical constants are .true. and .false., like in Fortran. The mathematical constant π (sometimes written as pi or PI) which is the ratio of any circle's circumference to its diameter is .PI. in LEFAR. For example, *y=sin(.PI./2)* assigns value 1.0 to *y*.

C. Intrinsic functions

Intrinsic functions in LEFAR have the same rules as those in Fortran. Currently implemented functions are: *abs(), exp(), log(), log10(), cos(), acos(), sin(), asin(), tan(), atan(), max(), min(), sqrt(), dble(),*and *int()*. Depending on demands other functions can be easily added to LEFAR.

D. Arrays

Arrays are specified and used the same way as in Fortran.

E. Operators, mathematical expressions and the assignment

All operators in Fortran work the same way in LEFAR. The assignment and mathematical expressions have the same rules as Fortran. For example, x(2,1)=2*(3.4+5)**2-5*x(1,2)**2 is valid in LEFAR and evaluated the same way as in Fortran. Additionally, “^” is also used for exponentiation, the same as “**”.

F. Do loop

There is only one construct of do looping:

```
do while(expr)
⋮
end do
```

which is equivalent to Fortran’s DO WHILE(expr) … END DO construct.

G. If construct

IF-THEN constructs are

```
if(expr)then
 ⋮
end if

if(expr)then
⋮
else
⋮
end if

if(expr)then
⋮
else if(expr)then
⋮
else if(expr)then
⋮
⋮
else
⋮
end if
```

They are equivalent to Fortran’s IF-THEN constructs. In LEFAR, however, the word “then” is not necessary, hence the following are also valid:

```
if(expr)
 ⋮
end if

if(expr)
⋮
else
⋮
end if

if(expr)
⋮
else if(expr)
⋮
else if(expr)
⋮
⋮
else
⋮
end if
```

H. data-end data construct

```
data(x)
⋮
end data
```

by which listed values between are read into x starting from the most right array index then the second array index from the right until the first array index from the left hand side.

I. datafile-end datafile construct

```
datafile(x)
'filename'
end datafile
```

by which values in the file 'filename' (a path can be included) are read into x starting from the most right array index then the second array index from the right until the first array index from the left hand side.

J. "print" statement

'print' outputs values to the monitor screen. For example,

```
print, "i, b(i):", i, b(i)
```

is similar to the following Fortran statement:

```
print*,'i, b(i):',i, b(i)
```

In 'print' statement, character strings must be put in double quotation marks "…", not single quotation marks.

K. Other statements and rules of LEFAR

- 'exit', 'cycle', 'stop', 'return' statements work the same as in Fortran.
- LEFAR statements use lower-case letter only. For example, 'print', but not 'PRINT'.
- LEFAR variables are case-sensitive.
- A LEFAR statement line can be up to 200 characters long.
- Like in Fortran, Any line starting with '!' is treated as a comment line.
- A function file can have up to 1000 lines.
- There is no way to continue a one-line statement (not a construct) in another line.
- All variables must be declared. There are no implicit rules. The way to declare variables and arrays are the same as in Fortran 95.

L. Rules of the runtime function

Generally the file may have the following structure:

```
function: function_name
real :: fmin
real, dimension(k) :: x
integer   :: if123
[declare other working variables]

if(if123==1)
⋮
end if

if(if123<=2)
⋮
end if

if(if123==3)
⋮
end if
end function
```

where 'fmin' is the value of the function to be returned to the main program, 'x' is the vector input of the function passed to and from the main program, and 'if123' is a working integer variable passed from the main program. **Important**: the first 'real ::' variable is the one to be returned, the first 'real, dimension(k) ::' vector is the input variable of the function, where 'k' is a positive integer like 2, 3, etc., which is the dimension of the function, while the first 'integer ::' variable is a special integer variable come from the main program (that is, the value of the variable is set in the main program and passed to the function for controlling which blocks to be executed). They can use different names such as 'f', 'b', 'iw'. Generally, within the block "if(if123==1) … end if" are statements to be processed only once. For example, 'data … end data' statements, to specify initial values for an optimization problem, etc. Within the block "if(if123<=2) … end if" are statements to be processed repeatedly like in optimization program. Within the block "if(if123==3) … end if" are statements to be processed only once in the final stage. For example, after a minimum x* has been found, it can be used in this block to evaluate values of other variables or functions depending on it. Two other rules are:

- In Fortran, ';' is used to put and separate two statements in one line. In LEFAR, however, there is no way to separate two one-line statements in one line.
- There should not be a ';' nor any other separator at the end of any statement.

A side note of the above two rules, they make the processing and running time of the codes shorter.

V. Resources and Future Work

More codes, examples and future work are available at http://sites.google.com/site/SoftSome. Future work may include a C computing language version of FEFAR for easy linking of optimization solver programs in C language to FEFAR.

Acknowledgment

The author thanks Professor Yosihiko Ogata, Professor Satoshi Ito, Professor and Director Tomoyuki Higuchi and The Institute of Statistical Mathematics, Tokyo, Japan, for supporting my research visit to the institute from the 5th of January to the 28th of February in 2012.

References

[1] F. Huang, Some test results of UNsolver: a solver for solving unconstrained non-smooth optimization problems, http://sites.google.com/site/VicSolver, 2012.

[2] F. Huang, Some test results of USsolver: a solver for solving unconstrained smooth optimization problems, http://sites.google.com/site/VicSolver, 2012.

[3] Wikipedia, "Fortran", http://en.wikipedia.org/wiki/Fortran, retrived on April 16, 2012.

[4] Wikipedia, “Compiled language”, http://en.wikipedia.org/wiki/Compiled_language, retrived on April 16, 2012.

[5] Wikipedia, “Interpreted language”, http://en.wikipedia.org/wiki/Interpreted_language, retrived on April 16, 2012.

[6] Wikipedia, “S (programming language”, http://en.wikipedia.org/wiki/S_(programming_language), retrived on April 16, 2012.

[7] Wikipedia, “MATLAB”, http://en.wikipedia.org/wiki/MATLAB, retrived on April 16, 2012.

[8] R. Fourer, D.M. Gay, and B.W. Kernighan, AMPL: A Modeling Language for Mathematical Programming. Scientific Press, San Francisco, CA, 1993.

[9] R. Fourer, J. Ma and Kipp Martin, “OSiL: An Instance Language for Optimization,” Computational Optimization and Applications, Computational Optimization and Applications, 2010.

[10] Wikipedia, “gfortran,” http://en.wikipedia.org/wiki/Gfortran, retrived on April 16, 2012.

[11] L. Luksan and J. Vlcek, “Test Problems for Nonsmooth Unconstrained and Linearly Constrained Optimization,” Technical report No. 798, Institute of Computer Science, Academy of Sciences of the Czech Republic, 2000.

AUTHORS PROFILE

Dr Fuchun Huang is a Senior Lecturer in the School of Engineering and Science at Victoria University, Melbourne, Australia. He was awarded a PhD degree by The Graduate University of Advanced Studies, Tokyo, Japan, and has published papers on computational statistics, in particular Monte Carlo methods, pseudo-likelihood and generalized pseudo-likelihood methods, and developed solver software for solving smooth and nonsmooth optimization problem. He is a member of The Japan Statistical Society.

Wavelet Based Image Retrieval Method

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Cahya Rahmad
Electronic Engineering Department
The State Polytechnics of Malang,
East Java, Indonesia

***Abstract*—A novel method for retrieving image based on color and texture extraction is proposed for improving the accuracy. In this research, we develop a novel image retrieval method based on wavelet transformation to extract the local feature of an image, the local feature consist color feature and texture feature. Once an image taking into account, we transform it using wavelet transformation to four sub band frequency images. It consists of image with low frequency which most same with the source called approximation (LL), image containing high frequency called horizontal detail (LH), image containing high frequency called vertical detail (HL), and image containing horizontal and vertical detail (HH). In order to enhance the texture and strong edge, we combine the vertical and horizontal detail to be other matrix. The next step is we estimate the important point called significant point by threshold the high value. After the significant points have been extracted from image, the coordinate of significant points will be used for knowing the most important information from the image and convert into small regions. Based on these significant point coordinates, we extract the image texture and color locally. The experimental results demonstrate that our method on standard dataset are encouraging and outperform the other existing methods, improved around 11 %.**

Keywords-component; Image retrieval; DWT; Wavelet; Local feature; Color; Texture.

I. INTRODUCTION

Image retrieval has been used to seek an image over thousand database images. In the web based search engine, the image retrieval has been used for searching an image based on text input or image. Once an input taking into account, the method will search most related image to the input. The correlation between input and output has been defined by specific role. With expansion in the multimedia technologies and the Internet, CBIR has been an active research topic since the first 1990's. The concept of content based retrieval (CBR) in image start from the first 1980s and serious applications started in the first 1990s. Retrieval from databases with a large number of images has attracted considerable attention from the computer vision and pattern recognition society.

Brahmi et al. mentioned the two drawbacks in the keyword annotation image retrieval. First, images are not always annotated and the manual annotation expensive also time consuming. Second, human annotation is not objective the same image may be annotated differently by different observers [1]. Unlike the traditional approach that using the keyword annotation as a method to search images, CBIR system performs retrieval based on the similarity feature vector of color, texture, shape and other image content. Comparing to the traditional systems, the CBIR systems perform retrieval more objectiveness [2]. A very basic issue in designing a CBIR system is to select the most effective image features to represent image contents (3). Global features related to color or texture are commonly used to describe the image content in image retrieval. The problem using global features is this method cannot capture all parts of the image having different characteristics [4].

In order to capture specific parts of the image the local feature is used. The proposed method uses 2D Discrete wavelet transform with Haar base function, combined the two high sub-band frequency to make significant points and edge then estimate the important point called significant point by threshold the high value. After the significant points have been extracted from image, the coordinate of significant points will be used for knowing the most important information from the image and convert into small regions. Based on these significant point coordinates, and then extract the image texture and color texture locally.

II. PROPOSED METHOD

A. Wavelet Transformation

The wavelet representation gives information about the variations in the image at different scales. Discrete Wavelet Transform (DWT) represents an image as a sum of wavelet functions with different locations (shift) and scales [5]. Wavelet is the multi-resolution analysis of an image and it is proved that having the signal of both space and frequency domain [6]. Any decomposition of an 1D image into wavelet involves a pair of waveforms: the high frequency components are corresponding to the detailed parts of an image while the low frequency components are corresponding to the smooth parts of an image.

DWT for an image as a 2D signal can be derived from a 1D DWT, implement 1D DWT to every rows then implement 1D DWT to every column. Any decomposition of an 2D image into wavelet involves four sub-band elements representing LL (Approximation), HL (Vertical Detail), LH (Horizontal Detail), and HH (Detail), respectively.

The wavelet transform may be seen as a filter bank and illustrated as follow, on a one dimensional signal x[n]. x[n] is input signal that contains high frequencies and low frequencies. h[k] and g[k] is channel filter bank involving sub sampling. c[n] is called averages contains low frequencies signal. d[n] is called wavelet coefficients contain high frequencies signal. c[n] and d[n] be sub sampled (decimated by 2: $\downarrow 2$) the next process for further decomposition is iterated on the low signal c[n].

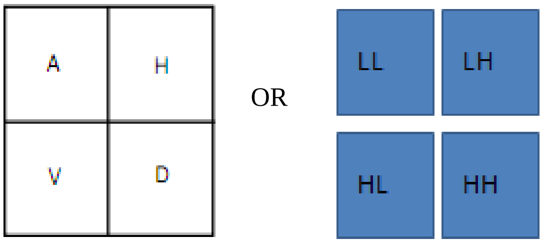


Figure 1. Level 1 of 2D DWT

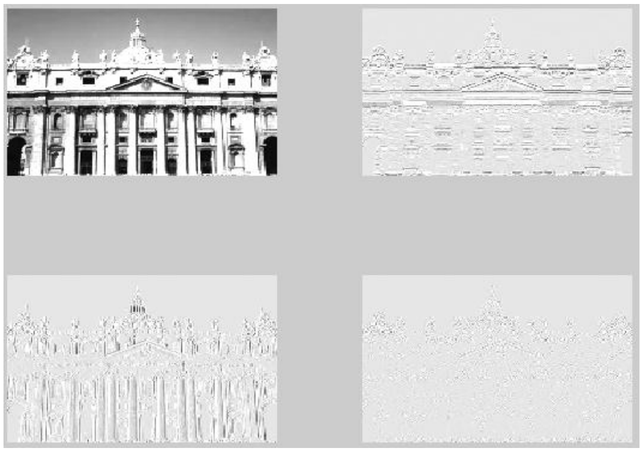


Figure 2. Example Level 1 of 2D DWT

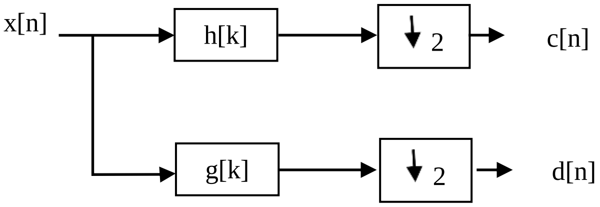


Figure 3. Two channel filter bank

For example, 1D Haar wavelet decomposition is expressed as follows, let x[n] be an input, x[n]= $X_0,X_1,X_2,\ldots X_{N-1}$ which contains N elements. Then output will consist of N/2 elements of averages over the input and is stored in c[n]. Also the other output contains N/2 elements wavelet coefficients values and is stored in d[n]. The Haar equation to calculate an average AV_i (See Eq.1) and a wavelet coefficient WC_i (See Eq.2) from pair data odd and even element in the input data are:

$$AV_i = \frac{X_i + X_{i+1}}{2} \quad (1)$$

$$WC_i = \frac{X_i - X_{i+1}}{2} \quad (2)$$

where:

AV = Average
WC =Wavelet coefficient

B. *Color and Texture*

Texture contain repeating pattern of local variations in image intensity also an area that can be perceived as being spatially homogeneous. Texture provides important characteristics for surface and object identification. Texture information extracted from the original image is typical features for image retrievals [7]. The texture is characterized by the statistical distribution of the image intensity using energy of Gabor filter on 7x7 pixels. Color is produced by spectrum of light that absorbed or reflected then received by the human eye and processed by the human brain. To extract the color feature, the first order statistical moments (See Eq.(3)) and the second order statistical moments (See Eq.(4)) HSV color space is similar to human perception color system so we used it to extract the color feature in the HSV color space on neighbor of significant points with size 3x3 pixels.

The first order statistical moments is expressed as follows,

$$\mu = \frac{1}{MxN}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j) \quad (3)$$

where:

P = Pixel value
MxN = Size of significant points and its neighbor.

The second order statistical moments is represented as follows,

$$\sigma = \sqrt{\frac{1}{MxN}\sum_{i=1}^{M}\sum_{j=1}^{N}(\,p(i,j) - \mu)^2} \quad (4)$$

where:

P = pixel value
μ = The first order statistical moments value
MxN = Size of significant points and its neighbor

C. *Image Retieval Algorithm*

The proposed image retrieval algorithm is as follows,

1. Read Query image and Convert from RGB image to gray image and HSV image then Decomposition using wavelet transformation.
2. Make absolute for every Wavelet coefficients, WCnew = | WCold |.
3. Combine Vertical Detail and Horizontal Detail, CVdHd(i,j) = Max(Vd(i,j),Hd(i,j)).
4. Choose significant points on CVdHd(i,j) by threshold the high value.
5. Choose points on HSV image and it neighbor (3x3 pixel) base on coordinate significant points on CVdHd(i,j) then Forming color feature vector by using The first order statistical moment and the second order statistical moment.
6. Forming texture feature vector by using Gabor transform on 7x7 pixel neighbor of significant points and Implement min/max normalization on all feature vector with range [0 1].
7. Measure the distance between feature vector image query and feature vector image in the dataset by using

Euclidean distance then display image results with X top ranking from the dataset.

III. Experiments

The retrieval result is not a single image but a list of image ranked by their similarity. The similarity measure is computed by using Euclidean distance (See Eq.(5)) between feature representation of image query and feature representation of image in dataset. The feature representation is image feature refer to the characteristics which describe the contents of an image.

$FQ = (Q_1, Q_2, \ldots, Q_{n)}$
$FD = (D_1, D_2, \ldots, D_{n)}$

$$\mathrm{dis}(FQ, FD) = \sqrt{\sum_{j=1}^{n} (Qj - Dj)^2} \qquad (5)$$

where :

FQ = Feature vector of query image.
FD = Feature vector of image in data set
n = Number element of feature vector

If the distance between feature representation of image query and feature representation of image in dataset small then it to be considered as similar.

The performance of the CBIR system is calculated by showing image a number of top ranking from the dataset. We used precision and recall to evaluate the performance of the CBIR system. Precision measures the retrieval accuracy; it is ratio between the number of relevant images retrieved and the total number of images retrieved (See Eq.(6)). Recall measures the ability of retrieving all relevant images in the dataset. It is ratio between the number of relevant images retrieved and the whole relevant images in the dataset (See Eq.(7)).

The performance of the CBIR system is calculated by using equation (6) and equation (7).

$$\mathbf{Precision} = \frac{NRRI}{XR} \qquad (6)$$

where :

NRRI = Number of relevant retrieved images
XR = X Top ranking of retrieved images

$$\mathbf{Recall} = \frac{NRRI}{TR} \qquad (7)$$

where:

NRRI = Number of relevant retrieved images
TR = Total number of relevant images in dataset

We used The Wang`s dataset [8] To evaluate the effectiveness of our approach and compared with the standard system SIMPLICITY, FIRM and also Color salient points by using the same dataset [8],[9],[10].

Figure 4, 5, and 6 shows query image and the retrieved image. Query image number is shown in the query image while retrieved image number is also indicated in the retrieved image. Figure 4 shows relatively good retrieval accuracy while Figure 5 and 6 for flower and horses shows strange retrieved results. Image number 236 and 334 for flower and image number 34 for horses are not correct.

The comparison average precision results between the proposed method and other method is showed in the Table 1 and the result by using the proposed method is improve. The improvement compared other method are 17%, 12%, 11%, respectively.

Image retrieval experiments with five query images including the aforementioned three query images are conducted. Figure 7 shows the query image and relevance image as results of this system. The relevance image results for bus, dinosaur, elephant, flower, and horse as a query image are 6, 10, 7, 8 and 7, respectively.

Table 2 shows the comparison average precision results between the proposed method and other methods. It shows that ability of system to retrieve relevance image from image in dataset is improve. The improvement compared other method are 12%, 17%, 11%, respectively. The graphic comparison of average precision can be seen in Figure 8.

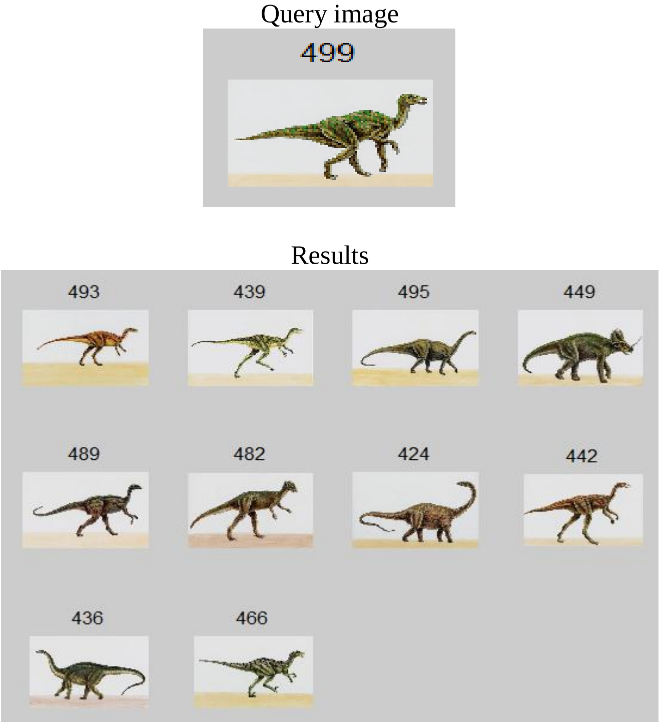


Figure 4 Example results for the dinosaur as query image

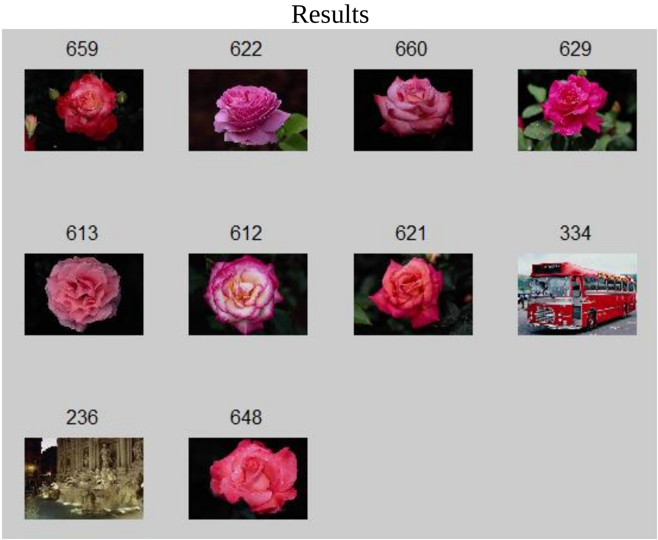


Figure 5. Example results for the flower as query image

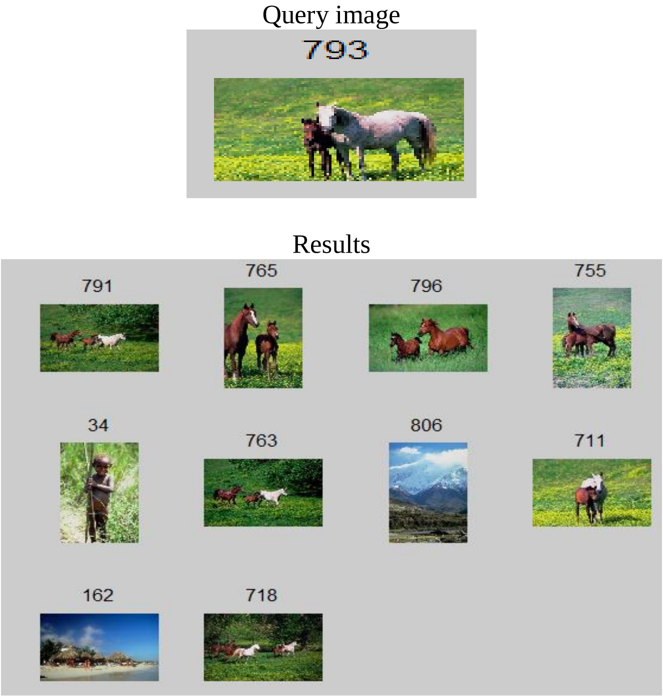


Figure 6. Example results for the horses as query image

TABLE 1. Comparation average precision results between the proposed method with other method

Method	Average Precision results
Simplicity	57%
Firm	62%
Color salient points	63%
Proposed method	74%

Query image number 2 shows almost same image retrieval accuracy and is more than 0.9 so that it is easy to retrieve this image. Meanwhile query image number 3 shows much poor image retrieval accuracy below 40% for the conventional methods with Firm, Simplicity, and Color salient points of gradient vector while the proposed method shows relatively high accuracy.

TABLE 2. Average precision results

Category	Firm	Simplicity	Color salient points	Proposed method
Bus	0.60	0.36	0.52	0.68
Dinosaur	0.95	0.95	0.95	0.94
Elephant	0.25	0.38	0.40	0.60
Flower	0.65	0.42	0.60	0.75
Horses	0.65	0.72	0.70	0.71
Average	0.62	0.57	0.63	0.74

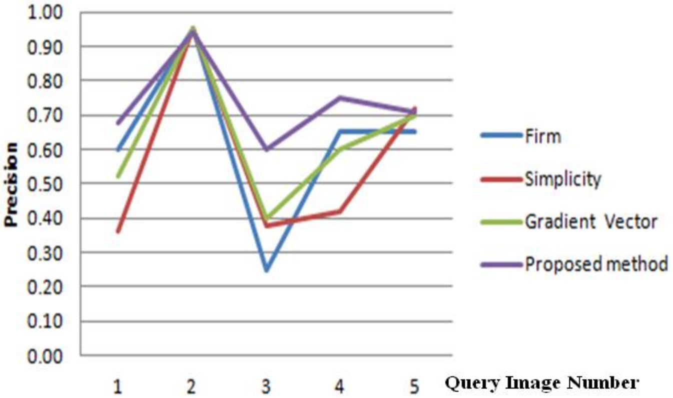


Figure 8. Comparison of average precision

Such this image need time-frequency components of image feature for image retrievals. The difference between the image retrieval accuracy of the proposed method and the conventional methods is around 20%, significant difference. On the other hand, both of spatial and color features are required for image retrievals of the query image number 5. The image retrieval accuracy of the conventional method with Simplicity and Color salient points of gradient vector is almost same as that of the proposed method so that these features work for image retrievals for this image.

Figure 9 shows the relation between image retrieval accuracy of the proposed method and those of the conventional methods with Firm, Simplicity, and Color salient points of gradient vector. In the figure, linear regressive equations are included with R-square values. The relation the image retrieval accuracy between the proposed method and the conventional method with Color salient points of gradient vector shows the highest R-square value of 0.909 followed by Firm and Simplicity. Therefore, the most significant feature for image retrievals is Color salient points of gradient vector followed by Firm and Simplicity for these retrieved images because the proposed method shows the highest image retrieval accuracy.

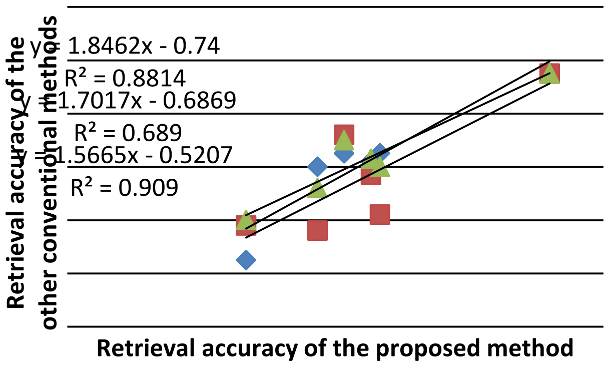


Figure 9 Relation between wavelet derived feature and the others

IV. Conclusion

In this research we proposed a method for image retrieval by using wavelet transformation. In order to enhance the texture and make strong edge, we combine the vertical and horizontal detail then estimate the important point called significant point by threshold the high value then by using it find the most important information from the image and convert it into small regions and extract the image texture and color locally. We proposed a method for image retrieval by using wavelet transformation. We combined the two high sub-band frequencies In order to make strong points and edge then detect the location of significant points. The experimental results demonstrate that our method on standard dataset is significantly improved around 11 %.

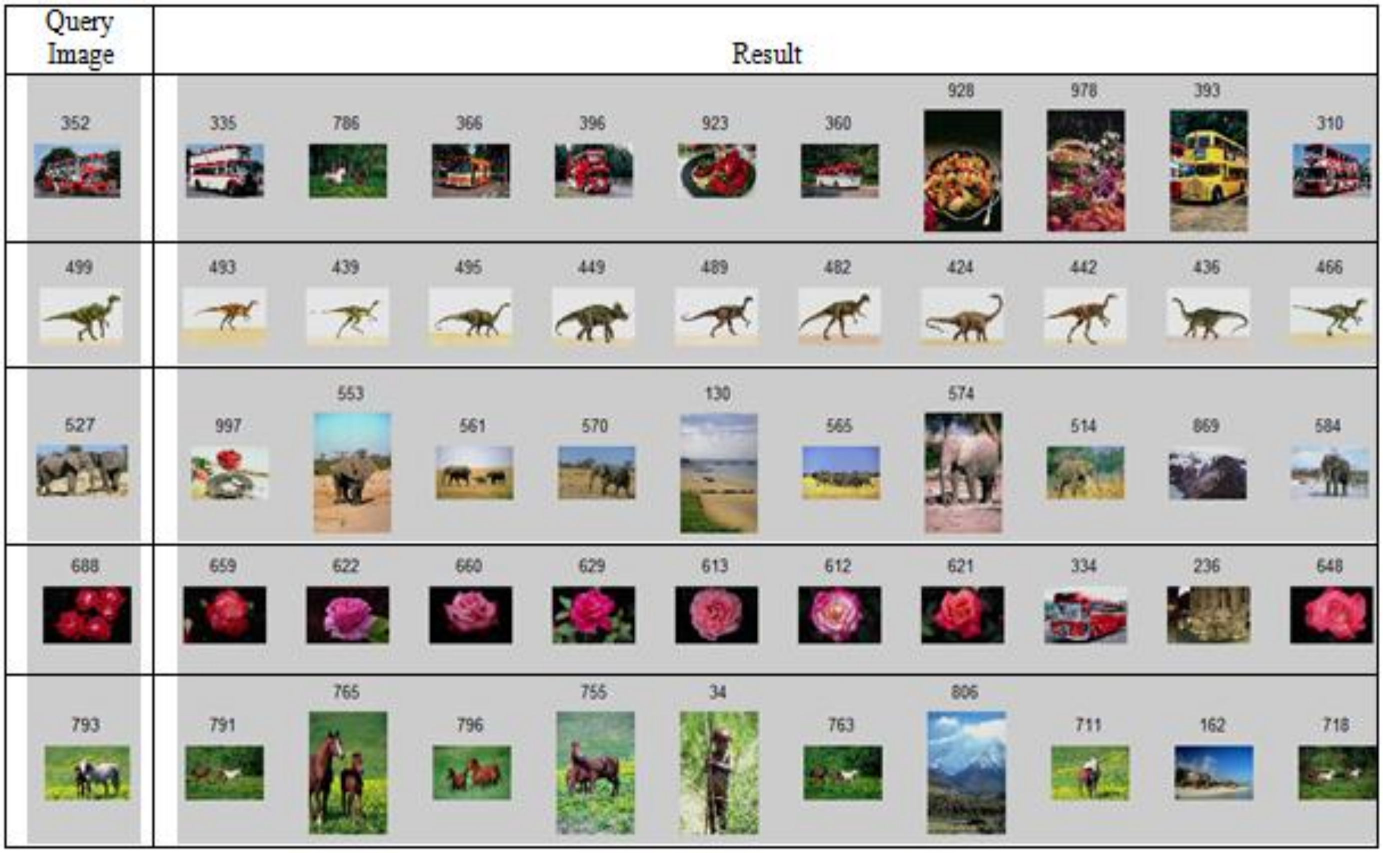


Figure 7. Example results for the bus, dinosaur, elephant, flower and horses as a query, respectively.

The experimental results with the world widely used images for evaluation of image retrieval performance shows that the proposed method is superior to the other conventional method in terms of retrieving accuracy.

References

[1] Kherfi, M., Brahmi, D. and Ziou D. "Combining Visual Features with Semantics for a More Effective Image Retrieval". ICPR ’04, vol. 2, 2004. pp. 961–964.

[2] H. Yu, M. Li, H.-J. Zhang and Feng, J. "Color texture moments for content-based image retrieval". International Conference on Image Processing, 2002. pp. 24–28.

[3] Hui yu, Mingjing Li, Hong-Jiang Zhang, and Jufu Feng. "Color Texture Moments For Content Based Image Retrieval".

[4] N. Sebe, Q. Tian, E. Loupias, M. Lew and T. Huang. "Evaluation of salient point techniques". International Conference, 2004.

[5] I.Daubechies. "Ten lecturer on wavelet". Philadelphia, PA:Sosiety for Industrial and Applied Mathematics Analysis, vol. 23, Nov. 1992. pp. 1544–1576.

[6] Stephane Mallet. "Wavelets for a Vision". Proceeding to the IEEE, Vol. 84, 1996. pp. 604-685.

[7] Kohei Arai and Y.Yamada. "Image retrieval method based on hue information and wavelet description based shape information as well as texture information of the objects extracted with dyadic wavelet transformation". Proceedings of the 11th Asian Symposium on Visualization, NIIGATA, JAPAN, 2011.

[8] http://wang.ist.psu.edu.

[9] J. Li, J.Z. Wang, and G. Wiederhold. "IRM: Integrated Region Matching for Image Retrieval". Proc. of the 8th ACM Int. Conf. on Multimedia, Oct. 2000. pp. 147-156.

[10] Hiremath P.S and Jagadeesh Pujari. "Content Based Image Retrieval using Color Boosted Salient Points and Shape features of an image". International Journal of Image Processing (IJIP) Vol.2, Issue 1, January-February 2008. pp. 10-17.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission A of ICSU/COSPAR since 2008. He wrote 30 books and published 307 journal papers

Cahya Rahmad, He received BS from Brawijaya University Indonesia in 1998 and MS degrees from Informatics engineering at Tenth of November Institute of Technology Surabaya Indonesia in 2005. He is a lecturer in The State Polytechnic of Malang Since 2005 also a doctoral student at Saga University Japan Since 2010. His interest researches are image processing, data mining and patterns recognition.

Analysis Method of Traffic Congestion Degree Based on Spatio-Temporal Simulation

Shulin He
Department of Management
Liaoning police Academy
Dalian 116036, China

***Abstract*—The purpose of this research is to design and implement a road traffic congestion and traffic patterns simulation (TPS) model and integrate it with extension-information model (EIM). The problems of road traffic simulation and control are studied according to the method of extension information model, and from the spatio-temporal analysis point of view. The rules of the traffic simulation from existence to evolution are analyzed using theories. Based on this study, the concept of traffic system entropy is introduced, and resulted in the establishment of a fundamental frame work for the road traffic simulation system based on extension spatio-temporal information system. Moreover, a practicable methodology is presented.**

Keywords- road traffic simulation; extension information model; degree of traffic congestion; traffic congestion entropy; traffic congestion control system.

I. Introduction

Traffic congestion has been a major problem on roads around the world for many years. In modern cities, traffic on major roads is abundant, and steps have to be taken to keep the traffic flowing at an acceptable speed. The volume of road traffic has increased rapidly in recent years. In the Dalian, total road traffic has almost doubled since 2001 (see Figure 2.1). Forecasts from the Department for Police show that the volume of road traffic will continue to increase at an alarming rate. These forecasts, which in the past have been conservative estimates, suggest that traffic levels will increase by approximately 50% between the years 2010 and 2020. If this is the case, then the causes and effects of traffic congestion need to be understood now or it is could become much worse a problem in the near future.

The purpose of this research is to design and implement a road traffic patterns and traffic congestion degree simulation (TPS) model and integrate it with extension-information model (EIM). Under current computer software development conditions, it is rather difficult to implement TPS, or other spatial-temporal based complex simulation models. Different approaches have been explored to build spatial-temporal simulation models of traffic system. Because of this technical barrier, spatial-temporal simulation modelers have to spend more time on technical issues, which complicates the application of classics model of traffic system and other spatial simulation theories. This research developed a dynamic traffic pattern simulation model from a static road traffic model [1].

This research has chosen commercial property robbery simulation as an example for integrating an extension spatial-temporal simulation model with EIM. The simulation model applied EIM to the traffic congestion likelihood evaluation formula from routine traffic activity theory (RAT). The simulation process runs through much iteration, each generating some individual traffic congestion. The accumulation of individual congestion reveals traffic patterns in space and time.

Another reason for choosing traffic congestion pattern simulation is to expand SP's application to invisible spatial-temporal processes. Most SP applications have occurred in ecology, urban planning and environment studies. One of the SP core elements is state variable. State variables represent the status of cells, which are the focus of the modeling. Land use type and number of pollutant particle are examples of state variable. Compared with visible phenomena, it is more difficult to simulate spatial-temporal changes of invisible phenomena with SP modeling. To simulation traffic congestion pattern developing process from a micro-level, first we need to find out the invisible phenomenon which broadcast over space and time, and set it as state variable. The other variables can then be related to the state variable directly or indirectly.

This paper is organized as follows: Section 2 introduces some basic concepts about traffic congestion and system simulation. Section 3 presents our new simulation model of traffic congestion degree, and section 4 explains our proposal for traffic congestion control by self-organizing theory. Finally, section 5 concludes the paper and draws some future work.

II. Traffic congestion and System Simulation

A. Traffic congestion

Physicists have been trying to describe the phenomena of traffic for at least half a century. In the 1950s, James Lighthill, an expert on the physics of fluid flow, suggested that the flow of traffic on a road was akin to the flow of liquid in a pipe. This theory (the Lighthill-Whitham-Richards model) represented the flow of traffic entirely with mathematical equations, and ignored the individual drivers. This sort of model is called macroscopic, and can often produce realistic output, but lacks the complexity to model realistic driver behaviours [2].

The complexity and uncertainty of traffic congestion make an ideal illustration of such research. Traffic congestion patterns develop over time period as a result of interactions between target and congestion over space. At its most basic, congestion is caused when the volume of traffic exceeds road capacity. This holds for most perceived causes of congestion; for example, accidents, breakdowns and road works decrease the available road capacity, while school-run and holiday traffic increase the volume of traffic. The next approach was to treat vehicles as individual units instead of a continuous flow, and see what behaviour emerges when the vehicles are given simple rules to follow. Each vehicle would move according to the vehicle ahead, speeding up or slowing down to match its speed while maintaining a safe distance between cars. This is a type of microscopic model, which can vary in complexity depending on the aims of the simulation. One well-known model is a cellular automata model designed by Nagel and Schreckenberg [5]. It was very simplistic and followed mainly the rules above, yet exhibits complex phenomena found in real traffic, as described below.

The results from these models and from traffic studies show that flow rate and traffic density are linked in an interesting way. Normally, flow rate increases as density increases, that is, more vehicles are on the road without any having to slow down. However, when the density reaches a so-called ‘critical density’, the flow rate begins to decrease and the traffic becomes congested. An interesting observation is a hysteresis effect that as the density increases above the critical density it is possible for the flow to continue to increase in a met stable or bi-stable state. In this state, any hiccup in the flow can cause the traffic to become congested [3].

B. *System Simulation*

We can find from the whole procedure of traffic simulation, bring forward traffic system property’s all kinds of believes basing on the information of traffic spatial-temporal. And then subjunctive prisoner and quality’s possibility basing on a variety of suppose. Recurrent subjunctive procedure may form constringency. Constringency fixes traffic system information. This traffic system information is called for proof. Because of it, traffic system law can be obtaining. Obtain depend on themselves feeling and recognize abilities to get traffic system information from all the directions. If in this conditions, obtain works can be done; otherwise, it will be shrivel. In reality, the work that finding traffic system information from expressional information is very difficult, mainly reason is that we can’t determine its resolution frame and information environment. Only depend on brain to finish quantity information development and constringe, to get to the purpose is very difficult.

To improve function and benefit of traffic system management and to build subjunctive traffic system model (TSM) is necessary. First definitude, TSM is tools dealing with traffic information in extension information space. TSM of extension information space is concurrence system that persons and computer are composed. The special complicacy and uncertainty of traffic system question are no having traffic’s internal rules. For example, traffic congestion when happened, where, what form, why, what change? These are named for 5CW problem. It is over the barrel basing on mathematics in hand for the question. The mainly content of traffic information reality system consisted of 5CW, its possibility space with mapping and inversion is consisting of element of crime simulation system.

In traffic system simulation, traditional ways are imitating inexact and uncertainty problem by building mathematics model, and then solve them by statistics and probability. But the model need a lot of believe and approximate, so at last, the model is different from reality largely. Obviously, traditional quantitative ways haven’t satisfied with needs. For being short of information and unstructured, it isn’t possible that building exact mathematical model; at the same time, decision aim of the question is miserable, so it is not necessary to building exact mathematics model. One practical way is building some qualitative models to analysis qualitatively. So make out some benefit analysis result, hand farthest knowledge’s effect. Canon computer is information conduct system building on number arithmetic. Therefore, in traffic reality system researching, for deleting expression and model from computer, usually building definite believe and data’s shortcut. In fact, those believes of subjunctive system aren’t being at all. So we need an information digging technology what is fit for reality question.

Resolution for traffic management system expressed that traffic information is a simulation system of time and space. But information’s unbending and asymmetric are formed by covert order of traffic properties causal order. Extension causal order searched relation order of causal order. And then it changed covert order into discoverable order in the traffic information, extended traffic expression information. We can get act information from expression information. Traffic system will reach to the purpose imitating and simulating under the causal mechanism. For example, in subjunctive traffic reality system, when computer gives us crime expressions, subjunctive system will appear causal relation images of traffic system properties. And then we can get traffic state expression information’s relation information by alternant feeling procedure of person with computer.

III. TRAFFIC SIMULATION BASED ON EXTENSION INFORMATION MODEL

A. *Methodologies*

As an introduction to traffic simulation analysis, this section provides the definition of traffic simulation analysis as a general concept as well as definition of four types of traffic analysis. These definitions are meant to enhance the understanding of traffic system simulation analysis and to help create commonly understood terminology, concepts and ideas in the field of traffic system analysis.

The quantitative and qualitative studies of traffic system and law enforcement information in combination with social demographic and spatial factors to apprehend traffic, prevent congestion, reduce accident, and evaluate organizational procedure. From the definition, a number of data are required in analyzing traffic system so as to come up with informed decisions in the apprehension of traffic and planning. In order to understand the definition of system analysis, major phrases and terms used are defined and discussed in detail in the following [5,6].

Traffic system simulation analysis uses both qualitative and quantitative data and it also uses analytical techniques. Qualitative data and analytical techniques refer to non-numerical data as well as the examination and interpretation of observations for the purpose of discovering underlying meanings and pattern of relationships. Quantitative data are data primarily in numerical or categorical format. Quantitative analysis consists of manipulations of observations for the purpose of describing and explaining the phenomena that those observations reflect and is primarily statistical.

Traffic system simulation analysis employs both types of data and techniques depending on the analysis and practical need. The information such as date, time, location, and type of traffic congestion is quantitative in that statistics can be used to analyze these variables. On the other hand, narrative of traffic congestion information are considered qualitative data in that a large number of narratives are nearly impossible to analyze statistically and are primarily examined to determine general themes and patterns. Three major elements of traffic simulation analysis emerge from the definition and these are traffic risk, spatial and temporal data as shown in table1 below.

TABLE1. MAJOR FACTORS OF TRAFFIC SIMULATION

	Factor		
Phase	Human	Vehicle	Environment
Pre-crash	Information Attitudes Impairment Police enforcement	Roadworthiness lighting Braking handling Speed management	Road design Road layout Speed limits Pedestrian facilities
Crash	Use of restrains Impairment	Occupant restraints Other safety devices Crash-protective design	Forgiving roadside
Post-crash	First-aid skill Access to hospital	Ease of access Fire risk	Rescue facilities Congestion

Traffic simulation analysis is performed for different purposes and because of this, it has been sub divided into different categories which have been given specific names for the purpose. The following are four type of analysis that fall under the umbrella of traffic simulation analysis. Each contains characteristics of traffic simulation analysis in general, but each is specific in the type of data and simulation analysis used as well as in its purpose.

Traffic congestion is an extremely complicated social phenomenon, and has the features of random mutation according to its occurrence, development and the trend of its evolution. From the viewpoint of government management, society security and traffic accident prevention is the aim of traffic management. According to the existing concept on traffic system control, the control process is determined by three factors:

(1) Determine the possible space and time of traffic congestion;

(2) select some states from the possible space and time as targets;

(3) Create the necessary conditions to make the traffic system control reach the preset aim.

It is known that the term “possible space” is the assembly of all the possibilities faced in the development process of an object. The possible space of traffic congestion is determined by the conditions leading to a congestion case. These factors have their characteristic possible spaces and time, differing from each other in the amounts and the forms, and may interchange from one to another. When a possible space and time of a congestion case is developing into a certain state, it may turn into a new possible space and time.

The occurrence of several possible spaces and times in the development of traffic congestion makes the traffic process appearing in different stages. In other words, the target of traffic control changes as the possible congestion space varies.

B. Simulation model of traffic congestion degree

The point of quantitative traffic study lies in the traffic congestion situation relationship structure. It establishes situation model and relationship model. By combining them together, we get relationship structure analysis of practical traffic congestion.

The form is that we use the mathematical model coming from the quantitative analysis to get the key, then based on which we make qualitative inference [7,8]. At the same time, according to the relationship structure set mapping which gets from qualitative inference, we establish mathematical model through all kinds of set mapping space hypothesis. Two results will be fed back through self-organized relationship structure and finally form a feasible result. If road traffic system is made up of variables C_1, C_2,…,$C_{n,,}$ and there are controlling parameters K_1, K_2,…,K_m, then traffic congestion dynamic equation is described as the following[6]:

$$\begin{aligned}
&\frac{dC_1}{dt} = f_1(\lambda,,C_1,C_2,\Lambda\ ,C_n) \\
&\frac{dC_2}{dt} = f_2(\lambda,,C_1,C_2,\Lambda\ ,C_n) \qquad (1) \\
&\Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda \\
&\frac{dC_n}{dt} = f_n(\lambda,,C_1,C_2\,\Lambda\ ,C_n)
\end{aligned}$$

In road traffic system, congestion situation variable is the key factor to form traffic system. The effective choice of situation variable is critical to reflect the real level of traffic management and describe social stability and the degree of the social development. The economic growth and consumption level reflect traffic system development. In economy study, if we lack the analysis of crime factors, our conclusion of the economy analysis will not be reliable. Among the former study of road traffic, there are many factors constructing traffic system, such as economy factor, education factor, social ethos, law factor, people relationship and management factor. Because these factors lie in different situations, the characteristics of traffic system are different. The evaluation of social stability and development depends on the choice of controlling parameter. For example unemployment rate, relative number of income difference, controlling proportion of informal social groups, management efficiency of road traffic and education quality.

$$\frac{dC_i}{dt} = f_i(\{C_j\},\{k_a\}) \quad (\ i,j\text{=1,2,…,}n\ a\text{=1,2,…,}m\ ) \qquad (2)$$

Suppose equation (2) is self-organization power traffic system, this system acts with a sudden change that has the statistical characteristics. But traffic action is invisible. The asymmetry of information makes it uncertain in the economic loss and investment when we solve cases and control traffic system. Therefore, the statistical prediction in traffic congestion degree and efficient estimate in solving traffic congestion cases are very important in macroeconomic analysis and growth, together with in the control the traffic stability.

Traffic congestion case is looked as the degree of congestion $n(t)$, which is decided by time (t)'s function. It happens randomly. Because congestion happening and solving congestion cases are two variables, and the degree of congestion is decided by these variables, we suppose expectation of congestion degree $E\{n(t)\} = n(t)$, then

$$E[n(t)] = N(t) = \sum_{n\to 0}^{\infty} nP_n(t)$$

In the formula mentioned above, $P_n(t)$ indicates that time (t) has n congestion rate. Above all, by instinct, the changing rate of the expectation congestion degree, whose rate changes with time, is $dN(t)/dt$. During the time (t), this happening rate explains the new congestion and repeated congestion. So in a certain area, the whole expected rate of congestion degree is that average rate $\lambda(t)$ counted on vehicle multiplies the expected congestion degree. If $\mu(t)$ is average rate of control degree of traffic congestion at time t, $\mu(t)\ n(t)$ is the general expected rate. So

$$\frac{dN(t)}{dt} = [\lambda(t) - \mu(t)]N(t) \qquad (3)$$

This is called simulation model of traffic congestion degree. By analyzing the congestion attribute in certain area, if the traffic congestion doesn't transfer between inside and outside, i.e. specific rate of congestion degree and cracking rate doesn't change with time, then $\lambda(t) = \lambda$, $\mu(t) = \mu$, formula (3) will change into

$$\frac{dN(t)}{dt} = (\lambda - \mu)N(t) \qquad (4)$$

In formula (4) result could be validated by replacing.

$$N(t) = N(0)e^{(\lambda-\mu)t} \qquad (5)$$

In the formula, $N(0)$) is congestion degree when time is $\lambda - \mu = 0$. When $\lambda - \mu$ >0, it explains that practical congestion cracking is less than practical congestion degree. From study of recessive situation, we know that the number that people didn't report the situation and traffic congestion haven't been found is great many. So controlling aim of traffic system is $\lambda - \mu \to$ 0. We find from above analysis that the indication of congestion degree has great effect on social development, especially on economy. If there is no traffic economy dynamic analysis in macroscopic traffic analysis, macroscopic traffic analysis will be half-baked. Here is congestion number interval [$N_{\min}$, $N_{\max}$]; the degree of congestion situation happening is controlled among interval. Supposing two trends, $N(t) \to N_{\min}$ and $N(t) \to N_{\max}$, traffic congestion controlling degrees k_α and k_β are replaced by $\lambda(t) - \mu(t)$ then Getting result of differential equation.

$$k_\alpha = [\lambda(t) - \mu(t)]_\alpha = [1 - N(t) \big/ N_{\min}](\lambda - \mu)$$

$$k_\beta = [\lambda(t) - \mu(t)]_\beta = [1 - N(t) \big/ N_{\max}](\lambda - \mu)$$

$$[N(t)]_\alpha = \frac{N_{\min} e^{(\lambda-\mu)t}}{(N_{\min} \big/ N(0)) - 1 + e^{(\lambda-\mu)t}},$$

$$[N(t)]_\beta = \frac{N_{\max} e^{(\lambda-\mu)t}}{(N_{\max} \big/ N(0)) - 1 + e^{(\lambda-\mu)t}}$$

Here supposing that the degree of traffic congestion controlling is chosen randomly between k_α and k_β and. logarithm [k_α , k_β] is the ideal key to traffic congestion controlling. It can be inferred from the following. Formula (6) shows the general number characteristics of traffic congestion controlling level over a certain period of time [9].

$$[N(t)]_c = \int_{t_0}^{t_n} \frac{[N(t)]_\alpha + [N(t)]_\beta}{2} dt$$

$$= \int_{t_0}^{t_n} [[\frac{N_{\min} e^{(\lambda-\mu)t}}{({N_{\min}}/{N(0)}) - 1 + e^{(\lambda-\mu)t}} + \frac{N_{\max} e^{(\lambda-\mu)t}}{({N_{\max}}/{N(0)}) - 1 + e^{(\lambda-\mu)t}}]]2^{-1} dt \qquad (6)$$

Through the above discussion we know that the key point determining [$N(t)$] is to get $\lambda(t)$ and $\mu(t)$, and the key point getting result is to establish equation (1) and (3). Establishing equation (1) is decided by analysis of congestion situation variable and equation (3)'s situation system structure. At the same time, congestion statistics and the methods of traffic information's measurement is very important.

IV. Traffic Congestion Control

A. Non-equilibrium traffic control system

In fact, the simulation system of traffic congestion is a self-organization system based on non-equilibrium system theory. Using non-equilibrium system theory to study the problems of traffic control is relatively new. Traffic congestion is a common social phenomenon. Traffic congestion prevention and control originate as congestion occur. The extent of the social stability is dependent on what level the traffic congestion is under control. Reducing the traffic congestion to the lowest extent means that the society advancing.

Philosophically, equilibrium exists temporarily and relatively. Non-equilibrium exists commonly and absolutely. The equilibrium achieved in the traffic congestion control system refers to the unified behavior of the control activity under the predictive traffic model. This is a static point of view for the traffic and their control. Equilibrium can be understood as a process of adapting with the traffic congestion and adjusting control, from the kinetic point of view. This reflects the characters of the non-equilibrium process. Since the traffic congestion appears to be kinetic in nature as the society is developing, traffic control system should be a non-equilibrium process.

However, the role of the traffic congestion information statistics should not be exaggerated. It has to be pointed out that the objectives processed by the statistics should be independent incidences and large numbers and have a random feature. Strictly speaking, the individual cases in a macroscopic traffic study do not fully fulfil the above mentioned assumptions. The statistics used before are not accurate and complete enough in describing the society traffic problem as a whole. Therefore, the needs are rising for a better theory to describe the social traffic thoroughly, and to take both microscopic and macroscopic viewpoints into consideration. From the view point of non-equilibrium system theory, the combination of statistics with kinetics should be useful in solving the problem mentioned before. The cooperation theory applies statistics for both microscopic and macroscopic investigations for the objectives under study.

The needs for a large number of objectives are no long a necessity. It is statistically meaningful on one hand, and it emphasizes the interactions in the social traffic on the other. Synergetic can be used to reveal the rules of the evolution of the traffic problems from the point of time, space and forms. Therefore, the social traffic problems can be described in a stricter and complete ways according to this theory, and the effective crime control can be found with the help of this comprehensive statistical method. The evolution of the traffic congestion problems is partly dependent on the changes of the traffic environment, and partly dependent on the control means and conditions of the government. According to the analysis of the mutation model, it is clear that the traffic congestion state model for studying the evolution pattern can be established by finding out the state variables representing the quantitative change, and the condition variables (i.e. controlling variables) which cause the change of the state of the crime-affected state in the social development.

For the analysis of the traffic congestion problem, the same principles apply. There are a number of factors, which affect the traffic congestion variations. The difficulty is to find out the dominant variables. All the traffic congestion variables can be divided into two groups: direct variables (directly affect the occurrence of traffic congestion). Once the relationship between the direct variables and the state variables is known, the mathematics model can be obtained, and the quantitative analysis can be carried out for the rules of the changes of the traffic state.

B. The Entropy Characteristic of the traffic congestion system

The traffic congestion system is non-equilibrium, so its feature is decided by the degree of the system disorder, which is determined by proper measurement. In physics, entropy denotes the amount of system probability, which quantitatively describes irreversible process. For traffic congestion system, any congestion situation couldn't happen again. It cannot restore. The macroscopic situation of traffic congestion system corresponds to many microcosmic traffic state. The more they correspond, the larger uncertainty of traffic system has. Thereby, it reflects the disorder of traffic system. It's obvious in traffic congestion problems. Because traffic action is complicated and changeable, and traffic action form and way are so various that we cannot predict. The characteristics in microcosmic traffic action situation can decide macroscopic congestion index of traffic system. Traffic system disorder can be measured by all kinds of congestion degree probability. Congestion entropy shows how much probability of each type of traffic action happens in traffic system. Supposing λ_c is happening rate of congestion degree of macroscopic traffic system, and k_c is traffic macroscopic controlling constant, then

$$S^c = k_c \sum_{j=1}^{n} \ln \lambda_c^j \qquad (7)$$

S^C is traffic congestion entropy, which is decided by adding trait and controlling constant of macroscopic traffic system λ_c .

In fact, traffic system is an open system. It shows that population flowing makes vehicle move between region inside and outside. Supposing the population in a region is *M*, in *M* there are M_{ic} people who probably commit a traffic action, then $dS_i = M_{ic} / M$ shows the change of inner entropy in traffic system. At the same time, due to the population flowing, the probability of traffic has changed, resulting in the change of exterior entropy. $dS_o = M_{oc} / M \quad M_{oc}$ is the number of flowing population in a region, and M_{oc} is the number of traffic congestion probably happening among the flowing population. Then the change of entropy in traffic system is described as the following [9].

$$dS^c = dS_i^c + dS_o^c$$

$$d[k_c \sum_{j=1}^{n} \ln \lambda_c^j] = \frac{M_{ic}}{M} + \frac{M_{oc}}{M} \qquad (8)$$

$$\prod_{j=1}^{l} \lambda_C^L = e^{\frac{1}{k_c}\int_T (\frac{m_{ic}}{m} + \frac{m_{oc}}{m})_c \, dt} \qquad (9)$$

$$(l \in (1,2,\Lambda \ , n))$$

The general expected number of a regional traffic congestion doesn’t always happen which must be combined with the rate of the congestion degree. We study the law of traffic system according to extension information changing model based on non-equilibrium system theory. At the same time, we will know how macroscopic traffic congestion control system affects social and economic system based on traffic congestion level and it’s solving control index, together with changes of order parameter.

V. Conclusions

In this paper we have extended and improved a previous model for evaluating traffic congestion using an extension information model and applying a self-organizing methodology. We are improving function and benefit of traffic simulation and to build extension traffic simulation system (ETSS) is necessary.

First definitude, ETSS is tools dealing with information in extension information space. ETSS of extension information space is concurrence system that persons and computer are composed. The special complicacy and uncertainty of road traffic question are no having congestion’s internal rules. For example, traffic congestion when happened, where, how, why, what change? These are named for 5CW problem. It is over the barrel basing on mathematics in hand for the question. The mainly content of road traffic information reality system consisted of 5CW, its possibility space with mapping and inversion is consisting of element of road traffic simulation system.

In future research, by incorporating the information variables involved in traffic congestion management such as the numbers of accidents, traffic information violations, and traffic policemen on duty into an artificial intelligence technique, it is possible to build a traffic decision making system to help decision makers for analysis of traffic laws and policies.

References

[1] Robert R，Theodore F. Contrasting the Use of Time-Based and Distance-Based Measures to Quantify Traffic Congestion Levels：An Analysis of New Jersey Counties. The 81th Annual Meetings of the Transportation Research Board, Washington，D.C, 2002.

[2] He Ping. Research on the Quantity Analysis of Social Crime. Journal of Liaoning Police Academy, 2004, vol. 37, pp.1-6.

[3] Kent Hymely. “Does Traffic Congestion Reduce Employment Growth?” Journal of Urban Economics, 2009,vol. 65/2, pp. 127-135.

[4] Kleijnen, J. P. C. Response surface methodology for constrained simulation optimization: An overview, Simulation Modelling Practice and Theory, 2008, vol.16, pp.50–64.

[5] Kleijnen, J. P. C., van Beers, W. and van Nieuwenhuyse, I. Constrained optimization in expensive simulation: Novel approach, European Journal of Operational Research, 2010, vol.202, pp.164–174.

[6] Marti, K. Stochastic optimization methods, Springer, Berlin. Mor´e, J. and Wild, S. (2009). Benchmarking derivative-free optimization algorithms, SIAM Journal on Optimization, 2009, vol.20, pp.172–191.

[7] Oeuvray, R. and Bierlaire, M. Boosters: a derivative-free algorithm based on radial basis functions, International Journal of Modelling and Simulation, 2009, vol. 29,pp.26–36.

[8] Osorio, C. Mitigating network congestion: analytical models, optimization methods and their applications, PhD thesis, Ecole Polytechnique F´ed´erale de Lausanne. Osorio, C. and Bierlaire, M. (2009a). An analytic finite capacity queueing network model capturing thepropagation of congestion and blocking, European Journal Of Operational Research, 2010, vol. 196, pp.996–1007.

Authors Profile

Shulin He was born in 1962, in Liaoyang, China. He is an assistant professor of the management department of the Liaoning Police Academy, Dalian, China. His academic interests are found in traffic management theory and applications, including uncertain traffic system analysis. He can be reached at e-mail: heshl888597@yahoo.com.cn.

Separability Detection Cooperative Particle Swarm Optimizer based on Covariance Matrix Adaptation

Sheng-Fuu Lin, Yi-Chang Cheng, Jyun-Wei Chang, and Pei-Chia Hung
Department of Electrical Engineering
National Chiao Tung University
Hsinchu, Taiwan

***Abstract*—The particle swarm optimizer (PSO) is a population-based optimization technique that can be widely utilized to many applications. The cooperative particle swarm optimization (CPSO) applies cooperative behavior to improve the PSO on finding the global optimum in a high-dimensional space. This is achieved by employing multiple swarms to partition the search space. However, independent changes made by different swarms on correlated variables will deteriorate the performance of the algorithm. This paper proposes a separability detection approach based on covariance matrix adaptation to find non-separable variables so that they can previously be placed into the same swarm to address the difficulty that the original CPSO encounters.**

Keywords- cooperative behavior; particle swarm optimization; covariance matrix adaptation; separability.

I. INTRODUCTION

The particle swarm optimizer (PSO) [1, 2] is a stochastic, population-based optimization learning algorithm. Its learning procedure is based on a population made of individuals with specific behaviors similar to certain biological phenomena. Individuals keep exploring the solution space and exploiting information between individuals while evolution proceeding. In general, by means of exploring and exploiting, the PSO is less likely to be trapped at the local optimum.

As with many stochastic optimization algorithms [1, 3-6], the PSO suffers from the "curse of dimensionality," which implies that its performance deteriorates as the dimensionality of the search space increases. To cope with this difficulty, Potter [3] proposed a cooperative coevolutionary genetic algorithm (CCGA) that partitions the search space by splitting the solution vectors into smaller ones. The mechanism proposed by Potter significantly improves the performance of the original GA. Van den Bergh [5] applies this technique to the PSO and presented several cooperative PSO models named CPSOs. In the CPSOs learning procedure, the search space can be arbitrarily partitioned into different number of subspaces. Each smaller search space is then searched by a separate swarm. The fitness function is evaluated by the context vector, which means the concatenation of particles found by each of the swarms. However, as with the CCGA algorithm, the performance of the CPSO deteriorates when correlated variables are placed into separate populations. In this paper, we call such variables "non-separable." A function f is said to be separable if

$$\begin{aligned}&\arg\min_{(x_1,\ldots,x_n)} f(x_1,\ldots,x_n)\\&=(\arg\min_{x_1} f(x_1,\ldots),\ldots,\arg\min_{x_n} f(\ldots,x_n))\end{aligned}, \tag{1}$$

and it is followed by a fact that f can be optimized in a sequence of n independent 1-D optimization processes. This paper proposes a variation on the original CPSO to detect the separability of the variables. To this end, we adopt a mechanism from evolution strategy with covariance matrix adaption (CMA-ES) [8, 9]. The performance of the CPSO after applying separability detection is compared with that of the traditional PSO and CPSO algorithm.

This paper is organized as follows. Section II presents an overview of the PSO and the CPSO. In section III, we describe the proposed separability detection cooperative particle swarm optimizer (SD-CPSO). This is followed by the experiment results presented in section VI. Finally, some directions for the future research are discussed in section V.

II. RELATED WORKS

The PSO is first introduced by Kennedy and Eberhart. It's one of the most powerful methods for solving global optimization problems. The algorithm searches an optimal point in a multi-dimensional space by adjusting the trajectories of its particles. The individual particle updates its position and velocity based on its personal best performance and the global best performance among all particles that denote y and respectively. The position $x_{i,d}$ and velocity $v_{i,d}$ of the d-th dimension of i-th particle are updated as follows:

$$\begin{aligned}v_{i,d}(t+1)&=v_{i,d}(t)+c_1\cdot rand_1\cdot(y_{i,d}(t)-x_{i,d}(t))\\&\quad+c_2\cdot rand_2\cdot(\hat{y}_d(t)-x_{i,d}(t)),\end{aligned} \tag{2}$$

$$x_i(t+1)=x_i(t)+v_i(t+1)\,, \tag{3}$$

where y_i represents the previous best position yielding the best performance of the i-th particle; c_1 and c_2 denote the acceleration constants describing the weighting of each particle been pulled toward y and $\hat{y}$ respectively; $rand_1$ and $rand_2$ are two random numbers in the range [0, 1].

Let s denote the swarm size and f() denote the fitness function evaluating the performance yielded by a particle. After (2) and (3) are executed, the personal best position y of each particle is updated as follows:

$$y_i(t+1) = \begin{cases} y_i(t+1), \text{ if } f(x_i(t+1)) \geq f(y_i(t)), \\ x_i(t+1), \text{ if } f(x_i(t+1)) < f(y_i(t)), \end{cases} \qquad (4)$$

and the global best position is found by:

$$\hat{y}(t+1) = \arg\min_{y_i} f(y_i(t+1)), \quad 1 \leq i \leq s \,. \qquad (5)$$

The CPSO [5, 6] is one of the most significant improvements to the original PSO. Van den Bergh presented a family of CPSOs, including CPSO-S, CPSO-S$_K$, CPSO-H, CPSO-H$_K$. Algorithm CPSO-H$_K$ is the hybrid from PSO and CPSO-S$_K$ and it is proposed to address the issue of "pseudominima." A discussion of pseudominima is outside of the scope of this article. The objective of this article is to propose a self-organized technique to assist the CPSO-S$_K$ in finding how the components on a context vector be related.

The concept of CPSO-S is that instead of trying to find an optimal n-dimensional vector, the vector is split into n parts so that each of n swarms optimizes a 1-D vector. The CPSO-S$_K$ is a family of CPSO-S, where a vector is split into K parts rather than n, where $K \leq n$. K also represents the number of swarms. Each of the K swarms acts as a PSO optimizer (2)-(5). The main difference between the PSO and the CPSO is that the fitness of a single particle of the CPSO has to be evaluated through global best particles of the other swarms. Let P_j denote the j-th swarm and $P_j \cdot x_i$ represents the i-th particle in the swarm j. The fitness of $P_j.x_i$ is defined as:

$$f(P_j \cdot x_i) = f(P_1 \cdot \hat{y}, \ldots, P_{j-1} \cdot \hat{y}, P_j.x_i, \ldots, P_K \cdot \hat{y}). \qquad (6)$$

The CPSO applies cooperative behavior to improve the PSO on find the global optimum in a high-dimensional space. This is achieved by employing multiple swarms to explore the subspaces of the search space separately to reduce the curse of dimensionality. However, there is no absolute criterion that the CPSO is superior than the PSO since independent changes made by different swarms on correlated variables will deteriorate its performance . In addition, in one generation of a n-dim CPSO-S operation, the computational cost is n times larger than that of a PSO operation.

III. Methodlogy

This paper proposes an approach to help the CPSO self-organize the swarms composed of non-separable variables. Consider a particular optimization task illustrated in Fig. 1, from which we can see a 2-dim function with a bar-shaped local optimal region and a global optimum lies in it. The task is to find its global optimum by particle swarm optimizer. At first, particles are uniformly distributed in the search space. At this moment, we expect particles to be divided into two swarms, performing separate 1-dim PSO operation on each dimension to speed up the process of particles gathering around the optimal region.

If by any chance particles gather around the optimal region as we expected, as shown in Fig. 2. At this point of time, we prefer particles performing 2-dim PSO operation on the whole search space to reduce the computational cost, which, in this case, represents the number of function evaluations.

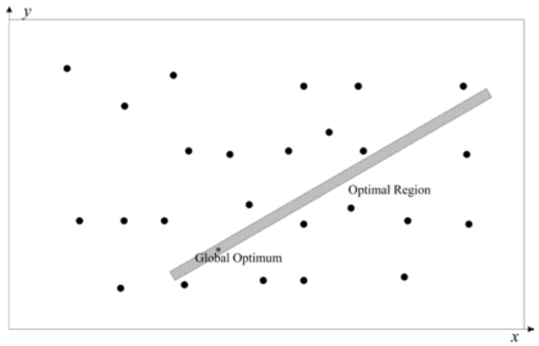


Figure 1. Case with particles uniformly distributed in the search space to find the global optimum lies in a bar-shaped local optimal region.

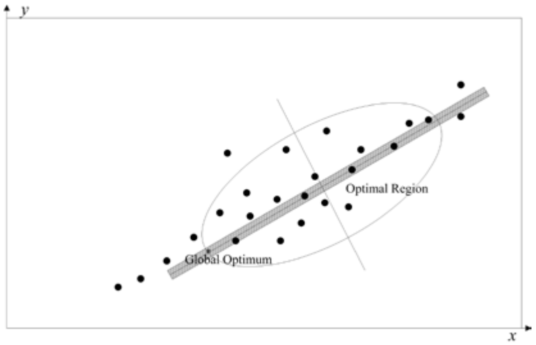


Figure 2. Case with particles gather around the bar-shaped optimal region to find the global optimum.

In order to implement the idea illustrated above, we have to determine the timing of switching between the PSO and the CPSO operation when dealing with a task. In this paper, we think this can be done by determining the separability between variables, and placing non-separable into the same swarm at each generation. If at certain moment, all variables are determined as non-separable, then the PSO operation is taken; otherwise, the CPSO operation is taken.

The separability between variables is found by estimating the covariance matrix of the distribution of particles. The method we adopt is called the covariance matrix adaptation proposed in [8, 9]. In the standard CMA-ES, a population of new search points is generated by sampling a multivariate normal distribution N with mean $m \in \mathbb{R}^n$ and covariance matrix $\mathbf{C} \in \mathbb{R}^{n\times n}$. The equation of sampling new search points, for each generation number g = 0, 1, 2, …, reads

$$x_i^{(g+1)} \sim m^{(g)} + \sigma^{(g)} N(0, \mathbf{C}^{(g)}) \quad \text{for } i = 1, \ldots, \lambda \,,$$
(7)

where ~ denotes the same distribution on the left and right hand side; $\sigma^{(g)}$ denotes the overall standard deviation, step-size, at generation g and λ is the sample size. The new mean $m^{(g+1)}$ of the search distribution is a weighted average of the μ selected points from λ samples $x_1^{(g+1)}, x_2^{(g+1)}, \ldots, x_\lambda^{(g+1)}$:

$$m^{(g+1)} = \sum_{i=1}^{\mu} w_i x_{i:\lambda}^{(g+1)} \,, \qquad (8)$$

with

$$\sum_{i=1}^{\mu} w_i = 1, \quad w_1 \ge w_2 \ge \cdots \ge w_1 > 0 ,$$

(9)

where w_i are positive weights, and $x_{i:\lambda}^{(g+1)}$ denotes the i-th rank individual out of λ samples from (8). The index i:λ denotes the i-th rank individual and

$$f(x_{1:\lambda}^{(g+1)}) \le f(x_{2:\lambda}^{(g+1)}) \le \cdots \le f(x_{\lambda:\lambda}^{(g+1)}) , \qquad (10)$$

where $f()$ is the objective function to be minimized. The adaption of new covariance matrix $\mathbf{C}^{(g+1)}$ is formed by a combination of rank-μ and rank-one update [10]

$$\mathbf{C}^{(g+1)} = (1-c_{\text{cov}})\mathbf{C}^{(g)} + \frac{c_{\text{cov}}}{\mu_{\text{cov}}} \underbrace{p_c^{(g+1)}\left(p_c^{(g+1)}\right)^T}_{\text{rank-one update}}$$

$$+ c_{\text{cov}}(1-\frac{1}{\mu_{\text{cov}}}) \times \underbrace{\sum_{i=1}^{\mu} w_i y_{i:\lambda}^{(g+1)}\left(y_{i:\lambda}^{(g+1)}\right)^T}_{\text{rank-}\mu\text{ update}} ,$$

(11)

where $\mu_{\text{cov}} \ge 1$ is the weighting between rank-μ update and rank-one update; $c_{\text{cov}} \in [0,1]$ is the learning rate for the covariance matrix update, and

$$y_{i:\lambda}^{(g+1)} = (x_{i:\lambda}^{(g+1)} - m^{(g)}) / \sigma^{(g)} ,$$

(12)

is a modified formula used to compute the estimated covariance matrix for the selected samples. The evolution path $p_c^{(g+1)}$ for rank-one update is described as follows:

$$p_c^{(g+1)} = (1-c_c)\, p_c^{(g)} + \sqrt{c_c(2-c_c)\mu_{\text{eff}}}\, \frac{m^{(g+1)} - m^{(g)}}{\sigma^{(g)}} ,$$

(13)

where $c_c \le 1$ denotes the backward time horizon and

$$\mu_{\text{eff}} = \left(\sum_{i=1}^{\mu} w_i^{\,2} \right)^{-1} , \qquad (14)$$

denotes the variance effective selection mass. The new step-size $\sigma^{(g+1)}$ is updated according to

$$\sigma^{(g+1)} = \sigma^{(g)} \exp\left(\frac{c_\sigma}{d_\sigma} \left(\frac{\left\| p_\sigma^{(g+1)} \right\|}{E\left\| N(0,\mathbf{I}) \right\|} - 1 \right) \right),$$

(15)

with

$$p_\sigma^{(g+1)} = (1-c_\sigma)\, p_\sigma^{(g)} + \sqrt{c_\sigma(2-c_\sigma)\mu_{\text{eff}}}\, \mathbf{C}^{(g)^{-1/2}} \frac{m^{(g+1)} - m^{(g)}}{\sigma^{(g)}} ,$$

(16)

where c_σ is the backward time horizon of evolution path, similar to c_c; d_σ is a damping parameter and $p_\sigma^{(g+1)}$ is the conjugate evolution path for step-size $\sigma^{(g+1)}$. The expectation of the Euclidean norm of a $N(0, \mathbf{I})$ reads

$$E\left\| N(0,\mathbf{I}) \right\| = \sqrt{2}\Gamma(\frac{n+1}{2}) / \Gamma(\frac{n}{2}) \approx \sqrt{n} + O(1/n) ,$$

(17)

where $O(\cdot)$ represents high-order terms.

Consider the estimated covariance matrix has the form shown as follows,

$$C = \begin{bmatrix} c_1^2 & c_{12} & \cdots & c_{1n} \\ \vdots & c_2^2 & \cdots & \vdots \\ \vdots & \vdots & & \vdots \\ c_{1n} & c_{2n} & \cdots & c_n^2 \end{bmatrix} , \qquad (18$$

)

where n is the number of dimensions, c_{jk} represents the weighted covariance between variables j and k. The separability between dimensions can be obtained from correlation coefficient matrix with its element defined as follows:

$$\rho_{jk} = c_{jk} \big/ c_j c_k ,$$

(19)

We define a parameter ρ_{thres} to determine whether dimension j and k are viewed as separable. If $\rho_{jk} < \rho_{\text{thres}}$ then we say variable j and k are separable. Conventionally, if |ρ|>0.8, it implies that there exists a very strong linear relationship between these two variables; 0.8>|ρ|>0.6 implies strong relationship, and 0.6>|ρ|>0.4 implies moderate relationship. So, in this paper, we avoid setting ρ_{thres} less than 0.6. The block diagram of the proposed method can be found in Fig. 3.

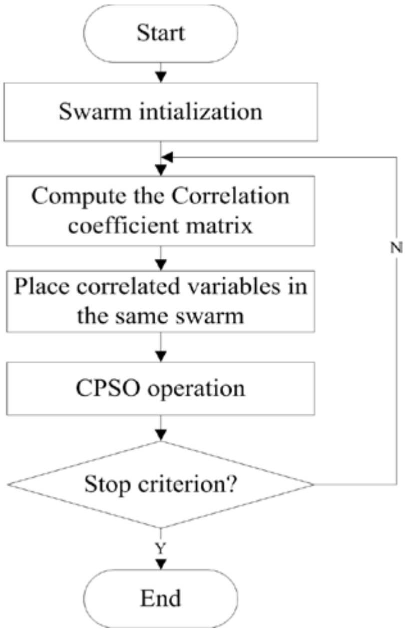

Start
Swarm intialization
Compute the Correlation coefficient matrix
Place correlated variables in the same swarm
CPSO operation
Stop criterion?
N
Y
End

Figure 3. Block diagram of SD-CPSO.

IV. Experiment Results

In order to compare the performance between different algorithms, a fair time measure must be selected. Here we use the number of function evaluations as a time measure following [5]. The performance of the proposed SD-CPSO is verified by real-parameter minimization tasks, which contains totally nine test functions. By their nature they can be divided into two parts: unimodal and multi-modal functions.

The first two functions are unimodal, followed by seven multimodal functions with three of them have simple global structures (single-funnel functions) and another four have complex global structures (multi-funnel functions). The difference between single- and multi-funnel functions can be illustrated by the following two figures, where Figure 4 shows a visualization of a 2-D Rastrigin's function, from which we can see that in spite of the large amount of local minima, there exists a trend to the global minimum. Figure 5 shows a visualization of a 2-D double Rastrigin's function, from which we can see that there are two funnel-type global trends and a large amount of noisy local minima.

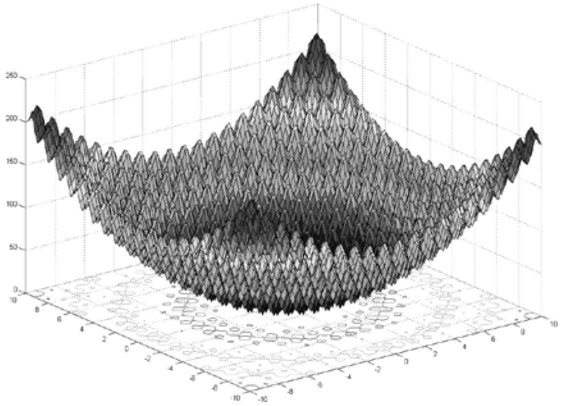


Figure 4. Visualization of a single-funnel, 2-D Rastrigin's function.

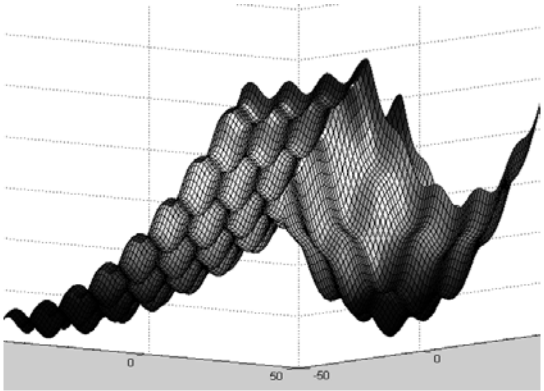


Figure 5. Visualization of a multi-funnel-funnel, 2-D double Rastrigin's function.

The types and names of functions are described in Table I. A detailed definition of test functions can be seen in [11, 12].

All functions are of 50 dimensions and have been adjusted to zero optimal solution respectively. To make sure that there was sufficient correlation between the variables, making it even harder for optimization, all the functions were further tested under 45 degree coordinate rotation.

In the following of this chapter, we will describe the configurations of the algorithms that we use to compare the performance with the proposed SD-CPSO in section 3.A. Experiment result and the discussion will be shown in section 3.B.

A. *Algorithms Configuration*

The three algorithms for comparison are listed as follows:

- PSO: the origin algorithm.
- CPSO-S: algorithm that splits swarm into each dimension.
- SD-CPSO: the proposed separability detection cooperative particle swarm optimization.

For each algorithm, experiments are executed for 50 times. Denote *n* the dimension of the optimization task and *s* the number of particles in one swarm. Parameters of the three algorithms are listed in Table II.

TABLE I. TYPE AND NAME OF THE TEST FUNCTION.

Unimodal Functions
F_1: Sphere Function
F_2: High Conditioned Ellipsoidal Function
Multimodal Functions
F_3: Rosenbrock Function
F_4: Rastrigin Function
F_5: Griewank Function
Multi-Funnel Functions
F_6: Schwefel Function
F_7: Double-Rastrigin Function
F_8: Weierstrass Function
F_9: Michalewicz Function

TABLE II. MS-CMA-ES AND CMA-ES PARAMETERS.

Parameters of Selection operator
$\lambda = 4 + \lfloor 3\ln n \rfloor$
$w_i = \frac{w_i^{'}}{\sum_{j=1}^{\lambda} w_j^{'}}, w_i^{'} = \begin{cases} \ln(\frac{\lambda}{2} + 0.5) - \ln i \text{ , for } i\text{=1,...}\left\lfloor \frac{\lambda}{2} \right\rfloor, \\ 0 \text{ , otherwise,} \end{cases}$
Parameters of Covariance adaptation:
c_{cov}=0.7
μ_{cov}=10
$c_c = \frac{4 + \mu_{\text{eff}} / n}{n + 4 + 2\mu_{\text{eff}} / n}$,
$c_\sigma = \frac{\mu_{\text{eff}} + 2}{n + \mu_{\text{eff}} + 5}$
$d_\sigma = 1 + 2\max\left(0 \text{ , } \sqrt{\frac{\mu_{\text{eff}} - 1}{n+1}} - 1\right) + c_\sigma$
Parameters of PSO operation:
c_1=c_2=1.49
ρ_{thres}=0.8
s=50

B. Experiment Result

This section presents optimization results. The number of maximum fitness calculation times, initial search range, initial search position and minimum fitness threshold are detailed in Table III. All particles are evenly distributed in the initial search range.

TABLE III. PARAMETERS OF THE EXPERIMENT.

	maximum fitness calculation times	Initial search range	Minimum fitness threshold
f_1	10000	$\mathbf{x}\in[0,100]^d$	1e-6
f_2	10000	$\mathbf{x}\in[0,100]^d$	1e-6
f_3	10000	$\mathbf{x}\in[0,100]^d$	1e-2
f_4	3000	$\mathbf{x}\in[0, 5]^d$	1e-2
f_5	8000	$\mathbf{x}\in[0,600]^d$	1e-2
f_6	4000	$\mathbf{x}\in[0,3]^d$	1e-2
f_7	2000	$\mathbf{x}\in[-20,20]^d$	1e-2
f_8	4000	$\mathbf{x}\in[0,0.5]^d$	1e-2
f_9	5000	$\mathbf{x}\in[0,5]^d$	1e-2

The experimental data is obtained by executing each 50 dimensional test functions until the stopping criterion is met. The procedure was repeated 50 times to compute the average fitness value. In the paper, instead of the actual numeric fitness value, the rank of the minimum average fitness value is defined as the standard of comparison. The reason is that we want to exclude the impact of the different degree of scale on the raw numeric difference between each test function. For example, some functions have very large fitness gap between the best and the second best local minimum, some of them don't even have local minima. Therefore, the numeric difference may not be a good performing index for evaluating algorithms. The experiment result is shown in Table IV as follows.

TABLE IV. AVERAGE FITNESS VALUE.

	CPSO-S	SD-CPSO	PSO
f_1	6.361e-99(1)*	**2.634e-062(3)**	9.653-76(2)
f_2	4.481e-84(1)*	**3.464e-033(3)**	2.876e-75(2)
f_3	18.8764 (3)	**0.8872 (1)***	1.4356(2)
f_4	11.871(1)	**17.721(2)**	26.65(1)*
f_5	9.6198(3)	**0.6893(1)***	6.3769(2)
f_6	469.9(3)	**288.3(2)**	87.36(1)*
f_7	12.57(2)	**7.659(1)***	95.03(3)
f_8	1.2287(2)	**0.6643(1)***	1.254(3)
f_9	5.75(3)	**7.864e-008(1) ***	4.08(2)

The results to be discussed are divided into three parts in accordance with the function types:

1) Unimodal Function:

Under the sphere function f_1, CPSO-S has the best performance, owing to its property of rapid convergence. As to ellipsoid function f_2, at first, PSO is better than the other two algorithms. As shown from the experiment result, all three algorithms are capable of solving unimodal optimization task, and no improvement of performance can be found by applying our method.

2) Multimodal Function:

The SD-CPSO is better than other algorithms under the f_3 and f_5 test functions except for f_4, the Rastrigin's function. We think it might due to the fact that Rastrigin's function is nearly the same after rotation, which makes our effort trying to find a special trend to the global optimum irrelevant. However, the superiority of the proposed SD-CPSO in finding global optima of multimodal functions can be seen in substance.

3) Multi-Funnel Function:

From Table IV we can see that in coping with multi-funnel function optimization tasks, the superiority of the proposed SD-CPSO is obvious. In general, the optimization of multi-funnel function is difficult as we can see especially from the optimization result of the f_6 function. Despite the proposed SD-CPSO has better performance on the optimization tasks of f_7 and f_8 function, the improvement is not very obvious. However, in the optimization of f_9, the Michalewicz's function, the improvement is remarkable. As a result, we will illustrate the optimization results of applying Michalewicz's function in both its unrotated and rotated form in Fig. 7.

Fig. 7(a) represents the result of applying unrotated Michalewicz's function. Michalewicz's function introduces many valleys into the plain, and the function values for points in the space outside the narrow valleys give very little information about the location of the global optimum. Thus, the swarms need to follow through these valleys to find minimums. In its rotated version, these narrow valleys are too correlated to follow through from the perspective of the CPSO. In Fig. 7(b), the SD-CPSO in evidence overcomes the drawback.

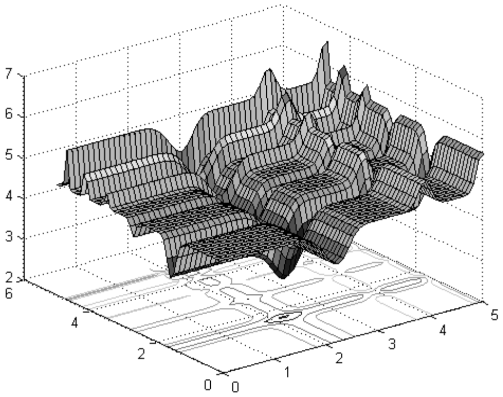
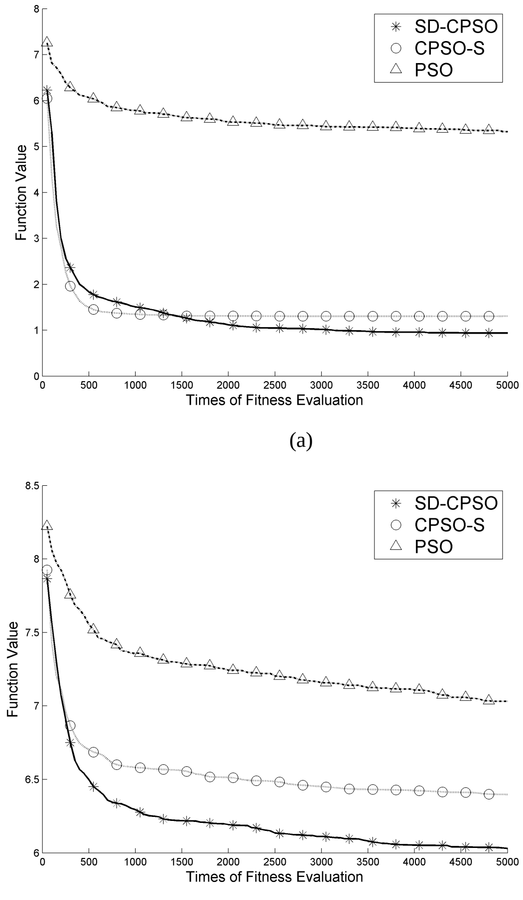


Figure 6. Visualization of a 2-D Michalewicz's function.



(a)

(b)

Figure 7. Experiment results of applying Michalewicz's function in its (a) unrotated form, (b) rotated form.

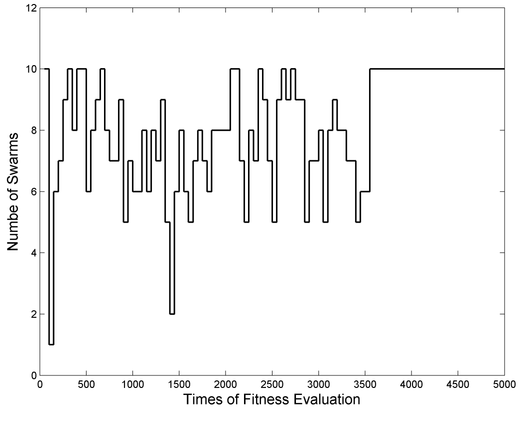


Figure 8. Results of the number of swarms of applying rotated Michalewicz's function.

Fig. 8, on the other hand, illustrates the ability of SD-CPSO self-organizes the decomposition of dimensions. We place the detected non-separable variables to the same swarm in the CPSO operation to alleviate the detrimental effect we encountered when placing independent variables into separate swarms. When particles waver in the valley, the number of swarm decreased for the sake of correlated dimension has being coupled, and when swarms step into the local minimum region, the number of swarm increased to adapt these uncorrelated sphere-liked region.

V. CONCLUSION

In this paper, we propose a self-organization approach to the CPSO. This approach determines the suitable swarm structure of the CPSO by estimating the correlations between variables. Experiments show reasonable performance. The combination of dimensions forming a swarm is detected by covariance matrix adaptation. Future research should be done to investigate the pseudominima caused by the split of swarm.

ACKNOWLEDGMENT

This work is supported in part by the National Science Council, Taiwan. R. O. C. under Grants NSC 99-2221-E-009-148.

REFERENCES

[1] J. Kennedy, “The particle swarm: social adaptation of knowledge,” in *Proc. of 1997 IEEE International Conference on Evolutionary Computation* (ICEC '97), pp. 303-308, 1997.

[2] M. Clerc and J. Kennedy, “The particle swarm - Explosion, stability, and convergence in a multidimensional complex space,” *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 58-73, 2002.

[3] M. A. Potter and K. A. De Jong, “A cooperative coevolutiouary approach to function optimization,” *in Parallel Problem Solving from Nature - Ppsn Iii - International Conference on Evolutionary Computation, Proceedings*, vol. 866, Y. Davidor, et al., Eds., ed Berlin: Springer-Verlag Berlin, pp. 249-257, 1994.

[4] N. Hansen, et al., “Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES),” *Evolutionary Computation*, vol. 11, pp. 1-18, 2003.

[5] F. van den Bergh and A. P. Engelbrecht, “A cooperative approach to particle swarm optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 8, pp. 225-239, 2004.

[6] Yi-Chang Cheng, Sheng-Fuu Lin, and Chi-Yao Hsu “Q-Value Based Particle Swarm Optimization for Reinforcement Neuro-Fuzzy System Design,” *International Journal on Computer Science and Engineering*, vol. 3, no. 10, pp. 3477-3489, 2011.

[7] C. K. Goh, et al., “A competitive and cooperative co-evolutionary approach to multi-objective particle swarm optimization algorithm design,” *European Journal of Operational Research*, vol. 202, pp. 42-54, 2010.

[8] N. Hansen and A. Ostermeier, “Completely derandomized self-adaptation in evolution strategies,” *Evolutionary Computation*, vol. 9, pp. 159-195, 2001.

[9] N. Hansen and S. Kern, “Evaluating the CMA evolution strategy on multimodal test functions,” *Parallel Problem Solving from Nature - Ppsn* Viii, vol. 3242, pp. 282-291, 2004.

[10] N. Hansen, “The CMA Evolution Strategy: A Tutorial.” 2008. (from http://www.bionik.tu-berlin.de/user/niko/cmatutorial.pdf)

[11] N. Hansen, A. Auger, S. Finck, and R. Ros, “Real-parameter block-box optimization benchmarking 2010: experimental setup.” INRIA Research Report RR-7215.

[12] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y. P. Chen, A. Auger, and S. Tiwari. “Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization.” Nanyang Technological University, Singapore and KanGAL Report Number 2005005.

AUTHORS PROFILE

Sheng-Fuu Lin was born in Tainan, the Republic of China, in 1954. He received the B.S and M.S. degree in mathematics from National Normal University in 1976 and 1979, respectively, the M.S. degree in computer science from the University of Maryland in 1985, and the Ph.D. degree in electrical engineering from the University of Illinois, Champaign, in 1988. Since 1988, he has been on the faculty of the Department of Electrical Engineering at National Chiao Tung University, Hsinchu, Taiwan, where he is currently a professor. His research interests include fuzzy systems, genetic algorithms, neural networks automatic target recognition, scheduling, image processing, and image recognition.

Yi-Chang Cheng received the B.S. degree in engineering science from the National Cheng Kung University, Taiwan, R.O.C., in 2005. He is currently pursuing the Ph. D. degree at the department of electrical engineering from the National Chiao Tung University, Taiwan, R.O.C. His research interests include neural networks, fuzzy systems, evolutional algorithms and genetic algorithms.

Jyun-Wei Chang received the B.S. and M.S. degree in electronic engineering from National Kaohsiung University of Applied Sciences, Taiwan, R.O.C. in 2005 and 2007, respectively. He is currently pursuing the Ph.D. degree at the department of electrical engineering from the National Chiao Tung University, Taiwan, R.O.C. His research interests include neural networks, fuzzy systems, and evolutional algorithms.

Pei-Chia Hung received the B.S. degree in engineering science from the National Chiao Tung University, Taiwan, R.O.C., in 2004. He is currently pursuing the Ph. D. degree at the department of electrical engineering from the National Chiao Tung University, Taiwan, R.O.C. His research interests include image processing and image compression.

Nearest Neighbor Value Interpolation

Rukundo Olivier 1
Department of Electronics and Information Engineering
Huazhong University of Science and Technology, HUST
Wuhan, China

Cao Hanqiang 2
Department of Electronics and Information Engineering
Huazhong University of Science and Technology, HUST
Wuhan, China

***Abstract*—This paper presents the nearest neighbor value (NNV) algorithm for high resolution (H.R.) image interpolation. The difference between the proposed algorithm and conventional nearest neighbor algorithm is that the concept applied, to estimate the missing pixel value, is guided by the nearest value rather than the distance. In other words, the proposed concept selects one pixel, among four directly surrounding the empty location, whose value is almost equal to the value generated by the conventional bilinear interpolation algorithm. The proposed method demonstrated higher performances in terms of H.R. when compared to the conventional interpolation algorithms mentioned.**

Keywords—neighbor value; nearest; bilinear; bicubic; image interpolation.

I. Introduction

Image interpolation is the process by which a small image is made larger by increasing the number of pixels comprising the small image [1]. This process has been a problem of prime importance in many fields due to its wide application in satellite imagery, biomedical imaging, and particularly in military and consumer electronics domains. At an early stage of research, non-adaptive methods such as nearest, bilinear and bicubic interpolation methods were developed for digital image interpolation purposes. Those traditional methods were markedly different in image resolution, speed, and theoretical assumptions (i.e. theory of spatial variability) [2], [3]. To the best of my knowledge, most of the assumptions applied today reduce interpolated image resolution due to the lowpass filtering process involved into their new value creative operations [4]. However, the nearest neighbor assumption does not permit to create a new value, instead set the value at the empty location by replicating the pixel value located at the shortest distance. The effect of this is to make image pixel bigger which results in heavy jagged edges thus making this algorithm more inappropriate for applications requiring a H.R image (to accomplish certain tasks). A solution to such jaggedness was achieved through the bilinear interpolation [5]. A bilinear based algorithm generates softer images but blurred thus making the algorithm inappropriate also for H.R. applications. The blurredness problem was reduced by introducing the convolution based techniques [6]. Such algorithms performed better than the two in terms of the visual quality but are also inappropriate to use where the speed is of the prime importance. Now, since the source image resolution is often reduced after undergoing the interpolation process, the easy way to generate a H.R. image using linear interpolation means is to reduce, at any cost, any operation that would underestimate or overestimate some parts of the image. In other words, it would be better if we could avoid 100% any operation leading to the new pixel value creation for image interpolation purposes. In this regards, one way to reduce using the newly created values, is based on supposing that one, of the four pixels, has a value that is appropriate enough to be assigned directly at an empty location. The problem, here, is to know which one is more appropriate than the others or their weighted average, etc. Therefore, we propose a scheme to be guided, throughout our H.R. interpolated image search, by the value generated by the conventional bilinear interpolator. The Fig.2 briefly explains how this can be achieved. More details are given in Part III.

This paper is organized as follows. Part II gives the background, Part III presents the proposed method, Part IV presents the experimental results and discussions and Part V gives the conclusions and recommendation.

II. Background

The rule in image interpolation is to use a source image as the reference to construct a new or interpolated/scaled image. The size of the new or constructed image depends on the interpolation ratio selected or set. When performing a digital image interpolation, we are actually creating empty spaces in the source image and filling in them with the appropriate pixel values [2]. This makes the interpolation algorithms yielding different results depending on the concept used to guess those values. For example, in the nearest neighbor technique, the empty spaces will be filled in with the nearest neighboring pixel value, hence the name [3].

This (nearest neighbor algorithm) concept is very useful when speed is the main concern. Unlike simple nearest neighbor, other techniques use interpolation of neighboring pixels (while others use the convolution or adaptive interpolation concepts - but these two are beyond the scope of this paper), resulting in smoother image. A good example of a computationally efficient basic resampling concept or technique is the bilinear interpolation. Here, the key idea is to perform linear interpolation first in one direction, and then again in the other direction. Although each step is linear in the sampled values and in the position, the interpolation as a whole is not linear but rather quadratic in the sample location [5]. In other words, the bilinear interpolation algorithm creates a weighted average value that uses to fill in the empty spaces. This provides better tradeoff between image quality and computational cost but blurs the interpolated image thus reducing its resolution.

III. Proposed Method

Assume that the letters A, K, P and G represent the four neighbors and E represents the empty location value as shown in Fig.1.

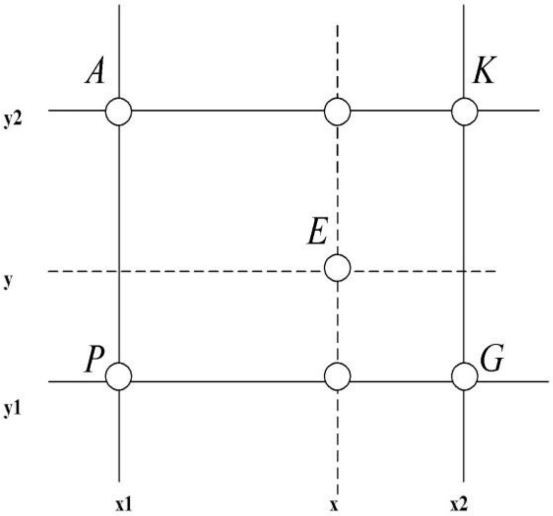


Fig. 1: Four neighbors locations around an empty location E

In order to implement successfully the proposed scheme, the following steps have been respected.

A. Neighbors mode calculation

Let us call $L=[A,K,P,G]$ a set of four neighbors or data. In statistics, the mode is the value that occurs most frequently in a data set or a probability distribution [7]. Here, the first step is to check whether in L there is a mode or not (i.e. if there exists a mode in L). If the mode exists then, the empty location will be assigned that mode. If the mode does not exist in L (i.e., when two data in L appear the same number of times or when none of the L data repeat) then, we proceed to performing the bilinear interpolation among L data in order to achieve a bilinearly interpolated value or bilinear value. Once the bilinear value is obtained, we do the subtraction operations as shown by Eq.(1), Eq.(2), Eq.(3) and Eq.(4). The letter B represents the bilinear value.

$$|A-B|=V_1 \tag{1}$$

$$|K-B|=V_2 \tag{2}$$

$$|P-B|=V_3 \tag{3}$$

$$|G-B|=V_4 \tag{4}$$

The values obtained, from the subtraction operations, are absolute values and can be represented by V_1, V_2, V_3 and V_4. Before, we proceed to finding the pixel value yielding the minimum difference, we must check that none of these absolute values is equal to another or simply occurs most frequently. In other words, we must check again the mode so that we can be able to end up with one neighbor whose value is nearly equal to the value yielded by the bilinear interpolator.

B. Absolute differences mode calculation

At this stage, the mode is calculated from a given set J containing all the absolute differences $J=[V_1,V_2,V_3,V_4]$. If there exists a mode in J then, we can find out that the mode is the minimum value or not, before we can proceed further. For instance, consider the following three examples.

Example 1: $V_1=0.2, V_2=0.2, V_3=0.2, V_4=0.8$

In this example, the mode is 0.2 and 0.2 is the minimum value. So, in order to avoid the confusion our algorithm will only consider/select the first value from J. The selection of the first value can be achieved based on the subscripted indexing theory [8]. Once this value is selected, we can calculate the absolute difference between this value and bilinear value and the difference obtained is assigned to the empty location.

Example 2: $V_1=0.2, V_2=0.8, V_3=0.8, V_4=0.8$

In this example, the mode is 0.8 and unfortunately 0.8 is not the minimum value therefore our concept, which is directed by the minimum difference value between the value yielded by the bilinear and one of the four neighbors value, cannot be respected. To solve this issue, first of all we find the value that is less or equal to the mode. In the matlab the *find* function returns indices and values of nonzero elements [9]. The obtained elements are presented in ascending order. In this example, the value that is less or equal to the mode would be 0.8 or 0.2 but since the two values cannot be selected at the same time, we can pick the first minimum value by applying again the subscripted index method.

Once the minimum value is obtained, we can find the neighbor that corresponds to that minimum value and calculate the absolute difference between that value and bilinear value, then assign it to the empty location.

C. When there is no 'absolute differences' mode

This case can also be regarded as an example number three of the B part (even though it is presented in C part).

Example 3: $V_1=0.2, V_2=0.2, V_3=0.8, V_4=0.8$

As shown, in this example, there is no mode when two data/elements of a set repeat the same number of times. The same when all J elements are different. In both cases, we have to find the minimum value in J using matlab *min* function so that we can be able to apply the subscripted indexing to get the first minimum value.

Once the minimum value is obtained, we must find the neighbor that corresponds to that minimum number. This can be achieved by subtracting the minimum difference from the bilinear value, because the minimum value is equal to the neighbor value minus the bilinear value (see Eq.(1), Eq.(2), Eq.(3) and Eq.(4)). Then, the obtained value is assigned to the empty location.

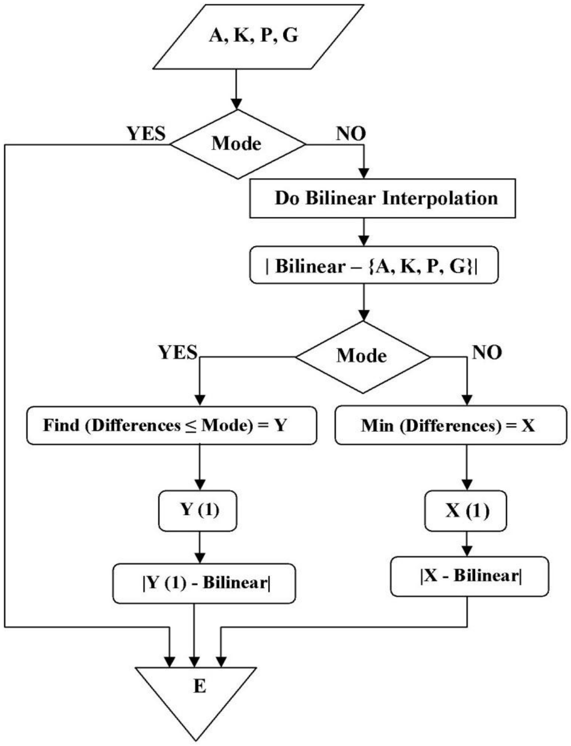


Fig.2: The summary of the proposed method

The Fig.2 shows four data input. These data are in fact the four neighbors surrounding the empty location. The E destination represents the final interpolated value that must be assigned to the empty location, accordingly. The aim, of finding this value in this way, is to minimize the underestimation or overestimation of some parts of image texture after undergoing the interpolation process because of the problems caused by the lowpass filtering processes involved in many linear interpolators, bilinear in particular.

IV. Experiments and Discussions

We tested the proposed NNV algorithm for image details quality (i.e. H.R), Matlab-lines Execution Time (MET) and Peak Signal to Noise Ratio (PSNR) against the conventional nearest, bilinear and bicubic interpolation algorithms using four full grayscale images shown in Fig.3. The interpolated images (ratio n = 4 and n = 2) are shown in Fig.[4-11]. The corresponding MET and PSNR results are shown in the Table 1 and Table 2. A higher peak signal to noise ratio would normally indicate the higher quality of the output image. The PSNR can easily be defined via the Mean Squared Error where one of the monochrome images *I* and *K* is considered as a noisy approximation of the other.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j) - K(i,j) \right]^2 \qquad (5)$$

The PSNR is defined as:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}} \right) \qquad (6)$$

where, MAX_I represents the maximum image pixel value. Typically, the PSNR values in lossy image and video compression range from 30 to 50 dB. When the two images are identical, the MSE=0 and consequently the PSNR is undefined.

A. Original full images - 128 x 128

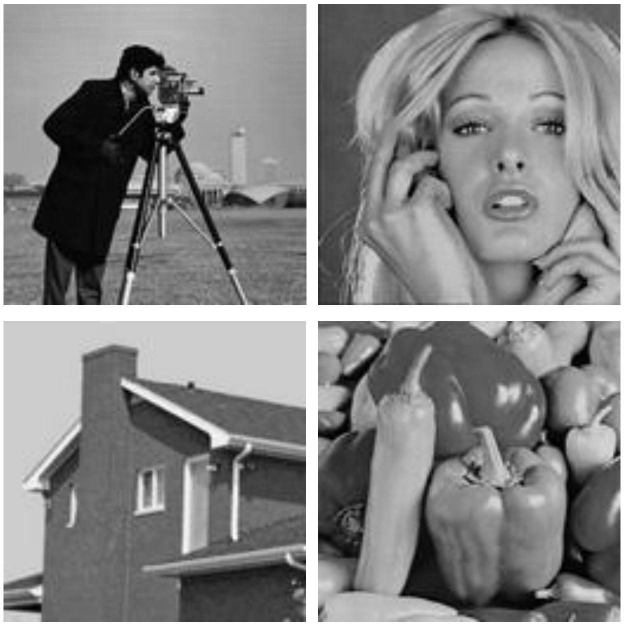


Fig.3: Source images

From left to right - top to bottom - there are Cameraman, Girl, House and Peppers grayscale images. Each image has the size of 128 x 128 and will be interpolated at the ratio n=4 and n=2 in part B and C, respectively.

B. Full image interpolation & Ratio = 4

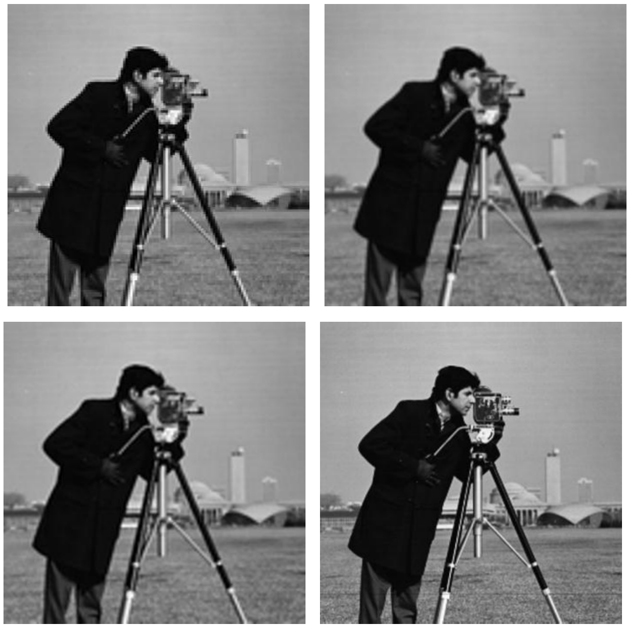


Fig.4: Cameraman: n=4

From left to right - top to bottom – 512 x 512 Cameraman-image interpolated by nearest neighbor (NN), bilinear (Bil.), bicubic (Bic.) and nearest neighbor value (NNV) algorithms, respectively.

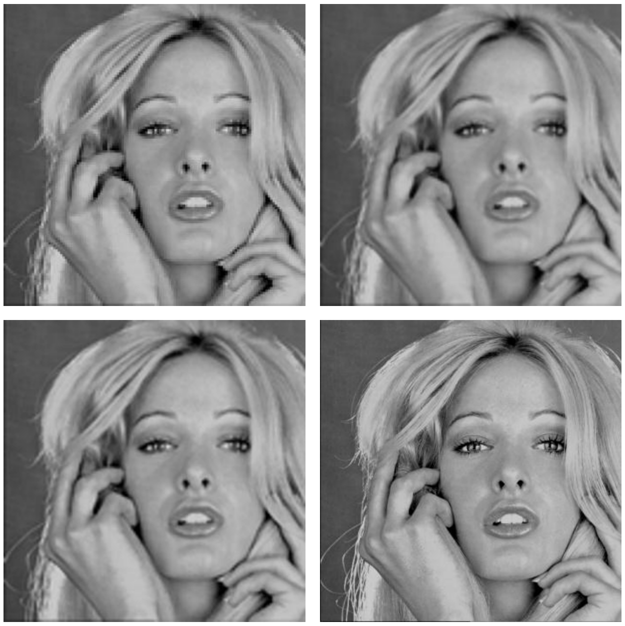


Fig.5: Girl: n=4

From left to right - top to bottom – 512 x 512 Girl-image interpolated by NN, Bil., Bic. and NNV algorithms, respectively.

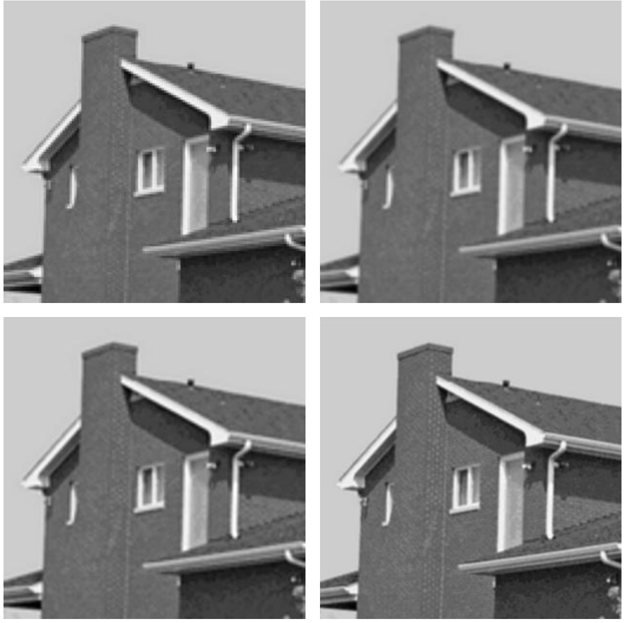


Fig.6: House: n=4

From left to right - top to bottom – 512 x 512 House-image interpolated by NN, Bil., Bic. and NNV algorithms, respectively.

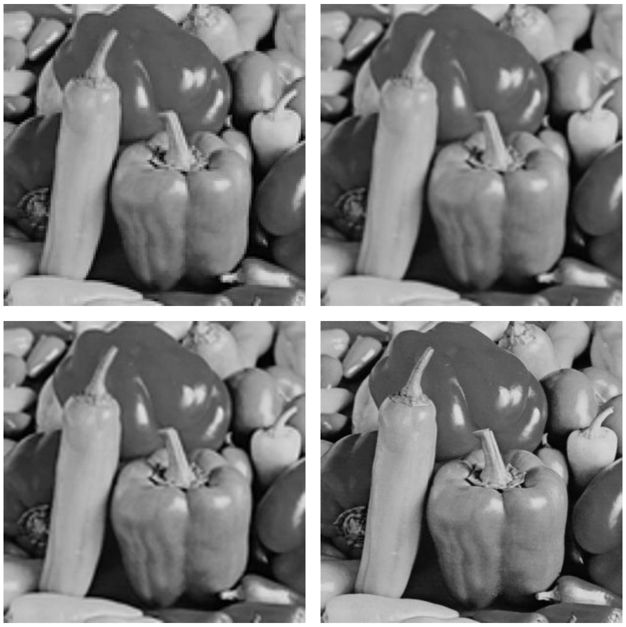


Fig.7: Peppers: n=4

From left to right - top to bottom – 512 x 512 Peppers-image interpolated by NN, Bil., Bic. and NNV algorithms, respectively.

C. Full image interpolation & Ratio = 2

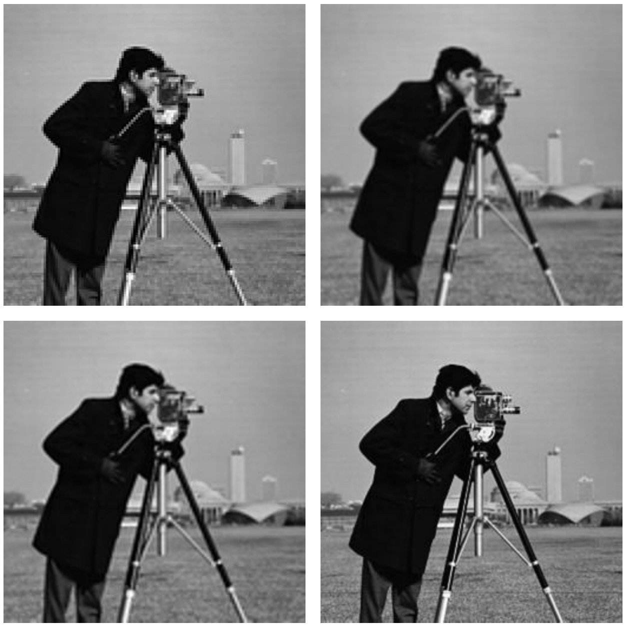


Fig.8: Cameraman: n=2

From left to right - top to bottom – 256 x 256 Cameraman-image interpolated by NN, Bil., Bic. and NNV algorithms, respectively.

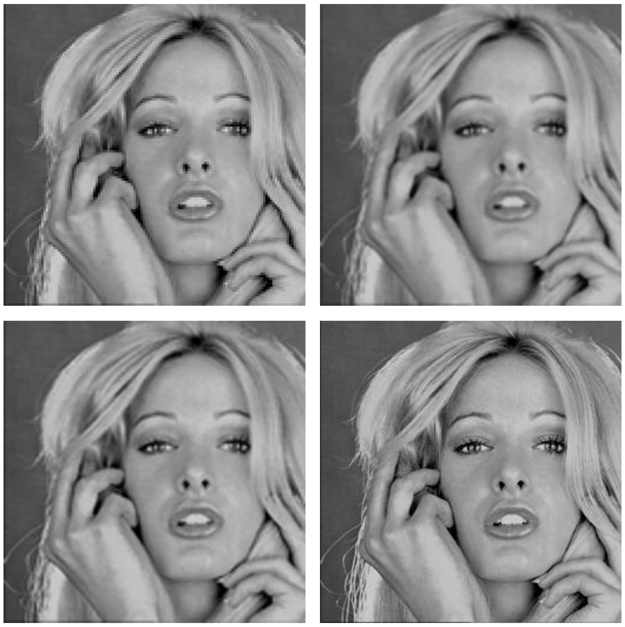


Fig.9: Girl: n=2

From left to right - top to bottom – 256 x 256 Girl-image interpolated by NN, Bil., Bic. and NNV algorithms, respectively.

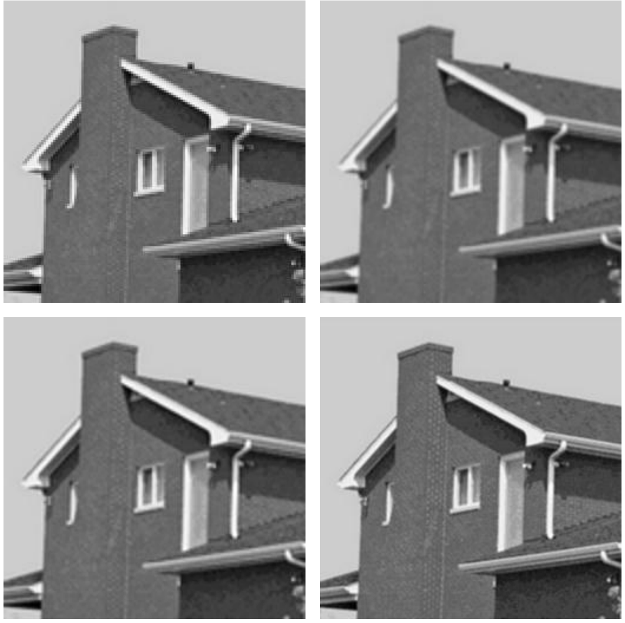


Fig.10: House: n=2

From left to right - top to bottom – 256 x 256 House-image interpolated by NN, Bil., Bic. and NNV algorithms, respectively.

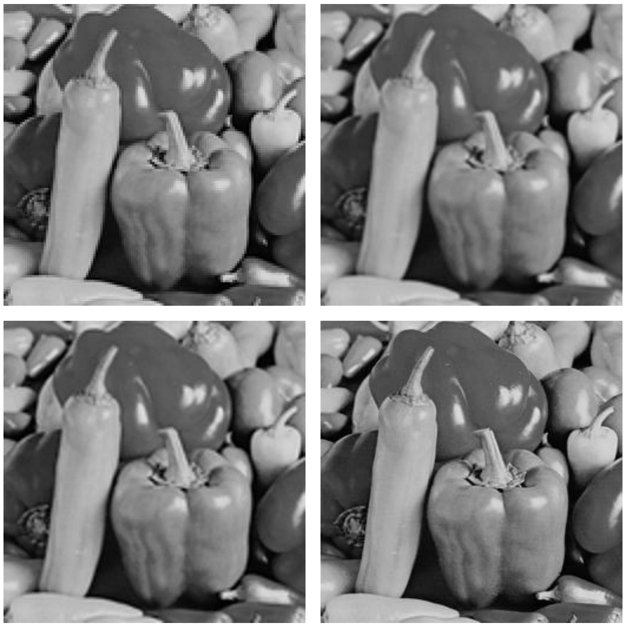


Fig.11: Peppers: n=2

From left to right - top to bottom – 256 x 256 Peppers-image interpolated by NN, Bil., Bic. and NNV algorithms, respectively.

V. Conclusion and Recommendations

Image interpolation based on the nearest neighbor value has been presented in this paper. The details on how it was developed (i.e. the scheme used) have been presented in part III and the working procedure has been summarized and shown in Fig.2. The MET and PSNR results have been presented in Table 1 and Table 2. Depending on the interpolation ratio selected or set (i.e. depending on the final size desired/targeted), the interpolation algorithms, mentioned here, gave different MET and PSNR as well as visual quality. For example, let us observe the interpolated images shown in part B (i.e. image that were interpolated at the ratio = 4). Starting from, the Cameraman image on the first row, the first image shows a texture with edge jaggedness (i.e. image interpolated using the NNI algorithm) while the second one (i.e. image interpolated using the bilinear algorithm) shows soft but blurred texture. The first image tends to look sharper than the second one. That look difference was noticed due to the lowpass filtering process involved in the algorithm used to interpolate the latter. On the second row, the first image (i.e. image interpolated using the bicubic algorithm) shows smoother but sharper texture so is the second one (i.e. image interpolated using the NNV algorithm), except that the latter shows more readily the image details. The same conclusions can be drawn for other image cases but with a slight change because the best interpolation method for an image may depend on the image itself.

In other words, one shoe may not fit all. For the interpolated images shown in part C (i.e. image that were interpolated at the ratio = 2). Except the first image on the first row (i.e. image interpolated using the NNI algorithm), it is difficult to notice the visual differences because the differences were minor (with the exception of the NNV) and it is often problematic as to which one looks the best.

TABLE 1: PSNR AND MET AFTER INTERPOLATION & (RATIO = 4) 1: CAMERAMAN, 2: GIRL, 3: HOUSE, 4: PEPPERS

	PSNR (dB)				**MET (s)**			
	NN	Bil.	Bic.	NNV	NN	Bil.	Bic.	NNV
1	34.0829	34.1135	34.1628	35.0154	0.036866	0.058984	0.060625	0.843483
2	32.9235	33.1043	33.1655	34.2262	0.037365	0.059842	0.060477	0.818222
3	35.4771	35.4563	35.4890	36.2054	0.038078	0.057881	0.074557	0.788410
4	33.6570	33.7862	33.8349	34.5064	0.040556	0.062336	0.063787	0.800693

Therefore, the PSNR was used to prove through the calculations which one has really higher quality than the others. As shown in Table 1 and Table 2, the PSNR value generated by the NNV has always been superior to other image interpolators'. This can be interpreted as an overwhelming sign showing the higher performances of the proposed algorithm with respect to other interpolation algorithms mentioned.

TABLE 2: PSNR AND MET AFTER INTERPOLATION & (RATIO = 2) 1: CAMERAMAN, 2: GIRL, 3: HOUSE, 4: PEPPERS

	PSNR (dB)				**MET (s)**			
	NN	Bil.	Bic.	NNV	NN	Bil.	Bic.	NNV
1	34.2996	34.0658	34.3385	36.0238	0.042512	0.053586	0.055415	0.700002
2	33.7806	33.8934	34.2070	35.9882	0.037528	0.060694	0.060948	0.780222
3	35.9944	35.6859	35.9841	36.3441	0.038178	0.061335	0.055840	0.776107
4	34.7829	34.8355	35.1103	35.7047	0.042090	0.055411	0.054192	0.768127

Comparing other methods against each other, we found that one could perform better than expected or not depending on the image interpolated and in all the cases none of them achieved a higher PSNR value than the proposed NNV. However, the mentioned conventional algorithms are all faster than the proposed NNV. For example, in case where the interpolation ratio = 4, the NNV was about 21.2 times slower than the fastest NNI whereas in the case of the ratio = 2, the NNV became approximately 18.8 times slower than the NNI. In fact, the best interpolation method for one size of enlargement may not necessarily be the best method for a different size, in terms of the visual resolution, PSNR value and MET value. Please note that the images presented, in the experimental part of this paper, have lost some of their quality when they were reduced to fit in the paper format. This fairly reduces the differences between the presented/interpolated images. With reference to the experimental results obtained, we suggest that the proposed NNV method be recommended for further applications, especially where some image tissues, or particular details, need to be seen in their richest and most pleasant way as well as where a balmy computational cost is not an issue. Future developments of the proposed approach may be guided by techniques using higher order polynomials to interpolate.

ACKNOWLEDGMENT

This work was supported by National Anti-counterfeit Engineering Research Center and National Natural Science Foundation of China (N0: 60772091). Rukundo Olivier and Cao Hanqiang would like to use this opportunity to thank the reviewers and editor for the helpful comments and decision, respectively.

REFERENCES

[1] What is photo interpolation resizing resampling. Americas Wonderlands, 2012 http://www.americaswonderlands.com/image_resizing.htm

[2] Digital Image Interpolation. Cambridge in colour, 2011http://www.cambridgeincolour.com/tutorials/image-interpolation.htm

[3] A Review of Some Image Pixel Interpolation Algorithms. Don Lancaster & Synergetics. 2007 http://www.tinaja.com

[4] G.Ramponi, Warped Distance for Space-Variant Linear Image Interpolation, IEEE Transactions on Image Processing, vol.8 n.5, May 1999, pp. 629–639

[5] Wikipedia free encyclopedia.(2012, March. 21).Bilinear intepolation.Available:http://en.wikipedia.org/wiki/Bilinear_interpolation

[6] Wikipedia free encyclopedia. (2012, March. 21). Bicubic intepolation.Available:http://en.wikipedia.org/wiki/Bicubic_interpolation

[7] B. Gregory . Mode. In Salkind, Neil. Encyclopedia of research design. Sage. pp. 140–142, 2010

[8] MATLAB Indexing (March 24, 2012).Mathworks.Subscripted Indexing.Available:http://www.mathworks.cn/support/tech-notes/1100/1109.html#referencing

[9] Mathworks,R2012a Documentation-MATLAB,(March 23, 2012).FIND. Available:http://www.mathworks.cn/help/techdoc/ref/find.html

AUTHORS PROFILE

Rukundo Olivier born December 1981 in Rwanda currently holds B.Sc. (2005) in electronics and telecommunication engineering and M.E. (2009) in communication and information system from Kigali Institute of Science and Technology and Huazhong University of Science and Technology, respectively. He is currently doing research in the area of signal processing at the Laboratory for Information Security and Identity, National Anti-Counterfeit Engineering Research Center and his previous published works included novel digital image interpolation algorithms as well as analog circuits test modes. Mr. Rukundo was employed for 18 months at Société Interbancaire de Monétique et de Télécompensation au Rwanda before moving on to Huazhong University of Science and Technology where he is currently a PhD candidate and expecting to graduate in June 2012.

Dr. Cao Hanqiang is professor currently in the department of Electronics and Information Engineering, Huazhong University of Science and Technology. His areas of expertise and interest are signal/image processing, information security.

A New 3D Model-Based Tracking Technique for Robust Camera Pose Estimation

Fakhreddine Ababsa
University of Evry, 40 Rue du Pelvoux,
91020 Evry, France

***Abstract*—In this paper we present a new robust camera pose estimation approach based on 3D lines features. The proposed method is well adapted for mobile augmented reality applications We used an Extended Kalman Filter (EKF) to incrementally update the camera pose in real-time. The principal contributions of our method include first, the expansion of the RANSAC scheme in order to achieve a robust matching algorithm that associates 2D edges from the image with the 3D line segments from the input model. And second, a new powerful framework for camera pose estimation using only 2D-3D straight-lines within an EKF. Experimental results on real image sequences are presented to evaluate the performances and the feasibility of the proposed approach in indoor and outdoor environments.**

***Keywords-Pose estimation; Line tracking; Kalman filtering; Augmented Reality*.**

I. INTRODUCTION

Estimation of relative 3D camera position and orientation (pose) is one of the most challenging problems in computer vision. Indeed, the knowledge of the camera pose is useful for numerous applications, including motion analysis, robotic control tasks, and Augmented Reality (AR) [1][2]. For example, Dasgupta and Banerjee [3] developed an AR system for autonomous navigation which uses a panoramic vision. Their system allows augmenting, via a laptop, the real video by simultaneously viewing an artificial 3D view of the scene. During the last years, several approaches based on natural features (points, planes, edges, etc.) extracted from the image scene have been developed. The mean idea of these techniques is to locate the camera in three dimensions given a sequence of 2D intensity images of the scene containing the 3D features whose positions and orientations are known relative to a reference frame. When the correspondences between 2D features extracted from the image and 3D features defined in the world frame are established, the problem is then solved using 2D-3D registration techniques. Numerical nonlinear optimization methods like the Newton-Raphson or Levenberg-Marquardt algorithm are generally used for the minimization [4][5][6]. Thus, Wang et al. [7] propose a 3D localization approach using Harris-Laplace interest points as natural landmarks. Their localization system is composed of two stages: coarse localization from a location vector space model (LVSM) and affine localization from the location database using a voting algorithm. The pose recovery of the camera is obtained from essential matrix decomposition. Line-based methods are more robust and more efficient than point-based techniques. Indeed, line features are present in many scenes and are more visible than points under a wider range of lighting and environmental conditions. Wuest et al. [8] present a model-based line tracking approach that can handle partial occlusion and illumination changes. The camera pose is computed by minimizing the distances between the projection of the model lines and the most likely matches found in the image. Drummond and Cipolla [9] propose a novel framework for 3D model-based tracking. Objects are tracked by comparing projected model edges to edges detected in the current image. Their tracking system predicts the edge locations in order to rapidly perform the edge search. They have used a Lie group formalism in order to transform the motion problem into simple geometrics terms. Thus, tracking becomes a simple optimization problem solved by means of iterative reweighed least squares. Recently, Comport et al. [10] propose a real-time 3D model-based tracking algorithm. They have used a visual servoing approach to formulate the pose estimation problem. A local moving edges tracker based on tracking of points normal to the object contours is implemented. In order to make their algorithm robust, they have integrated a M-estimator into the visual control law. Other approaches have also been applied where different features have been combined to compute the camera pose. Ababsa and Mallem [11] propose to combine point and line features in order to handle partial occlusion. They integrated a M-estimator into the optimization process to increase the robustness against outliers. Koch and Teller [12] describe an egomotion estimation algorithm that takes as input a coarse 3D model of an environment. Their system uses a prior visibility analysis to speed initialization and accelerate image/model matching.

For real-time applications, extended Kalman filtering (EKF) is the most widely used method for recursive estimation. EKF allows also to achieve noise and disturbance rejection and to enhance estimation accuracy. EKF was used for the first time to compute the 3D location by Wilson et al. [13]. Yu et al. [14] propose an EKF based method to recover structure and motion from image sequences. The EKF is used to estimate the object’s pose and to refine the positions of the model points in the 3D space. More recently the same authors propose an EKF algorithm for pose tracking using the trifocal tensor [15]. Their approach is also based on natural points and incorporates the tri-focal constraint into the measurements model. Yoon et al. [16] present a model-based object tracking to compute the 3D camera pose. Their algorithm uses an EKF to provide an incremental pose-update scheme in a prediction-verification framework. In order to enhance the accuracy and the robustness of the tracking against occlusion, they take into account the measurement uncertainties associated with the location of the extracted image straight-lines. Lippiello et al. [17] developed an algorithm for pose estimation of a moving object. In order to enhance the robustness with respect to measurement noise and modeling error, they proposed an adaptive version of the EKF customized for visual applications. Jiang and Neumann [18] propose an EKF-based extendible tracking method that can extend tracking from prepared to unprepared regions of the environment by auto-calibrating a 3D structure of a priori unknown 3D lines.

Other approaches use Simultaneous Localization And Mapping (SLAM) to track the camera pose while building a 3D map of the unknown scene. William et al. [19] proposed a recovery module that finds the pose of a camera using a map point features created by a single-camera SLAM system. They used an efficient version of RANSAC to achieve pose hypotheses verification and to reject outliers. Davison et al. [20] used the EKF to estimate the 3D trajectory of a monocular camera moving through an unknown scene. The core of their system, named MonoSLAM, is the online creation of a sparse map of natural landmarks within a probabilistic framework. However, the main problem with most existing monocular SLAM techniques is a lack of robustness when rapid camera motions, occlusion and motion blur occur.

In this paper we present an original robust camera pose tracking using only straight lines and which differs from existing works. We propose to combine an EKF with a RANSAC scheme in order to achieve a robust 2D-3D lines matching. This gives an efficient solution for outlier’s rejection. To our knowledge such solution has not been explored before. Furthermore, we have combined the 2D-3D lines correspondence constraints for object pose estimation, developed by Phong et al. [21], with an EKF in order to update recursively the camera pose. We have compared our results with classical approaches where pose estimation is solved using least square approaches [22][23][24]. Our method requires no training phase, no artificial landmarks, and uses only one camera.

The rest of the paper is structured as fellows: In section II, we describe the camera pose estimation problem formulation when using straight lines, and we also give a complete implementation of the Extended Kalman Filter to update the camera pose recursively over the time using 2D and 3D lines features. In section III, we explain how we have expended the RANSAC scheme [25] in order to achieve robust 2D-3D lines matching. In section IV, we show experimental results and evaluations; we discuss also the merits and the limitations of the proposed approach. Conclusion and future work are presented in section V.

II. CAMERA POSE ESTIMATION ALGORITHM

In any Kalman Filter implementation, the system state is stored as a vector. In our algorithm the state is represented by the position and the orientation of the camera with respect to the world coordinate system. For computational we use a unit quaternion to represent the rotation [26]. Thus, the state vector is given by:

$$X = \lfloor q_0 \quad q_x \quad q_y \quad q_z \quad t_x \quad t_y \quad t_z \rfloor \tag{1}$$

where $\left(q_0^2 + q_x^2 + q_y^2 + q_z^2 = 1\right)$

We denote the camera state at time t by the vector X_t. The EKF is used to maintain an estimate of the camera state X in the form of a probability distribution $P\left(X_t \middle| X_{t-1}, Z_t\right)$, where Z_t is the measurement vector at time t. The Kalman filter models the probability distribution as Gaussian, allowing it to be represented by a covariance matrix Σ. In an extended Kalman filter, the nonlinear measurements and motion models are linearised about the current state estimate as Jacobian matrices [27][28]. Our algorithm follows the usual predict-refine cycle, whereby the state X is predicted at timestep t, and measurements are used to refine the prediction.

A. *Time update*

The time update model is employed in order to predict the camera pose at the following time step. In our case, the time update is simple because of the fact that we estimate the camera pose at each frame of the images sequence. Therefore the 3D cameras pose between two successive frames changes very little. The time update equation is then given by :

$$X_t^- = A \cdot X_{t-1} \tag{2}$$

Where A is 7×7 identity matrix.

The time update step also produces estimates of the error covariance matrix Σ from the previous time step to the current time step t. To perform this prediction we use the general update equation of the Kalman filter:

$$\Sigma_t^- = A \cdot \Sigma_t \cdot A' + Q_{t-1} \tag{3}$$

Where Q_t represents the covariance matrix of the process noise. Σ reflects the variance of the state distribution.

B. *Measurement model and estimate update*

The measurement update model relates the state vector to the measurement vector. Since our goal is to estimate the camera pose using only straight lines, we will first describe the constraint equation which relates the state vector to the 3D model lines and their corresponding 2D image edges. We choose to base our technique on line features, rather than points, because this approach is relatively unexplored in the vision literature. We consider a pin-hole camera model and we assume that the intrinsic camera parameters are known. The world coordinate frame is a reference frame. All the 3D model lines are defined with respect to it. Let Li be a 3D line. Li is represented with the Cartesian coordinates of its two end-points P_1^i and P_2^i (see figure 1). The points P_1^i and P_2^i in world coordinates can be expressed in the camera frame as well :

$$\begin{cases} P_{1/C}^i = R \cdot P_{1/W}^i + T \\ P_{2/C}^i = R \cdot P_{2/W}^i + T \end{cases} \tag{4}$$

Where the 3×3 rotation matrix R and the translation vector T describe the rigid body transformation from the world coordinate system to the camera coordinate system and are precisely the components of the camera state vector.

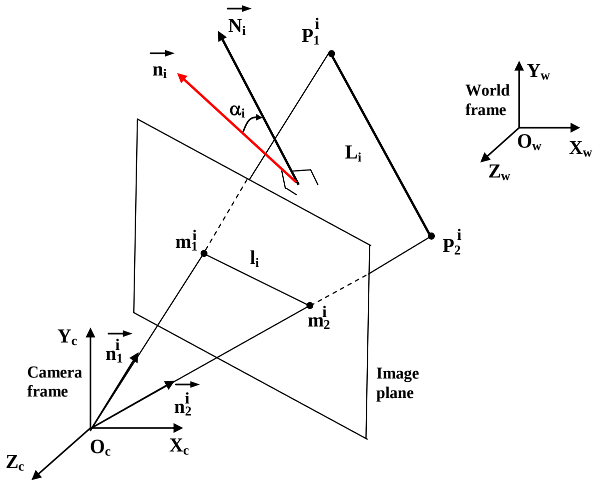


Figure 1. Projection plane. The model line, its projection onto the image and the center of projection OC are coplanar.

We can see that the points $P^i_{1/C}$, $P^i_{2/C}$ and the center of projection OC are coplanar. $\vec{N}_i$ is the unit vector normal to this plane. $\vec{N}_i$ can be expressed in the camera coordinates frame as follows :

$$\vec{N}_i = \frac{\overrightarrow{O_C P^i_{1/C}} \times \overrightarrow{O_C P^i_{1/C}}}{\left\| \overrightarrow{O_C P^i_{1/C}} \times \overrightarrow{O_C P^i_{1/C}} \right\|} \qquad (5)$$

Furthermore, a measurement input of the normal vector $\vec{N}_i$ can be obtained from the image data. Indeed, image line matched with model line belongs also to the projection plane defined above. Let l_i be a 2D image line corresponding to the 3D line L_i. In similar manner l_i is represented by its two extremities $m_1^i = [u_1^i \quad v_2^i]^T$ and $m_2^i = [u_2^i \quad v_2^i]^T$ defined in the 2D image coordinates frame. The points m_1^i and m_2^i can be expressed in the camera coordinate frame as follows:

$$\begin{cases} m^i_{1/C} = K^{-1} \cdot \begin{bmatrix} u_1^i & v_1^i & 1 \end{bmatrix}^T \\ m^i_{2/C} = K^{-1} \cdot \begin{bmatrix} u_2^i & v_2^i & 1 \end{bmatrix}^T \end{cases} \qquad (6)$$

Where the matrix K contains camera calibration parameters, such as focal length, aspect ratio and principal point coordinates.

A measurement $\vec{n}_i$ of the unit vector $\vec{N}_i$ normal to the projection plane is thus given by (see figure 1):

$$\vec{n}_i = \vec{n}_1^i \times \vec{n}_2^i \qquad (7)$$

Where

$$\vec{n}_1^i = \frac{\overrightarrow{O_C m^i_{1/C}}}{\left\| \overrightarrow{O_C m^i_{1/C}} \right\|} \quad et \quad \vec{n}_2^i = \frac{\overrightarrow{O_C m^i_{2/C}}}{\left\| \overrightarrow{O_C m^i_{2/C}} \right\|}$$

Combining equations (5) and (7), a measurement equation can be written, for each matching event $L_i \rightarrow l_i$:

$$z_t = h(X_t) + v_t \ (8)$$

Where

$$\begin{cases} z_t = \vec{n}_i = \begin{bmatrix} n_x^i & n_y^i & n_z^i \end{bmatrix}^T \\ h(X_t) = \vec{N}_i = \begin{bmatrix} N_x^i & N_y^i & N_z^i \end{bmatrix}^T \end{cases} \qquad (9)$$

v_t represent the noise term in the measurement input with covariance R_t. The noise is due to the uncertainty in the measured image position of the end points of the extracted 2D lines. The nonlinear function $h(X)$ in measurement equation (8) relates the state to the measurement input. Three 2D-3D line correspondences are sufficient in theory to recover 6-DOF camera pose [29] through in practice mores line may be required to increase accuracy.

The state estimate and covariance are refined after each feature measurement z_t using the standard equation of the EKF as follows:

$$\begin{aligned} K_t &= \Sigma_t^- \cdot H_t^T \cdot \left(H_t \cdot \Sigma_t^- \cdot H_t^T + R_t\right)^{-1} \\ X_t &= X_t^- + K_t \cdot \left(z_t - h\left(X_t^-\right)\right) \\ \Sigma_t &= \Sigma_t^- - K_t \cdot H_t \cdot \Sigma_t^- \end{aligned} \qquad (10)$$

Where H_t is the Jacobian matrix defined by:

$$H_t = \left.\frac{\partial h(X)}{\partial X}\right|_{X = X_t^-} \qquad (11)$$

The measurement update model is executed once a set of 2D-3D matched lines become available.

C. Iterated EKF

The standard EKF method does not consider errors due to the linearization of the nonlinear function $h(X)$ in the vicinity of X_t^-. However, these errors can lead to wrong estimates and/or divergence of the camera pose. Since the nonlinearity is only in measurement equation, the Iterated Extended Kalman Filter (IEKF) is the best technique to deal with it [30][31]. The IEKF uses the same prediction equation as EKF, namely (2) and (3). The measurement update relations are replaced setting $X_t^0 = X_t^-$ and doing iteration on :

$$\begin{aligned} H_t^k &= \left.\frac{\partial h(X)}{\partial X}\right|_{X = X_t^k} \\ K_t^k &= \Sigma_t^- \cdot H_t^{k^T} \cdot \left(H_t^k \cdot \Sigma_t^- \cdot H_t^{k^T} + R_t\right)^{-1} \\ X_t^{k+1} &= X_t^k + K_t^k \cdot \left(z_t - h\left(X_t^k\right)\right) \end{aligned} \qquad (12)$$

For iteration number $k = 0,1,\Lambda\ , N-1$. At the end of all iterations, $X_t = X_t^N$. The covariance matrix is then updated based on X_t^N according to :

$$\Sigma_t = \Sigma_t^- - K_t^N \cdot H_t^N \cdot \Sigma_t^- \qquad (13)$$

The iteration could be stopped when consecutive values X_t^k and X_t^{k+1} differ by less than a defined threshold.

III. Robust 2D-3D Lines Matching Algorithm

In this section we explain the expansion of the RANSAC scheme that we have developed in order to achieve a robust matching algorithm that associates 2D edges from the image with the 3D line segments from the input model, and without using any verification algorithm.

Let $\{l_i\}, i = 1,\ldots,N$ be a set of 2D edges extracted from the image and $\{L_j\}, j = 1,\ldots,M$ a set of 3D model lines. Our robust lines matching algorithm is summarized as follows:

1. Randomly sample subsets of four {li ↔ L_j} pairs of 2D and 3D lines. In theory a minimum three pairs of lines are sufficient to compute an accurate rigid transformation.

2. For each sample, compute the camera pose Π(R,T) using the IEKF algorithm described in section II.

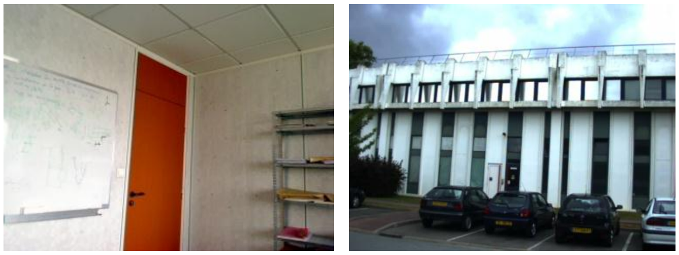
3. Each candidate Π is tested against all the correspondences $l_i \rightarrow L_j$ by computing, in the camera frame, the angle between the normal vector $\vec{n}_i$ (see figure 1) associated with the image line l_i and the transformed line $R \cdot L_j$. If this match is wrong with respect the pose Π, then the co sinus of the angle should be significantly larger than zero.

4. We choose the pose Π which has the highest number of inliers, i.e the Π for which all the pairs are within a fixed angle threshold.

Hence, the obtained camera pose for the current image is robustly updated using only inliers of correspondences.

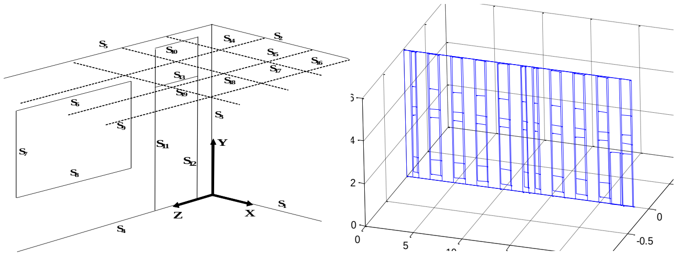
IV. Experimental Results

The proposed camera pose estimation algorithm has been tested in real scenes and the registration accuracy was analyzed. The results are presented for both indoor and outdoor images (Figure 2). The indoor scene consists of a camera moving in an office room whereas the outdoor one corresponds to a moving camera pointing towards one frontage of a building. The frame rate of the recorded image sequences is about 25 frames/s and the resolution of the video images is 320×240 pixels. The 3D models of the office room and the building frontage are known, they are composed, respectively, of 19 lines and 120 line defined by the 3D coordinates of their end points within the world coordinates frame (Figure 3).



(a) (b)

Figure. 2. Two frames from the recorded images sequence. (a) indoor scene. (b) outdoor scene.



(a) (b)

Figure. 3. 3D models of indoor and outdoor scenes used for experiments. (a) the office room model. (b) The frontage building model.

In order to estimate the camera pose accuracy we defined, in the camera space, the registration error ξ.

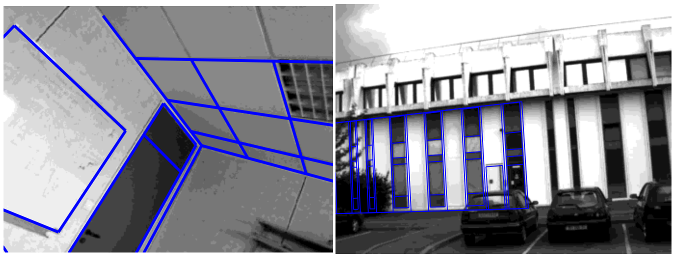
Given a set of correspondences between image edges and model segments the error ξ corresponds to the normalized square sum of the sinus of the angular disparities α_i for each correspondence between image edge and the re projected model segments (Figure 1):

$$\xi = \frac{1}{M}\sum_{i=1}^{M} \sin(\alpha_i)^2 = \frac{1}{M}\sum_{i=1}^{M} \left\| \vec{n}_i \times \vec{N}_i \right\|^2 \qquad (14)$$

Where M is the number of correspondences and α_i the angle between the two planes spanned by the camera center, the observed image edge l_i and the model segment L_i (see figure 1).

In our experiment we took M=10 correspondences. We have first considered that the data set has no outliers (100 % inliers or good matching) and we computed the registration error for several frames of the image sequence. Our algorithm was run on both the indoor and outdoor scenes. Similar results were obtained in both situations.

Thus, for the indoor scenario, the mean error is about $\xi_m = 4.01 \times 10^{5}$ which corresponds to the mean angular disparities $\alpha_m = 0.28^\circ$. For the outdoor scene, we obtained $\xi_m = 3.43 \times 10^{5}$ and $\alpha_m = 0.28^\circ$. Figure (4) shows the projection of the office model using the camera pose estimated by our algorithm, and as can see it is quite skewed. All the lines are fairly well aligned.



(a) (b)

Figure 4.. Projection of the 3D models using the final camera pose estimate when the input data has no outliers. (a) indoor scene (b) outdoor scene

In the second experiment, we have evaluated the capacity of our robust algorithm to reject outliers in observed data. In our case, an outlier corresponds to a wrong feature matching between 2D and 3D line.

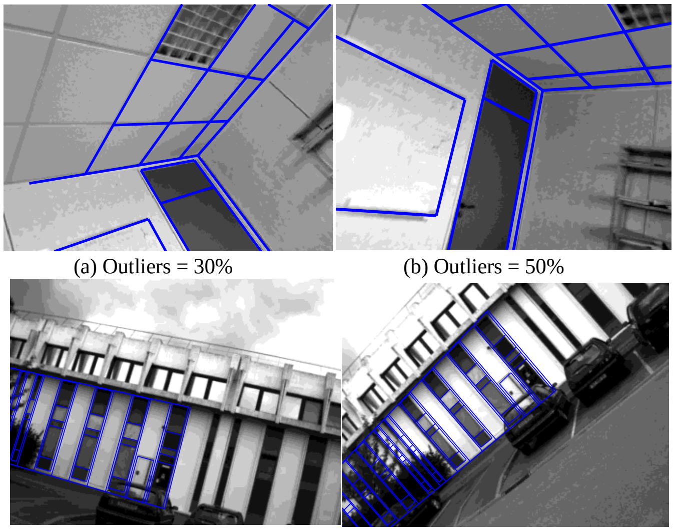
For that, we have contaminated the input data set with different percentage of outliers, for both indoor and outdoor scenes, and have computed the corresponding registration error. The obtained results are summarized in table 1.

TABLE I : EXPERIMENTAL RESULTS OF THE ROBUST ALGORITHM

Indoor scene			
Outliers (%)	ξ_m	α_m(°)	Number of trials
30%	4.93×10^{-5}	0.37	4
40%	6.95×10^{-5}	0.39	12
50%	7.84×10^{-5}	0.42	50
60%	8.18×10^{-5}	0.44	240
Outdoor scene			
Outliers (%)	ξ_m	α_m(°)	Number of trials
25 %	3.81×10^{-5}	0.34	4
37 %	5.57×10^{-5}	0.37	6
50 %	8.64×10^{-5}	0.44	56

As can be seen, our robust algorithm succeeded in all the cases to detect and delete the outliers. The camera pose is then estimated using the final data consensus which contains only the good 2D-3D lines correspondences. For example, in the worst case when 60% of the input data for the indoor scene (i.e. 6 lines correspondences among 10) are labeled as outliers, our algorithm was been able to identify the four inliers in the data. The camera pose returned using this inliers gives a registration error about 8.18×10^{-5}.

This result demonstrates the robustness and the accuracy of the proposed approach. Furthermore, we note that the number of trials needed to get the best solution increase with number of outliers (for example 240 trials for 60% of outliers for the indoor scenario and 56 trails for 50% of outliers for the outdoor scene). This means more processing time and will decrease the real time performance of the algorithm. 40% of outliers (15 trials on average) represents a good compromise. Figure 5 shows the projection of the models using the pose estimated by our algorithm for different frames of the indoor and outdoor images sequence and when run with different percentage of outliers. We can see that all lines are well aligned, and in this cases, the outliers have not affected the registration performance of our robust algorithm.



(a) Outliers = 30% (b) Outliers = 50%

(c) Outliers = 37% (d) outliers = 50%

Figure. 5. Camera pose estimation results

Another advantage of our approach is its robustness to severe lines occlusion. Indeed, as the line constraint equation (see section II-B) for the cameras pose parameters was developed in the case of “infinite image line”. Any image points on the 2D line can be used to construct the corresponding projection plane. So, when partial occlusion occurs, it is enough to detect only small parts of the image edges to estimate the camera pose. In figure 6 we can see that several image edges are partially occluded (table and door for the indoor scene and because of the trees and the cars for the outdoor scene), in spite of that, the camera pose was successfully estimated.

We analyzed the processing time needed for camera pose estimation on a Pentium IV with 3GHz. All computations were performed in Matlab. The pose estimation process using IEKF does not take much time. Indeed, we have tuned the parameters of the IEKF in such manner so that it converges in a few iterations (20 at the maximum).

The processing time strongly depends only on the number of outliers in the current field of view. For example, the average time is about 28 millisecond per frame when having 40% of outliers in 10 input data for indoor scene. 3 milliseconds are used to estimate the camera pose with the IEKF and 25 milliseconds are measured for the time needed to reject outliers.

We have also evaluated the 3D camera localization accuracy. For that, we considered the outdoor scene, and we compared the computed camera trajectory obtained with our algorithm with the ground truth data. The ground truth for position is provided by a Trimble Pathfinder ProXT GPS receiver giving a submeter accuracy, The covered distance is about 30 meters. The ground truth for rotation was obtained using an Xsens MTx gyro mounted rigidly with the camera [32].

The results of the experiments are given in figure 6 which shows the error between the ground truth data and the recovered position and orientation of the camera for different frames of the image sequence. Table 2 reports the mean values and the standard deviation of the position and the orientation errors. It can be seen that our algorithm allows a good 3D localization performance, especially for orientation components.

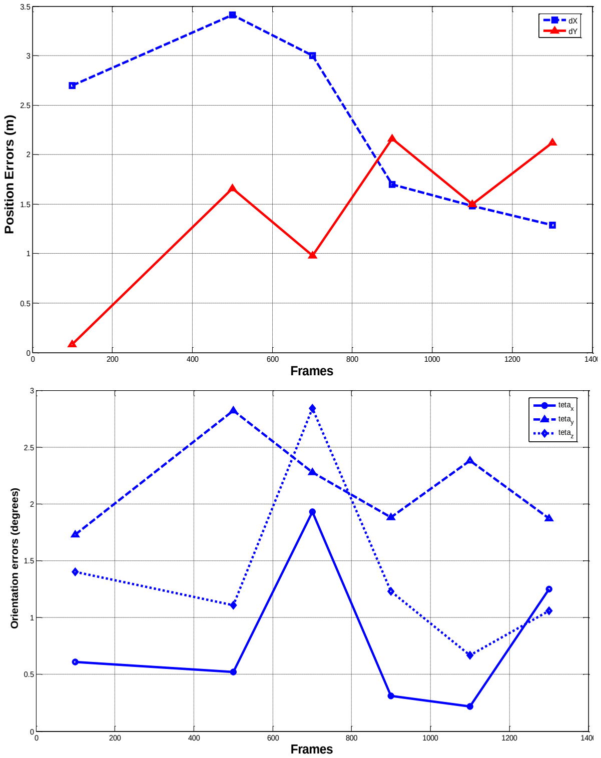


Figure. 6. Position and orientation accuracy with respect to ground truth

TABLE 2 : 3D LOCALIZATION PERFORMANCE

	Position errors (m)		Rotation errors (deg)		
	ΔX	ΔY	θ_x	θ_y	θ_z
Mean	2.26	1.42	0.81	2.61	1.38
Std	0.88	0.77	0.66	0.41	0.75

V. CONCLUSION

In this paper, we proposed a new approach for 6-DOF camera localization based on matching between 2D image edges and 3D model segments. We performed a generic camera pose estimation framework based only on lines features using an Iterated Extended Kalman Filter. We also achieved significant improvements on robust 2D/3D lines matching scheme by adapting the well-know RANSAC algorithm to our application. The experimental results confirm the robustness of our approach against severe occlusion and outliers for both indoor and outdoor applications. We also showed the good performance of our algorithm to localize a moving camera in 3D environment. Future work will be devoted to extend our approach by using other outdoor sensors (e.g. an inertial sensor and a GPS). Thus, the system could be used for navigation and localization in large-scale outdoor environments. An hybrid algorithm will the fuse the data provided by the three sensors in order to refine the camera pose.

REFERENCES

[1] M. Haller, M. Billinghurst, and B. Thomas, "*Emerging Technologies of Augmented Reality : Interfaces and Design*". Idea Group Publishing, USA, 2007.

[2] F. Ababsa, M. Maidi, J-Y. Didier, and M. Mallem, *Vision-based Tracking for Mobile Augmented Reality*. Studies in Computational Intelligence (SCI). Berlin: Spriger-Verlag, 2008, ch. 12.

[3] S. Dasgupta, and A. Banerjee, "An augmented-reality-based real-time panoramic vision system for autonomous navigation". *IEEE Trans.* Systems, Man and Cybernetics, Part A, vol. 36, no 1, pp. 154 – 161, 2006

[4] D.G. Lowe, "Fitting Parameterized Three-Dimensional Models to Images". *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, pp. 441-450, 1991.

[5] R.M. Haralick, "Pose Estimation From Corresponding Point Data". *IEEE Trans. Systems, Man, and Cybernetics*, vol. 19, no°6, pp. 1426-1446, 1989.

[6] D. F. DeMenthon and L. S. Davis, "Model-based object pose in 25 lines of code". *International Journal of Computer Vision*, vol. 15, pp. 123–141, 1995.

[7] J Wang, H Zha, and R Cipolla, "Coarse-to-Fine vision-based localization by indexing scale-invariant features". *IEEE Trans. on System, Man, Cybernetics, Part B*, vol. 36, no. 2, pp. 413-422, 2006.

[8] H. Wuest, F. Vial, and D. Stricker, "Adaptive Line Tracking with Multiple Hypotheses for Augmented Reality". In *Proc. of ACM/IEEE Int. Symp. on Mixed and Augmented Reality (*ISMAR 2005*)*, Vienna, Austria, 2005, pp. 62-69.

[9] T. Drummond and R. Cipolla, "Real-Time Visual Tracking of Complex Structures," *IEEE Trans. Patt. Anal. and Mach. Intell*, vol. 24, no. 7, pp. 932-946, July 2002.

[10]A.I. Comport, E. Marchand, M. Pressigout, and F. Chaumette, "Real-time markerless tracking for augmented reality: the virtual visual servoing framework". *IEEE Trans. on Visualization and Computer Graphics*, vol. 12, no. 6, pp. 615-628, July/August 2006.

[11]F. Ababsa and M. Mallem, "Robust camera pose estimation combining 2D/3D points and lines tracking". In *Proc. of the 2005 IEEE International Symposium on Industrial Electronics (ISIE'08)*. Cambridge, UK, 2008, pp. 774-779.

[12]O. Koch and S. Teller, "Wide-Area Egomotion Estimation from Known 3D Structure". In *Proc. of the 2007 IEEE International Conference on Vision and Pattern Recognition (CVPR'07)*, Minneapolis, USA, 2007, pp.1-8.

[13]W.J. Wilson, C.C. Hulls, and G.S. Bell, "Relative End-Effector Control Using Cartesian Position Based Visual Servoing". *IEEE Trans. Robotics and Automation*, vol.12, pp.684-696, 1996.

[14]Y.K.Yu, K.H.Wong and M.M.Y.Chang, "Recursive three dimensional model reconstruction based on Kalman filtering". *IEEE Trans. Syst., Man, Cybern. Part B*, vol. 35, no. 3, pp. 587-592, Jun. 2005.

[15]Y. K. Yu, K. H. Wong, and M.M.Y. Chang, "Recursive Camera-Motion Estimation With the Trifocal Tensor". *IEEE Trans. Syst., Man, Cybern. Part B*, vol. 36, no. 5, pp. 1081-1090, Oct. 2006.

[16]Y. Yoon, A. Kosaka, J. B. Park and A. C. Kak, "A New Approach to the Use of Edge Extremities for Model-based Object Tracking". In *Proc. of the 2005 IEEE International Conference on Robotics and Automation (ICRA'05)*. Barcelonna, Spain, 2005, pp. 1883-1889.

[17]V. Lippiello, B. Siciliano, and L. Villani, "Adaptive Extended Kalman Filtering for Visual Motion Estimation of 3D Objects". *Control Engineering Practice*, Vol. 15, pp. 123-134, 2007.

[18]B. Jiang, and U. Neumann, "Extendible Tracking by Line Auto-Calibration". In *Proc. of ACM/IEEE Int. Symp. on Augmented Reality (*ISAR 2001*)*, New-York, USA, 2001, pp. 97-103

[19]B. Williams, P. Smith, and I. Reid, "Automatic Relocalisation for a Single-Camera Simultaneous Localisation and Mapping System". In *Proc. of the 2007 IEEE International Conference on Robotics and Automation (ICRA'07)*. Roma, Italy, 2007, pp. 2784-2790.

[20]A. J. Davison, I. D. Reid, N. D. Molton and O. Stasse, “MonoSLAM: Real-Time Single Camera SLAM”, *IEEE Trans. Patt. Anal. and Mach. Intell*., vol. 2, no. 6, pp. 1052-1067, June 2007.

[21]T. Q. Phong, R. Horaud, A. Yassine and P. D. Tao,”Object Pose from 2D to 3D Point and Line Correspondences”. *International Journal of Computer Vision*, vol. 15, pp. 225-243, 1995.

[22]C. P. Lu, G. Hager, and E. Mjolsness. “Fast and globally convergent pose estimation from video images”. *IEEE Trans. Patt. Anal. and Mach. Intell*, vol. 22, no 6, pp. 610–622, June 2000.

[23]R. Kumar, and A.R. Hanson, “Robust Methodsfor Estimating Pose and a Sensitivity Analysis”. Computer Vision and Image Understanding, vol. 66, pp.313-342,1994.

[24]M. Dhome, M. Richetin, J.T. Lapreste, and G. Rives, “Determination of the Attitude of 3D Objects from Single Perspective View”. *IEEE Trans. Patt. Anal. and Mach. Intell*, vol. 11, pp. 1265-1278, 1989.

[25]M. A. Fischler and R.C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography,” *Comm. ACM*, vol. 24, no. 6, pp. 381- 395, June 1981.

[26]B.K. Horn, “Closed-Form Solution of Absolute Orientation Using Quaternions”. *J. Optical Soc. Am. A*, vol. 4, pp.629-642, 1987.

[27] R. E Kalman, "A New Approach to Linear Filtering and Prediction Problems". *Transaction of the ASME--Journal of Basic Engineering*, pp. 35-45, March 1960.F.L. Lewis, “*Optimal Estimation”*. John Wiley & Sons, Inc., New York, 1986.

[28]O. Faugeras, “Three-Dimentional Computer Vision: a Geometric Viewpoint”. MIT Press, 1993.

[29]Y. Bar-Shalom, and X.R. Li, “Estimation and Tracking: Principles, Techniques, and Software”. Artech House, Boston, MA, 1993.

[30]B.M. Bell, “The Iterated Kalman Smoother as a Gauss-Newton Method,” SIAM J. on Optimization, Vol. 4, 1994, p. 626.

[31] I.M. Zendjebil, F. Ababsa, J-Y. Didier, and M. Mallem, “On the Hybrid Aid-Localization for Outdoor Augmented Reality Applications”. In proc. Of ACM Symposium on Virtual Reality Software and Technology (VRST’08), Bordeaux, France, 2008, pp. 249-250.

Multi-input Multi-output Beta Wavelet Network: Modeling of Acoustic Units for Speech Recognition

Ridha Ejbali, Mourad Zaied and Chokri Ben Amar
REsearch Groups on Intelligent Machines, University of Sfax
Sfax, Tunisia

Abstract—**In this paper, we propose a novel architecture of wavelet network called Multi-input Multi-output Wavelet Network MIMOWN as a generalization of the old architecture of wavelet network. This newel prototype was applied to speech recognition application especially to model acoustic unit of speech. The originality of our work is the proposal of MIMOWN to model acoustic unit of speech. This approach was proposed to overcome limitation of old wavelet network model. The use of the multi-input multi-output architecture will allows training wavelet network on various examples of acoustic units.**

Keywords-wavelet network; multi-input multi-output wavelet network MIMOWN; speech recognition; modeling of acoustic units; wavelet network.

I. INTRODUCTION

The development of robust systems for speech recognition is now one of the main issues of language processing. Many systems were based on concepts that were relatively close. However, despite those progresses, the performances of current systems remain much lower than the capacity of the human auditory system. In this respect, the prospects for improvement are strong. Economic issues related to speech recognition make this area a sector in constant evolution and pushes us to constantly imagine new approaches in order to be able to at least match or exceed the performance of our hearing.

To solve some problems of modelling and recognition of speech, we suggest a new method based on wavelet networks [21] that exploit the intrinsic properties of the speech signal, while the statistical models are concerned with statistical properties of the speech signal [2]. These models are similar to neural network for the structure and the training approach. But, training algorithms for wavelet network require a smaller number of iterations when compared with neural network.

Wavelet network model, the single-input single-output wavelet network, was introduced, firstly, by Zhang and Benveniste in 1992 [8]. This model was inspired from neural network architecture as a combination of neuronal contraption and wavelets as activation functions [20]. Those models have supplied an access to all frequency of signal thanks to the use of wavelets in the hidden layer of each neurone. It has more advantages than common networks such as faster convergence, avoiding local minimum, easy decision and adaptation of structure [4].

Despite the contribution of these models in different fields of pattern recognition [4], [5], [7], they remained limited in the field of modelling. These prototypes cannot instil entities at different occurrences. To overcome these limitations, we advanced a new model for the training of several instances of a single entity at the same time. These models are called multi-input multi-output wavelet network MIMOWN [23][24].

The idea of this architecture was presented for the first time by Zhao in [23]. The training algorithm of this model is identical to the training algorithm of the old version of wavelet networks. It is based on the forward backward algorithm. Identically to the old version, this approach takes as input a unitary random vector and according to an error defined by the characteristics of the networks, it estimate the original vector output.

Gutierrez in [25] has used the MIMOWN for the Voltammetric Electronic Tongues. In his works, the architecture is implemented with feed forward one-hidden layer architecture. In the hidden layer, he uses wavelet functions. The network training is performed using a back propagation algorithm, adjusting the connection weights along with the network parameters.

In [22], Fengqing has presented other architecture of wavelet network called multi-input single-output wavelet network. In his work, he has shown that his architecture can eliminate the useless wavelet on the network when training a data.

An improvement was made to the approach of Zhao [23] in this article. Improvement concerns the training algorithm. In this case, the input vector of the network corresponds to the vector model and not a random unitary vector. The new architecture allows training several multivalve examples.

The proposed algorithm car train a MIMOWN on original vector and not on unitary random vector as in [23][24][25].

Thinks to those models, the training system can model a variety of occurrences of a single entity by a single acoustic wavelet network [22]. The newel approach can be seen as a superposition of finite number of single-input single-output wavelet network.

Those prototypes are used to model acoustic unit of speech to be used on speech recognition system.

In this paper, we propose a newel approach to model acoustic units of speech for the recognition of isolated words based on MIMOWN. These new models are similar to multilayer neural network for the structure and the training approach. Compared to old wavelet network structure, those prototypes can represent several acoustic units by one network.

Our systems, based on wavelet network have adaptable parameters; dilation, translation, and weights. Initial values of weights are made randomly. Then, parameters are optimized automatically during training [6]. To update parameters, gradient method has been applied by using momentum. Quadratic cost function is used for error minimization.

This paper includes four parts. In the first part, we present the techniques used to construct the newel prototype MIMOWN. In the second part, we detail the training approach of MIMOWN. The third part presents the recognition part of our system. The finale part illustrates the results that crown our approach.

II. Technical Background

A. Beta wavlet

According to works published in [17], [18] and [19] Beta function is a wavelet and from it we can build a network called Beta wavelet.

It satisfies the properties of a wavelet namely:

- ✓ Admissibility : $0 < C_\psi = 2\pi \int \frac{\|\hat{\psi}(w)\|^2}{\|w\|} dw < \infty$
- ✓ Localization: wavelet is a function of $L^2(IR)$ with the property of location; it decreases rapidly on both sides of its domain.
- ✓ Oscillation : wavelet is a function of $L^2(IR)$, integrated and oscillating enough to have zero as integral: $\grave{o}\, y(t)dt = 0 \ \hat{U} \ TF(y(0)) = 0$
- ✓ The translation and dilation : The wavelet analysis involves a family of copies of itself, translates and dilated: $y_{ab}(t) = \frac{1}{\sqrt{a}} y(\frac{t - b}{a})$ *with a,b* $\hat{I}$ *IR,a* > 0

B. Wavelet Network

The field of wavelet network is new, although some attempts have recently been conducted to build a theoretical basis and several applications in various fields. The use of wavelet network began with the use of Gabor wavelets in classification and image recognition [1], [3], [8].

As shown in the figure below, wavelet network can be seen as a neural network with wavelet in the hidden layers [9], [10], [11]:

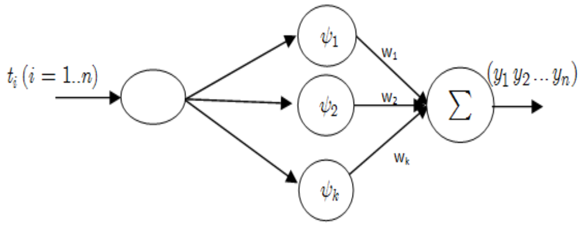


Figure 1. Example of a figure caption.

The reconstruction formula of a signal after processing by a continuous wavelet is given by:

$$f(x) = \frac{1}{C_\psi}\int_{IR_+}\int_{IR} W_f(a,b)\frac{1}{\sqrt{a}}\psi\ (\frac{x-b}{a})dadb \qquad (1)$$

This formula gives the expression of the signal f in the form of an integral over all possible translations and dilations of a mother wavelet ψ . Suppose we have only a finite number N_w wavelet ψ_j obtained from the mother wavelet ψ . We can then reformulate the equation (1) as following:

$$f(x) \approx \sum_{j=1}^{N_w} w_j \psi_j(x) \qquad (2)$$

We can consider the expression of the equation (2) as a decomposition of the function f to a sum of weights w_i and wavelets y_i .

To define a wavelet network, we start with taking a family of N_w wavelets $y = \{y_1 ... y_{N_w}\}$ with different parameters of dilatation and translation that can be chosen randomly at this point.

The old architecture of wavelet networks can learn more examples but this example must be mono-valued. That's why we thought to extend this architecture for new models of wavelet network that can learn several examples which are multi-valued.

C. Idea of multi-input multi-ouput wavelet network

1) Architetcure: The old version of wavelet network has certain proximity of the multilayer perceptron network MLP architecture with one hidden layer. The essential difference between these two networks is the nature of transfer functions used by the hidden cells.

The new proposed architecture is a wavelet network with multi-input multi-output shown in figure 2. This network is seen as the superposition of a finite set of wavelet network with one input and one output.

2) Limits: The superposition of three patterns (a), (b) and (c) constitute the new MIMOWN. This new architecture will, unlike the old, take as input the vector of learning and not an impulse vector. In addition to the shift from a single-input single-output architecture provides a new model to multiple outputs.

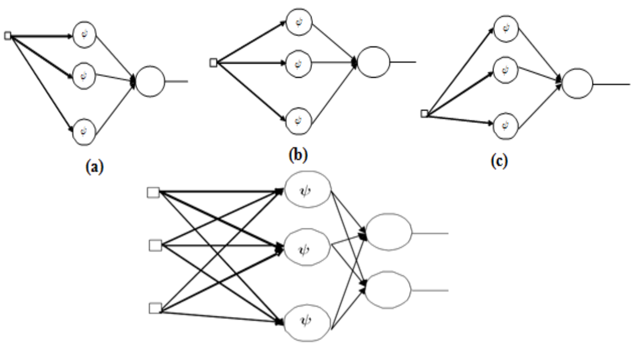


Figure 2. Building of multi-input multi-output network of wavelet.

D. Trainig of a milti-input multi-outout wavelet network

1) Introduction: In this section, we will present Zhao architecture [23] and our architecture considered as generalization of the training algorithm of Zhang's networks to train MIMOWN.

2) Approach proposed by Zhao to train multi-input multi-output wavelet network: In his work, Zhao has changed the functions of activations of a neural network with wavelet. For training architecture, Zhao uses the method of wavelet network learning single-input single-output.

The following figure illustrates the Zhao wavelet network architecture:

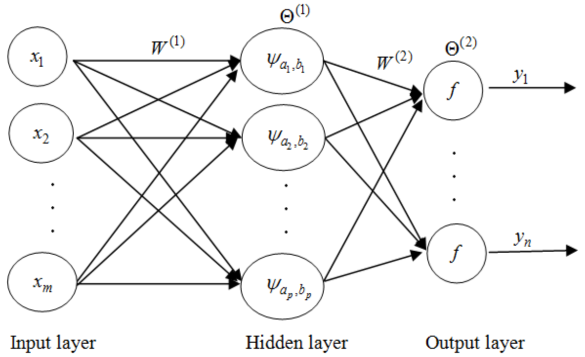


Figure 3. Zaho architecture of wavelet network.

The training phase starts with the preparation of unitary random vectors of size equal to the base vector. These vectors constitute the input of the network. The calculation of the weights of networks is based on a calculated distance between the output vector of the network and the original vector of training. The error between these two vectors is determined by the characteristics of the network such as the number of wavelets in the hidden layer of the network. If the error set is reached, the learning process stops otherwise it will change the parameters of the wavelets in the hidden layer to improve the value of the output vector.

3) Mathematical concepts:

✓ *Periodic Function*

A function f is periodic if there is a strictly positive real T such that $\forall x \in D_f$, we have

$$x+T \in D_f \text{ and } f(x+T) = f(x) \qquad (3)$$

We call the period of the function f the smallest real T verifying the property **(3)**. If f is a function of period T, then "x Î D_f, we have $f(x+kT) = f(x)$ with k Î IN.

The study interval of a periodic function can be reduced to an interval covering a single period.

✓ *Taylor formula for polynomial functions*

P is a polynomial function, if degP $=$ n with n $\in IN^*$. Using the Taylor formula for the polynomial functions, we have

$$P(x) = \sum_{k=0}^{n} \frac{P^{(k)}(x_0)}{k!}(x-x_0)^k = P(x_0) + P'(x_0)(x-x_0) + P''(x_0)\frac{(x-x_0)^2}{2} + ... + P^{(n)}(x_0)\frac{(x-x_0)^n}{n!} \qquad (4)$$

If p f 0 and q f 0

$$\beta(x) = \begin{cases} \left(\frac{x-x_0}{x_c-x_0}\right)^p \left(\frac{x_1-x}{x_1-x_c}\right)^q & \text{if } x \in]x_0, x_1[\text{ with } x_c = \frac{px_1-x_0}{p+q} \\ 0 & \text{else} \end{cases} \qquad (5)$$

Applying (4) for the Beta function (polynomial for p and q integers)[16]

$$\begin{aligned} \beta(x) &= \sum_{k=0}^{n} \frac{\beta^{(k)}(x_0)}{k!}(x-x_0)^k \quad \text{with n=p+q} \\ &= \sum_{k=0}^{n} \frac{Q_k(x_0)\beta(x_0)}{k!}(x-x_0)^k \quad \text{using (5)} \\ &= \left(\sum_{k=0}^{n} \frac{Q_k(x_0)}{k!}(x-x_0)^k\right)\beta(x_0) \end{aligned} \qquad (6)$$

4) General idea of the algorithm: As shown in figure 4, it is a multi-layer wavelet network.

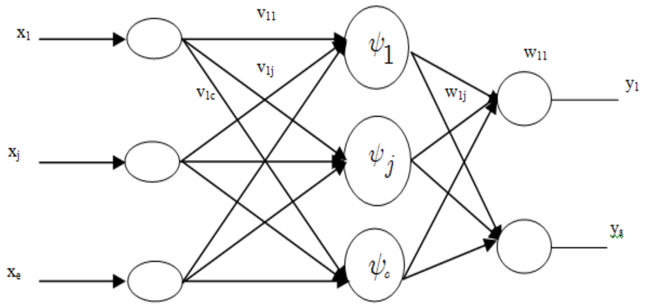


Figure 4. Multi-input multi-ouput wavelet network.

The network presented in the previous figure is constructed by a superposition of networks Zhang. For signal: $Y = (y_1, y_2, ..., y_s)$ The network of Zhang expressed as follows:

$$(I)\begin{cases} y_1 = \alpha_1\psi_1(A_1) + \alpha_2\psi_2(A_1) + ... + \alpha_c\psi_c(A_1) \\ .... \\ y_s = \alpha_1\psi_1(A_s) + \alpha_2\psi_2(A_s) + ... + \alpha_c\psi_c(A_s) \end{cases} \qquad (7)$$

While the MIMOWN expressed by:

$$(II)\begin{cases} y_1 = w_{11}\psi_1(A_1) + w_{12}\psi_2(A_2) + ... + w_{1c}\psi_c(A_c) \\ \text{..} \\ y_s = w_{s1}\psi_1(A_1) + w_{s2}\psi_2(A_2) + ... + w_{sc}\psi_c(A_c) \end{cases} \qquad (8)$$

With A_i the entry of the neuron number i, α_i represent the weights of the connections of the network of Zhang and w_{ij} are the weights of connections of the MIMOWN.

In each scale, we can use a single wavelet instead of 2^{j-m} because:

On the scale m in the library, there are 2^{j-m} wavelets $(\psi_i)_{1\le i\le 2^{j-m}}$. The graphical representations of wavelet scales are adjacent and have the same size. We suppose that they form a single function, so the latter is periodic with period T. The size, the support, of wavelet is $T = \frac{N}{2^{j-m}}$ with $j = \log_2(N)$ is the number of scales to cover the whole signal. N is the size of the signal.

We have $\forall\ 1 \le i \le j \le 2^{\text{j-m}}$ for $\text{x''-x'=(j-i)T}, \psi_i(x') = \psi_j(x'')$

For a signal of size N = 256, we can analyze up to the scale N = 8 (j = 8). In the 6th scale, there are 4 wavelets.

III. Training process

Let $X = (x_1, x_2, ..., x_e)$ the input vector, $\psi = (\psi_1, \psi_2, ..., \psi_c)$ the vector containing the wavelet to be used in the hidden layer and $Y = (y_1, y_2, ..., y_s)$ the output vector.

- ✓ For the weights connecting the input layer and hidden layer $(V_{ij})_{\substack{1\le i\le e \\ 1\le j\le c}}$

We calculated $V_{i,j}$ in order to have as input of the jème wavelet $A_j = j$.

Hence $V_{i,j} = \frac{j}{\text{length of the signal*X(i)}}$

- ✓ For the weights connecting the hidden layer and the input layer $(W_{ij})_{\substack{1\le i\le c \\ 1\le j\le s}}$

Let g be a signal, f_i is a family of functions that form a basic family of dual functions $\tilde{f}_i$ then there exist α_i weights such as:

$$g = \sum_i \alpha_i f_i \qquad (9)$$

Since the wavelet used is Beta (is a polynomial for p and q integers) using Taylor's theorem, from equation **(9)**, we can write $\psi_i(A_j) = c_{ij}\psi_i(A_i)$ with $c_{ij} = \sum_{i=1}^{p+q} \frac{Q_k(A_i)}{k!}(A_j - A_i)^k$.

$$(I')\begin{cases} y_1 = \alpha_1 c_{11}\psi_1(A_1) + \alpha_2 c_{12}\psi_2(A_2) + ... + \alpha_c c_{1c}\psi_c(A_c) \\ \text{..} \\ y_s = \alpha_1 c_{s1}\psi_1(A_1) + \alpha_2 c_{s2}\psi_2(A_2) + ... + \alpha_c c_{sc}\psi_c(A_c) \end{cases} \qquad (10)$$

And by identifying with the system (II) we have: $w_{ij} = \alpha_i c_{ij}$

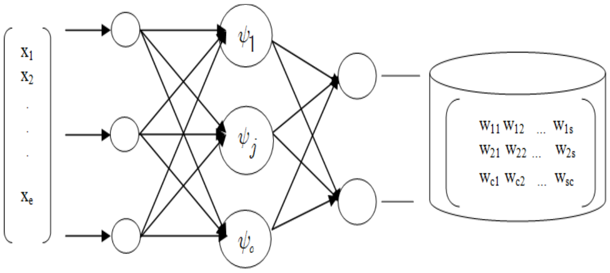


Figure 5. Training phase.

The performance of approximation networks is the basis of parameters estimated by measuring the mean squared error, expressed by the following formula:

$$EQM = \frac{1}{M * N} \sum_{i=1}^{N} \sum_{j=1}^{M} (A(i,j) - B(i,j))^2 \qquad (11)$$

Such as A and B represent the coefficients of input and output network, while M and N dimensions.

IV. Recognition process

During the recognition phase, the system will construct a new vector for each element of the test. Every element of the test will be the entry of all networks-based learning. The system will modify the weights of the network learning approach to maximize the input test vector.

At the end of the phase, we will save the weight from each network after approximating the test vector.

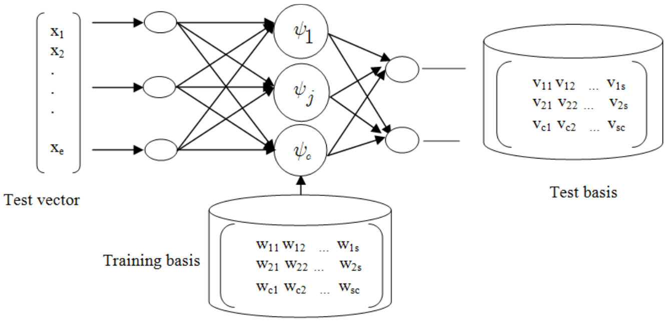


Figure 6. Reocgnition phase.

V. Corpus

We have tested our modelling approach on two corpuses:

The first corpus was recorded by 12 speakers (6 women and 6 men) from works of (Boudraa and Boudraa, 1998) [12]. We segmented this corpus manually by PRAAT to Arabic words and we have chosen 19 different words. We have added sound effects for each word to test robustness of our approach. Finally, we got 912 words. The second corpus was recorded by 14 speakers. It was about Tunisian city name (24 cities). We have added the same sound effects to each word. Finally, we got 1344 words.

We have chosen Mel-Frequency Cepstral Coefficients MFCC and Perceptual Linear Predictive (PLP) coefficients to represent acoustic data. Our choice of acoustic representation is based on the specificities of these types.

- ✓ MFCC coefficients contain features around Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT) converted on Mel scale. This method used mostly to represent a signal in speech recognition, because of its robustness. Its advantage is that the coefficients are uncorrelated.
- ✓ PLP technique aims at estimating the parameters of an autoregressive filter and all-pole modelling to better auditory spectrum. It takes into account many limitations and characteristics of human speech production and hearing to reduce a number of direct waveform samples into a few numbers that represent the perceived frequency concentrations and widths.

VI. Tools

We used PRAAT to record the speech and to handle signals and words segmentation. Recognition system of Arabic word using HMM has been implemented with the HTK library version 4.0 [13]. Recognition system of Arabic word using WN was implemented using MATLAB.

VII. Results

To test the new architecture of modelling of acoustic units of speech, we have implemented a recognition system of Arabic word [4]. The results of this approach, MIMOWN, of modelling was compared to results given by a recognition system of Arabic words based on Hidden Markov Models HMM [14], [15] and the old architecture of Wavelet Network WN. The following figures illustrate the recognition rate of the three modelling approaches using the same type and number of MFCC coefficient describing units.

Figure 7 shows the recognition rate of the three modeling approaches applied to the corpus of Tunisian cities names on MFCC coefficients.

According to Figure 7, we deduce that our new approach based on MIMOWN has produced good results compared to WN and HMM based system.

The recognition rates of the same systems are also compared using PLP coefficients. Those results are illustrated on figure 8.

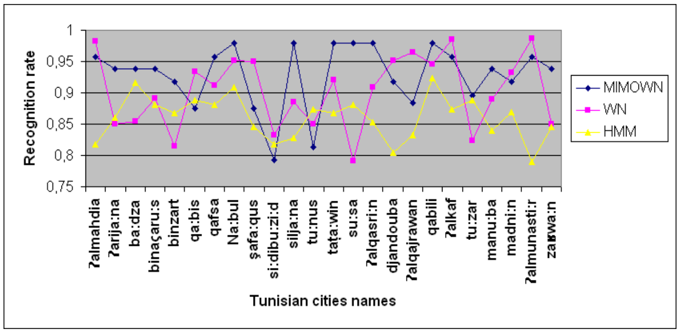


Figure 7. Recognition rates of the Tunisian cities names

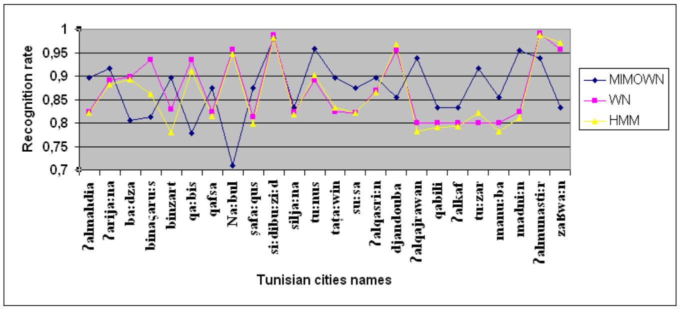


Figure 8. Recognition rates of the Tunisian cities names

According to figure 8, we can say that our approach gives good results comparing to the old wavelet network and HMM approach.

Figure 9 shows the recognition rate of the three modelling approaches applied to the corpus of Arabic words on MFCC coefficients.

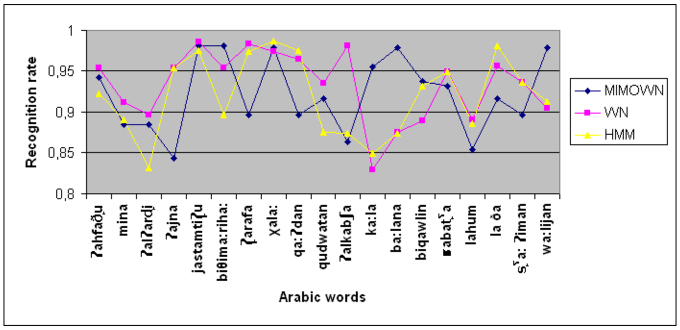


Figure 9. Recognition Rates of Arabic words

According to figure 9, the results given by multi-input multi-output approach are comparable to the wavelet network and hidden Markov model approach.

Increasing the number of training data over networks will refine multi-input multi-output. Each new sound unit will add a set of nodes in the hidden layer networks thus a better identification of test units.

Figure 10 illustrate recognition rate of the tree systems on PLP coefficients.

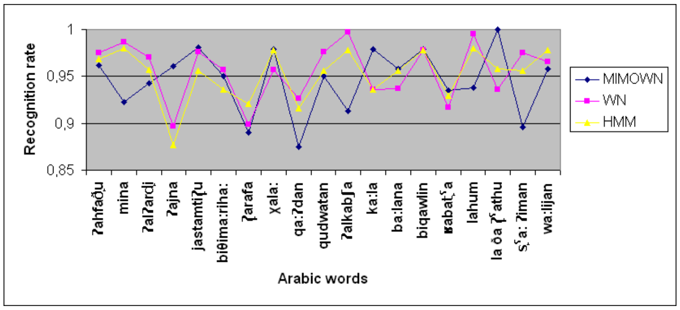


Figure 10. Recognition Rates of Arabic words

Based on the previous figures, we can realize that our approach based on MIMOWN gave better results compared to the approach based on WN and HMM.

These results can be improved by increasing the size of the training base, getting a better quality of recordings, choosing other types of wavelet…etc.

This new architecture can learn several examples with multiple values. It will allow the modeling of instances of an acoustic unit by a wavelet network. Thanks to this architecture, each unit will be represented by a matrix of weight. The updating of the training base will not require the creation of a new network for each new entry, but the update of the associated network. In case the new unit does not exist in the database, it is evident to create a new model that is network wavelet.

VIII. Conclusion

Given the success of the old architecture of wavelet networks, single-input single-output architecture, it was interesting to study the possibility of generalizing this prototype to remedy the gaps of the old version. This work focuses on suggesting a new architecture of wavelet networks, MIMOWN. This new architecture has been proposed for modelling the acoustic units of speech. This new prototype will allow the modelling of several examples. It takes as input the coefficients of each unit and not a vector pulse.

Tests made on cities words corpus and Arabic word corpus show that our approach based on MIMOWN gives good results compared to WN and HMM model based system.

Interpreting the findings of this new modelling technique, we came to the conclusion that its adoption in large vocabulary applications and real application will improve performance of the speech recognition system

Acknowledgment

The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

References

[1] C. Ben Amar and O. Jemai, "Wavelet networks approach for image compression," ICGST International Journal on Graphics, Vision and Image Processing, SI1 37--45 (2005).

[2] J.T. Chien and C.H. Chueh, "Joint acoustic and language modeling for speech recognition, " Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 70101, Taiwan, ROC, Speech Communication xxx (2009) xxx–xxx.

[3] J. Daugman, "Demodulation by complex-valued wavelets for stochastic pattern recognition," Int'l Journal of Wavelets, Multi-resolution and Information Processing, vol. 1, no. 1, pp 1-17 (2003).

[4] R. Ejbali, M. Zaied and C. Ben Amar, "Wavelet network for recognition system of Arabic word," International Journal of Speech Technology Volume 13 Number 3p.163--174 (2010).

[5] R. Ejbali, M. Zaied and C. Ben Amar, "Intelligent approach to train wavelet networks for Recognition System of Arabic Words," KDIR International Joint Conference on Knowledge Discovery and Information Retrieval, Valencia Spain, 25-28 October 2010, 518--522 (2010).

[6] M. Zaied, O. Jemai and C. Ben Amar, "Training of the Beta wavelet networks by the frames theory: Application to face recognition," The international Workshops on Image Processing Theory, Tools and Applications, Tunisia November (2008).

[7] M. Zaied, C. Ben Amar and M.A. Alimi, "Beta Wavelet Networks for Face Recognition," Journal Of Decision Systems - New Trends in the Design of Intelligent Decision Systems, vol. 14 pp 109 à 122 (2005).

[8] Q. Zhang and A. Benveniste, "Wavelet Networks," IEEE Trans. on Neural Networks, 889-898 (1992).

[9] Q. Zhang, "Using wavelet network in nonparametric estimation," IEEE Trans. Neural Networks, 227- 236 (1997).

[10] Z. Zhang, "Learning algorithm of wavelet network based on sampling theory," Neurocomputing , 244-269 (2007).

[11] Z. Zhang, "Iterative algorithm of wavelet network learning from non uniform data,".Neurocomputing, 2979-2999 (2009).

[12] M. Boudraa, B. Boudraa, "Twenty list of ten arabic Sentences for Assessment," ACUSTICA acta acoustica. Vol. 86, no. 43.71, pp. 870-882, (1998).

[13] S. Young and al. "The HTK Book (for HTK version 3.3)," Cambridge University Engineering Department, (2005).

[14] H. Bahi, A. Benouareth and M. Sellami, "Application of HMMs for Arabic speech recognition," Proceeding of Maghreb conference MCSEAI'2000, Fes, Maroc, pp. 379-388, (2000).

[15] H. Bahi and M. Sellami, "Combination of vector quantization and Hidden Markov Models for Arabic speech recognition," Proceeding ACS/IEEE International conference on computer systems and applications, Beirut, Liban, pp. 96-100, (2001).

[16] C. Ben Amar, M. Zaied and M.A. Alimi, "Beta wavelets. Synthesis and application to lossy image compression," Journal of Advances in Engineering Software, Elsevier Edition, Vol. 36, N7, PP 459 - 474. (2005).

[17] W. Bellil, C. Ben Amar and M.A. Alimi, "Beta wavelet networks for function approximation," 7th International Conference on Adaptative and Natural Computing Algorithms ICANNGA'05, Coimbra, Portugal, March 21-23, , Riberio et al. (eds), Springer Computer Science, SpringerWien NewYork, vol. 1, pp 18-21. (2005).

[18] O. Jemai, M. Zaied, C. Ben Amar and M.A. Alimi, "Pyramidal Hybrid Approach: Wavelet Network with OLS Algorithm Based-Image Classification," International Journal of Wavelets, Mutiresolution and Information Processing,Vol. 9, No. 1, pp. 111-130, (2011).

[19] O. Jemai, M. Zaied, C. Ben Amar and M.A. Alimi, "FBWN: an architecture of Fast Beta Wavelet Networks for Image Classification," 2010 IEEE World Congress on Computational Intelligence (IEEE WCCI 2010), the 2010 International Joint Conference on Neural Networks (IJCNN 2010), p. 1953-1960, July, 18-23, CCIB, Barcelona, Spain, (2010).

[20] S.S. Iyengar, E.C. Cho and V. Phoha, "Foundations of Wavelet Networks and Applications," Chapman and Hall/CRC Press, June (2002).

[21] S. Postalcioglu and Y. Becerikli, "Nonlinear System Modelling Using Wavelet Networks," Lecture Notes in Computer Science (LNCS), Vol.3497, pp.411-417, June, (2005).

[22] H. Szu, B. Telfer and S.Kadambe, “Neural network adaptative wavelets for signal representation and classification,” Optical Engineering 31:1907-1961, (1992).

[23] J. Zhao, W. Chen and J. Luo, “Feedforward Wavelet Neural Network and Multi-variable Functional Approximation,” CIS 2004: 32-37.

[24] A.B. Stephen and W.Hua-Liang, “A New Class of Wavelet Networks for Nonlinear System Identification,” IEEE Transactions on Neural Networks vol. 16, no. 4, July 2005.

[25] J.M. Gutièrrez, L. Moreno-Baron, M. Del Valle, L. Leija and R. Munoz, ” Wavelet Neural Network as a Multivariable Calibration Method in Voltammetric Electronic Tongues,” Neural Network World 1/09, January 15, 2009, 53-64.

[26] H. Fengqing, W. Dacheng and L. Jianping, “Research for multi-input wavelet neural network,” International Computer Conference on Wavelet Active Media Technology and Information Processing, 2006, Vol 2, 664-669.

AUTHORS PROFILE

Ridha Ejbali is a PhD student in the REsearch Group on Intelligent Machines, National Engineering School of Sfax, university of Sfax Tunisia. He received the Master degree of science from the National Engineering School of Sfax (ENIS) in 2006. He obtained the degree of Computer Engineer from the National Engineering School of Sfax (ENIS) in 2004. He was assistant technologist at the Higher Institute of Technological Studies, Kebili Tunisia since 2005. He joined the faculty of sciences of Gabes (FSG) where he is an assistant in the Department computer sciences since 2012. His research area is now in pattern recognition and machine learning using Wavelets and Wavelet networks theories. He is IEEE Graduate Student Member, SPS society and Assistant in the Faculty of Sciences of Gabes Tunisia.

Dr. Mourad ZAIED received the Ph.D degree in Computer Engineering and the Master of science (DEA : Diploma in Higher Applied Studies) from the National Engineering School of Sfax (ENIS) respectively in 2008 and in 2003. He obtained the degree of Computer Engineer from the National Engineering School of Monastir (ENIM) in 1995. Since 1997 he served in several institutes and faculties in the university of Gabes as a teaching assistant. He joined the National Engineering School of Gabes (ENIG) where he is an assistant professor in the Department of Electrical Engineering since 2007. His research area is now in pattern recognition and machine learning using Wavelets and Wavelet networks theories. He is an IEEE member and SMC society member.

Prof. Chokri BEN AMAR received the B.S. degree in Electrical Engineering from the National Engineering School of Sfax (ENIS) in 1989, the M.S. and PhD degrees in Computer Engineering from the National Institute of Applied Sciences in Lyon, France, in 1990 and 1994, respectively. He spent one year at the University of "Haute Savoie" (France) as a teaching assistant and researcher before joining the higher School of Sciences and Techniques of Tunis as Assistant Professor in 1995. In 1999, he joined the Sfax University (USS), where he is currently an associate professor in the Department of Electrical Engineering of the National Engineering School of Sfax (ENIS), and the Vice director of the Research Group on Intelligent Machines (REGIM). His research interests include Computer Vision and Image and video analysis. These research activities are centered on Wavelets and Wavelet networks and their applications to data Classification and approximation, Pattern Recognition and image and video coding, indexing and watermarking. He is a senior member of IEEE, and the chair of the IEEE SPS Tunisia Chapter sinc 2009.

Real-Time Fish Observation and Fish Category Database Construction

Yi-Haur Shiau
Data Computing Division
National Center for High-Performance Computing
Hsinchu, Taiwan

Sun-In Lin
Data Computing Division
National Center for High-Performance Computing
Hsinchu, Taiwan

Fang-Pang Lin
Data Computing Division
National Center for High-Performance Computing
Hsinchu, Taiwan

Chaur-Chin Chen
Department of Computer Science
National Tsing Hua University
Hsinchu, Taiwan

***Abstract*—This paper proposes a distributed real-time video stream system for underwater fish observation in the real world. The system, based on a three-tier architecture, includes capture devices unit, stream processor unit, and display devices unit. It supports variety of capture source devices, such as HDV, DV, WebCam, TV Card, Capture Card, and video compression formats, such as WMV, FLV/SWF, MJPEG, MPEG-2/4. The system has been demonstrated in Taiwan for long-term underwater fish observation. CCTV cameras and high-definition cameras are deployed on our system. Video compression methods and image processing methods are implemented to reduce network transfer flow and data storage space. Marine ecologists and end users can browse these real-time video streams via the Internet to understand the ecological changes immediately.**

These video data is preserved to form a resource base for marine ecologists. Based on the video data, fish detection is implemented. However, it is complicated in the unconstrained underwater environment, due to the water flow causes the water plants sway severely. In this paper, a bounding-surrounding boxes method is proposed to overcome the problem. It efficiently classifies moving fish as the foreground objects and the swaying water plants as the background objects. It enables to remove the irrelevant information (without fish) to reduce the massive amount of video data. Moreover, fish tracking is implemented to acquire multiple species of fish images with varied angles, sizes, shapes, and illumination to construct a fish category database.

Keywords-Real-time streaming; Fish observation; Fish detection; Distributed architecture.

I. INTRODUCTION

Video stream over the Internet is a hot research topic recently. It can broadcast live events from a server, over the Internet, to end users. In recent years, owing to the advance of video stream technology and the booming of network bandwidth, live video stream is getting more and more popular. In this paper, we develop a distributed unmanned underwater video stream system for the long-term fish observation [1, 2, 3]. CCTV and high-definition cameras are set up as test cases that are installed on the Southern-most coast of Taiwan. Presently, real-time video streams are accessible online via the Internet broadcasting. Worldwide marine ecologists and end users can now perform comparative studies between sites and attempt to understand the behavior of fish. The system facilitates marine ecologists to closely observe the ecosystem of fish, and understand the immediate phenomena of the underwater environment. It enables to enhance the public's awareness of the marine conservation. [4].

Although many applications for object detection and tracking have been proposed, application in uncontrolled conditions, i.e. in real-life underwater systems, remains a challenge [5]. Fish detection and tracking is complicated by the variability of the underwater environment. The water plants may be regarded as foreground objects as result of the severe sway from interference of the water flow, which is able to result in the complexities and difficulties to discriminate moving fish and swaying water plants. In this paper, we propose a bounding-surrounding boxes method, which effectively achieves the purpose that classifies moving fish as the foreground objects and swaying water plants as the background objects. Then, we implement the object tracking method for multiple species of fish from the stored video data to acquire fish images with varied angles, sizes, shapes, and illumination. Furthermore, we construct a fish category database by using image resizing method to let all of fish images with the same resolution.

This paper is organized as follows: Section 2 describes the distributed underwater observation system architecture details. Fish category database construction method is presented in Section 3. Section 4 shows the implemental results and the conclusion is drawn in Section 5.

II. DISTRIBUTED REAL-TIME UNDERWATER VIDEO STREAM SYSTEM

A. Distributed System Architecture

In this paper, a distributed real-time underwater video stream system is developed. The system is loose coupling and three-tier architectures that includes capture devices unit, stream processor unit, and display devices unit. Figure 1 illustrates the distributed video stream system architecture and stream pipeline.

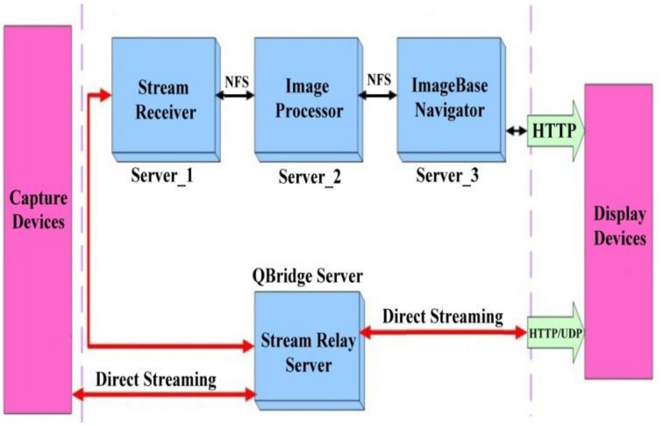


Figure 1. Architecture blocks and stream pipeline.

The left part of Figure 1 is capture devices unit. It receives signals from multiple capturing devices, such as HDV, DV, Webcam, TV Card, and automatically identifies formats of the signals. Identification of signal formats, video information and conversion is implemented by modified the functions of VideoLAN Client (VLC) [6] and FFMPEG.

The received signal can be converted to multiple video formats, such as MJPEG, MPEG-2/4, SWF/FLV, WMV, and multiple bit-rates for different bandwidths [7]. The benefit is it doesn’t have to bind the specific hardware devices and the video formats. Table 1 shows the relationship between video formats and the corresponded video displayers.

TABLE I. THE RELATIONSHIP BETWEEN THE VIDEO FORMAT AND THE VIDEO PLAYER

Video format	Video player
WMV	Window media player
FLV/SWF	Wimpy player
MJPEG	Axis plug-in
MPEG-2/4	VLC media player

The center part of Figure 1 is stream processor unit. This unit is in charge of post-processing of the video stream and two modes are supported. One is direct streaming to display devices unit and there is a stream relay server that bridges the video stream between in unicast and multicast. The other is the video stream is stored for further implementing image processing methods, such as object detection and tracking.

The right part of Figure 1 is display devices unit. It supports multiple display devices handy to end users. Two protocols, HTTP and UDP, are adapted to transmit streaming to display devices. Figure 2 shows multiple display devices, include web-based interface, mobile display interface, Google Earth based interface, and a 4x3 Tiled Display Wall (TDW), a versatile, large, and high-resolution display system that was constructed by National Center for High-Performance Computing (NCHC) [8].

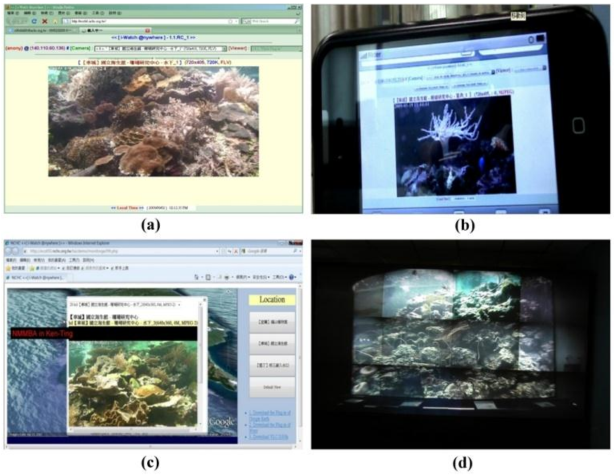
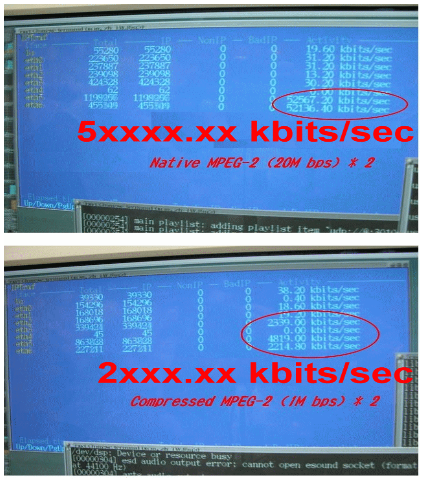


Figure 2. Multiple display interfaces (a) web-based, (b) mobile, (c) Google Earth based, and (d) TDW.

B. Video Compression Methods

The quantities of acquired raw data of these video streams can approach 1-2 gigabyte per hour. The massive amount of video data is huge for network transfer and storage space. Two video compression concepts are implemented to reduce the network transfer flow and data capacity. After receiving signals, a compress method that converts video streams to different bit-rates is implemented for decreasing network traffic. Meanwhile, the raw video data is converted to a variety of compressive video formats such as MJPEG, MPEG-2/4, SWF/FLV, and WMV. Figure 3 shows the comparison with/without using video compression methods. The top image shows the native MPEG-2 bit-rate and the bottom image shows the compressed MPEG-2 bit-rate. The compressed ratio advances to about 25 times.



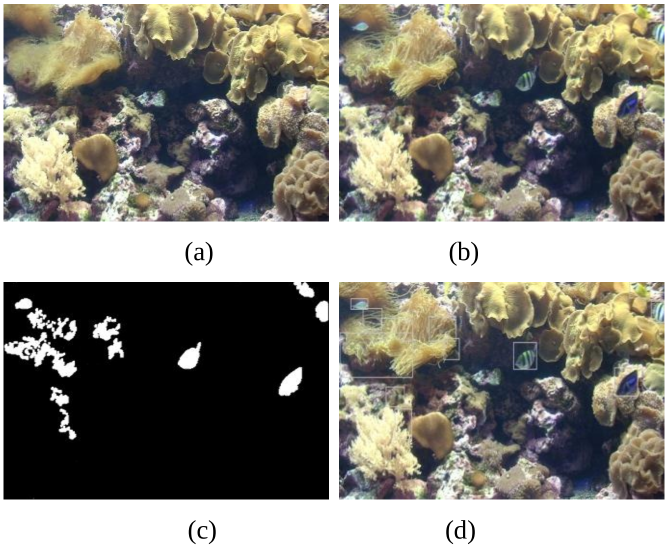
```
IPTraf
 Iface ───── Total ───── IP ── NonIP ── BadIP ── Activity
 lo          55280      55280     0       0       19.60 kbits/sec
 eth0       223650     223650     0       0       31.20 kbits/sec
 eth1       237887     237887     0       0       31.20 kbits/sec
 eth2       239098     239098     0       0       13.20 kbits/sec
 eth3       424328     424328     0       0       30.20 kbits/sec
 eth4           62         62     0       0        0.00 kbits/sec
 eth5      1199256    1199256     0       0    52567.20 kbits/sec
 eth6       455909     455909     0       0    52136.40 kbits/sec
 Elapsed ti
 Up/Down/PgUp
[00000254] main playlist: adding playlist item `udp://@:2010
[00000254] main playlist:
```

```
IPTraf
 Iface ───── Total ───── IP ── NonIP ── BadIP ── Activity
 lo          39330      39330     0       0       38.20 kbits/sec
 eth0       154296     154296     0       0        0.40 kbits/sec
 eth1       168018     168018     0       0       18.60 kbits/sec
 eth2       168696     168696     0       0       19.20 kbits/sec
 eth3       339424     339424     0       0     2339.00 kbits/sec
 eth4           45         45     0       0        0.00 kbits/sec
 eth5       863828     863828     0       0     4819.00 kbits/sec
 eth6       227211     227211     0       0     2214.80 kbits/sec
 Elapsed ti
 Up/Down/PgUp
/dev/dsp: Device or resource busy
[00000304] esd audio output error: cannot open esound socket (format
at 44100 Hz)
[00000304] arts audio
```

Figure 3. The bandwidth with/without video compression methods.

III. FISH CATEGORY DATABASE CONSTRUCTION

For the stored video data, background subtraction [9, 10, 11, 12], foreground segmentation and object tracking methods are implemented for fish detection and tracking. In this paper, Gaussian Mixture Matrix (GMM) method is adopted for background subtraction [13]. The highest color histogram similarity and the shortest distance are used for feature extraction to track the foreground objects. Figure 4(a) shows the background model and the current frame is illustrated in Figure 4(b). Figure 4(c) illustrates the foreground objects and Figure 4(d) shows the bounding boxes of these foreground objects.

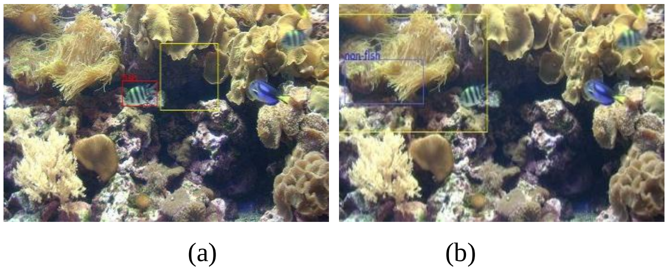

(a) (b)

(c) (d)

Figure 4. (a) The background model (b) the current frame (c) the foreground objects (d) the bounding boxes of foreground objects.

A. Bounding-Surrounding Boxes Method

The underwater environment in the real world is unconstrained, owing to the interference of the water plants sway severely. It raises the difficulty and complexity to discriminate moving fish and swaying water plants. However, the water plants always sway in a fixed field, but fish can free move to anywhere. Based on the concept, we propose a bounding-surrounding boxes method to discriminate fish as the foreground objects and water plants as the background objects. The foreground object is circumscribed by its bounding box with width w_1 and height h_1. Let (c_x, c_y) be the center point of the bounding box and the upper-left point is $(c_x\text{-}0.5{*}w_1, c_y\text{-}0.5{*}h_1)$. Then, the surrounding box is set to T times the size of the bounding box with the same center point. Let B_t and S_t be the bounding box and surrounding box observed at time t. The location of S_t is fixed in the image, and the location of bounding box of the object is observed in a period of time τ. If the location of the bounding box from time t to time t+τ is always inside the range of S_t, the object is classified as a non-fish object (water plants). It is not only identified as a background object, but also eliminated from the tracked object. On the other hand, if the location of the bounding box has left the range of S_t, the object is classified as a foreground object (fish). The detecting results are shown in Figure 5. The yellow box represents the fixed surrounding box of the object. The red box in Figure 5(a) represents the object is classified as "fish", and the blue box in Figure 5(b) represents the objects is classified as "non-fish" object.



(a) (b)

Figure 5. (a) The object (red box) is classified as fish (b) the object (blue box) is classified as non-fish (water plant).

B. Image Resizing Method

In order to for further implementing fish recognition and verification, the image size of each detected fish must be identical. In this case, the width and height of the resized image is the maximum width W_{Max} and height H_{Max} of all the images of mulitple species of fish. In order to avoid the deformation of the images, we resized the fish images by pasting the original image to a $W_{Max} \times H_{Max}$ black image, furthermore, the center of the black image was aligned. The result of an image resizing is shown in Figure 6.

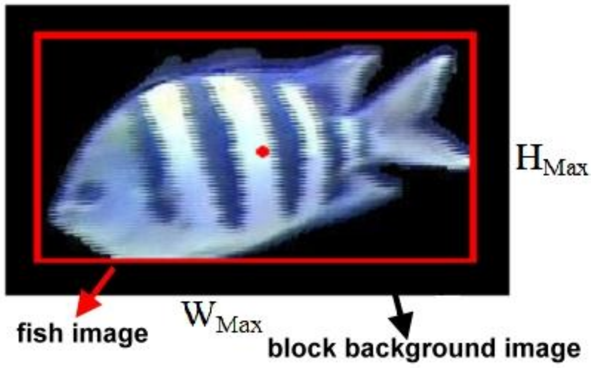


Figure 6. The result using image resizing method.

After fish detection is implemented using our proposed method, we can only record the video data that contains fish and remove the irrelevant information (without fish) to reduce the stored data volume. We also acquire multiple species of fish images with varied angles, sizes, shapes, and illumination. For each species of fish, we select some images that are almost different to construct a fish category database in the real world.

IV. IMPLEMENT RESULTS

Some cases are implemented to test the above-mentioned distributed real-time underwater video stream system. It is briefly introduced as followed.

A. Underwater Fish Observation With CCTV Cameras

Kenting, located on the southernmost tip of Taiwan, is famous for its abundant marine resources. Setting up cameras here can help marine ecologists observe fish behavior and the hydrological environment. Figure 7 illustrates the distributed architecture of the real-time underwater video stream system in Kenting.

It includes four underwater CCTV cameras with the resolution of 640x480 pixels and a sensor of water temperature and pressure. A video server that converts analog video signals into digital video streams is installed within a steel casing located on the dike. It delivers video streams to a media server via optical network with wireless network as a back-up. The stream data is transferred back to NCHC's multicasting pool, located 300km North in Hsinchu, through four ADSL lines. Figure 8 illustrates the real-time underwater video streams that the left side is the four CCTV cameras and the right side is the water temperature. These long-term continuous recordings can help marine ecologists in elucidating the ecological processes, and the real-time underwater observation system is able to enhance public's awareness of marine conservation.

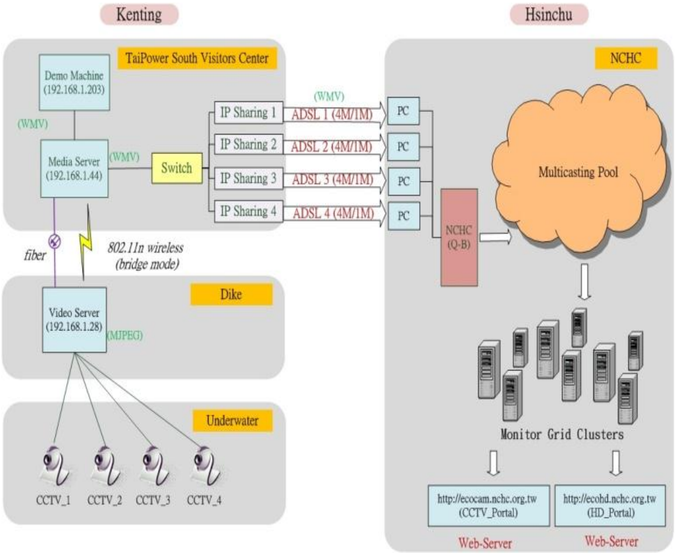
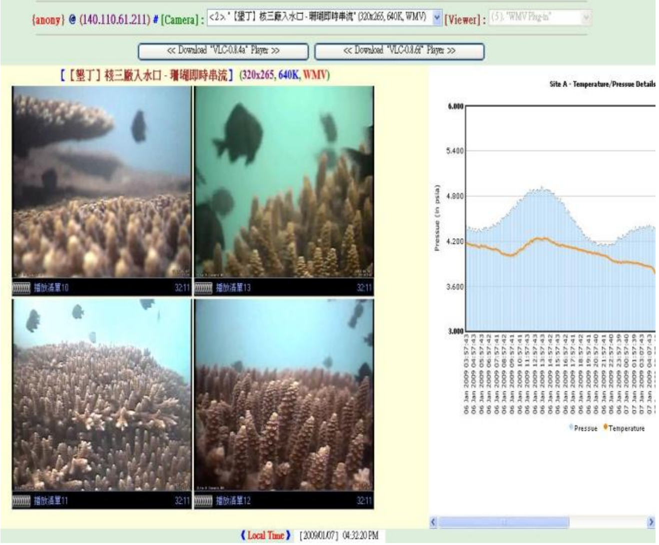


Figure 7. The distributed architecture of the real-time underwater video stream system in Kenting.


{anony} @ (140.110.61.211) # [Camera] : <2>."【墾丁】核三廠入水口 - 珊瑚即時串流" (320x265, 640K, WMV) [Viewer] : (5). "WMV Plug-in"

【【墾丁】核三廠入水口 - 珊瑚即時串流】 (320x265, 640K, WMV)

【Local Time】 [2009/01/07] 04:32:20 PM

<< Download "VLC-0.8.4a" Player >>
<< Download "VLC-0.8.6f" Player >>
Site A - Temperature/Pressue Details
6.000
5.400
4.800
4.200
3.600
3.000
Pressure (in psia)
Pressure
Temperature

Figure 8. The real-time underwater video streams with CCTV cameras.

B. Underwater Fish Observation with High-Definition Cameras

In order to support marine ecologists to get more detail data, the high-definition case is adopted. Two high-definition cameras with the resolution of 1280x1080 pixels are located on two different sites inside a fairly large lagoon in Taiwan. We set up a waterproof case to protect the high-definition camera, 1394 repeater and optical fiber. The challenge of this case is network bandwidth and to decode a high-definition video stream in real-time. The video compression method is implemented to compress the raw data from 20mb to 4mb and 1mb per second. Our experimental result shows that the proposed distributed real-time video stream system is successful for high-definition camera with only about 3~5 seconds latency. Figure 9 illustrates the distributed architecture of the high-definition real-time underwater video stream system. Figure 10 shows the high-definition real-time underwater video stream.

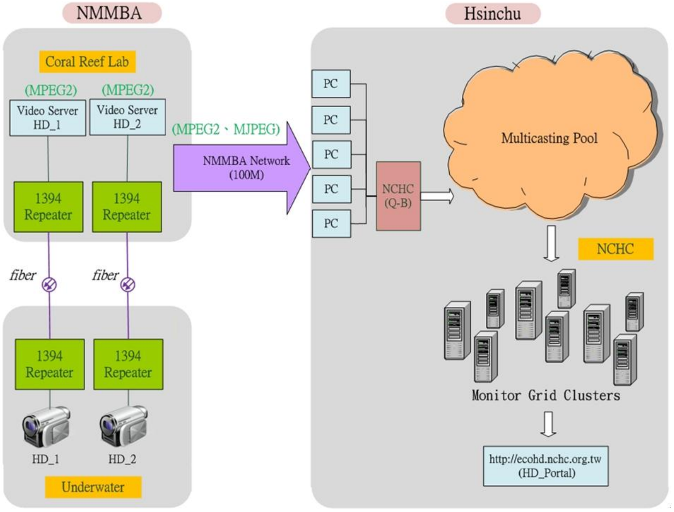
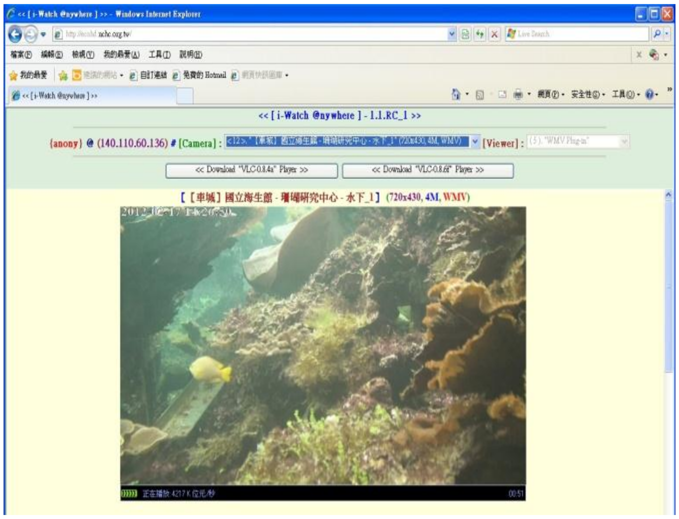


Figure 9. The distributed architecture of the high-definition real-time underwater video stream system.


<< [i-Watch @nywhere] >> - Windows Internet Explorer

http://ecohd.nchc.org.tw/

<< [i-Watch @nywhere] - 1.1.RC_1 >>

{anony} @ (140.110.60.136) # [Camera] : <12>."【車城】國立海生館 - 珊瑚研究中心 - 水下_1" (720x430, 4M, WMV) [Viewer] : (5). "WMV Plug-in"

【【車城】國立海生館 - 珊瑚研究中心 - 水下_1】 (720x430, 4M, WMV)

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)
我的最愛
自訂連結
免費的 Hotmail
<< [i-Watch @nywhere] >>
網頁(P) 安全性(S) 工具(O)
Live Search
<< Download "VLC-0.8.4a" Player >>
<< Download "VLC-0.8.6f" Player >>

Figure 10. Figure 10. The real-time underwater video stream with high-definition cameras.

C. Fish Category Database Construction

After implementing our proposed method, we enabled to obtain multiple species of fish images with varied angles, sizes, shapes, and illumination to construct a fish category database. The fish category database that we constructed is composed of 1,000 fish images of 180 rows and 130 columns with JPEG file format. Totally, there are 25 different species of fish. Each one contributed 40 images. The 5 images of 6 species of fish are illustrated in Figure 11. The total 40 images of 2th species of fish are illustrated in Figure 12. The fish category database can further for fish recognition and verification.

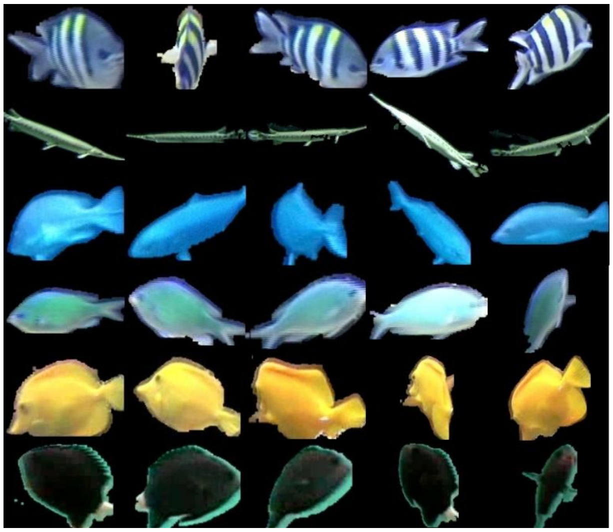


Figure 11. The 5 images of 6 species of fish.

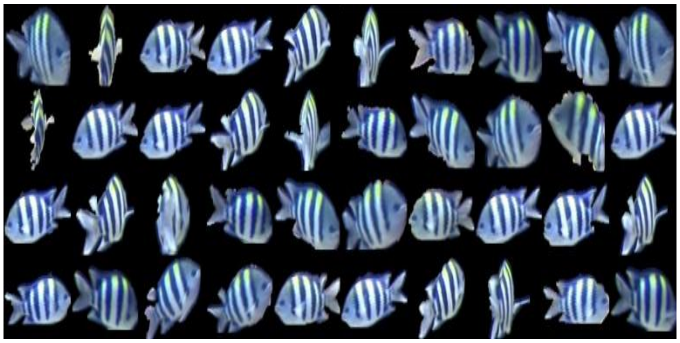


Figure 12. Examples of total 40 images of 2th species of fish.

V. Conclusions

In this paper, a distributed architecture for real-time underwater video stream system was developed. The system had been demonstrated in Kenting and NMMBA, Taiwan, for long-term fish observation in the real world. Four CCTV cameras and two high-definition cameras were set up to test our developed video stream system. The distributed servers are located on Kenting and NMMBA, and video streams are transmitted back to NCHC's multicasting pool for real-time observation. Experimental results showed that the proposed distributed video stream system is robust, adaptive, and powerful.

In this paper, a bounding-surrounding boxes method had been proposed to reduce the stored video data capacity. It efficiently discriminated moving fish as the foreground objects and swaying water plants as the background objects. Then, it enabled to remove the irrelevant information (without fish) and only save the data containing fish. It reduced the massive amount of the video data greatly. After that, we implemented fish tracking to acquire multiple species of fish images with varied angles, sizes, shapes, to construct a fish category database.

Acknowledgment

This research was funded from Taiwan Power Company. We thank the Third Nuclear Power Plant of Taipower for logistical support. We also thank the National Museum of Marine Biology & Aquarium and Ecogrid team at National Center for High-Performance Computing for video stream data supply.

References

[1] H. M. Chou, Y. H. Shiau, S. W. Lo, S. I. Lin, F. P. Lin, C. C. Kuo, and C. L. Lai, "A Real-Time Ecological Observation Video Streaming System based on Grid Architecture," In HPC Asia, Taiwan, 2009.

[2] Y. H. Shiau, J. S. Cheng, S. I. Lin, Y. H. Chen, K. T. Tseng, H. M. Chou, and S. W. Lo, "A Distributed Architecture for Real-Time High-Resolution Video Streaming," International Conference on Parallel & Distributed Processing Techniques & Applications, pp. 345-349, 2009.

[3] Y. H. Shiau, Y. H. Chen, K. t. Tseng, J. S. Cheng, S. I. Lin, S. W. Lo, and H. M. Chou, "A Real-Time High-Resoulution Underwater Ecological Observation Streaming System," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Japan, pp. 517-521, 2010.

[4] E. Strandell, S. Tilak, H. M. Chou, Y. T. Wang, F. P. Lin, P. Arzberger, T. Fountain, T. Y. Fan, R. Q. Jan, and K. T. Shao, "Data management at Kenting's underwater ecological observatory," The third International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Melbourne, Australia, pp. 715-720, 2007.

[5] D. R. Edgington, I. Kerkez, D. E. Cline, J. Mariette, M. Ranzato, and P. Perona, "Detecting, tracking and classifying animals in underwater video," IEEE International Conference on Computer Vision and Pattern Recognition, pp. 634-638, 2007.

[6] H. Nguyen, P. Duhamel, J. Brouet, and D. Rouffet, "Robust vlc sequence decoding exploiting additional video stream properties with reduced complexity," In IEEE International Conference on Multimedia and Expo, Taiwan, pp. 375-378, 2004.

[7] C. Traiperm, and S. Kittitomkun, "High-performance MPEG-4 multipoint conference unit," In Networks and Communication Systems, Thailand, 2005.

[8] S. I. Lin, F. P. Lin, C. Chang, S. W. Lo, Y. C. Mai, P. W. Chen, and Y. Shia, "Development of grid-based tiled display wall for networked visualization," 9th International Workshop on Cellular Neural Networks and Their Applications, pp. 315-318, 2005.

[9] D. S. Lee, J. J. Hull, and B. Erol, "A Bayesian framework for gaussian mixture background modeling," IEEE International Conference on Image Processing, vol.3, pp. 973-976, 2003.

[10] M. Piccardi, "Background subtraction techniques: a review," IEEE International Conference on System, Main, and Cybernetices, pp. 3099-3104, 2004.

[11] V. Mahadevan, and N. Vasconcelos, "Background Subtraction in Highly Dynamic Scenes," IEEE Computer Society Conf. Computer Vision and Pattern Recognition, pp. 1-6, 2008.

[12] C. Stauffer, and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," IEEE Computer Society Conf. Computer Vision and Pattern Recognition, vol. 2, pp. 246-252, 1999.

[13] S. Y. Yang and C. T. Hsu, "Background Modeling from GMM Likelihood Combined with Spatial and Color Coherecny," IEEE International Conference on Image Processing, pp. 2801-2804, 2006.

Intelligent Joint Admission Control for Next Generation Wireless Networks

Mohammed M. Alkhawlani
Faculty of Science and Engineering
University of Science and Technology
Sanaa, Yemen

Fadhl M. Al-Akwaa
Biomedical Engineering Department
University of Science and Technology
Sanaa, Yemen

Abdulqader M. Mohsen
Computer Science Department
University of Science and Technology
Sanaa, Yemen

***Abstract*—The Heterogeneous Wireless Network (HWN) integrates different wireless networks into one common network. The integrated networks often overlap coverage in the same wireless service areas, leading to the availability of a great variety of innovative services based on user demands in a cost-efficient manner. Joint Admission Control (JAC) handles all new or handoff service requests in the HWN. It checks whether the incoming service request to the selected Radio Access Network (RAN) by the initial access network selection or the vertical handover module can be admitted and allocated the suitable resources. In this paper, a decision support system is developed to address the JAC problem in the modern HWN networks. This system combines fuzzy logic and the PROMETHEE II multiple criteria decision making system algorithm, to the problem of JAC. This combination decreases the influence of the dissimilar, imprecise, and contradictory measurements for the JAC criteria coming from different sources. A performance analysis is done and the results are compared with traditional algorithms for JAC. These results demonstrate a significant improvement with our developed algorithm.**

Keywords- Heterogeneous Wireless Network; Radio Access Network; PROMETHEE; Joint Admission Control

I. INTRODUCTION

Heterogeneous Wireless Network (HWN) is defined as a new type of wireless networks where anyone can communicate with anyone else, anywhere and anytime, or enjoy any service of any network operator, through any network of any service provider in the most efficient and optimal way according to the user criteria. The current Radio Resource Management (RRM) solutions and mechanisms for the wireless networks consider only the case of a single Radio Access Technology (RAT) where mobile users can only access that RAT and co-existed sub-networks can only be operated independently. The needs for supporting various applications and services and for providing ubiquitous coverage in the HWN require more complex and intelligent RRM techniques that enable the co-ordination among the different RATs.

Joint Admission Control (JAC) handles all new or handoff service requests in the HWN. It checks whether the incoming service request to the selected RAT by the initial access network selection algorithm or the vertical handover algorithm selection can be admitted. Then, it allocates the required resources and guarantees the QoS constraints for the service. The relationship between JAC and local admission controls of the involved RATs is highly dependent on the level of coupling and type of relationship between the Common RRM (CRRM) entity and RRM entities of the coupled networks [2].

The most important related work of the JAC problem is presented in the next section. A brief overview for PROMETHEE and FLC is presented in section III. The proposed JAC algorithm for HWN environments is presented in Section IV. The simulation models and performance metrics are presented in section V. The performance evaluation of the proposed algorithm is carried out in Section VI. The conclusions and future works are presented in Section VII.

II. RELATED WORK

O. E. Falowo et al. in paper [1] review the recent call admission control algorithms for heterogeneous wireless networks. The benefits and requirements of JAC algorithms are discussed. The authors examine eight different approaches for selecting the most appropriate RAT for incoming calls in HWN and classify the JAC algorithms based on these approaches. The advantages and disadvantages of each approach are discussed. The same authors in [3] propose a JAC algorithm which considers the users preference in making an admission decision and a specific case where the user prefers to be served by the RAT which has the least service cost is modeled and evaluated. In [4] a JAC scheme for multimedia traffic that maximizes the overall network revenue with QoS constraints over coupled WLAN and CDMA cellular network is considered. X. G. Wang et al. [5] propose an adaptive call admission control for integrated cellular and WLAN network. In the proposed scheme, call admission decisions are based on requested QoS and availability of radio resources in the considered RATs. D. Karabudak et al. [6] propose a call admission control scheme for the heterogeneous network using genetic algorithm. The objectives of the scheme are to achieve maximum wireless network utilization and meet QoS requirements. A network capacity policy based joint admission controller is presented by K. Murray et al. [7], [8]. D. Qiang et al. in [9] propose a joint admission control scheme for multimedia traffic that exploits vertical handoffs as an effective tool to enhance radio resource management while guaranteeing handoff users QoS requirements. The network resources utilized by the vertical handoff user are captured by a link utility function. X. Li et al. in [10] propose an efficient joint session admission control scheme that maximizes overall network revenue with QoS constraints over both the WLAN and the TD-SCDMA cellular networks.

In paper [11], the authors propose a call admission control reservation algorithm that takes resource fluctuations into consideration. They consider two types of applications denoted by wide-band and narrow band. The performance of the algorithm is modeled through a queuing theory approach and its main performance measures are compared with a conventional algorithm through simulation. The authors in paper [12], propose an algorithm, which incorporates traditional Admission Control (AC) and Wiener Process (WP)-based prediction algorithms to determine when to carry out access service network gateway relocation. The authors further develop an analytical model to analyze the proposed algorithm. Simulations are also conducted to evaluate the performance of the proposed algorithm.

The main contribution of this paper is the development of a new class of JAC algorithms that are based on hybrid parallel Fuzzy Logic (FL) based decision and PROMETHEE II (Preference Ranking Organization Method for Enrichment Evaluation II) MCDM systems. This class of algorithms represents the first attempt to develop adaptive, flexible, and scalable JAC algorithms that are utilizing the advantages of hybrid parallel FL decision making systems and PROMETHEE II method. FL helps out in reducing the complexity involved on the JAC decision in several ways. First, the data, information, and measurements that have to be taken into account in the JAC are in general very dissimilar, imprecise, contradictory, and coming from different sources. As a result of that, a FL based solution has been thought to be a good candidate for reaching suitable JAC decisions from such imprecise and dissimilar information. Second, JAC solution has to be able to response to the changing conditions of the heterogeneous environments and the accumulated experience of the operators and users. FL based solution is easy to modify by tuning and adjusting the inference rules and membership functions. The application of parallel FL rather than traditional FL achieves more advantages for the JAC solution. The idea of the parallel Fuzzy Logic Control (FLC) reduces the number and complexity of the inference rules used in the FL based solution, which helps out in achieving more scalable solutions. In a very complex and uncertain decision environments, MCDM can sufficiently reduce the uncertainty and doubt about the alternatives and allows a reasonable choice to be made from among them.

III. PROMETHEE AND FLC

The PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluation) method was developed by Brans and Vincke in 1985 [13]. The PROMETHEE I method can provide the partial ordering of the decision alternatives, whereas, PROMETHEE II method can derive the full ranking of the alternatives. In this method, pair-wise comparison of the alternatives is performed to compute a preference function for each criterion. Based on this preference function, a preference index for alternative i over alternative i' is determined. This preference index is the measure to support the hypothesis that alternative i is preferred to alternative i'. The PROMETHEE method can classify the alternatives which are difficult to be compared because of a trade-off relation of evaluation standards as non-comparable alternatives. It is quite different from Analytic Hierarchy Process (AHP) method in that there is no need to perform a pair-wise comparison again when comparative alternatives are added or deleted.

Fuzzy Logic (FL) is a problem solving method based on the theory of fuzzy sets, where variables can have different degrees of membership in different sets. Fuzzy Logic Control (FLC) is based on the principles of FL. FLC is a non-linear control method, which attempts to apply the expert knowledge of an experienced user to the design of a controller. Any FLC system contains three stages, the input stage, the processing stage and the output stage. The input stage maps the real valued numbers into fuzzy sets and defines their membership functions. The processing stage maps the input fuzzy sets into output fuzzy sets by combining a set of IF-THEN rules that represents the human knowledge about the problem. The output stage maps the output fuzzy sets into real valued numbers. Mamdani style fuzzy inference system has been used in our work. The idea behind using a Mamdani style is that the rules of the system can be easily described by the humans in terms of fuzzy variables. Thus we can effectively model a complex non-linear system with common sense rules on fuzzy variables [15].

IV. JAC SOLUTION

A novel JAC algorithm is developed in this section. The algorithm has two main components, the FL based control component and the MCDM component. The input criteria values of the MCDM are the outputs of the FL based control subsystems in the first component. The criteria with more importance to the operator and user can be assigned higher weight. Our algorithm considers five different decision criteria. It consider the Received Signal Strength (RSS), the Signal-to-Noise Ratio (SNR), the Resources Availability (RA), the Service Type (ST), and the Mobile Station Speed (MSS) criteria.

A. FLC Component

Our JAC algorithm contains five FL based subsystems. Each subsystem considers one of the JAC criteria mentioned above. Every subsystem has x output variables, where x is the number of existing RATs. Every output variable describes the probability of acceptance for the admission request in one of the existing RATs. Figure 1 shows a sample for an output variable with its membership functions.

For simplicity, only ST Subsystem is considered as an example. The Figures 2 and 3 show the membership functions of the *DelayReqc* and *RateReqc* input variables. In case of three RATs, the output variables will be *STc1* , *STc2,* and *STc3*

The subsystem has nine rules as shown in Table I.

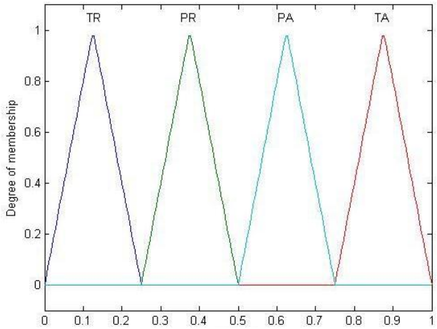


Fig. 1. The output variable

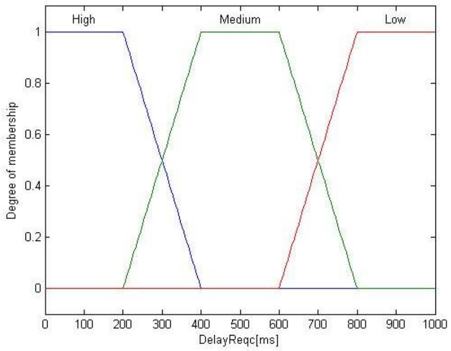


Fig. 2. The input variable *DelayReqc*

TABLE I
THE INFERENCE RULES OF THE ST FUZZY BASED SYSTEM

Rule No.	DelayReqc	RateReqc	STc1	STc2	STc3
1	H	L	TA	TR	PR
2	H	M	PA	PR	PA
3	H	H	PA	PA	PA
4	M	L	PA	PR	PA
5	M	M	PA	PA	PA
6	M	H	PR	PA	PA
7	L	L	PA	PA	PA
8	L	M	TR	TA	TA
9	L	H	TR	TA	PA

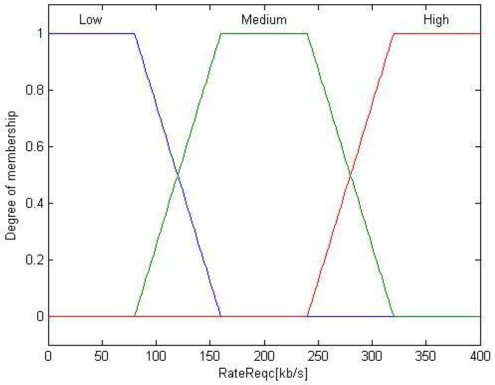


Fig. 3. The input variable *RateReqc*

B. MCDM Component

The MCDM system takes the outputs of the FL subsystems as its input and calculates the total ranking value for all alternatives. The procedural steps as involved in PROMETHEE II method can be summarized as follows [14]:

Step 1: The decision matrix for the different alternatives against the set of criteria can be written as shown in equation 1.

$$NDM = \begin{pmatrix} RSS_{c1} & RSS_{c2} & RSS_{c3} \\ SNR_{c1} & SNR_{c2} & SNR_{c3} \\ RA_{c1} & RA_{c2} & RA_{c3} \\ MSS_{c1} & MSS_{c2} & MSS_{c3} \\ ST_{c1} & ST_{c2} & ST_{c3} \end{pmatrix} \qquad (1)$$

Step 2: Normalize the decision matrix using a suitable normalization method. In our approach, since all the outputs of FL subsystems are in the range [0, 1], there is not any need to scale and normalize the criteria performance against alternatives.

Step 3: the evaluative differences of *ith* alternative with respect to other alternatives are calculated. This step involves the calculation of differences in criteria values between different alternatives pair-wise.

Step 4: the preference function $P_j(i, i')$ is calculated. Many types of generalized preference functions are proposed so far. In our algorithm, the following simplified preference function is adopted.

$$P_j(i,i') = \begin{cases} 0 & \text{if } R_{ij} \leq R_{i'j} \\ (R_{ij} - R_{i'j}) & \text{if } R_{ij} > R_{i'j} \end{cases} \qquad (2)$$

Step 5: the next step is to decide on the relative importance of each of the attributes involved in the decision about admission control. For this purpose, each of the attributes is assigned a specific weight, such that

$$TW = W_{rss} + W_{snr} + W_{ra} + W_{mss} + W_{st} = 1 \qquad (3)$$

where W_{rss} is the assigned weight for the received signal strength criterion. W_{ra} is the assigned weight for the resource availability criterion. W_{snr} is the assigned weight for the signal to noise ratio criterion. W_{mss} is the assigned weight for the mobile station speed criterion. W_{st} is the assigned weight for the service type criterion. *TW* is the total weight and is calculated using 3.

Step 6: the aggregated preference function is calculated as follows

$$APF(i,i') = \frac{\sum_{j=1}^{m} W_j P_j(i,i')}{\sum_{j=1}^{m} W_j} \qquad (4)$$

where W_j is the relative weight of *jth* criterion.

Step 7: Determine the Leaving Outranking Flow (LOF) and Entering Outranking Flow (EOF) as follows:

$$LOF_i = \frac{1}{1-n} \sum_{i'=1}^{n} APF(i,i') \qquad (5)$$

$$EOF_i = \frac{1}{1-n} \sum_{i'=1}^{n} APF(i',i) \qquad (6)$$

where *n* is the number of alternatives. The leaving (positive) flow expresses how much an alternative dominates the other alternatives, while the entering (negative) flow denotes how much an alternative is dominated by the other alternatives.

Step 8: the Net Outranking Flow (NOF)for each alternative

$$NOF_i = LOF_i - EOF_i \qquad (7)$$

The higher value of NOF, the better is the alternative. Thus, the best alternative is the one having the highest NOF value.

V. The Performance Evaluation

Our proposed solution is evaluated using the simulation approach. This section presents the used performance metrics and simulation models.

A. The performance metrics

The performance of the proposed JAC algorithm is evaluated using three performance evaluation metrics. The used metrics can be described briefly as follows.

- Blocking probability (P_b) is defined as the ratio of the number of blocked users to the total number of new users requesting admission. A user is blocked if at the session start the JAC algorithm assigns a bit rate of 0 kb/s.
- Outage probability (P_o) is calculated as the ratio of the number of users not fulfilling their Guaranteed Bit Rate (GBR) requirement, to the total number of admitted users.
- Unsatisfied user probability (P_u) that could be calculated based on P_b and Po as shown in equation 8

$$P_u = 1 - (1 - P_b)(1 - P_o) \qquad (8)$$

B. The simulation environment

A modified version of MATLAB based simulator called RUNE [16] has been used. Our models developed in [17] have been updated to be used in this work. The simulation environment defines a system model, a mobility model, a propagation model, and services model. The system model specifies the type of networks and the number and characteristics of the cells. Our system model considers the coexistence of six types of RATs. The characteristics of the RATs are summarized in Table II.

TABLE II.
SYSTEM MODELS DETAILS

RAT No.	Multiple Access Type	Antenna Type	Cell Radius	Number of Cells
RAT1	CDMA/WWAN	Omni-directional	750m	7
RAT2	CDMA/WMAN	Omni-directional	375m	12
RAT3	CDMA/WLAN	Omni-directional	100m	27

The mobility model simulates the mobility of the users in the system environment. In our mobility model, the mobiles are randomly distributed over the system. In every slot each mobile is moved a random distance in a random direction at defined time steps.

The propagation model simulates the different losses and gains during the signal propagation between the transmitter and the receiver in the system environment. The wireless propagation model used in this paper is described in a logarithmic scale as in equation 9.

$$G = G_D + G_F + G_R + G_A \qquad (9)$$

Equation 9 contains four components, the first component is the distance attenuation *GD* that is calculated by Okumura-Hata formula. The second component is the shadow fading *GF* that is modeled as a log-normal distribution with standard deviation of 6 dB and 0 dB mean. The third component is the Rayleigh fading *GR* that is modeled using a Rayleigh distribution. The forth component is the antenna gain *GA* that adds the antenna gain in dB.

The services model specifies the type of services and their percentages of use in the system environment. The *ith* service is mainly characterized by its bit rate requirement *RateReqc* and delay requirement *DelayReqc*. The users are generated according to Poisson process. The service holding time is exponential distribution with mean holding time equals to 150 seconds.

VI. THE RESULTS STUDY

Three different alternative algorithms are simulated and evaluated against our proposed solution. The first alternative does not take into account the JAC concept (It is denoted as NJAC) and the local RRM entities take the full responsibility to admit or reject the users. The second reference algorithm is denoted as Load-based JAC (LJAC) and it selects the least loaded RAT for new or handoff request. Finally, the third algorithm selects the RAT in which the mobile measures the strongest received signal strength, and it is denoted as Signal Strength JAC (SSJAC). In all the three cases, once the RAT has been selected, the bandwidth assigned to each user is the maximum bandwidth considered for this RAT for this type of service. Some simulation results for different sets of users are presented in this section.

Table III and Figure 4 illustrate some numerical results for the P_b values in all algorithms. The results show that our solution achieve good performance enhancement over all algorithms. On average, our algorithm achieves around 18%, 7%, and 12% enhancement over NJAC, LJAC, and SSJAC algorithms respectively. Better results can be gained if more suitable weights are used.

Table IV and Figure 5 illustrate some numerical results for the P_o values in all algorithms. The results show that our solution achieve good performance enhancement over all algorithms. On average, our algorithm achieves around 15%, 4%, and 9% enhancement over NJAC, LJAC, and SSJAC algorithms respectively. Better results can be gained if more suitable weights are used.

Table V and Figure 6 illustrate some numerical results for the P_u values in all algorithms. The results show that our solution achieve good performance enhancement over all algorithms. On average, our algorithm achieves around 25%, 9%, and 17% enhancement over NJAC, LJAC, and SSJAC algorithms respectively. Better results can be gained if more suitable weights are used.

TABLE III
P_b VALUES IN ALL ALGORITHMS

No. of Users	Our solution	NJAC solution	LJAC solution	SSJAC solution
50	0.01	0.09	0.04	0.07
150	0.04	0.18	0.09	0.14
250	0.09	0.29	0.18	0.22
350	0.14	0.37	0.24	0.3
450	0.19	0.48	0.31	0.41

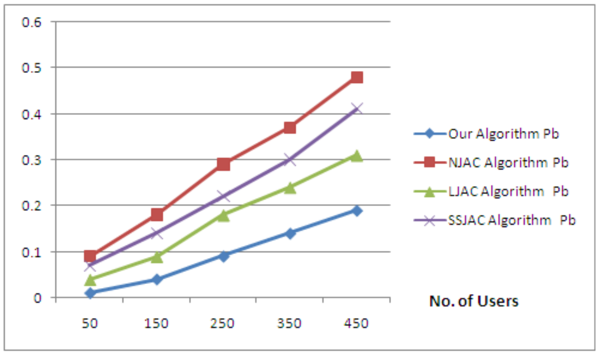


Fig. 4. P_b values for all algorithms

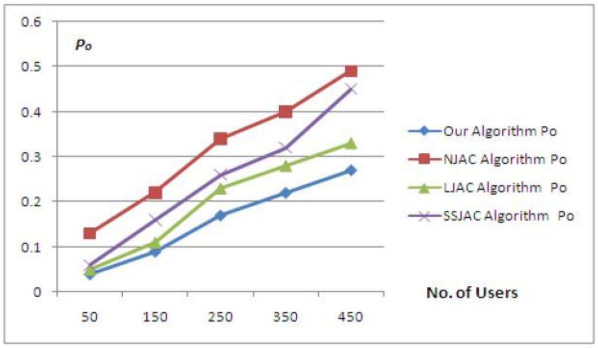


Fig. 5. Po values for all algorithms

TABLE IV
P_O VALUES IN ALL ALGORITHMS

No. of Users	**Our solution**	**NJAC solution**	**LJAC solution**	**SSJAC solution**
50	0.04	0.13	0.05	0.06
150	0.09	0.22	0.11	0.16
250	0.17	0.34	0.23	0.26
350	0.22	0.4	0.28	0.32
450	0.27	0.49	0.33	0.45

TABLE V
P_U VALUES IN ALL ALGORITHMS

No. of Users	**Our solution**	**NJAC solution**	**LJAC solution**	**SSJAC solution**
50	0.0496	0.2083	0.088	0.1258
150	0.1264	0.3604	0.1901	0.2776
250	0.2447	0.5314	0.3686	0.4228
350	0.3292	0.622	0.4528	0.524
450	0.4087	0.7348	0.5377	0.6755

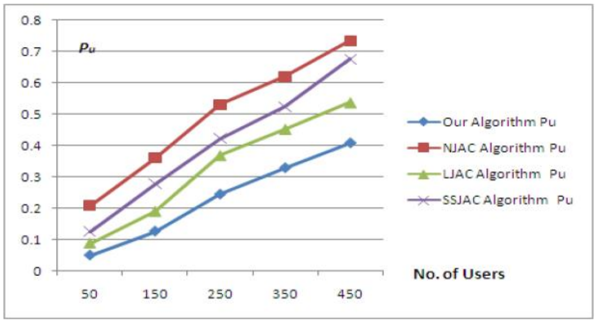


Fig. 6. P_u values for all algorithms

In one hand and for reasons of simplicity and computational complexity, the simulation cannot be carried out at higher number of users. In the other hand, the achieved simulation results show that our algorithm outperforms the reference algorithms, and a clear monotonic increasing relationship could be directly observed between the number of users and the performance metrics. To check if there is any linear relationship between the number of users and the achieved performance metrics, the Pearsons Correlation Coefficient (PCC) is used. PCC investigates the strength and direction of a linear relationship between two random variables. PCC = +1 means very strong positive linear relationship. PCC = -1 means very strong negative linear relationship. PCC = 0 means no linear relationship is existed between both variables. The results shows that the values of the PCC are all around +1 which means very strong positive linear relationship and we hence expect that our algorithm will keep outperforming the other algorithms at very high number of users.

VII. CONCLUSIONS AND FUTURE WORK

A suitable design for the JAC algorithms is a highly important issue to achieve the aims of the HWN. This paper explores the issue of JAC in the HWN. A novel JAC algorithm has been designed, implemented, simulated and evaluated. The developed JAC solution attempts to increase the user satisfaction, and decrease the blocking and outage probability.

Our future works can be extended in several directions. An optimum values for the weights of the different criteria can be found using a global optimization method. The developed JAC algorithm can be integrated with other CRRM mechanisms such as Access Network selection (ANS), Joint Congestion Control (JCC), and Vertical Handover (VHO). A joint optimization of these mechanisms can enhance overall system performance. In addition, this study has developed generic JAC algorithms. The algorithms can be tailored to specific wireless standards such as UMTS, IEEE802.16, and IEEE802.11.

REFERENCES

[1] O. E. Falowo and H. A. Chan, "Joint call admission control algorithms: Requirements, approaches, and design considerations", Computer Communications, vol. 31, no. 6, pp. 1200-1217

[2] J. Prez-Romero, O. Sallent, and R. Agust, "On evaluating beyond 3G radio access networks: architectures, approaches and tools," IEEE 61st Vehicular Technology Conference (VTC 2005-Spring), vol. 5, pp. 2964 - 2968, June 2005

[3] O. E. Falowo, H. A. Chan, "Joint call admission control For next generation wireless network," IEEE Conference on Electrical and Computer Engineering, pp.1151 - 1154, May 2006

[4] Yu. Fei and V. Krishnamurthy, "Optimal joint session admission control in integrated WLAN and CDMA cellular networks with vertical handoff," IEEE Transactions on Mobile Computing, vol. 6, no. 1, pp. 126 - 139, Jan. 2007

[5] X. G. Wang, G. Min, and J. E. Mellor Adaptive QoS control in cellular and WLAN interworking networks, Second International Working Conference on the Performance Modeling and Evaluation of Heterogeneous Networks (HET-NETs '04), West Yorkshire, U.K., July 2004

[6] D. Karabudak, C. Hung, and B. Bing, A call admission control scheme using genetic algorithms, Symposium on Applied Computing (SAC 04), pp.1151 1158, Cyprus, March 2004

[7] K. Murray and D. Pesch, "Policy based access management and handover control in heterogeneous wireless networks," The IEEE Vehicular Technology Conference, vol.5, pp.3319- 3323, Sep. 2004

[8] K. Murray and D. Pesch,"Intelligent network access and inter-system

handover control in heterogeneous wireless networks for smart space environments," 1st International Symposium on Wireless Communication Systems, pp. 66-70, 2004

[9] D. Qiang, H. Bo, S. Yan, and C. Shan-zhi,"Joint admission control through vertical handoffs in heterogeneous wireless network," Global Mobile Congress (GMC), 2010, pp. 1-5, Nov. 2010

[10] X. Li, H. Ji, P. Si, and L. Zhang, "Joint session admission control scheme in integrated WLAN and 3G networks," 5th International ICST Conference on Communications and Networking in China (CHINACOM 2010), pp.1-5, Aug. 2010

[11] M. Khabazian, O. Kubbar, H. Hassanein, "Call Admission Control with Resource Reservation for Multi-service OFDM Networks," 2012 International Conference on Computing, Networking and Communications (ICNC), pp.781-785, Feb. 2012

[12] Zong-Hua Liu, Jyh-Cheng Chen,"Design and Analysis of the Gateway Relocation and Admission Control Algorithm in Mobile WiMAX Networks," IEEE Transactions on Mobile Computing, pp.5-18, Jan. 2012

[13] M. Doumpos, and C. Zopounidis, "A multi-criteria classification approach based on pair-wise comparison," European Journal of Operational Research, pp. 378-389, 2004

[14] V. Manikrao Athawale, S. Chakraborty, "Facility location selection using PROMETHEE II method," The 2010 International Conference on Industrial Engineering and Operations Management, pp. 18-22, Jan. 2010

[15] J.S.R. Jang, C.T. Sun, E. Mizutani, Neuro-Fuzzy and soft Computing, Prentice Hall, 1997.

[16] J. Zander and S. Kim, Radio Resource Management for Wireless Networks, Artech House, 2001

[17] Mohammed M. Alkhawlani, "Intelligent Vertical Handover for Heterogonous Networks Using FL And ELECTRE," International Journal of Computer Networks and Communications (IJCNC), Vol.3, No.4, July 2011

Comparison of OpenMP & OpenCL Parallel Processing Technologies

Krishnahari Thouti
Department of Computer Science & Engg.
Visvesvaraya National Institute of Technology,
Nagpur – 440010 (India)

S.R.Sathe
Department of Computer Science & Engg.
Visvesvaraya National Institute of Technology,
Nagpur – 440010 (India)

Abstract—**This paper presents a comparison of OpenMP and OpenCL based on the parallel implementation of algorithms from various fields of computer applications. The focus of our study is on the performance of benchmark comparing OpenMP and OpenCL. We observed that OpenCL programming model is a good option for mapping threads on different processing cores. Balancing all available cores and allocating sufficient amount of work among all computing units, can lead to improved performance. In our simulation, we used Fedora operating system; a system with Intel Xeon Dual core processor having thread count 24 coupled with NVIDIA Quadro FX 3800 as graphical processing unit.**

Keywords- OpenMP; OpenCL; Multicore;Parallel Computing; Graphical processors.

I. INTRODUCTION

Nowadays, Quad-core, multi-core & GPUs [1] have already become the standard for both workstations and high performance computers. These systems use aggressive multi-threading so that whenever a thread is stalled, waiting for data, the thread can efficiently switch to execute another thread. Achieving good performance on these modern systems requires explicit structuring of the applications to exploit parallelism and data locality.

Multi-core technology offers very good performance and power efficiency and OpenMP [2] has been designed as a programming model for taking advantage of multi-core architecture. The problem with GPU is that, their architecture is quite different to that of a conventional computer and code must be (re)written to explicitly expose algorithmic parallelism. A variety of GPU programming models have been proposed in [3 - 5].

The most popular development tool for scientific GPU computing has proved to be CUDA (Compute Unified Device Architecture) [6], provided by the manufacturer NVIDIA for its GPU products. However, CUDA is not designed for heterogeneous systems, while OpenCL programming model, by the Kronos Group [7] supports cross-platform, parallel programming of heterogeneous processing systems. The architectural details of multi-core and GPUs are explained in next section.

Given, a diversity of high-performance architectures, there is a question of which is the best fit for a given workload and extent to which an application benefit from these systems, depends on availability of cores and other workload parameters. This paper addresses these issues by implementing parallel algorithms for the four test cases and compares their performance in terms of time taken to execute and percentage of speed-up factor achieved.

In Section II, we present parallel computing paradigm. We then present architectural framework for Multi-core and GPU architectures in Section III & IV.

Experimental results are presented in Section V. Section VI presents related work done and conclusion and future scope are discussed in Section VII.

II. PARALLEL COMPUTING PARADIGM

Parallel computing [8] depends on how the processors are connected to memory. The way of system connection can be classified into shared or distributed memory systems, each of these two types are discussed as follows:-

A. Shared Memory System

In such a system, a single address space exists, within it every memory location is given a unique address and the data stored in memory are accessible to all processing cores. The processor P_i reads the data written by processor P_j. Therefore, in order to enforce consistency, it is necessary to use synchronization.

The OpenMP is one of the popular programming languages for the shared memory systems. It provides a portable, scalable and efficient approach to run parallel programs in C/C++ and FORTRAN.

In OpenMP, a sequential programming language can be parallelized with pre-processor compiler directive #pragma omp in C and $OMP in FORTRAN.

B. Distributed Memory System

In such a system, each processor has its own memory and can only access its local memory. The processors are connected with other processors via high-speed communication links. MPI (Message Passing Interface) [9] provides a practical, portable, efficient and flexible standard for message passing across distributed memory systems.

We limit our discussion to shared memory systems. Based on above classification, we classify systems as Multi-core systems and Many-core systems or GPGPU devices [1].

III. ARCHITECTURAL FRAMEWORK – MULTICORE & OPENMP PROGRAMMING MODEL

The present typical multi-core architecture is shown in Figure 1. It consists of a number of processing cores, each having a private level one (L1) data and instruction cache, L2 cache, which are attached via a bus interconnect to shared level three (L3) cache.

Each core supports multi-threading, which allows sharing of several micro-architectural resources between threads e.g. L1 caches, physical registers, and execution units.

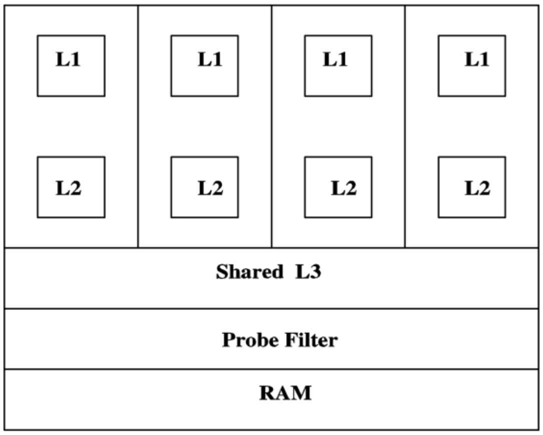


Figure 1: Multi-core Architecture with Cache Hierarchy

Benefits of multi-core systems can be obtained by Open MP Programming model. OpenMP is a specification of compiler directives, library routines, and environmental variables that provides an easy parallel programming model portable across shared memory architecture.

OpenMP is a set of compiler directives (# pragma) and callable runtime library routines that express shared memory parallelism [10]. The directive itself consists of a directive name and followed by clauses. OpenMP programs execute serially until they encounter the “parallel” directive. This directive is responsible for creating group of threads. The exact number of threads can be specified in the directive, set using an environmental variable, or at run-time using OpenMP functions. The main thread that encounters the “parallel” directive becomes the “master” of this group of threads and is assigned the thread id 0. There is an implicit barrier at the end of parallel region. The master thread with thread id 0 collects results from other threads and executes serially from that point on.

IV. ARCHITECTURAL FRAMEWORK – GPU & OPENCL PROGRAMMING MODEL

GPUs i.e. Graphic Processing units are the basic building blocks for high performance computing but its programming complexity pose a significant challenge for developers. To improve the programmability of GPUs, the Open CL (Open Computing Language) [7] has been introduced. Open CL is an industry standard for writing parallel programs to execute on the heterogeneous platforms like GPU devices. OpenCL (Open Computing Language) is a low-level API for heterogeneous computing that runs on CUDA architecture.

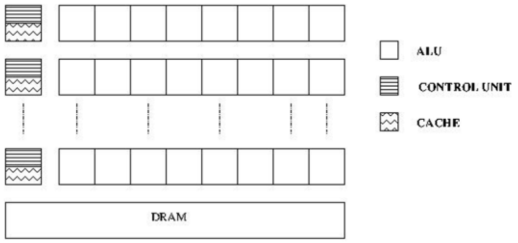


Figure 2: GPU Architecture

NVidia GPUs comprises of array of multithreaded Streaming Multiprocessors (SMs) and each one consists of multiple Scalar Processor (SP) cores, a multithreaded instruction unit, and on-chip shared memory. The SMs creates, manages, and execute concurrent threads in hardware with zero scheduling overhead.

In short, we say, following are the steps to initialize an OpenCL Application.

Set Up Environment – Declare OpenCL context, choose device type and create the context and a command queue.

Declare Buffers & Move Data – Declare buffers on the device and enqueue input data to the device.

Runtime Kernel Compilation – Compile the program from the kernel array, build the program, and define the kernel.

Run the Program – Set kernel arguments and the work-group size and then enqueue kernel onto the command queue to execute on the device.

Get Results to Host – After the program has run, read back result array from device buffer to host memory.

V. EXPERIMENTAL RESULTS

In this section, we present experimental results. For experimental setup, we have tested our system on four test cases. We compare the performance of these test cases with the OpenCL code on the GPU and on a multi-core CPU with Open MP support.

The host machine used has Intel Xeon 2.67GHz Dual processors with 12Mb L3 cache. Each processor has Hyper-Threading [11] technology such that, each processor can execute simultaneously instructions from two threads. Overall numbers of cores are 12 and because of hyper threading thread count of host machine equal 24.

Each core of two processors has 32KB L1 Data cache, 256 KB L2 cache shared between 2 threads of that core. In addition to that, there is 12MB L3 cache shred among all the threads.

The GPU device used in our experiment was NVidia Quadro FX3800. The device has 192 processing cores with 1 GB 256 bit memory interface and memory bandwidth of 51.2 GB/sec. The GPU device was connected to CPU through X58 I/O Hub PCI Express. The environment used was Fedora-x86_64 and kernel version is - 2.6. Gcc version is 4.6.

OpenCL-1.0 was used for compiling OpenCL programs by providing “-lOpenCL” as compile time option for gcc compiler.

We have used “gettimeofday ()” library routine to measure time taken to execute test problems. “start” and “end” time recorded and execution time is calculated as shown below in Program Listing-1:

```
gettimeofday (&start, NULL );
    Backtrack (0, 0, 0, 0);
    gettimeofday (&end, NULL );
    time = (end.tv_sec-start.tv_sec) + (float) (end.tv_usec - start.tv_usec) * 0.000001;
    printf ("Time = %f\n", time );
```

Program Listing 1: Measuring CPU Time

A. *Matrix Multiplication*

We consider the problem of computing the product C = A*B of two large, dense, matrices. A straight forward matrix multiplication performs scalar operations on data items. We choose matrix multiplication, because of following two reasons:

- Matrix Multiplication is widely used in Image processing applications
- Matrix Multiplication is a fundamental parallel algorithm with respect to data locality, cache coherency etc.

Matrix Order	***Sequential***	***OpenMp***	***OpenCL***
1024	6.04	0.71	1.64
2048	136.14	18.39	2.05
3072	345.06	40.84	2.45
4096	1261.29	177.79	3.66
5120	2819.54	328.52	5.53
6144	5023.87	593.13	8.4

Table 1: Execution Time for Matix Multiplication of Sequential, OpenMp, OpenCL Version

Matrix Order	***Seq-to-OpenMp***	***Seq-to-OpenCL***
1024	8.51	3.68
2048	7.40	66.41
3072	8.45	140.85
4096	7.09	344.62
5120	8.58	509.86
6144	8.47	598.08

Table 2: Speed-Up of Sequential to openMP & OpenCL

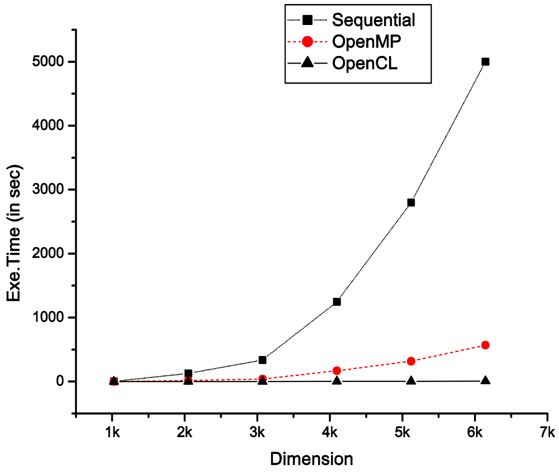


Figure 3: Execution Time v/s Dimension of Matrix

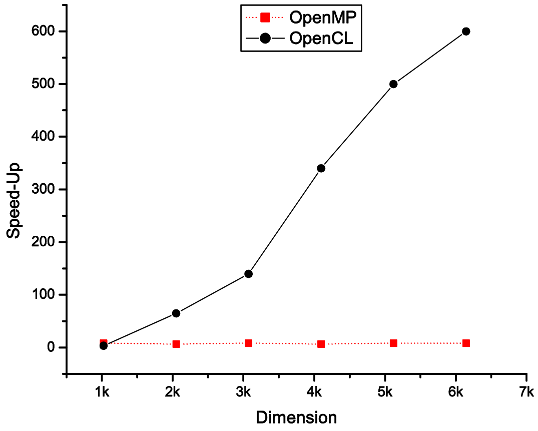


Figure 4: Speed-Up Comparison

As the dimensions of the matrix increase, the execution time for sequential algorithm also increases by manifold as shown on Figure 3. After analyzing Table -1 and 2, Figure 3 and 4, we conclude that, given the Multi-core architecture, OpenMP shows good improvements for smaller matrix dimensions but as the matrix dimension increases OpenCL gives very good Speed-Up factor and very less execution time. This can be evident from Table 1 & 2 that for matrix dimension 1024 OpenMP is much better than OpenCL. But for matrix dimension above 1024, OpenCL gives very good performance. Note that values shown in Table 1 & 2 are obtained after performing the experiment for nearly 10 – 20 runs.

B. *N-Queens Problem*

The n-queen problem is a classic problem of placing n-chess queens on chessboard so that no two queens attack each other.

The most obvious way to solve this problem consists of trying systematically all ways of a placing N-Queens on a chessboard, checking each time to see whether a solution has been obtained. But this approach will take very large time to arrive at solution. Backtracking is an approach to solve this problem. But backtracking takes exponential time-complexity. Because of this reason, it is very interesting to parallelize this problem.

Recursion is a peculiar property of backtracking. The earlier version of OpenMP doesn't support recursion. Support for recursion is introduced in OpenMP 3.0 specifications by "task "clause. However, we find that there is no significant improvement in performance, since most of the code to be parallelized is kept in critical section region as shown below:

```
int put(int Queens[], int row, int column)
{
   Queens[row]=column;
   if(row==N-1) {
   pragma omp critical
   {       solutions++;   }
   }
   else{
   for(i=0; i<N; i++){ put(Queens,row+1,i);}
   }
 return solutions;
}
```

Program Listing 2: OpenMP code for N-Queens

Therefore, we have taken into consideration of sequential v/s OpenCL code to evaluate N-Queen problem. For larger values of N (> 23), the number of solutions and time to solve the given problem is still not known [12, 13].

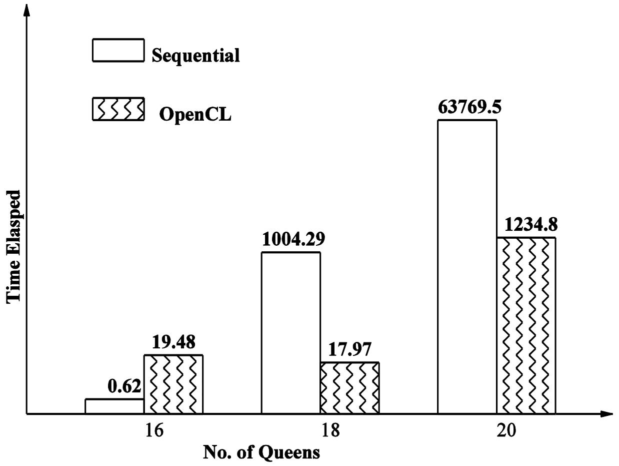


Figure 5: Performance of N-Queen Problem

As shown in Figure 5, we have taken only practical cases where solutions are available within stipulated time. For N=18, sequential took nearly 16.75Minutes (1004.29sec) whereas OpenCL took only 17.97 seconds to generate all correct results. The power of OpenCL can be observed in cases for N >=20. For N=20, sequential program took nearly 17.72Hrs (63769.5sec) whereas OpenCL took only 20.58Min (1234.88sec). Speed-up achieved is enormous. The graph shown in Figure 5 is not up to the scale.

C. Image Convolution

The convolution of images is a commonly used technique for image filtering. It is best described as a combining process that copies one image into another. Any number of filters may be applied to an image by convolving the filter mask with the original image. The equation for image convolution is given by

$$Out(i,j)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} In(m,n)Mask(i-m,j-n)$$

Where In is the input image, Mask is the convolution mask, and Out is the output image. The dimension of the image is M x N, the Mask image is smaller than image size; may be padded with zeroes to allow for consistency in indexing.

The convolution technique consists of the following steps:

- Select a pixel in original image to convolute
- Apply mask to the pixel by reading the selected pixel's neighbor
- Write the new values to the out image

The convolution algorithm will generate results that are greater than the range of original values of the input image. For this, scaling operation is performed to restore the result to same gray level range of original picture.

Prog. Type	***Sequential***	***OpenMP***	***OpenCL***
Time (in sec)	0.51	0.05	0.96

Table 3: Time Elapsed in Image Convolution

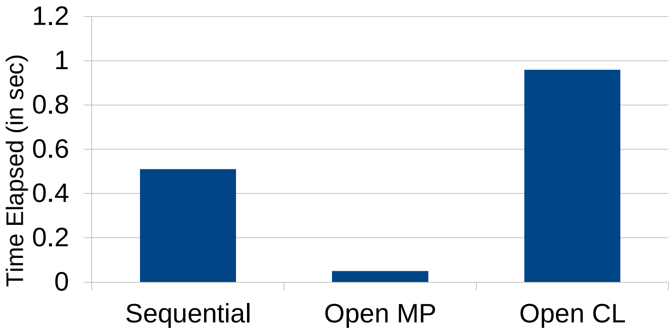


Figure 6: Image Convolution

We took 600 x 400 image, 10 x 10 mask and applied convolution. Table 3 shows time required to process convolution. Performance of OpenMP is better compared to the OpenCL, as evident from the Figure 6.

The speed-up achieved is (Seq/MP) = 0.51/0.05 = 10.2 whereas no speed-up is achieved w.r.t OpenCL as (Seq/CL) = 0.05/0.96 = 0.53.

The convolution algorithm computes the two-dimensional discrete correlation between an image and a template and leaves the result in output image. As a result, OpenMP is much faster compared to OpenCL, as OpenCL is busy in doing background of kernel creation and other things, than actual execution. The actual GPU device execution time can be found by profiling [14] the gpu device which will be very less as compared to OpenMP.

D. String Reversal

For the comparison purpose, we have taken a string reversal problem. We have considered a huge file in mega-bytes and tried to reverse it using OpenCL.

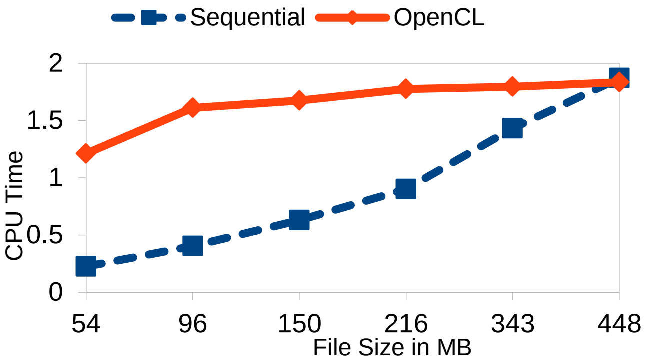


Figure 7 : String Reversal

File Size	*Sequential*	*OpenCL*
54MB	0.22	1.22
96MB	0.41	1.62
150MB	0.63	1.68
216MB	0.91	1.78
343MB	1.43	1.79
448MB	1.87	1.84

Table 4: String Reversal

String reversal problem is straight forward. Just read the entire file and start copying values from end of the file. As there are no dependencies [15] in this operation, we have not considered OpenMP Programming model. Even if we take OpenMP into consideration, performance will be same as reversing of read string falls in critical section.

From Figure 7 and Table 4, it can be concluded that, OpenCL Programming model is not suitable for this kind of applications.

VI. Related Work

The present studies dealing with multi-core and GPGPUs have been accelerated by parallelizing matrix-matrix multiplication on a CPU and GPU [16, 17]. In multiple-core clusters systems without GPU accelerators, some contributions have been made to improve the computing power by developing hybrid models [18].

All these studies mentioned above have focused on parallel usage of CPUs and GPUs and have demonstrated significant performance improvement. However, because of the issues related to cache hierarchy, memory and bandwidth in CPU and GPUs, obtaining effective performance evaluation results is an open issue.

VII. Conclusion & Future Scope

We studied the behavior of parallel algorithms with respect to OpenMP and OpenCL. The initial results we found were not satisfactory. But, as the number of input data size increased OpenCL gives good performance.

Latest systems are equipped with multi-core architecture. So, OpenMP will be a viable option for cases such as matrix multiplication, image convolution, and other applications. But OpenCL scores well with matrix multiplication.

OpenCL involves a lot of background work like memory allocation, kernel settings and loading, getting platform, device information, computing work-item sizes etc. All this adds overhead in OpenCL. However, we find that, in spite of this overhead, OpenCL gives very good performance. But OpenCL fails in application where it has less scope of work; this can be seen from the string reversal example.

Another finding is that critical section is too expensive. We have implemented OpenMP version of N-Queen problem, but, we find that it has no improvement as only one thread is running at a time. However, we can take advantage of “task” directive in application such as tree traversal.

Overall, we sum up our conclusion as

OpenCL > OpenMP > Sequential

Where > indicates performance. As a future work, we will find algorithms, where OpenMP is more preferable over OpenCL.

Future research work is required in the following problem areas: given an application program, we must check how useful OpenMP or OpenCL is in heterogeneous environment consisting of multiple GPUs and multi-cores. Secondly, a library routine can be developed, which will port application program to CPU using OpenMP or to GPU using OpenCL or combination of these two technologies.

References

[1] General Purpose Computations Using Graphics Hardware, http://www.gpgpu.org/

[2] OpenMP Architecture Review Board: The OpenMP Specifications for Parallel Programming, http://openmp.org/

[3] Advanced Micro Devices, Technical Report, http://developer.amd.com/gpu/ATIStreamSDK/assets/ATI_Intermediate_Language_(IL)_Specification_v2d.pdf, 2010

[4] I. Buck, et al., Brooks for GPUS: Stream Computing on graphics hardware, in: Proceedings of ACM SIGGRAPH 2004, ACM Trans. Graph, 23 (2004) 777 – 786.

[5] The Portland Group, PGI Fortran & C accelerator programming model, Technical report, http://www.pgroup.com/lit/whitpapers/pgi-accelerator.pdf, 2009.

[6] NVIDIA CUDA, http://developer.nvidia.com/object/cuda.html/

[7] OpenCL – The Open Standard for Parallel Programming of Heterogeneous Systems, http://www.khronous.org/opencl/

[8] Introduction to Parallel Computing, Second Edition, http://www.scribd.com/doc/42971558/an-Introduction-to-Parallel-Computing.

[9] OpenMPI: Open Source High Performance Computing, http://www.open-mpi.org/

[10] http://siber.cankaya.edu.tr/ParallelComputing/ceng471/node96.html

[11] Intel Hyper-Threading Technology, http://www.intel.com/

[12] http://en.wikipedia.org/wiki/Eight_queens_puzzle

[13] http://www.ic-net.or.jp/home/takaken/e/queen/

[14] Intel. Intel® VTune^{TM} Performance Analyzer. http://www.intel.com/

[15] F. Irigoin and R. Triolet, "Computing Dpendence Direction Vectors and Dependnce Cones with Linear Sytems," ENSMP-CAI-87-E94, Ecole des Mines de Paris, France, (1987).

[16] M. Fatica: Accelerating Linpack with CUDA on Heterogeneous Clusters, ACM International Conference Proceeding Series, pp.46 – 51 (2009).

[17] S. Ohshima, K. Kise, T.Katagiri, and T.Yuba: Parallel Processing of Matrix Multiplication in a CPU and GPU Heterogeneous Environment, High Performance Computing for Computational Science, pp.305 – 318 (2007).

[18] T.Q.Viet, T.Yoshinaga, and B.A.Abderazek: Performance Enhancement for Matrix Multiplication on a SMP PC Cluster, IPSJ SIG technical Report, pp.115-120 (2005).

Authors Profile

S.R.Sathe received M.Tech in Computer Science from IIT, Bombay (India) and received Ph.D fronm Nagpur University (India). He is currently working as Professor in Department of Computer Science & Engineering at Visvesvaraya National Institute of Technology, Nagpur (India). His research includes parallel processing algorithms, computer architecture.

Krishnahari Thouti, Research Scholar, pursuing Ph.D in Department of Computer Science Engg, at Visvesvaraya National Institute of Techology, Nagpur (India). His area of interest includes parallel processing, compiler and computer architecture.

An Authentication Protocol Based on Combined RFID-Biometric System

Noureddine Chikouche
Department of Computer Science
University of M'sila
Algeria

Foudil Cherif
LESIA Laboratory
University of Biskra
Algeria

Mohamed Benmohammed
LIRE Laboratory
University of Constantine
Algeria

Abstract — **Radio Frequency Identification (RFID) and biometric technologies saw fast evolutions during the last years and which are used in several applications, such as access control. Among important characteristics in the RFID tags, we mention the limitation of resources (memory, energy, …). Our work focuses on the design of a RFID authentication protocol which uses biometric data and which confirms the secrecy, the authentication and the privacy. Our protocol requires a PRNG (*Pseud-Random Number Generator*), a robust hash function and Biometric hash function. The Biometric hash function is used to optimize and to protect biometric data. For Security analysis of protocol proposed, we will use AVISPA and SPAN tools to verify the authentication and the secrecy.**

Keywords-component; RFID; authentication protocol; biométric; security.

I. INTRODUCTION

At present, the problem of access control is very important in several applications. Physical access control consists in verifying if a person asking to reach a zone (e.g. building, office, parking, laboratory, etc.), has the right necessities to make it. The protocols of identity verification which allow access are called the authentication protocols. They answer the following two questions: "Who am I?" and "Am I really the person who is proceeding?". Answer to this first question is based on the recognition or the identification of the user which consists in associating an identity to a person, such as a smartcard or a RFID tag. Concerning the second question which articulates on the verification or the authentication of the user, it gives permission to a proclaimed identity. In other terms, it consists in identifying a user from one or several physiological characteristics (fingerprints, face, iris, etc.), or behavioural (signature, measure, etc.). These techniques are called Biometric Methods [14].

Among techniques and systems of identification which were developed quickly during the last years, we can notice that Radiofrequency identification (RFID) that is used in different domains (health, supply chain, access control, etc.). The RFID systems consist of three entities: (1) the tag (or the label) is a small electronic device, supplemented with an antenna that can transmit and receive data, (2) the reader which communicates with the tag by radio waves and (3) the server (or database, back-end) which uses information obtained from the reader for useful purposes. The main characteristic of a RFID system is the limitation of resources (memory, the processor, the consumption of energy, etc. …); on the other hand, these systems are necessary to assure security in all the levels of the system. Major difference between a RFID tag and a contactless smartcard is the limitation of computer resources.

In RFID systems, several authentication protocols have been developed [4,5,6,7]. Difference between these techniques lies in the realized properties of security and the complexity of implementation. Most of these protocols answer the first question only "Who am I? ". On the contrary systems with smartcards there are several authentication protocols based on the biometric technology, we mention here [8,9,10].

This paper, we propose an authentication protocol based on the combination between a RFID system and a biometric system. We verify secrecy, authentication of the tag and authentication of the reader by AVISPA&SPAN tools [1,2]. The conceived protocol protects the privacy of the user. To estimate these performances, we will compare it with the other RFID protocols and the biometric protocols of smart cards.

This paper is organized as follows: section II presents related work. Section III presents the system and hypotheses. Section IV presents the proposed protocol. The section V presents a check of the protocol automatically. An analysis of privacy side is then presented. Section VI presents a comparison of performances with existing works. We end by a general conclusion.

II. RELATED WORK

In the protocols using identifier ID, two mechanisms are used: static and dynamic. The characteristic of the mechanism of static ID is that the identifier of the tag remains the same during the complete authentication, but that of the dynamic mechanism, the identifier of tag is modified. Every mechanism has his advantages and inconveniences. Here we present mainly the mechanism of static ID used in this article.

In the RHLS protocol (Randomized Hash Lock scheme) [5], information passed on with the tag every time when it is interrogated consists of random value nt and value H1 = h(ID, nt). RHLS which discovers two types of attacks: replay attack and tracing attack.

Concerning the protocol which is proposed with Chien and Huang (CH protocol) [4], the reader R and the tag T share secrets k and ID. Launch by the reader which sends a nonce nr. The tag produces an unpredictable nonce nt and calculates the hash function *g*, such as $g = h(nr \oplus Nt \oplus ID)$. This hash function and ID are used as parameters for the function rotate. The value of ID is returned; it depends on the value of g. The tag calculates the xor of the returned ID and g, before the sending of the left half of the results and nt to the reader. The reader calculates every pair of ID and k until it finds the corresponding tag. It sends then the right half of the xor of ID returned and g to the tag. CH protocol which discovers an attack of the type algebraic replay attacks, its cause is the false use of the algebraic operator xor in messages passed on by the function g. This attack is discovered also in LAK protocol [19].

Lee and al. [7] propose a protocol improved to avoid two types of attack: tracing and spoofing attack by use of various values of the hash function h during every authentication. These objectives are realized after analysis of this protocol. The cost of a hash function operation in tag is four, what is incompatible with the storage space and lower capacity to calculate. Therefore, excessive calculation will affect inevitably the efficiency of the protocol.

Biometry is widely used in the authentication protocols of the smart cards applications [8, 9, and 10]. The use of these protocols in RFID systems will depend on the availability of computer resources (memory, complexity, performance …) in the constituents of RFID systems and especially the RFID tag. The recent protocol [10] requires the calculation of seven operations of the function h in the phases of login and authentication and requires 4l as storage space in the tag. This number of calculations and this storage space influences negatively on the efficiency of a RFID protocol. Another difficulty concerns “Matching” treatment. In the biometric authentication protocols, this part is made in the smart card with the technique Match-on-card.

Concerning the material implementation of combined systems biometric-RFID, we shall quote two recent works. Rodrigues and al. [15] propose a decentralized authentication solution for embedded systems that combines both token-based and biometric-based mechanism authentication. Aboalsamh [16] propose a compact system that consists of a CMOS fingerprint sensor (FPC1011F1) is used with the FPC2020 power efficient fingerprint processor; which acts as a biometric sub-system with a direct interface to the sensor as well as to an external flash memory for storing finger print templates. An RFID circuit is integrated with the sensor and fingerprint processor to create an electronic identification card (e-ID card). The e-ID card will pre-store the fingerprint of the authorized user. The RFID circuit is enabled to transmit data and allow access to the user, when the card is used and the fingerprint authentication is successful.

III. System and Hypotheses

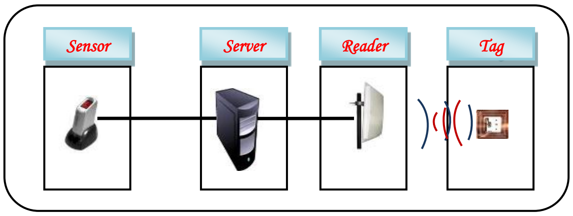


Figure 1. RFID-Biometric System

A. *System modeling*

The proposed system of authentication is based on the combination of two sub-systems: a RFID system and a biometric system. RFID system consists of: a tag ($\mathcal{T}$), a reader ($\mathcal{R}$) and a server ($\mathcal{S}$). Used biometric system consists of two entities, a sensor ($\mathcal{SR}$) and a server ($\mathcal{S}$), see the Figure 1.

1) *Biometric data*

Biometric data can be stored in the tag or in the data base. The biometric template will be stored in the tag. It offers a greater privacy and the mobility of the user. This assures also that information will always be with the user’s tag. Storing the raw biometric data typically requires substantially more memory. For example, a complete fingerprint image will require 50 to 100 Kbytes, while a fingerprint template requires only 300 bytes to 2 Kbytes [14]. This condition is not always sufficient especially for the type of passive RFID tags. In our system, a practicable solution to optimize and to protect biometric data is the hash function. This function of template allows pressing the biometric template in an acceptable size.

The problem which lies with the hash functions standard (e.g. SHA-1, MD5, SHA-256, …) is comparison between two templates: the template which is protected in the tag a *h(B)* and the template which is generated from the capture *h(B’)*. Equality $h(B) = h(B')$ for the same person is not always assured, because *B'* is a dynamic template where the person never keeps the same biometric features, (e.g. movement of the finger during the purchase), which implies the existence of a rate of error. We will quote two research works:

Sutcu and *al.* [12] propose a secure biometric based authentication scheme which fundamentally relies on the use of a robust hash function. The robust hash function is a one-way transformation tailored specifically for each user based on their biometrics. The function is designed as a sum of properly weighted and shifted Gaussian functions to ensure the security and privacy of biometric data. They also provide test results obtained by applying the proposed scheme to ORL face database by designating the biometrics as singular values of face images.

A. Nagar and al. [13] propose six different measures to evaluate the security strength of template transformation schemes. Based on these measures, they analyze the security of two well-known template transformation techniques, namely, Biohashing and cancelable fingerprint templates based on the proposed metrics.

2) *Tag and Reader*

The tag stores the identity (ID) and the biometric hash function of the template of the person (GB). This ID is strictly confidential and is shared between the database of the back-end server ($\mathcal{S}$) and the tag ($\mathcal{T}$). The tag can generate random numbers and calculation of the hash function *h* of a number. Standard ISO and EPC GEN2 (*Electronic Product Code, Generation 2*) support to produce the random numbers (nonces) in the tag. The reader $\mathcal{R}$ can generate also the random numbers.

3) *Server*

The server has two main functionalities:

- For the biometric system: extraction of the characteristics of a biometric modality to create a model or template ($\mathcal{B}$),
- Concerning RFID system: it contains the database which includes the list of the identity of tags ($\mathcal{ID}$).

4) Sensor

A biometric sensor is an electronic device used to capture a biometric modality of a person (fingerprint, face, voice, etc.).

B. Security and privacy requirements:

Our protocol strives to achieve four requirements: secrecy, authentication of the tag, authentication of the reader and untraceability.

- Secrecy: or confidentiality, the verification that the identity of the tag ID is never passed on clearly to air on the interface radio frequency which can be spied.
- Authentication of tag: A reader has to be capable of verifying a correct tag to authenticate and to identify a tag in complete safety.
- Authentication of reader: A tag has to be capable of confirming that it communicates with the legitimate reader (a single reader exists in communications between the constituents of the RFID system).
- Untraceability: We consider the notion of untraceability as defined in [17] which captures the intuitive notion that a tag is untraceable if an adversary cannot tell whether he has seen the same tag twice or two different tags.

C. Intruder Model

Besides modelling security protocols, it is also necessary to model the intruder, that is to say, to define its behaviour and limit. For this, the assumptions used are gathered under the name “Dolev-Yao model” [6]. This intruder model is based on two important assumptions that are *the perfect encryption* and the *intruder is the network.*

Perfect encryption ensures in particular that an intruder can decrypt a message *m* encrypted with key *k* if it has the opposite of that key. The second hypothesis which is "*the intruder is the network*" means that, the intruder has complete control over the network and he can derive new messages from his initial knowledge and the messages received from honest principals during protocol runs. To derive a new message, the intruder can compose and decompose, encrypt and decrypt messages, in case he knows the key.

For the assumption “*the intruder is the network*”, the RFID network system in this case is wireless, it is based on communication by radio waves. Communication among the server and the reader and between the server and the sensor is secure. Contrary to this, communication between the tag and reader is not assured and based on radio frequencies waves. We assume that the adversary can observe, block, modify, and inject messages in any communication between a reader and a tag.

IV. Proposed Protocol

The proposed Protocol is divided into two processes: the phase of registration and the phase of mutual authentication. We, afterward, use the following notations:

T	RFID tag or transponder
R	RFID reader or transceiver
S	Back-end server
Nt	random number (nonce) generated by tag T
Nr	random number (nonce) generated by reader R
H()	One-way hach function
G()	BioHash (Biometric hash function)
‖	Concatenation of two inputs
B	Biometric template
ID	Identity of tag
GB	Biohashed value of B
⊕	Or-exclusif
H_R	The right half of H
H_L	The left half of H
X≈Y	mean X=Y±E (E : rate of error)

Steps detailed by two processes are described below.

A. Registration Processus

This initial phase called also registration. The objective is to create a template biometric and stored in related to a declared identity (see the figure 2.). In this phase, it has to execute the following steps to obtain the RFID tag.

Step 1: the authorized user inputs his/ her personal biometrics, to pass it on to the server of the trusted registration center ($\mathcal{RC}$).

Step 2: the $\mathcal{RC}$, after extraction of biometric characteristics, creates a biometric template B, and computes the biometric hash function GB such as GB = g (B).

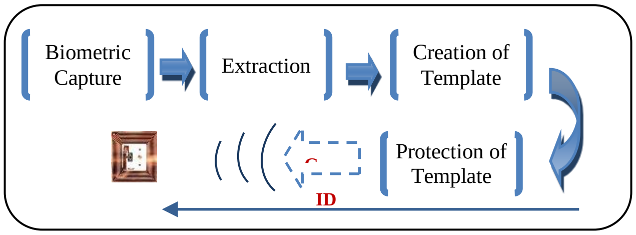


Figure 2. Registration Processus

Step 3: Then, the registration center stores the information {ID, GB} in the user’s tag and sends it to the tag through a secure channel.

$$\mathcal{RC} \xrightarrow{ID,\ GB} \mathcal{T}$$

B. Mutual Authentication Processus

According to the order of the passed on messages, the process of authentication takes place as follows

(to see Figure 3):

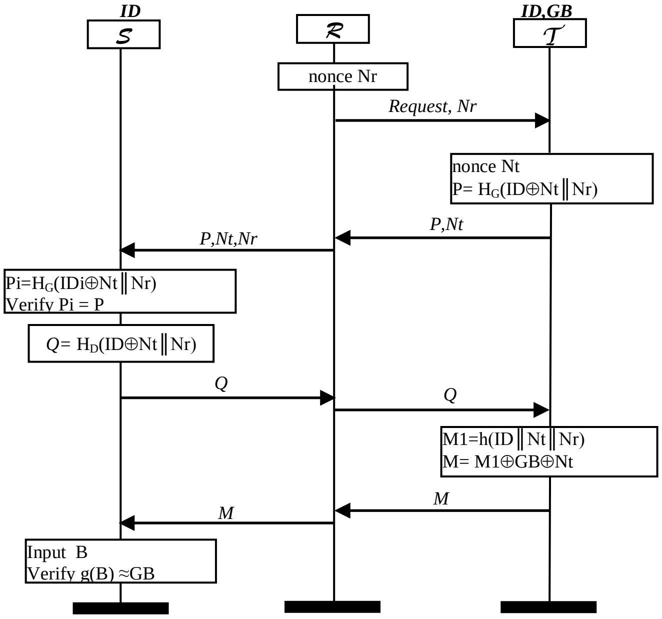


Figure 3. Proposed Protocol

Step 1: Challenge

The reader RFID produces a nonce Nr and sends it then, and a request to the tag. Three cases can occur: *1)* No tag answers, *2)* A tag answers, *3)* Many tags answer at the same time. In our protocol, the last case is not approved because one requires the capture of only one biometry for every person and for every tag in every process of authentication.

Step 2: Authentication of the tag

Step 2.1: the tag found in the step 1 generates a nonce Nt and computes $P = H_G(ID \oplus Nt \| Nr)$

Step 2.2: the tag sends P with the nonce Nt to the reader RFID,

Step 2.3: the reader resends the successful P message, Nt and the nonce Nr to the server.

Step 2.4: from the database, the server looks for certain IDi (such as 1 ≤ i ≤ n, n the number of tags) to compute $Pi=H_G(IDi \oplus Nt \| Nr)$, and make the following comparison:

$$Pi \ ?= P$$

If it is found, the tag crosses the authentication of the tag and is considered as legitimate, otherwise to end.

Step 3: Authentication of the server

Step 3.1: the server computes and sends to the reader Q;

$$Q = H_D(ID_i \oplus Nt \| Nr) \text{ such as } ID_i = ID$$

Step 3.2: the reader sends the Q message in the tag.

Step 3.3: the tag computes $H_D(ID \oplus Nt \| Nr)$ and verifies if:

$$Q \ ?= H_D(ID \oplus Nt \| Nr)$$

If they are equal, the authentication of the reader is successful; otherwise the authentication of the reader has failed.

Step 4: Verification of biometry

Step 4.1: the tag computes $M1 = h\ (ID \| Nt \| Nr)$ and makes operation or-exclusive of M1 with GB and Nt. The resultant message is $M = M2 \oplus GB \oplus Nt$.

Step 4.2: the tag sends M to the reader RFID, and the reader resends received message to the server.

Step 4.3: after acquiring of the biometry of the user from the sensor, it sends it to the server. The server extracts biometric characteristics and generates the template B. the server computes the biometric hash function of the template g(B).

Step 4.4: from the database, the server computes $M2=h\ (IDi \| Nt \| Nr) \oplus Nt$, such as $IDi= ID$ (of the step 2.4), and extracts GB from:

$$M2 \oplus GB = M$$

Step 4.5: to make the comparison of type 1:1 of g(B) ≈ GB, if it is confirmed, the person is a trusted user, otherwise, the bearer of the tag is illegitimate, the information of failure will be sent to the reader, the protocol is interrupted.

V. SECURITY VERIFICATION OF PROTOCOL

A. Automatic Verification

There are several tools of automatic verification of protocols. We chose tools AVISPA (*Automated Validation of internet Security Protocols and Applications*) [1] and SPAN (*Security Protocol ANimator*) [2] for the following reasons: four tools are available using various techniques of validation (Model-checking, automate trees, resolution of constraints, Solver SAT). These tools are based on the same language of specification: language HLPSL (*High-Level Protocol Specification Language*) [18]. The platform AVISPA is the analyzer which models a big number of protocols (more than 84 protocols). Among these four tools, two tools OFMC and CL-ATSE which can verify protocols requiring the operator or-exclusive (xor). Concerning our protocol, we verify the properties of the confidentiality of the identity ID (sec_id_TR and sec_id_RT respectively), the confidentiality of the template B (sec_b), the authentication of the tag (aut_tag) and the authentication of the reader (aut_reader). These properties are specified in HLPSL as follows:

```
goal
    secrecy_of sec_b, sec_id_TR, sec_id_RT
    authentication_on aut_reader
    authentication_on aut_tag
end goal
```

Concerning the authentication, there are two possible attacks: the replay attack and the attack Man-in-the-Middle. For it, we uses two types of specification in the role HLPSL's environment.

1) Replay Attack

In the replay attack, the intruder can listen to the message of answer of the tag and to the reader. It will broadcast the message listened without modification to the reader later.

Specification below of the role environment in HLPSL depends on the treatment of two identical sessions between the same tag and the same reader (t and r). This scenario allows discovering the attacks of the type replay attack if it exists.

```
role environment() def=
const t,r : agent,
      id,b : text,
      h,g,left,right : hash_func
intruder_knowledge = {t,r,h,g,hright,hleft}
composition
session(t,r,id,b,h,g,hright,hleft) /\
session(t,r,id,b,h,g,hright,hleft)
end role
```

After the verification of this protocol by AVISPA tools, result is as follows:

```
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  UNTYPED_MODEL
PROTOCOL
  C:\progra~1\SPAN\testsuite\results\BioMRFID.if
GOAL
  As Specified
BACKEND
  CL-AtSe
STATISTICS
  Analysed   : 600 states
  Reachable  : 188 states
  Translation: 0.01 seconds
  Computation: 0.02 seconds
```

This result means in clearly that there is no replay attack. We can thus deduct that the diagnosis of AVISPA&SPAN tools for this protocol is secure.

2) Main-in-the-midlle Attack

The scenario of the role `environment` below allows discovering the attacks of this type if it exists.

```
role environment() def=
const t,r : agent,
      id,b,idti,idri,bti,bri : text,
      h,g,hright,hleft : hash_func
intruder_knowledge = {t,r,h,g,
hright,hleft,idti,idri,bti,bri}
composition
     session(t,r,id,b,h,g,hright,hleft)
 /\  session(t,i,idti,bti,h,g,hright,hleft)
 /\  session(i,r,idri,bri,h,g,hright,hleft)
end role
```

The result of the check with this scenario is the same as with the scenario a). We can thus deduct that this protocol is resistant in the attack of the "man in the middle".

B. Security Analysis

We now analyze the security properties of the proposed protocol as follows: untraceability, desynchronization resistance and with Denial of service (DOS) attack prevention.

1) Untraceability :

During every session of authentication, an opponent can observe only the values of (Nt, Nr, M1, P, Q), where, Nt and Nr are random numbers and M1 and Q messages are calculated the right/ left part of the function H(ID⊕Nt‖Nr). The P message = H(ID‖Nt‖Nr)⊕GB⊕Nt. The opponent cannot deduce the value of ID because function H(ID‖Nt‖Nr) is very effective as is shown in the paper of [11]. In M1 messages, P and Q, the opponent cannot correlate ID and B because these two values are secret and Nt and Nr are random numbers changed in every authentication. So, an opponent cannot track tags.

2) Desynchronization Resistance :

Our protocol belongs to the static mechanism ID where the identifier of the tag is fixed. So, in the case of the loss of message, failing of energy or the loss of connection with the server during the authentication, it will not affect the database of the server and will not become an obstacle to the protocol.

3) DOS attack Prevention:

There are several categories of Dos attacks, one is to desynchronize the internal states of two entities, and the other is to exhaust the resources of the parties involved. For RFID authentication protocols, researchers are concerned about desynchronization.

For our protocol, the internal state ID is kept static and not changed during authentication process. So, it can resist the attack of denial of service.

In the Table I below, a comparison of the security with protocols mentioned early is given [4, 5, 6, and 7].

TABLE I : ANALYSIS OF SECURITY

RFID Protocol *(static ID)*	RLHS [5]	LCAP [6]	CH [4]	LHYC [7]	**Our Protocol**
Mutual Authentification	+	+	+	+	+
Replay attack prevention	-	+	-	+	+
Non traceability	-	+	+	+	+
DoS attack prevention	-	-	+	+	+
Desynchronization Resistance	+	+	+	+	+

VI. PERFORMANCE ANALYSIS

As compared with what follows in Table II. This table illustrates the storage cost, the communication cost, and the computation cost of entities. The computation cost is a function of the number of operations of the hash function in phase's login and the authentication on the smartcard for the biometric protocols, as well as of the number of operations of the hash function on the tag in RFID protocols.

Computation Cost: the tag used in the protocol proposed by Lee and al. (Protocol LHYC) [7] and the smart cards of the biometric protocols require an important number of operations for the hash function. On the contrary, in the protocol of Chien and Huang [4], it requires a random numbers generator with an input number, but it is necessary not to forget the replay algebraic attack.

In our protocol, we require two operations of calculation of function h in the tag, so these calculations are effective for RFID tags.

Communication Cost: Communication cost between a tag and a reader consists of: the number of message exchanges, and the total bit size of the transmitted messages, per each communication. Concerning our protocol, the total of the bits of the messages of communication tag to the reader is: 2½l and for the message of communication reader to tag is: ½l. With regard to the other protocols of smart cards the performance of the communication of our protocol is more effective.

TABLE II : PERFORMANCE ANALYSIS

Protocol		**Computation Cost Tag/Card**	**Storage Cost**	**Communication Cost**		
				R →T	**T →R**	**∑**
RFID	[4]	*1g*	*2l*	½*l*	1½*l*	*2l*
	[5]	*1h*	*1l*	-	*2l*	*2l*
	[6]	*2h*	*1l*	1*l*	*2l*	*3l*
	[7]	*4h*	*2l*	*1l*	*2l*	*3l*
Smart Card	[8]	*4h*	*3l*	*2l*	*3l*	*5l*
	[9]	*4h*	*3l*	*2l*	*3l*	*5l*
	[10]	*3h*	*4l*	*2l*	*3l*	*5l*
Our protocol		***2h***	***2l***	***½l***	***2½l***	***3l***

Notations: *h* - the cost of a hash function operation,
g - random number generator with an input number,
l: size of required memory.

Storage Cost: The amount of storage needed on the back-end server is also another important issue. In the biometric protocols [8, 9], the smart card requires 3l bit and 4l for the protocol [10]. In our protocol, the tag requires 2l bit to store the identity (ID) and the function h of template (GB). Consequently, in the implemented protocols, the tag requires only 2l bits at most of the memory, which is adapted to tags with weak cost.

We can conclude that our protocol is effective and adapted to RFID tags as far as the computation cost; the storage cost and the communication cost are concerned.

VII. CONCLUSION

We proposed in this article a new RFID authentication protocol which uses biometric data. Our protocol is compatible with the constrained computational and memory resources of the RFID tags. Concerning the problem of the size of biometric data, we applied the hash function to the biometric template, which allows to optimize and to protect these data. Our protocol realizes the secrecy private data, the authentication of the tag and the authentication of the reader. Experimental tests (with AVISPA and SPAN tools) proved it. We made an analysis of security on the efficiency of our protocol for untraceability, resistance for the denial of service (DOS) attack prevention and the desynchronization resistance.

The advantage of our protocol is that it can be used in decentralized applications since we have no need of biometric database of the users in the system.

Future research includes additional work in regards to the biometric hash function. There are many researches on the implementation of the robust hash function in RFID tags. But researches on the implementation of Biometric hash function are limited.

REFERENCES

[1] A. Armando, D. Basin, Y. Boichut, Y.Chevalier, L. Compagna, J. Cuellar, P. Hankes Drielsma, P.-C. H_eam, J. Mantovani, S. Mödersheim, D. von Oheimb, M. Rusinowitch, J. Santos Santiago, M. Turuani, L. Viganò, and L. Vigneron, "The AVISPA Tool for the automated validation of internet security protocols and applications," In K. Etessami and S. Rajamani, Eds. 17th International Conference on Computer Aided Verification, CAV'2005, vol. 3576, pp. 281-285, Edinburgh, Scotland, 2005.

[2] Y. Glouche, T. Genet, O. Heen, E. Houssay and R. Saillard, "SPAN (a Security Protocol ANimator for AVISPA) version 1.6," http://www.irisa.fr/celtique/genet/span/, 2009.

[3] D. Dolev and A. C. Yao, "On Security of Public Key Protocols," In proceding IEEE transactions on Information Theory, vol. 29, pp. 198-208, 1983.

[4] H.-Y. Chien, C.-W. Huang, "A lightweight RFID protocol using substring," in: EUC, pp. 422–431, 2007.

[5] S. Weis, S. Sarma, R. Rivest, and D. Engels. "Security and privacy aspects of low-cost radio frequency identification systems," In D. Hutter, and all., editors, International Conference on Security in Pervasive Computing – SPC 2003, vol. 2802 of LNCS, pp.454–469, Boppard, Germany, Springer-Verlag, March 2003.

[6] S.M. Lee, Y.J. Hwang, D.H. Lee, and J.I. Lim. "Efficient Authentication for Low-Cost RFID Systems," International Conference on Computational Science and its Applications - ICCSA 2005, May 2005.

[7] Y.C. Lee, Y.C. Hsieh, P.S. You and T.C. Chen, "An Improvement on RFID Authentication Protocol with Privacy Protection," Third 2008 International Conferences on Convergence and Hybrid Information Technology, South Korea: Busan, vol.2, pp. 569–573, 2008.

[8] M.K. Khan, J. Zhang, and X. Wang. "Chaotic hash-based fingerprint biometric remote user authentication scheme on mobile devices," Chaos, Solitons and Fractals, Vol. 35, No. 3, pp. 519-524, 2008.

[9] C.T. Li and M.S. Hwang. "An efficient biometrics-based remote user authentication scheme using smart cards," Journal of Network and Computer Applications, Vol. 33, pp. 1-5, 2010.

[10] Y.W. Lai , S.-C. Chang ,and C. Chang." An Improved Biometrics-based User Authentication Scheme without Concurrency System," International Journal of Intelligent Information Processing Vol. 1, N° 1, Sep 2010.

[11] A. Juels and S.A. Weis. "Defining strong privacy for RFID", In Proceedings of PerCom'07, pp. 342–347, http://eprint.iacr.org/2006/137, 2007.

[12] Y. Sutcu, H-T. Sencar, and N. Memon. "A Secure Biometric Authentication Scheme Based on Robust Hashing," MM-SEC'05, New York, USA, August 1–2, 2005.

[13] A. Nagar, K. Nandakumar, and A.K. Jain, "Biometric template transformation: a security analysis", in Proc. Media Forensics and Security, 2010.

[14] SmartCard Alliance. "Smart Cards and Biometrics," available to: wwww.smartcardalliance.org , mars 2011.

[15] Joel J.P.C. Rodrigues, F.D. Heirto, and B. Vaidya "Decentralized RFID authentication Solution for embedded Systems," 4th Int. Conference on Systems and Networks Communications, IEEE, pp. 174-178, 2009.

[16] H.A. Aboalsamh. "A Potable Biometric Access device using Dedicated Fingerprint Processor", WSEAS Transaction on Computers, Issue 8, Vol. 9, pp. 878-887, August 2010.

[17] T. van Deursen, S. Mauw, and S. Radomirović, "Untraceability of RFID protocols", In Information Security Theory and Practices. Smart Devices, Convergence and Next Generation Networks, Seville, Spain, Springer, vol. 5019, pp.1-15, 2008.

[18] Y. Chevalier, L. Compagna, J. Cuellar, P. Hankes Drielsma, J. Mantovani, S. Modersheim, and L. Vigneron, "A high level protocol specification language for industrial security-sensitive protocols," In Proceedings of Workshop on Specification and Automated Processing of Security Requirements (SAPS), Linz, Austria, September 2004.

[19] S. Lee, T. Asano, and K. Kim. "RFID mutual authentication scheme based on synchronized secret information," In Symposium on Cryptography and Information Security, Hiroshima, Japan, January 2006.

ATM Security Using Fingerprint Biometric Identifer: An Investigative Study

Moses Okechukwu Onyesolu
Department of Computer Science
Nnamdi Azikiwe University, Awka
Anambra State, Nigeria.

Ignatius Majesty Ezeani
Department of Computer Science
Nnamdi Azikiwe University, Awka
Anambra State, Nigeria.

***Abstract*—The growth in electronic transactions has resulted in a greater demand for fast and accurate user identification and authentication. Access codes for buildings, banks accounts and computer systems often use personal identification numbers (PIN's) for identification and security clearances. Conventional method of identification based on possession of ID cards or exclusive knowledge like a social security number or a password are not all together reliable. An embedded fingerprint biometric authentication scheme for automated teller machine (ATM) banking systems is proposed in this paper. In this scheme, a fingerprint biometric technique is fused with the ATM for person authentication to ameliorate the security level.**

Keywords- ATM; PIN; Fingerprint; security; biometric.

I. Introduction

Rapid development of banking technology has changed the way banking activities are dealt with. One banking technology that has impacted positively and negatively to banking activities and transactions is the advent of automated teller machine (ATM). With an ATM, a customer is able to conduct several banking activities such as cash withdrawal, money transfer, paying phone and electricity bills beyond official hours and physical interaction with bank staff. In a nutshell, ATM provides customers a quick and convenient way to access their bank accounts and to conduct financial transactions. Personal identification number (PIN) or password is one important aspect in ATM security system. PIN or password is commonly used to secure and protect financial information of customers from unauthorized access [1]. An ATM (known by other names such as automated banking machine, cashpoint, cash machine or a hole in the wall) is a mechanical system that has its roots embedded in the accounts and records of a banking institution [1]-[2]. It is a computerized machine designed to dispense cash to bank customers without need of human interaction; it can transfer money between bank accounts and provide other basic financial services such as balance enquiries, mini statement, withdrawal and fast cash among others [3].

The paper is arranged as follows. Section II provided the background of ATM security and the need for biometrics. Section III introduced the related works on biometric identifiers. Section IV described the materials and methods employed to conduct the survey. Section V presented the results obtained and the discussions on the results. Section VI concluded the paper.

II. Research Background

Crime at ATMs has become a nationwide issue that faces not only customers, but also bank operators and this financial crime case rises repeatedly in recent years [4]. A lot of criminals tamper with the ATM terminal and steal customers' card details by illegal means. Once users' bank card is lost and the password is stolen, the users' account is vulnerable to attack. Traditional ATM systems authenticate generally by using a card (credit, debit, or smart) and a password or PIN which no doubt has some defects [5]. The prevailing techniques of user authentication, which involves the use of either passwords and user IDs (identifiers), or identification cards and PINs (personal identification numbers), suffer from several limitations [6]. Passwords and PINs can be illicitly acquired by direct covert observation. When credit and ATM cards are lost or stolen, an unauthorized user can often come up with the correct personal codes. Despite warning, many people continue to choose easily guessed PIN's and passwords - birthdays, phone numbers and social security numbers. Recent cases of identity theft have heightened the need for methods to prove that someone is truly who he/she claims to be. Biometric authentication technology may solve this problem since a person's biometric data is undeniably connected to its owner, is nontransferable and unique for every individual. The system can compare scans to records stored in a central or local database or even on a smart card.

Biometrics can be defined as a measurable physiological and behavioral characteristic that can be captured and subsequently compared with another instance at the time of verification. It is automated methods of recognizing a person based on a physiological or behavioral characteristic [7]. It is a measure of an individual's unique physical or behavioral characteristics to recognize or authenticate its identity [8]. Common physical biometrics characteristics include fingerprint, hand or palm geometry, retina, iris and face while popular behavioral characteristics are signature and voice. Biometrics technologies are a secure means of authentication because biometrics data are unique, cannot be shared, cannot be copied and cannot be lost.

III. Related Works

Shaikh and Rabaiotti [9] analyzed the United Kingdom identity card scheme. Their analysis approached the scheme from the perspective of high volume public deployment and described a trade-off triangle model. They found that there is a trade-off between several characteristics, i.e., accuracy, privacy and scalability in biometric based identity management system, where emphasis on one undermines the other.

Amurthy and Redddy [6] developed an embedded fingerprint system, which is used for ATM security applications. In their system, bankers collect customers' finger prints and mobile numbers while opening accounts, then customer only access ATM machine. The working of the ATM machine is such that when a customer place a finger on the finger print module it automatically generates every time different 4-digit code as a message to the mobile of the authorized customer through GSM modem connected to the microcontroller. The code received by the customer is entered into the ATM machine by pressing the keys on the touch screen. After entering it checks whether it is a valid one or not and allows the customer further access. Das and

Schouten and Jacobs [8] presented an evaluation of the Netherlands' proposed implementation of a biometric passport, largely focusing on technical aspects of specific biometric technologies (such as face and fingerprint recognition) but also making reference to international agreements and standards (such as ICAO and the EU's ''Extended Access Control'') and discussed the privacy issue in terms of traditional security concepts such as confidentiality. Debbarma [1] proposed an embedded Crypto-Biometric authentication scheme for ATM banking system.

IV. Materials and Methods

The target population of this study was customers and staff of some commercial banks in Awka, Anambra State, South-Eastern Nigeria. The customers and students were randomly selected. The instrument used for this study was a 16-item questionnaire developed by the researchers. The items in the questionnaire were derived from extensive survey of relevant literature and oral interview. The instrument has three sections. The first section deals with participants' profile. The second section deals with participants' use and reliability of ATM. The third section deals with the reliability of fingerprint biometric characteristic. Of the 200 copies of the questionnaire administered, 163 usable copies were returned. This represented 82 percent return rate. This study was carried out over a period of four months. The items in the instrument were analyzed using descriptive statistical methods [10]-[13]. The secondary sources of data were obtained from journals, the Internet and textbooks. Expert judgments were used to ascertain the validity of the items in the questionnaire. Two experts face-validated all the items in the questionnaire. The wordings of items were also checked for clarity. Two items in the questionnaire were deleted for irrelevance while three ambiguously worded items were restructured to reflect clarity. After the corrections, the two experts found the items to be suitable for administration on the subjects. The reliability co-efficient of the instrument was tested by using the Cronbach alpha which is adequate for reliability measure. The instrument yielded a reliability coefficient of 0.81.

V. Results and Discussion

The summary of the results obtained is presented (Tables I – III). Table I shows the profile of participants. The range of age of participants was 20-53 years. 85 males and 78 females took part in the study.

TABLE I. Profile of Participants

No	Profile	Description
1.	Age	20-53 years old
2.	Sex (male: female)	85:78
3.	Bank account and ATM card	Respondents own different types of account depending on the bank, bank products and types of services rendered.

Each of the participants own at least one type of bank account. This depended on the bank, the products offered and services provided by the bank. Banks in Nigeria continually churn out new products and services to have competitive advantage. This resulted to the introduction of ATM and the services it provides.

Table II shows the use and reliability of ATM. 139 respondents representing some customers and staff of some banks, representing 85 percent of the population use the ATM while 15 percent of the population is yet to use the machine. This 15 percent of the population is still skeptical about using ATM because of the issues associated with it. Such issues as inability of the machine to return a customer's card after transaction which may take days to rectify, debiting a customer's account in a transaction even when the customer is not paid and cash not dispensed, and "out of service" usually displayed by the machine which most of the time is disappointing and frustrating among others. 100 percent of the population is aware of one form of ATM fraud or another. Most banks in Nigeria constantly update their customers on ATM frauds and measures to be taken to avert them. 89 percent of the population thinks that ATM transactions are becoming too risky this necessitated 93 percent of the population affirming that they will continue the use of ATM because of security issues associated with the machine. Hence, 100 percent of the population preferred a third authentication aside the use of ATM card and PIN and this population believed that with the infusion of biometrics characteristics to the existing ATM card and PIN, ATM security will be improved drastically.

Table III shows the reliability and popularity of fingerprint biometric character. 74 percent of the population is familiar with fingerprint biometric. 63 percent of the population strongly believed that with the incorporation of fingerprint to the existing ATM card and PIN, will provide a better security to the ATM.

TABLE II. USE AND RELIABILTY OF ATM

No	Question	Responses		Total	Percentage (%)	
		Yes	No		Yes	No
1.	Do you use ATM?	139	24	163	85	15
2.	Is password (PIN) secured in using ATM?	60	103	163	37	63
3.	How Long have you been using ATM?			163		
	a. Less than a year	23			14	
	b. Greater than one year but less than 3 years	37			23	
	c. More than 3 years	103			63	
4.	Have you ever heard of any ATM fraud?	163	0	163	100	00
5.	Is anything being done about ATM fraud?	150	13	163	92	08
6.	Are ATM transactions becoming especially risky?	145	18	163	89	11
7.	Will you discontinue the use of ATM because of the security issues associated with it?	152	11	163	93	07
8.	Would you prefer a third level security aside card and PIN?	163	0	163	100	00
9.	Have you heard of biometrics as a means of authentication?	143	20	163	88	12
10.	Do you think the use of biometrics can improve ATM security?	163	0	163	100	00

TABLE III. RELIABILTY OF BIOMETRICS CHARACTERISTICS

S/N	Question	Biometric Characteristic	Responses	Percentage (%)
1.	Which of these biometric characteristics have you heard of?	a. Fingerprint	120	74
		b. Iris	9	06
		c. Face Recognition	7	04
		d. Signature	12	07
		e. DNA	2	01
		f. Retina	8	05
		g. voice	5	03
		Total	163	100
2.	Which of the biometrics characteristic will provide better security when fused with the ATM?	a. Fingerprint	101	62
		b. Iris	13	08
		c. Face Recognition	9	06
		d. Signature	12	07
		e. DNA	12	07
		f. Retina	9	06
		g. voice	7	04
		Total	163	100

A. Fingerprint Biometrics

The use of fingerprints as a biometric is both the oldest mode of computer-aided, personal identification and the most prevalent in use today [14].

In the world today, fingerprint is one of the essential variables used for enforcing security and maintaining a reliable identification of any individual. Fingerprints are used as variables of security during voting, examination, operation of bank accounts among others. They are also used for controlling access to highly secured places like offices, equipment rooms, control centers and so on [15]. The result of the survey conducted by the International Biometric Group (IBG) in 2004 on comparative analysis of fingerprint with other biometrics is presented in Fig. 1.

The result shows that a substantial margin exists between the uses of fingerprint for identification over other biometrics such as face, hand, iris, voice, signature and middleware [16].

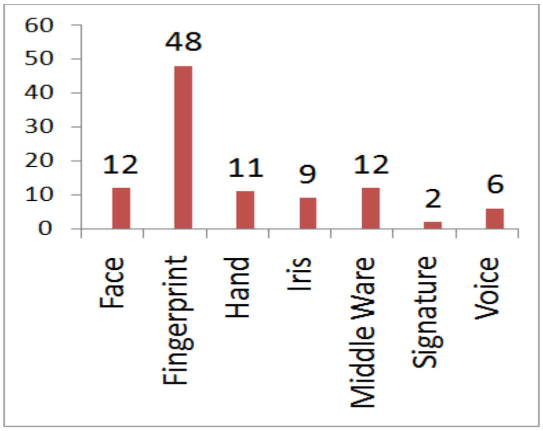


Figure 1. Comparative survey of fingerprint with other biometrics

References [16]-[19] adduced the following reasons to the wide use and acceptability of fingerprints for enforcing or controlling security:

a. Fingerprints have a wide variation since no two people have identical prints.
b. There is high degree of consistency in fingerprints. A person's fingerprints may change in scale but not in relative appearance, which is not the case in other biometrics.
c. Fingerprints are left each time the finger contacts a surface.
d. Availability of small and inexpensive fingerprint capture devices.
e. Availability of fast computing hardware.
f. Availability of high recognition rate and speed devices that meet the needs of many applications
g. The explosive growth of network and Internet transactions
h. The heightened awareness of the need for ease-of-use as an essential component of reliable security.

VI. Conclusion

The growth in electronic transactions has resulted in a greater demand for fast and accurate user identification and authentication. Access codes for buildings, banks accounts and computer systems often use PIN's for identification and security clearances. Conventional method of identification based on possession of ID cards or exclusive knowledge like a social security number or a password are not all together reliable. ID cards can be lost, forged or misplaced; passwords can be forgotten or compromised, but ones' biometric is undeniably connected to its owner. It cannot be borrowed, stolen or easily forged. Using the proper PIN gains access, but the user of the PIN is not verified. When credit and ATM cards are lost or stolen, an unauthorized user can often come up with the correct personal codes. Despite warning, many people continue to choose easily guessed PIN's and passwords - birthdays, phone numbers and social security numbers. Recent cases of identity theft have heightened the need for methods to prove that someone is truly who he/she claims to be. Biometric authentication technology using fingerprint identifier may solve this problem since a person's biometric data is undeniably connected to its owner, is nontransferable and unique for every individual. Biometrics is not only a fascinating pattern recognition research problem but, if carefully used, could also be an enabling technology with the potential to make our society safer, reduce fraud and lead to user convenience by broadly providing the following three functionalities (a) positive identification (b) large scale identification and (c) screening.

References

[1] S.S, Das and J. Debbarma, "Designing a Biometric Strategy (Fingerprint) Measure for Enhancing ATM Security in Indian e-banking System", International Journal of Information and Communication Technology Research, vol.1, no. 5, pp.197-203, 2011.

[2] W.W.N. Wan, C.L. Luk, and C.W.C. Chow, "Customers Adoption of Banking Channels in Hong Kong", International Journal of Bank Marketing, vol. 23, no. 3, pp. 255-272, 2005.

[3] Wikipedia the free encyclopedia, "Biometrics", Downloaded March 20, 2012 from http://en.wikipedia.org/wiki/Biometrics.

[4] B. Richard and M. Alemayehu, "Developing E-banking Capabilities in a Ghanaian Bank: Preliminary Lessons. Journal of Internet Banking and Commerce, vol. 11, no. 2, 2006. Downloaded March 15, 2012 from http://www.arraydev.com/commerce/jibc/

[5] P.K. Amurthy and M.S. Redddy, "Implementation of ATM Security by Using Fingerprint recognition and GSM", International Journal of Electronics Communication and Computer Engineering vol.3, no. 1, pp. 83-86, 2012.

[6] N.K. Ratha, J.H. Connell, and R.M. Bolle, "Enhancing Security and Privacy in Biometrics-based Authentication Systems, IBM Systems Journal, vol. 40, no. 3, pp. 614-634, 2001.

[7] N.K. Ratha, S. Chikkerur, J.H. Connell and R.M. Bolle. "Generating Cancelable Fingerprint Templates", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 29, no. 4, 2007.

[8] B, Schouten and B. Jacobs, "Biometrics and their use in e-passport", Image and Vision Computing vol. 27, pp. 305–312. 2009,

[9] S.A. Shaikh and J.R. Rabaiotti,. "Characteristic trade-offs in designing large-scale biometric-based identity management systems". Journal of Network and Computer Applications vol. 33, pp. 342–351, 2010.

[10] C.A. Oyeka, An Introduction to Applied Statistical methods. Enugu, Nigeria: Modern Avocation Publishing Company. Pp. 4, 36, 56. 1990.

[11] E.O. Akuezilo, and N. Agu, Research and Statistics in Education and Social Sciences: Methods and Applications. Awka, Nigeria: Nuel Centi Publishers and Academic Press Ltd, 1993.

[12] J.I. Eze, M.E. Obiegbu, and E.N. Jude-Eze, Statistics and Qualitative Methods for Construction and Business Managers. Lagos, Nigeria: The Nigerian Institute of Building, 2005.

[13] F.H. Zuwaylif, General Applied Statistics. 3rd. Ed., California: Addison Wesley Publishing Company, 1999.

[14] L. O'Gorman, "Overview of fingerprint verification technologies", Elsevier Information Security Technical Report, vol. 3, no. 1, 1998.

[15] G.B. Iwasokun, O.C. Akinyokun, B.K. Alese, and O. Olabode. "Fingerprint Image enhancement: Segmentation to thinning", International Journal of Advanced Computer Science and Applications, vol. 3, no. 1, pp. 15-24, 2012.

[16] C. Roberts, "Biometrics". Downloaded February 13, 2012 from http://www.ccip.govt.nz/newsroom/informoationnotes/2005/biometrics .pdf,

[17] C. Michael and E. Imwinkelried, "Defence practice tips, a cautionary note about fingerprint analysis and reliance on digital technology", Public Defense Backup Centre Report, 2006

[18] M. J. Palmiotto, Criminal Investigation, Chicago: Nelson Hall, 1994

[19] D. Salter, "Fingerprint: An Emerging Technology", Engineering Technology, New Mexico State University, 2006

AUTHORS PROFILE

Moses Okechukwu Onyesolu. Ph.D. in Modeling and Simulation (Nnamdi Azikiwe University, Awka, Nigeria, 2011). Researcher/Lecturer, Nnamdi Azikiwe University, awka, Nigeria.

Ignatius Majesty Ezeani. B.Sc. in Computer Science (Nnamdi Azikiwe University, Awka, Nigeria), M.Sc. Software Engineering (University of Bournemouth, Bournemouth, UK, 2006). Researcher/Lecturer, Nnamdi Azikiwe University, Awka, Nigeria.

A Globally Convergent Algorithm of Variational Inequality

Chengjiang Yin
Linyi University at Feixian
Feixian, Shandong, P.R.China

Abstract- **The algorithm of variational inequality is the important and valuable question in real life all the time. In this paper, a globally convergent algorithm of variational inequality is proposed. The method ensures that the corrector step sizes have a uniformly positive bound from below. In order to prove convergence of algorithm, we first establish some definitions, properties and theorem, and then we prove its global convergence under appropriate conditions.**

Keywords-Iinfinite The variational inequality; Uniformly positive bound from below; Global convergence.

I. Introduction

Consider the generalized variational inequalities: we seek $u^* \in R^n$,that satisfies $g(u^*) \in \Omega$ and

$$(g(u)-g(u^*))^T F(u^*) > 0, \forall g(u) \in \Omega$$
(1)

Here $F: R^n \to R^n$, $G: R^n \to \Omega$ are given nonlinear operators and $\Omega \subseteq R^n$ is non-empty closed convex set. Noor had introduced problem (1)in his literature[1], and generalized variational inequalities is called GVI. GVI have important applications in economics and nonlinear analysis etc.

When $g(u) \equiv u$, GVI degenerate into VI: seek a vector $u^* \in \Omega$,that Satisfies $(u-u^*)^T F(u^*) \ge 0$ and $\forall u \in \Omega$.

Sun studied VI in his literature [2],and gave a new method to find the descent direction in the following way:

$$d(u,\beta) = e(u,\beta) - \beta F(u) + \beta F(p_\Omega[u-\beta F(u)])$$

Where $e(u,\beta) = u - p_\Omega[u-\beta F(u)]$ and p_Ω is a rectangular projection from R^n to Ω .

Assuming that the underlying mapping F is pseudo-monotone, the author proves that this new method is globally convergent and gives a necessary and sufficient condition when the solution set is nonempty. Using literature thoughts [2-6], we propose a modified projection-type method of GVI.

II. Definition and properties

In order to prove convergence of algorithm, we need someconclusions and assumptions, review that as fellows:

For a nonempty closed convex-set $\Omega \subseteq R^n$, where

$u \in R^n$, set $p_\Omega(u)$ is a rectangular projection from u to Ω , that means $p_\Omega(u) = \arg\min\{\|v-u\|, u,v \in \Omega\}$.

Some important properties of projection are as follows:

Lemma1 If $\Omega \subseteq R^n$ is a nonempty-closed convex set,for all $u \in R^n$ and $v \in \Omega$ .then we have that

$$\langle p_\Omega(u)-u, v-p_\Omega(u)\rangle \ge 0 \qquad (2)$$

According to Lemma1 we have a conclusion that projection operator is non-expansive,and we also obtain an equivalent proposition on solution set of GVI.

Lemma2 Vector $u^* \in R^n$ is a solution of GVI if and only if $g(u^*) = p_\Omega\left[g(u^*) - \beta F(u^*)\right]$, for all $\beta > 0$.

Assumptions:

1) Ω^* is non-empty set : and $\Omega^* \in R^n$ is a solution set of GVI.

2) $F: R^n \to R^n$ is g- monotonous projection mapping, such that: $\langle F(u)-F(v), g(u)-g(v)\rangle \ge 0$,for all $u,v \in R^n$.

3) $g: R^n \to R^n$ is nonsingular mapping, there is a positive constant φ satisfies $\|g(u)-g(v)\| \ge \varphi\|v-u\|$.for all $u,v \in R^n$

In this paper, we define residual vector $e(u,\beta)$ as fellows:

$$e(u,\beta) := g(u) - p_\Omega[g(u)-\beta F(u)] .$$

Lemma3 Suppose that $\Omega \subseteq R^n$ is a nonempty-closed convex set, we choose optional $u \in R^n$ and $\beta_1 > \beta_2 > 0$,then we have

1) $\|e(u,\beta_1)\| \ge \|e(u,\beta_2)\|$；

2) $\|e(u,\beta_1)\|/\beta_1 \ge \|e(u,\beta_2)\|/\beta_2$

III. Algorithm and Convergence

A. Algorithm

Initial step:

set $u^0 \in R^n$ and $g(u^0) \in \Omega$ and δ , $\mu \in$ (0,1), $\gamma \in$ (0,2). If $e(u^0,1) = 0$, then stop it； set k:=0,otherwise.

Iterative step:Set $\beta_k = \mu^{m_k}$ and m_k is a minimal nonnegative integer m and satisfies formula below

$$\mu^m \| F(g(u^k)) - F(g(u^k) - e(u^k, \mu^m)) \| \le \delta \| e(u^k, \mu^m) \|$$

(3)

Set $g(u^{k+1}) = p_\Omega\left[g(u^k) - \gamma\rho(u^k, \beta_k) d(u^k, \beta_k) \right]$,

where

$$d(u^k, \beta_k) = e(u^k, \beta_k) - \beta_k F(g(u^k)) + \beta_k F(p_\Omega\left[g(u^k) - \beta_k F(u^k) \right]),$$

$$\rho(u^k, \beta_k) = (1-\delta) \| e(u^k, \beta_k) \|^2 / \| d(u^k, \beta_k) \|^2 .$$

If $e(u^k, \beta_k) = 0$, then stop; the iterative process continues otherwise.

Lemma4 Suppose $u \in R^n$, $g(u) \in \Omega$,and u isn`t a solution of GVI, choose arbitrary $\delta \in (0,1)$,exist $\tilde{\beta}(u) \in (0,1]$,for all $\beta \in \left(0, \tilde{\beta}(u)\right]$,we have

$$\beta \| F(g(u)) - F(g(u) - e(u,\beta)) \| \le \delta \| e(u,\beta) \| \qquad (4)$$

PROOF. Suppose that there is $\delta \in (0,1)$, for arbitrary $\tilde{\beta} \in (0,1]$, exist $\beta \in \left(0, \tilde{\beta}\right]$,it satisfies such that

$$\beta \| F(g(u)) - F(g(u) - e(u,\beta)) \| > \delta \| e(u,\beta) \|$$

If $\tilde{\beta}_1 = 1$, there is $\beta_1 \in \left(0, \tilde{\beta}_1\right]$, we obtain that

$$\beta_1 \| F(g(u)) - F(g(u) - e(u,\beta_1)) \| > \delta \| e(u,\beta_1) \|.$$

If $\tilde{\beta}_2 = \min\{ 1/2, \beta_1 \}$, there is $\beta_2 \in \left(0, \tilde{\beta}_2\right]$,we obtain that

$$\beta_2 \| F(g(u)) - F(g(u) - e(u,\beta_2)) \| > \delta \| e(u,\beta_2) \|.$$

....

If $\tilde{\beta}_n = \min\{ 1/n, \beta_{n-1} \}$, there is $\beta_n \in \left(0, \tilde{\beta}_n\right]$,we have that

$$\beta_n \| F(g(u)) - F(g(u) - e(u,\beta_n)) \| > \delta \| e(u,\beta_n) \|$$

....

when $n \ge 1$,we obtain a sequence $\{\beta_n\}$,such that $\beta_n \le \frac{1}{n}$ and

$$\beta_n \| F(g(u)) - F(g(u) - e(u,\beta_n)) \| > \delta \| e(u,\beta_n) \|$$

(5)

From $g(u) \in \Omega$ and the expression of $e(u,\beta)$ and $\beta_n \le 1/n$, we have $\lim_{n\to\infty} e(u,\beta_n) = 0$ .According to formula(5)and the continuity of F and the conclusion of Lemma3,we obtain that

$$0 = \lim_{n\to\infty} \| F(g(u)) - F(g(u) - e(u,\beta_n)) \| \ge$$

$$\delta \lim_{n\to\infty} (\| e(u,\beta_n) \| / \beta_n) \ge \delta \| e(u,1) \|$$

By combining Lemma2, obviously, we have $u \in \Omega^*$,this is contradiction for the known .and the result is proved.

□

Lemma5 Suppose that $F(x)$ is g- monotonic mapping, we choose $u^* \in \Omega^*$ arbitrarily, when $u \in R^n$ and $g(u) \in \Omega$,we have that

$$(g(u) - g(u^*))^T d(u,\beta) \ge \| e(u,\beta) \|^2 -$$

$$\beta e(u,\beta)^T (F(g(u) - F(g(u) - e(u,\beta)))$$

PROOF. From (2),we obtain that

$$\left\{ g(u) - \beta F(u) - p_\Omega[g(u) - \beta F(u)] \right\}^T \left\{ p_\Omega[g(u) - \beta F(u)] - g(u^*) \right\} \ge 0$$

i.e. $(e(u,\beta) - \beta F(u))^T (g(u) - g(u^*) - e(u,\beta)) \ge 0$,

so we have that

$$e(u,\beta)^T (g(u) - g(u^*)) \ge \| e(u,\beta) \|^2 +$$

$$\beta F(u)^T (p_\Omega[g(u) - \beta F(u)] - g(u^*)) \qquad (6)$$

From g- monotonic properties of F, we have

$$\beta(F(p_\Omega[g(u) - \beta F(u)]) - F(g(u^*)))^T (p_\Omega[g(u) - \beta F(u)] - g(u^*)) \ge 0 \qquad (7)$$

From $u^* \in \Omega^*$ and $p_\Omega[g(u) - \beta F(u)] \in \Omega$, we have that

$$\beta(p_\Omega[g(u) - \beta F(u)] - g(u^*))^T F(g(u^*)) \ge 0 \qquad (8)$$

Add(6),(7),(8),we obtain that

$$\{ e(u,\beta) + \beta F(g(u) - e(u,\beta)) \}^T (g(u) - g(u^*)) \ge$$

$$\| e(u,\beta) \|^2 + \beta F(u)^T (g(u) - g(u^*) - e(u,\beta)) +$$

$$\beta e(u,\beta)^T F(g(u) - e(u,\beta)) .$$

By transposing and sorting , the proposition is founded.

According to algorithm design and Lemma5,For all u^k and $u^* \in \Omega$,we have

$$(g(u^k) - g(u^*))^T d(u^k, \beta_k) \ge (1-\delta) \| e(u^k, \beta_k) \|^2 \qquad (9)$$

The result is proved.

Theorem1 Suppose that $F(x)$ is a continuous and g-monotonic mapping projection, so the sequence $\{u^k\}$ generate by algorithm is bounded.

PROOF. Suppose $u^* \in \Omega^*$, then we have

$$\| g(u^{k+1}) - g(u^*) \|^2 \le \| g(u^k) - g(u^*) -$$

$$\gamma\rho(u^k, \beta_k) d(u^k, \beta_k) \|^2 - \| g(u^k) - g(u^*) \|^2$$

$$-2\gamma\rho(u^k, \beta_k)(g(u^k) - g(u^*))^T d(u^k, \beta_k) +$$

$$\gamma^2 \rho(u^k, \beta_k)^2 \| d(u^k, \beta_k) \|^2 \le$$

$$\| g(u^k) - g(u^*) \|^2 - \gamma(2-\gamma)(1-\delta)\rho(u^k, \beta_k) \| e(u^k, \beta_k) \|^2$$

For $\gamma \in (0,2)$ and $\delta \in (0,1)$,use the formula of above,we have that

$$\| g(u^{k+1}) - g(u^*) \| \le \| g(u^k) - g(u^*) \| \le \cdots \le \| g(u^0) - g(u^*) \|$$

By combining assumption3),the sequence $\{u^k\}$ which generated by algorithm is bounded.The result is proved.

□

From theorem1,we obtain this algorithm is a projection-type and contraction method.

Lemma6 $\rho(u^k, \beta_k)$ which is generated by algorithm is a uniformly positive bound from below. i.e. there is $\tau > 0$,such that $\rho(u^k, \beta_k) \ge \tau$.

PROOF. From the expression of $d(u^k, \beta_k)$ and(3),we obtain that

$$\|d(u^k,\beta_k)\|^2 \le 2\|e(u^k,\beta_k)\|^2 + 2\beta_k{}^2\|F(g(u^k)) - F(p_\Omega\left[g(u^k)-e(u^k,\beta_k)\right])\|^2 \le 2(1+\delta^2)\|e(u^k,\beta_k)\|^2$$

According to the expression of $\rho(u^k,\beta_k)$, we have

$$\rho(u^k,\beta_k) \ge \frac{1-\delta}{2(1+\delta^2)} := \tau$$

So the step length of this algorithm is a uniformly positive bound from below. the result is proved. □

From Lemma6 and Theorem1,we have that

$$\|g(u^{k+1})-g(u^*)\|^2 \le \|g(u^k)-g(u^*)\|^2 - \tau\gamma(2-\gamma)(1-\delta)\|e(u^k,\beta_k)\|^2$$

Such that

$$\sum_{k=0}^{\infty}\|e(u^k,\beta_k)\|^2 \le \infty \text{ i.e. } \lim_{k\to\infty}\|e(u^k,\beta_k)\| = 0 \qquad (10)$$

2.2 Proof of the global convergence

Theorem2 Suppose that the conditions of theorem1 holds ,we have that

1) $\lim_{k\to\infty}\|e(u^k,\beta_k)\|/\beta_k = 0$; 2) $\{u^k\}$ converges to the solution of GVI.

PROOF. 1) Suppose that there is a boundless index subset K_0 ,we have

$$\|e(u^k,\beta_k)\|/\beta_k \ge \varepsilon > 0 \quad \text{for all } k\in K_0,$$

by combining $\lim_{k\to\infty}\|e(u^k,\beta_k)\| = 0$,we obtain that

$\lim_{x\to\infty,k\in K_0}\beta_k = 0$,because $\{u^k\}$ is bounded , so $\{F(u^k)\}$ is bounded too, according to non-expansivity of projection algorithm, we obtain that

$$\lim_{x\to\infty,k\in K_0}\|g(u^k)-p_\Omega\left[g(u^k)-\beta_kF(u^k)/\mu\right]\| \le \lim_{x\to\infty,k\in K_0}\beta_k\|F(u^k)\|/\mu = 0,$$

by combining the continuity of F and the conclusion of lemma3 and liner search (3),for all $k\in K_0$ and $k\to\infty$ we obtain that

$$0 \leftarrow \|F(g(u^k)) - F(p_\Omega\left[g(u^k)-\beta_kF(u^k)/\mu\right])\| > \delta\frac{\|e(u^k,\beta_k/\mu)\|}{\beta_k/\mu} \ge \mu\delta\frac{\|e(u^k,\beta_k)\|}{\beta_k}$$

this is contradiction for the assumptions ,then the result is proved.

2) Divide two kinds of case

i) $\limsup_{k\to\infty}\beta_k > 0$ i.e. there is $\varepsilon_0 > 0$ and K_1 ,such that $\beta_k \ge \varepsilon_0$, $k\in K_1$.

From $\|e(u^k,\beta_k)\| \ge \|e(u^k,\varepsilon_0)\|$ and $k\in K_1$ and (10) we have that $\|e(u^k,\varepsilon_0)\| \to 0$.

Because $\{u^k\}$ is bounded ,so we ensure that there is a convergent subset $\{u^{k_j}\}$,and suppose its limiting point is $\bar{u}$ .By combining the continuity of $\|e(u,\varepsilon_0)\|$,we obtain that

$$\|e(\bar{u},\varepsilon_0)\| = 0 \text{ , means } \bar{u}\in\Omega^* .$$

Proof of global convergence is below. From theorem1,we have

$$\|g(u^{k+1})-g(\bar{u})\| \le \|g(u^k)-g(\bar{u})\|,$$

we choose $\hat{u}$ arbitrarily , $\hat{u}$ is an accumulation point of $\{u^k\}$, and $\{u^{k_l}\}$ is a convergence subsequence, there is $k_l > k_j$ for all k_j ,such that

$$\|g(u^{k_l})-g(\bar{u})\| \le \|g(u^{k_j})-g(\bar{u})\|,$$

set $k_j\to\infty$,by combining $k_l > k_j$,we have

$$\|g(\hat{u})-g(\bar{u})\| \le \|g(\bar{u})-g(\bar{u})\| = 0 .$$

Then $g(\hat{u}) = g(\bar{u})$, by combining 3) ,there is $\hat{u} = \bar{u}$ .Thus sequence $\{u^k\}$ global converge to $\bar{u}$.

ii) $\lim_{k\to\infty}\beta_k = 0$,from the conclusion 2) of lemma3, when k is big enough, we obtain that

$$\frac{\|e(u^k,\beta_k)\|}{\beta_k} \ge \|e(u^k,1)\|,$$

from the conclusion 1) of theorem2 and the formula of above, we have that $\|e(u^k,1)\| \to 0$.

From the similarly analysis as i) we proved theorem. □

IV. Conclusions And Prospect

As everyone knows, variational inequality originated in the mathematical physics problem with the nonlinear programming, In this paper, we get a new globally convergent algorithm of variational inequality, it can be widely used in the physical, mechanical, engineering, economic and other fields ,how to popularize It better in Practical Engineering Application is a topic for future research.

Acknowledgment

The authors wish to give their sincere thanks to the anonymous referees for their valuable suggestions and helpful comments which improved the presentation of the paper.

References

[1] NOOR M.A. General Variational Inequalities[J]. Appl Math Lett,1988,1(2):119-121.

[2] SUN D. A Class of Iterative Methods for solving Nonlinear Projection Equations[J]. J Optim Theory Appl,1996,91:123-140

[3] XIU N H,ZHANG J Z. Global projective-type error Bound for generalized variational inequalities [J]. J Optim Theory and Appl,2002,112(1:213-228)

[4] SOLODOV M V. Convergence Rate Analysis of Iteractive Algorithms for solving Variational Inequality Problems[J].Math Programming,2003,96:513-528

[5] Cheng X D. In Banach space reflexive solutions of a class of variational inequalities are given[J].Chongqing normal college journals.1991,16(2):44-48

[6] Wang C W. A variational inequalities problem solving the projection algorithm[J].Chongqing normal college journals.2005,22(1):6-1

Reversible Anonymization of DICOM Images using Cryptography and Digital Watermarking

Youssef ZAZ*
Department of Computer Sciences
Faculty of Sciences, Abdelmalek Essaadi University
Tetuan, Morocco

Lhoussain ELFADIL
FPO
Ibn Zohr University
Ouarzazate, Morocco

Abstract—**Digital Imaging and Communications in Medicine (DICOM) is a standard for handling, storing, printing, and transmitting information in medical images. The DICOM file contains the image data and a number of attributes such as identified patient data (name, age, insurance ID card,…), and non-identified patient data (doctor's interpretation, image type,…). Medical images serve not only for examination, but can also be used for research and education purposes. For research they are used to prevent illegal use of information; before authorizing researchers to use these images, the medical staff deletes all the data which would reveal the patient identity to prevent patient privacy. This manipulation is called anonymization. In this paper, we propose a reversible anonymization of DICOM images. Identifying patient data with image digest, computed by the well-known SHA-256 hash function, are encrypted using the proposed probabilistic public key crypto-system. After compressing the Least Significant Bit (LSB) bitplan of the image using Hofmann coding algorithm, the encrypted data is inserted into a liberated zone of the LSB bitplan of the image. The proposed method allows researchers to use anonymous DICOM images and keep to authorized staff -if necessary- the possibility to return to the original image with all related patient data.**

Keywords-DICOM images; watermaking; Hofmann coding; reversible anonymization, public key cryptosystem.

I. Introduction

DICOM images contain different kind of information, intermixing identifying patient data (I-Data) and non-identifying patient data (M-data) in a single file. To use these images by scientific researchers or for teaching purposes, hospitals proceed to the image anonymization by deleting all I-Data to ensure the patient privacy. Several software and web based applications were proposed to ensure this anonymization, as proposed by [5], [4] and [6].

For research purposes, sometimes the return to some I-Data in order to explain typical phenomenon is inescapable, but, the images are already anonymized and there is no way to use those information. To deal with this problem, [2] proposed to substitute I-Data by a unique anonymous token. In case that later an authenticated user needs full access to an image, the token can be used for re-linking separated I-Data and M-Data. [7] proposes to extract and save identifying data in another database and non-identifying data is stored in the archive.

When data is requested, the proposed system resolves the correlating and gathers the person-identifying information from the separate database. Another web-based separation is proposed in [8]. All above methods circumvent the main objective of DICOM images, which is to keep the image and the related data in the same file. To ensure the anonymization with keeping I-Data in the same file, the watermarking techniques are unavoidable. [3] proposed embedding the digest computed by SHA-256, in the Region of Non-Interest (RONI) of the image LSB bitplan. This method presents some difficulties to determine the RONI.

In this paper, we propose a reversible anonymization of DICOM images based on cryptography and watermarking. After liberating a space in LSB bitplan of the host image by compressing the original LSB bitplan using the Hoffmann coding, the I-Data and the image digest computed by SHA digital signature algorithm, are encrypted using the proposed public key crypto-system and inserted in liberated zone.

This paper is organized as follows: Section 2 gives a brief review of DICOM standard. Section 3 explains the security requirement for medical data storage. Section 4 exposes the proposed public key crypto-system. Section 5 exposes the global algorithm of reversible anonymization, and the last section concludes the paper.

II. DICOM Images

Introduced in 1993, DICOM (Digital Imaging and Communications in Medicine) a technology standard that is used virtually in Hospitals, clinics, imaging centers and specialists. Its structure is designed to ensure the interoperability of systems used to produce, store, display, send, …, and retrieve medical images and derived structured documents as well as to manage related workflow.

DICOM is required by all Electronic Health Records Systems that include imaging information as an integral part of the patient record.

DICOM is used in radiology, cardiology, radiotherapy, oncology, ophthalmology, dentistry, and so on.

For more description about DICOM, see the official web site [13].

III. Security Requirements for Medical Data Storage

To preserve patient privacy, all medical data are considered as sensitive. To read the content of an image, a user should be authorized. To prevent data infiltration, the anonymization prevents the exposure of identified patient data to unauthorized users. Many techniques are available to ensure the storage of medical data:

A. *File access control*

Under operating systems, the administrator defines access restrictions (read, write and execute) to file owner, the stuff members, and public users.

B. *Data access control*

Medical databases are stored in local servers and can be consulted remotely for tele-diagnostic for example. Access or denial to medical data should be adequately granted.

C. *File encryption and signature*

To reduce considerably the risk of disclosure, the use of crypto-system is a great solution. Encrypt medical data before transmission upon open networks, like Internet, ensures the confidentiality of patient identity. Adding digital signature ensure the data integrity also.

D. *File anonymization*

The identified patient data or de-identified patient data - information that does not identify the individual and for which there is no reasonable basis to believe the individual can be identified from it - must be kept confidential. Several software and web based applications can ensure the DICOM anonymization by deleting certain attributes like (Name, Address, Social card ID,…) .

IV. Proposed Public Key Cryptosystem

A. *Overview*

Formally, PKE = three efficient (probabilistic) algorithms:

KeyGen( ):

Outputs: public key pk and secret key sk

Enc(pk , m):

Outputs: a ciphertext c

Dec(sk , c):

Outputs: a message m

And always, we assume that the communication is exchanged in insecure channel, and then always, we assume that there are pirates (adversaries).

The scheme is called semantically secure if the probability $probability("A"wins) - \frac{1}{2}$ is negligible for every efficient adversary $"A"$.

Our proposed scheme is based on third order linear sequences.

In [10], P. Smith and M. J. J. Lennon proposed using Lucas sequences cryptosystems, and they proved that the computation cost by using Lucas sequences is half reduced instead of using exponentiation in the standard RSA. Moreover, from [12], the security of Lucas sequences is polynomial-time equivalent to the generalized discrete logarithm problem. In [11], Gong and L. Harn introduced cryptosystems based on third order linear sequences, and they show that the computation cost of the proposed scheme is reduced by $\frac{2}{3}$ instead of using exponentiation in the standard RSA. All these given variants have a weak point on semantic security. In this paper, a probabilistic variant is given, together with the security analysis. Moreover, as the crucial property of Lucas sequences is that cryptosystem are not formulated in terms of exponentiation, this would make them unsusceptible to various well known attacks that threaten the security of more traditional exponentiation based cryptosystems like RSA.

B. *Mathematical foundation*

Remind that for two integers a , b and a polynomial $f(X) = X^3 - aX^2 + bX - 1$, a third order linear characteristic sequence generated by (a,b) is denoted by $s(a,b)$ and defined by the following recurrence:

$$s_{k+3}(a,b) = as_{k+2}(a,b) - bs_{k+1}(a,b) + s_k(a,b) \qquad (1)$$

(a,b) is called the generator of $s(a,b)$ and k is its exponent.

It is well known that if α_1, α_2 and α_3 are the complex roots of $f(X)$. Then there exist three rational numbers (a_1, a_2, a_3) such that for every integer k .

$$s_k(\ \mathrm{a,b}) = a_1\alpha_1{}^k + a_2\alpha_2{}^k + a_3\alpha_3{}^k \quad (2)$$

Note that the tuple (a_1, a_2, a_3) depends only on the choice of $s_0(a,b)$, $s_1(a,b)$, $s_{-1}(a,b)$ and conversely. If $s_0(a,b), s_1(a,b)$ and $s_{-1}(a,b)$ are integers, then a_1, a_2, a_3 are integers too.

Through the paper, let p is an odd prime integer, $s(a,b)$ is a third order linear sequence such that $s_0(a,b), s_1(a,b)$, $s_{-1}(a,b)$ are integers and $a_1 \equiv a_2 \equiv a_3 \equiv 1 (\mathrm{mod}\ p)$.

Then $s_{-1}(a,b) \equiv b (\mathrm{mod}\ p)$, $s_0(a,b) \equiv 3 (\mathrm{mod}\ p)$, $s_1(a,b) \equiv a (\mathrm{mod}\ p)$ and for every k , $s_k(a,b)$ is an integer. Since $\overline{f(X)}$ is irreducible in $F_p[X]$, then $f(X)$ is irreducible in $Q[X]$ too. In that case, let $K = Q[\alpha_1]$, Z_K its ring of integers, $N_{K/Q}$ and $T_{K/Q}$ the norm and trace of K . Then for every integer k , $s_k(a,b) \equiv T_{K/Q}(\alpha_1{}^k) \mathrm{mod}\ p$. Since $\alpha_1{}^p (\mathrm{mod}\ p)$ and

${\alpha_1}^{p^2} (\bmod p)$ are the conjugate of $\alpha_1 (\bmod p)$, we have: $N_{K/Q}(\alpha_1) \equiv {\alpha_1}^{p^2+p+1} \equiv 1(\bmod p)\cdot$

Thus, $T = p^2 + p + 1$ is a period of $s(a,b)$ modulo p.

The following cryptographic properties are well known modulo p (see [GH 99]). We give theme modulo p^2 without proof.

To simplify, for every integer k, let us denote $s_k := s_k(a,b)(\bmod p^2)\cdot$

If $s_0(a,b) \equiv 3(\bmod p^2)$, $s_1(a,b) \equiv a(\bmod p^2)$ and

$s_{-1}(a,b) \equiv b(\bmod p^2)$. Then for every integer k,

if $f_k(X) \equiv X^3 - s_k X^2 + s_{-k} X - 1(\bmod p^2)$,

then $f_k(X) \equiv (X - {\alpha_1}^k)(X - {\alpha_2}^k)(X - {\alpha_3}^k)(\bmod p^2)\cdot$

In particular, for every integers k and e,

$$s_e(s_k(a,b), s_{-k}(a,b)) \equiv (s_{ek}, s_{-ek})(\bmod p^2). \qquad (3)$$

C. Infrastructure

The algorithms of the proposed public key schemes are based on the result given in Proposition.1, given in this paragraph.

Let $n = pq$ be an RSA, (a,b) two integers such that $f(X) = X^3 - aX^2 + bX - 1$ is irreducible modulo p (resp. modulo q), $s(a,b)$ the third order linear sequence modulo n^2 generated by (a,b) such that $s_{-1}(a,b) \equiv b(\bmod n^2)$, $s_0(a,b) \equiv 3(\bmod n^2)$,

$s_1(a,b) \equiv a(\bmod n^2)\cdot$

Let

$$T = LCM(p^2 + p + 1, q^2 + q + 1) \quad (4)$$

be the least common multiple of $p^2 + p + 1$ and $q^2 + q + 1\cdot$

Let $\Gamma := \{(x,y) \in Z^2, s_T(x,y) \equiv 3(\bmod n)\}$

and $L : \Gamma^2 \to \dfrac{Z}{nZ}$ defined by

$$\mathrm{L}(x,y) = \frac{s_T(x,y) - 3}{n}(\bmod n)\cdot$$

Since for every $(x,y) \in \Gamma$, $s_T(x,y) \equiv 3(\bmod n)$, L is well defined and we have the following proposition:

Proposition.1

If $L(a,b)$ is invertible modulo n, then for every integer k,

$$\frac{L(s_k(a,b), s_{-k}(a,b))}{L(a,b)} \equiv k(\bmod n) \qquad (5)$$

Proof.

For every $i := 1,2,3$,

let $\Gamma^i := \{x \in Z[\alpha_i],\ x \equiv 1(\bmod\ n\, Z[\alpha_i])\}$

and $L_i : \Gamma^i \to \dfrac{Z[\alpha_i]}{nZ[\alpha_i]}$ defined by $\mathrm{L}_i(x) = \dfrac{x-1}{n}(\bmod n)\cdot$

Then for every $(x,y) \in \Gamma^{i^2}$,

$$\mathrm{L}_i(xy) \equiv \frac{xy-1}{n} \equiv \frac{x(y-1) + x - 1}{n} \equiv x\frac{(y-1)}{n} + \frac{x-1}{n}(\bmod n)\cdot$$

Since $x \in \Gamma^i$, we have $\mathrm{L}_i(xy) \equiv \dfrac{(y-1)}{n} + \dfrac{x-1}{n}(\bmod n)$

and then $\mathrm{L}_i(xy) = \mathrm{L}_i(x) + \mathrm{L}_i(y)$.

Let $\mathrm{T} = k_p(p^2 + p + 1)$. Since for every $i := 1,2,3$,

${\alpha_i}^T \equiv ({\alpha_i}^{p^2+p+1})^{k_q} \equiv (N_{K/Q}(\alpha_i))^{k_q} \equiv 1(\bmod p)$ (resp. ${\alpha_i}^T \equiv 1(\bmod q)$), it follows that ${\alpha_i}^T \equiv 1(\bmod n)$. Thus for every $i := 1,2,3$, ${\alpha_i}^T \in \Gamma^i$ and for every integer k, $L_1(k{\alpha_i}^T) \equiv kL_1({\alpha_i}^T)(\bmod n)$. So, $s_{kT}(a,b) - 3 \equiv ({\alpha_1}^{kT} - 1) + ({\alpha_2}^{kT} - 1) + ({\alpha_3}^{kT} - 1)(\bmod n)$ and

$$L(\mathrm{s_k}(a,b), \mathrm{s_{-k}}(a,b)) \equiv \frac{\mathrm{s_k}(a,b) - 3}{n} \equiv \sum_{i=1}^{3} \frac{{\alpha_i}^{kT} - 1}{n} \equiv k\sum_{i=1}^{3} L^i({\alpha_i}^T)(\bmod n)$$

Finally, $\mathrm{L}(\mathrm{s_k}(a,b), \mathrm{s_{-k}}(a,b)) \equiv kL(a,b)(\bmod n)$, and if $L(a,b)(\bmod n)$ is invertible,

then $\dfrac{\mathrm{L}(\mathrm{s_k}(a,b), \mathrm{s_{-k}}(a,b))}{L(a,b)} \equiv k(\bmod n)\cdot$

The solely problem that remains to establish our proposed scheme, is to describe a method How to choose a couple (a,b) such that $L(a,b)(\bmod n)$ is invertible: if $L(a,b)(\bmod\ n)$ is not invertible, we describe a method which allows us to choose a couple (a,b) of integers that $L(a,b)(\bmod\ n)$ is invertible.

If $\dfrac{s_T(a,b) - 3}{n}(\bmod n)$, then we will keep $s_{-1}(a,b) \equiv b(\bmod n^2)$, $s_0(a,b) \equiv 3(\bmod n^2)$ and $s_1(a,b) \equiv a(\bmod n^2)$. Else, i.e., if p or q divides $\dfrac{s_T(a,b) - 3}{n}$, then let $E := \{\mathrm{n, q, p}\}$ and $y = \dfrac{n}{x} \in E$ the largest element of E dividing $\dfrac{s_T(a,b) - 3}{n}$. (If $x \neq 1$, then x does not divide $\dfrac{s_T(a,b) - 3}{n}$). Let $A := a + nx$, and consider $s(A,b)$ the characteristic sequence generated by (A,b) modulo n^2:

$s_{-1}(A,b) \equiv b(\bmod n^2)$, $s_0(A,b) \equiv 3(\bmod n^2)$

and $s_1(A,b) \equiv A(\bmod n^2)$.

Let β_1, β_2 and β_3 be the complex roots of $f(X) = X^3 - AX^2 + bX - 1$. Since $\overline{f(X)} \equiv \overline{g(X)}(\bmod p)$ (resp. $\overline{f(X)} \equiv \overline{g(X)}(\bmod q)$), then up to a permutation for every $i := 1,2,3$, there exists an integral complex t_i such that $\beta_i = \alpha_i + nt_i$. Thus, for every integer k,

$$s_k(A,b) = \sum_{i=1}^{3} \beta_i^{\,k} = \sum_{i=1}^{3} (\alpha_i + nt_i)^k = s_k(a,b) + nk(t_1 + t_2 + t_3) + n^2 u_k \qquad ,$$

where u_k is an integral complex. For $k = 1$, we have $t_1 + t_2 + t_3 \equiv x(\bmod n)$.

Since $\frac{s_T(A,b) - 3}{n} \equiv \frac{s_T(a,b) - 3}{n} + x(\bmod n)$, $x \in \{1, \mathrm{p}, \mathrm{q}\}$ and if $x \neq 1$, then x does not divide $\frac{s_T(a,b) - 3}{n}$,

then $\frac{s_T(A,b) - 3}{n}$ is invertible.

Finally, without loss of generality, up to replace a by $a + nx$, we can assume that $\frac{s_T(a,b) - 3}{n}(\bmod n)$ is invertible.

Now we are ready to establish our proposed schemes.

D. The deterministic version

Algorithm of encryption and decryption:

1- Public parameters: $pk = (n, a, b)$.

2- Private parameters: $sk = (p, q)$.

3- Encryption: For a message $0 \le m \le n-1$, Bob calculates the block ciphertext (c_1, c_2) such that $c_1 = s_m(a,b)(\bmod n^2)$ and $c_2 = s_{-m}(a,b)(\bmod n^2)$.

4- Decryption: For a given block ciphertext(c_1, c_2), Alice can decrypt, by calculating $\frac{L(c_1, c_2)}{L(a,b)}(\bmod n)$.

Indeed, since (c_1, c_2) is a ciphertext, let $0 \le m \le n-1$ such

$c_1 = s_m(a,b)(\bmod n^2)$ and $c_2 = s_{-m}(a,b)(\bmod n^2)$. Therefore by using (3), we have

$$s_T(c_1, c_2) \equiv s_T(s_m(a,b), s_{-m}(a,b)) \equiv (s_{Tm}(a,b), s_{-Tm}(a,b)) \equiv (\bmod n^2) \qquad .$$

Moreover as $L(a,b)$ is invertible, using (5) of proposition.1, we have this equality:

$$\frac{L(c_1, c_2)}{L(a,b)} \equiv \frac{L(s_m(a,b), s_{-m}(a,b))}{L(a,b)} \equiv m(\bmod n).$$

E. The probabilistic version

Algorithm of encryption and decryption:

1- Public parameters: $pk = (n, a, b)$.

2- Private parameters: $sk = (p, q)$.

3- Encryption: For a message $0 \le m \le n-1$, Bob chooses a random integer r and calculates the block ciphertext (c_1, c_2) such that

$$c_1 = s_{m+rn}(a,b)(\bmod n^2) \text{ and}$$

$$c_2 = s_{-(m+rn)}(a,b)(\bmod n^2).$$

4- Decryption: For a given block ciphertext (c_1, c_2), Alice can decrypt, by calculating $\frac{L(c_1, c_2)}{L(a,b)}(\bmod n)$.

First, for the same plaintext 0, if $r \neq s$, then for $c_r = s_{rn}(a,b)(\bmod\ n^2)$, $c_s = s_{sn}(a,b)(\bmod n^2)$,

we have $c_r \neq c_s (\bmod n^2)$. Thus, as r is randomly chosen, then this scheme is probabilistic.

On the other hand, as in the proof of the last scheme, let $0 \le m \le n-1$ such $c_1 = s_{m+rn}(a,b)(\bmod n^2)$

and $c_2 = s_{-(m+rn)}(a,b)(\bmod n^2)$. By using (5) of Proposition.1, we have this equality :

$$\frac{L(c_1, c_2)}{L(a,b)} \equiv m + rn \equiv m(\bmod\ n).$$

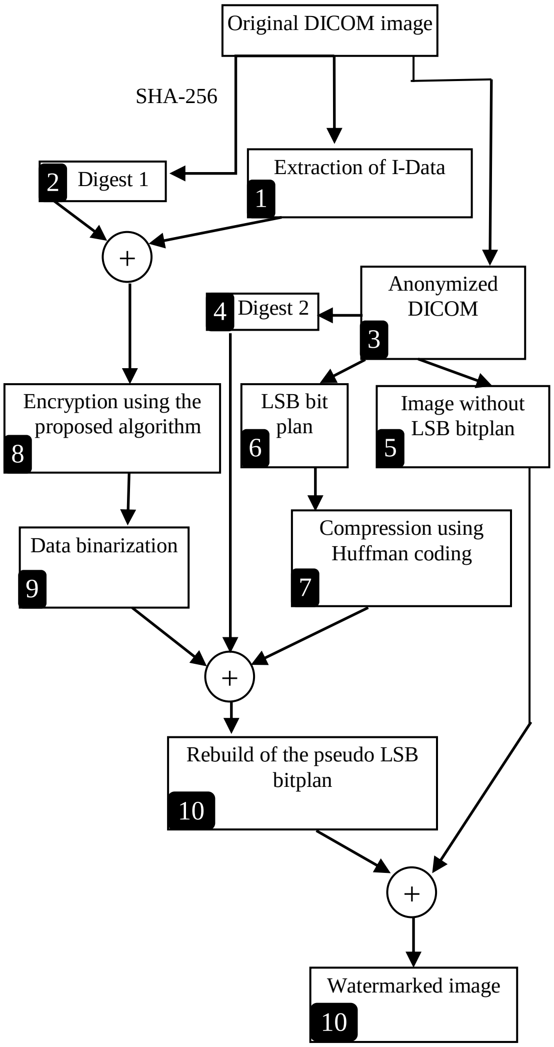
V. New DICOM Reversible Anonymization Mechanism

In order the keep the possibility to return to I-Data by using the secret key, we proposed to hide these data inside the image by watermarking it. To respect data integrity, the LSB bitplan of the image is compressed using Hoffmann coding and the liberated zone is used to embed I-Data. (See [9] for explanations about LSB bitplan compression). The I-Data and the digest1, computed from the original DICOM image using the well-known SHA-256 hash function, are encrypted using the proposed cryptosystem (see paragraph IV). The encrypted data is converted to binary form and inserted in the liberated zone of the image LSB bitplan.

The proposed method ensures that the anonymized DICOM images treated by our method are authentic, and when the return to the patient's identity is paramount, an authorized user (who has the secret key) can reveal the patient identified data to have more information about the patient and explain certain cases.

The algorithm is schematized below:

Emission side:



1– Extraction of identified patient data (I-Data) from the original DICOM image.

2– Computation of the original image digest, using SHA-256 hash function (digest 1).

3- Anonymization of the original DICOM.

4– Computation of the anonymized image digest, to be used by researchers to ensure the originality of the image (digest 2).

5– Extraction of the image without LSB bitplan.

6– Extraction of the LSB bitplan.

7– Compression of the LSB bitplan using Hofmann coding algorithm (used in [9]).

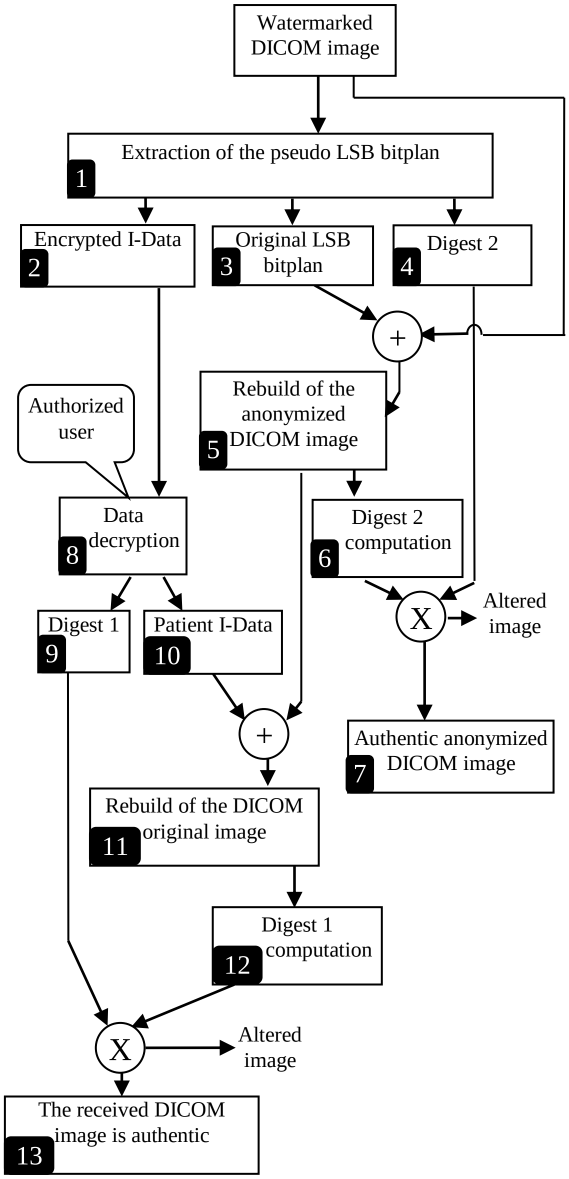
8– Encryption of I-Data and Digest 1 using the proposed crypto-system (see the algorithm in paragraph (IV - E).

9– Conversion of the encrypted data to binary format.

10– Rebuild of the pseudo LSB bitplan composed by: original LSB and binarized data (I-Data and digest 1)

11– Rebuild of the watermarked image to be used by researchers. The new image contains hidden patient I-Data patient.

Reception side:



Extraction of the LSB bitplan from the watermarked image.

2- Extraction of the encrypted data.

3– Extraction of the compressed LSB bit plan, to be decompressed using Hofmann decoding algorithm.

4- Extraction of Digest 1.

5- Rebuild of the anonymized image using the extracted LSB bitplan.

6- Computation of the digest from the rebuilt image. (digest 2).

7- Verification of the authenticity of the anonymized image by comparing the saved and the computed digests.

8- Authorized user, having secret key, can decrypt Data using the proposed decryption algorithm (see paragraph IV-E).

9– Extraction of the saved digest (digest 1).

10 – Extraction of the patient I-Data.

11 –Rebuild of the original DICOM image by combining anonymized image and I-Data.

12 – Computation of Digest 2.

13- Verification of the DICOM image originality, by comparing the saved and the computed digests.

VI. Conclusion

In this paper, we proposed a mechanism to perform a reversible anonymization to DICOM images. The identity patient data and image digest are encrypted using a new public key crypto-system and then watermarked in liberated zone of the image LSB bitplan obtained by compressing the original LSB. The proposed algorithm efficiently ensures the data confidentiality (encryption and watermarking), the reversibility (original data may be re-obtained), the authenticity (only the authorized user can access to identified patient data) and the timeliness by using a new scheme of public key crypto-system.

References

[1] S. Hidenobu, N. Mami, S. Shinsuke, K. Mitsuru, K. Yoshiki, N. Noboru, N. Hiromu. Anonymization System to Protect the Personal Data in Secondary Use of DICOM Images. Institute of Electronics, Information and Communication Engineers, Vol.105, NO.579, pp. 71-74 (2006).

[2] M. Onken, J. Riesmeier, M. Engel, A. Yabanci, B. Zabel, S. Després. Reversible Anonymization of DICOM Images Using Automatically Generated Policies. Medical Informatics in a United and Healthy Europe. pp 861-865 (2009).

[3] J. M. Zain and M. Clarke. Reversible Region of Non-Interest (RONI) Watermarking for Authentication of DICOM Images. International Journal of Computer Science and Network Security, Vol.7, No.9, pp 19-28 (2007).

[4] Peyton H. Bland, Gary E. Laderach, and Charles R. Meyer. A Web-based Interface for Communication of Data between the Clinical and Research Environments without Revealing Identifying Information. Academic Radiology Journal, Vol 14, Issue 6, pp 757-764 (2007).

[5] A P Toms B Kasmai, S Williams, and P Wilson. Building an anonymized catalogued radiology museum in PACS: a feasibility study. British Journal of Radiology 79, pp 666-671 (2006).

[6] M. G. Ruiz, A. G. Chaves, C. R. Ibañez, J. M. Gutierrez Mazo, et al.. . mantisGRID: A Grid Platform for DICOM Medical Images Management in Colombia and Latin America. Springer - Journal of Digital Imaging. Vol 24, Number 2, pp 271-283 (2010).

[7] D. Abouakil, J. Heurix, and T. Neubauer. Data Models for the Pseudonymization of DICOM Data. The 44th Hawaii International Conference on System Sciences. pp 1-11 (2011).

[8] P. H. Blanda, G. E. Laderach, and C. R. Meyer. Implementation and use of a web-based interface for confidential communication of data between the clinical and research environments. Proceedings of the Society of Photo-Optical Instrumentation Engineers. PMC. pp. 1-16 (2009).

[9] Y. Zaz and L. Elfadil. Enhanced EPR Data Protection using Cryptography and Digital Watermarking. The 2nd International Conference on Multimedia Computing and Systems (IEEE Xplorer). Ouarzazate, Morocco (2011).

[10] P. Smith and M. J. J. Lennon, LUC : A new public key system. Ninth IFIP Int. Symp. on Computer Security, pp. 103-117 (1993).

[11] G. Gong and L. Harn. Public-Key Cryptosystems Based on Cubic Finite Field Extensions. In IEEE Trans. Inform. Theory, vol. 45, pp. 2601-2605 (1999).

[12] Chi-Sung Laih, Fu-Kuan Tu, and Wen-Chun Tai. On the security of the Lucas function, Information Processing Letters 53, pp 243-247 (1995).

[13] Official DICOM web site http://medical.nema.org/

Authors Profile

Youssef Zaz is an assistant professor in University of Abdelmalek Essaadi (Morocco). He currently teaches and conducts research in computer related and image processing. He is also interested in the e-commerce and GNSS. Before joining Abdelmalek Essaadi in Sptember 2011, he was in Ibn Zohr University (Morocco) from 2006 to 2011. He worked as IT project Manager with National Department of Post Telecommunication and IT from 2000 to 2006. His main scientific interests lay in the fields of image watermarking and pattern recognition. Dr Youssef has authored numerous scientific publications and communications in international conferences. He is a chairman of the IEEE conference (International Conference on Multimedia Computing and Systems). He is also a reviewer for several international conferences.

Lhoussain El Fadil is currently assistant professor of mathematics in the Polydisciplinary Faculty of Ouarzazat (FPO), Ibn Zohr University (Morocco). In 2004, he was awarded his PhD from Department of Mathematics and Computer Sciences in FST of Fez (Morocco). He was a post doc for one year in CRM of Barceleno (2008) and a fellow for one year in NTNU (Norway) (2009). He has authored several research papers in algebra, number theory and cryptography.

Secure Digital Cashless Transactions with Sequence Diagrams and Spatial Circuits to Enhance the Information Assurance and Security Education

Dr. Yousif AL-Bastaki
ACMSIG Teaching & learning
University of Bahrain
P.O. Box -32038
Kingdom of Bahrain

Dr. Ajantha Herath
ACMSIG Teaching & learning
University of Bahrain
P.O. Box -32038
Kingdom of Bahrain

Abstract— Often students have difficulties mastering cryptographic algorithms. For some time we have been developing with methods for introducing important security concepts for both undergraduate and graduate students in Information Systems, Computer Science and Engineering students. To achieve this goal, Sequence diagrams and spatial circuit derivation from equations are introduced to students. Sequence diagrams represent progression of events with time. They learn system security concepts more effectively if they know how to transform equations and high level programming language constructs into spatial circuits or special purpose hardware. This paper describes an active learning module developed to help students understand secure protocols, algorithms and modeling web applications to prevent attacks and both software and hardware implementations related to encryption. These course materials can also be used in computer organization and architecture classes to help students understand and develop special purpose circuitry for cryptographic algorithms.

Keywords-e-cashless; transactions; cryptographic; algorithms; Sequence diagrams, Spatial circuits.

I. INTRODUCTION

During the last decade Postal mail became E-mail, face-to-face Banking became Online Banking and Commerce transformed to E-Commerce. An electronic transaction is an agreement made using internet between a buyer and a seller. The user immediately becomes vulnerable to attacks or infiltration as soon as a computer starts to share the resources available on the web or local network. Confidentiality guarantees privacy, no loss of information from client or the server. Integrity assures no modifications of data, messages or impersonation. Authentication helps identify the user. The validation is provided by an authentication factor which is used to validate or authenticate the communicating person's identity. Confidentiality, Integrity and Authentication is achieved through encryption of the message. Authentication is implemented through encryption, signatures and certificates [1]. The Kerberos authentication service restricts access to authorized users all the time with single sign-on. It is secure and scalable to support a large number of clients and servers. Kerberos ticket generation resembles social systems such as an airline system where a user purchases a ticket to receive the service. Figure 1 illustrates online airline ticket purchase. Symmetric key cryptography consists of a private key that is used for both encryption and decryption. Faster symmetric key encryption algorithms like Advanced Encryption Standard, AES, are popular for larger data encryption.

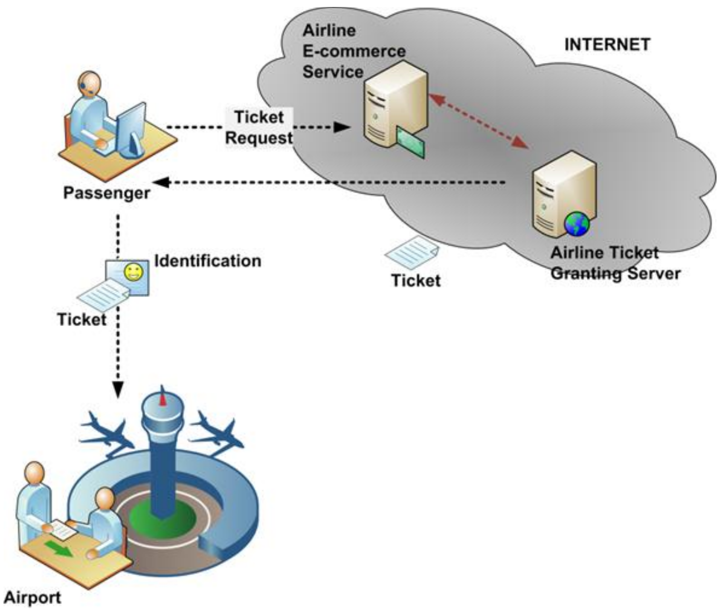


Figure 1: Airline Ticket Processing

The Kerberos authentication scheme consists of a client, Kerberos Authentication Server, Ticket Granting Service and a service provider. Kerberos communications are represented using a sequence diagram as shown in Figure 2. Encrypted keys and tickets help sharing symmetric keys.

Non-repudiation makes sure of the security of the E-Commerce transaction. It ensures participants online actions undeniable and no back out of their transaction later. Hence, the seller cannot change the agreed price or delivery time frame and the customer cannot change his/her mind of the product by considering the low price of other vendors after confirming the transaction. The digital certificate, encryption and signature methods are useful in this matter because each party can validate the sender of the message and information.

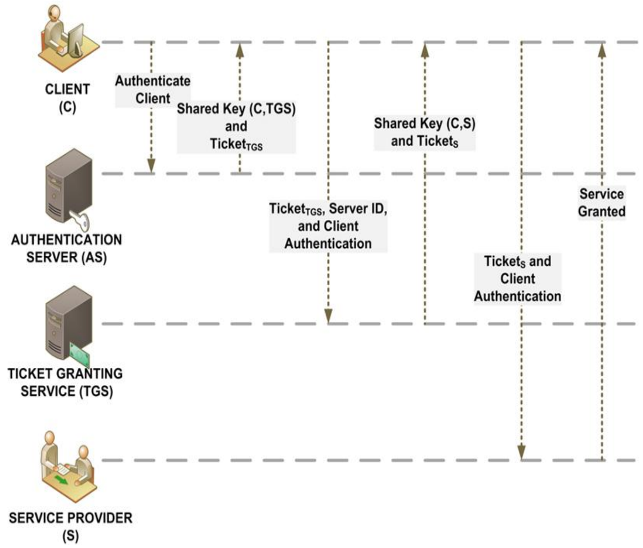

CLIENT (C)
Authenticate Client
Shared Key (C,TGS) and Ticket_TGS
Shared Key (C,S) and Ticket_S
AUTHENTICATION SERVER (AS)
Ticket_TGS, Server ID, and Client Authentication
Service Granted
Ticket_S and Client Authentication
TICKET GRANTING SERVICE (TGS)
SERVICE PROVIDER (S)

Figure 2: Kerberos Authentication

Availability ensures that the system responses promptly and the service and information is available when needed to authorize persons. Flooding machine with requests or filling up memory threatens the Availability. Denial of service is the consequence of such an attack. Availability of the service can be improved by providing fast, reliable and efficient service. The network security is the key feature of ensuring availability of the service. Therefore, deploying network security devices such as firewalls and configuring them along with associated protocols properly is the key to ensuring service availability.

Asymmetric key cryptography consists of a pair of public and private keys. The private key is kept secret whereas the public key is distributed for use by multiple parties.

Digital Signatures are used to provide authenticity. A message signed with merchant's private key can be verified by any consumer who has access to merchant's public key. David Chaum proposed the blind signature scheme based on RSA digital signature and its application for online electronic cash system. Thereafter, Okamoto developed the first practical divisible electronic cash system. A. Chan and Frankel further improved the divisible electronic cash system. This verifies that the signed message has not been tampered with by any unauthorized party.

A public key certificate contains the identity of the certificate holder such as name, public key and the digital signature of the certificate issuing authority. Public key certificate is used to validate the sender's identity. The certification authority attests that the public key indeed belongs to the sender.

Section 2 of this paper provides a brief description of an e-commerce transaction, derivation of a sequence diagram from the transaction that could be used in software system implementations and major threats that might be seen in an e-commerce transaction. Also, it discusses five major security concepts that can be used to avoid those threats. Section 3 presents some details of integration of confidentiality, integrity and authentication to the transaction. Section 4 describes the transformation of security equations in secure electronic transactions [2-4] to spatial circuits that could be used in hardware implementations. It also illustrates the working of dual signature. Section 5 describes other related work in electronic transactions.

II. E-COMMERCE TRANSACTIONS

Major players of electronic cashless transactions are clients, internet service providers, merchant's servers, client's and merchant's banks, warehouses and deliver services. In a transaction diagram major players are represented by nodes and directed arcs present messages transferred. Purchase of goods from the internet can be represented using a sequence diagram as shown in Figure 3.

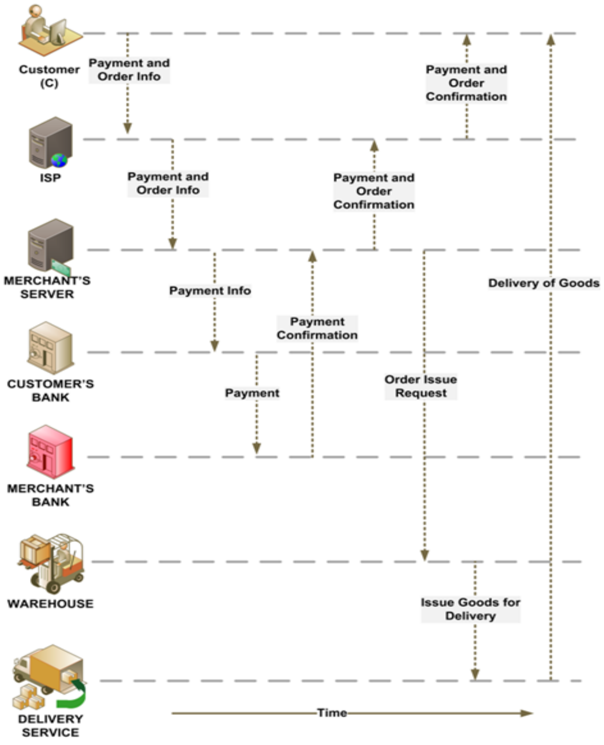


Figure 3 Sequence Diagram for an E-Commerce Transaction

The sequence diagram illustrates the snapshot of the sequence of events taking place represented in the vertical axis progressing from top to bottom and the particular time slot of the event taking place is shown in the horizontal axis from left to right. The Client first sends payment and order information to merchant's server via internet service provider. Then the Merchant's server sends payment information to client's bank. The client's bank then sends the payment to merchant's bank. Payment confirmation will be issued by the merchant's bank to the Merchant's server. Thereafter the payment and order confirmation will be sent to the client by the merchant's server via ISP. The Merchant's server sends the order issue request to the warehouse. Warehouse issues goods for delivery. The delivery service delivers the goods to the client.

Figure 3 also depicts the insecure e-commerce transaction. In this transaction any one can read or modify the payment and order information. An intruder can interrupt, modify or initiate the transaction. Client's bank information can be stolen by a third party. Particularly, E-commerce transactions involve with client's and merchant's secure information such as credit/debit card numbers and private information. Most of the communications among the client, merchant and banks are done through the internet. Much of the message passing, billing and payments are done by electronic message transfers. There is a higher possibility of stealing, loosing, modifying, fabricating or repudiating information. Such systems and messages transmitted need extra protection from the eavesdroppers.

Many threats such as Denial of Service, DoS, Distributed Denial of Service, DDoS, Trojans, phishing, Bot networks, data theft, identity theft, credit card fraud, and spyware can be seen in these systems. These attacks might cause the loss of private information or revelation of sensitive information such as credit card numbers and social security numbers, misinterpretation of users, gaining unauthorized access to sensitive data, altering or replacing of data. Sniffing can take place at vulnerable points such as ISP, Merchant's server, client's bank, merchant's bank or at the internet back bone.

III. Consolidation of Integrity, Confidentiality and Authenticity in Applications

Providing confidentiality is vital in e-commerce. Figure 4 shows the transaction with confidentiality. The transaction can be made secure by converting the plain text message to cipher text so that the holders of the keys can decrypt and read the messages. Common algorithms used to achieve this encryption and decryption goal are AES, DES with single symmetric keys and RSA with public/private asymmetric key pairs.

Encryption will prevent strange third party to have client's credit/ debit card numbers, passwords, pin numbers or personal details. But in the internet world there are many possibilities that an unauthorized third party can obtain this sensitive and private information and violate the privacy of the people, particularly in e-commerce service, the privacy of the consumer and the merchant. Thus, this e-commerce system needs to be assured that the information is not to be spread to the unauthorized people in order to provide a genuine and reliable service. The symmetric encryption plays a key role in assuring confidentiality of the data because even though an unauthorized third party intercepts the message, usage of the unique session key, which can be accessed only by the two parties involved, prevents that person from viewing the message. Hence, the encryption of the information is not only guaranteed by the authentication of the information but also it assures confidentiality of the information.

To make the transaction secure the data need to be received free from modification, destruction and Repetition. When we consider the security of the electronic transaction, data integrity is another significant feature, because changing address, order information, or payment information may have possibly happened in this system. Therefore, to get the message free from modifications the e-commerce system should provide protection to the message during transmission. This can be achieved by using encryption and message digesting.

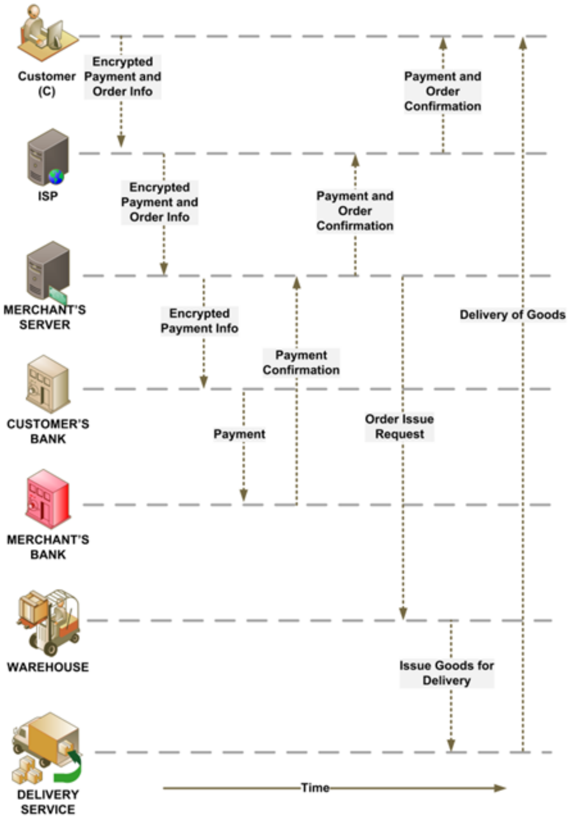


Figure 4 E-commerce Transaction with Confidentiality

A unique message digest can be used to verify the integrity of the message. Hash functions take in a variable length input data and produce a fixed length unique outputs that are considered as the fingerprint of an input data/message. Thus, it is very likely that if two hashes are equal, the messages are the same. Hash functions are often used to verify the integrity of a message.

The sender computes hash of the message, and concatenates the hash and the message, and sends it to the receiver. The receiver separates the hash from the message and then generates the hash of the message using the same hash function used by the sender. The integrity of the message is said to be preserved if the hash generated by sender is equal to the Hash generated by the receiver. This implies that the message has not been altered or fabricated during the transmission from sender to receiver.

Encryption algorithms such as AES, DES could be used to generate message digests. In addition there are special purpose hash functions such as SHA-3 [5] for this purpose. SHA-3 is the message-digest algorithm developed by the National Institute of Standards and Technology and the National Security Agency. SHA-3 will be selected from five new Hash functions, BLAKE, Groestel, Skein, JH and Keccek. Grooestel is similar to AES. SHA-1 is secure but slower than MD5. MD5 produces the digest of 128 bits whereas SHA-1 produces a 160-bit message digest and is resistant to brute force attacks. It is widely used for digital signature generation.

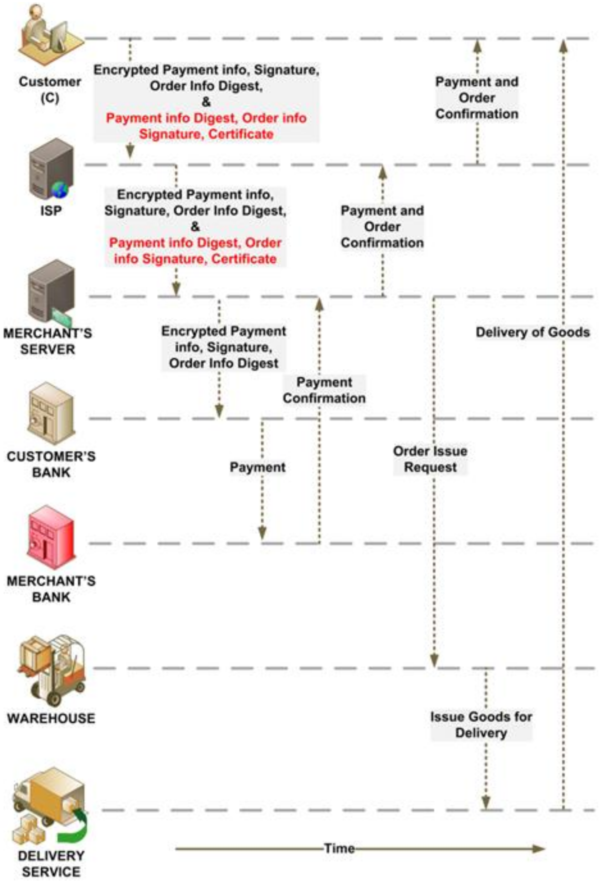

Customer (C)
ISP
MERCHANT'S SERVER
CUSTOMER'S BANK
MERCHANT'S BANK
WAREHOUSE
DELIVERY SERVICE
Encrypted Payment info, Signature, Order Info Digest, & Payment info Digest, Order info Signature, Certificate
Payment and Order Confirmation
Encrypted Payment info, Signature, Order Info Digest, & Payment info Digest, Order info Signature, Certificate
Payment and Order Confirmation
Encrypted Payment info, Signature, Order Info Digest
Payment Confirmation
Delivery of Goods
Payment
Order Issue Request
Issue Goods for Delivery
Time

Figure 5 Secure Transaction

The Figure 5 shows how the authenticity, confidentiality and integrity can be used in our example. It uses the encryption, message digest, digital signature and digital certificate to ensure the authenticity, confidentiality and integrity of the order and payment information. Fig. 6 represents the transaction with symbols.

One of the most important aspects of the security of the transaction is authenticating that the suppliers and consumers are who they say they are and assure the trustworthiness of the sources they are exchanging. This is really important in cashless e-commerce transactions because of the supplier and consumer never meet face to face. Authentication can be presented in different ways. Exchanging digital certificates helps seller and buyer verify each other's identity so that each party knows who is at the other end of the transaction. The digital signature is another method to be certain that the data is indeed from a trusted party. In addition, symmetric encryption can also be used in certifying the authenticity. In this way, the receiver of the information can make sure that the information that they have received is sent by a trusted party, because the key that is used to encrypt and decrypt the information is shared only by the sender and the receiver.

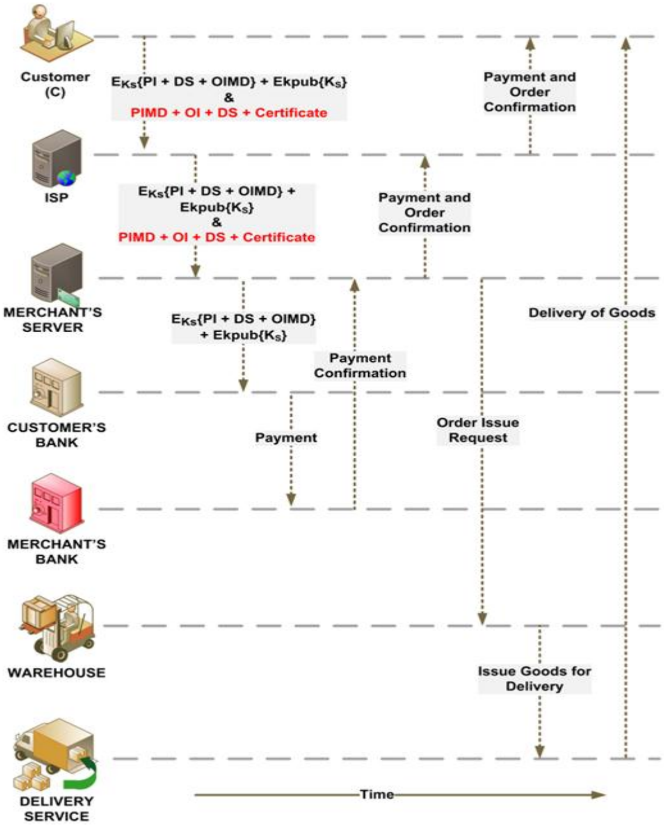

Customer (C)
ISP
MERCHANT'S SERVER
CUSTOMER'S BANK
MERCHANT'S BANK
WAREHOUSE
DELIVERY SERVICE
E_Ks{PI + DS + OIMD} + Ekpub{K_S} & PIMD + OI + DS + Certificate
Payment and Order Confirmation
E_Ks{PI + DS + OIMD} + Ekpub{K_S} & PIMD + OI + DS + Certificate
Payment and Order Confirmation
E_Ks{PI + DS + OIMD} + Ekpub{K_S}
Delivery of Goods
Payment Confirmation
Payment
Order Issue Request
Issue Goods for Delivery
Time

Figure 6 Secure Transactions with Symbols

In Figures 5, 6 and 7
PI = Payment Information
DS = Dual Signature
OIMD = OI message digest
Ks = Temporary symmetric key
PIMD = PI message digest
OI = Order Information
Certificate = Cardholder Certificate.

IV. Symbolic Representations and Algorithms to Spatial Circuits Transformation

The equation in Figure 5

$$\mathbf{E_{ks}\ \{PI + DS + OIMD\} + \ E_{k\ pub\ B\{}\ \{K_s\}\ \ \&\ PIMD + OI + DS + Certificate}$$

summarizes the message generation in Secure Electronic Transaction protocol, an application of hashing and encryption algorithms in providing integrity, confidentiality and authentication for messages.

This message consists of two parts: one for the client's bank and the other for the merchant. The request message part **{PI + DS + OIMD}** is encrypted by using the session key $\mathbf{K_s}$**.** The Digital Envelope consists of the session key encrypted by using the public key of the Bank $\mathbf{K_{pubB}}$**.** Secure transactions use both public and private key encryption methods for message exchange between the merchant and the consumers. The DES – Data Encryption Standard algorithm is used by most financial institutions to encrypt Personal Identification

Numbers. Light-weight-crypto algorithms such as Simplified-DES take an 8-bit block of plaintext and a 10-bit key as input to produce an 8-bit block of ciphertext. A spatial circuit can be easily drawn from this representation as shown in the Figure 6:

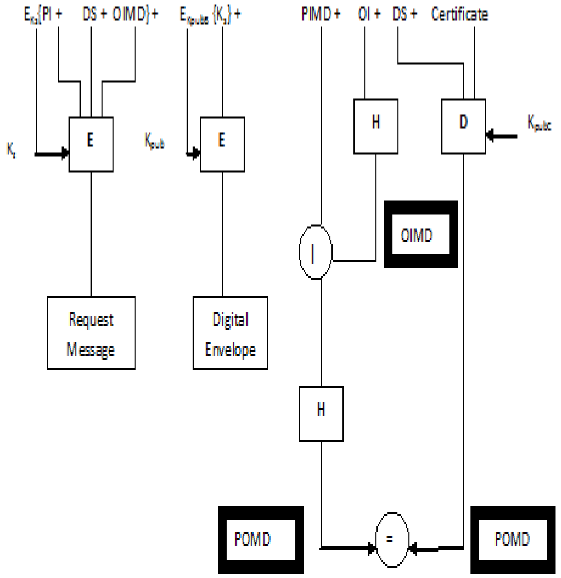


Figure 7 Cardholder Sends Purchase Request / Merchant Verifies

In Figure 7
POMD = Payment order message digest
D = Decryption (RSA)
H = Hash function
E = Encryption (RSA for asymmetric and DES for symmetric)
KpubB = Bank's public key-exchange key
KpubC = Customer's public signature key.

The goal of dual signature generation and use is to send a message that is intended for two different recipients. Each recipient has access to the message, however only a part of the message can be read by each. In case of SET protocol, the customer sends the order information (OI) and payment information (PI) using dual signature. The merchant can only see the OI and the bank can only access PI. Figure 6 shows how the order information and payment information is securely delivered to the two recipients – merchant and bank using Dual Signature, DS.

In Figure 8
PI = Payment Information
OI = Order Information
PIMD = PI message digest
OIMD = OI message digest
POMD = Payment order message digest
H = Hash function (SHA-1)
|| = Concatenation
E = Encryption (RSA for asymmetric and DES for symmetric)

K_{priC} = Customer's private signature key

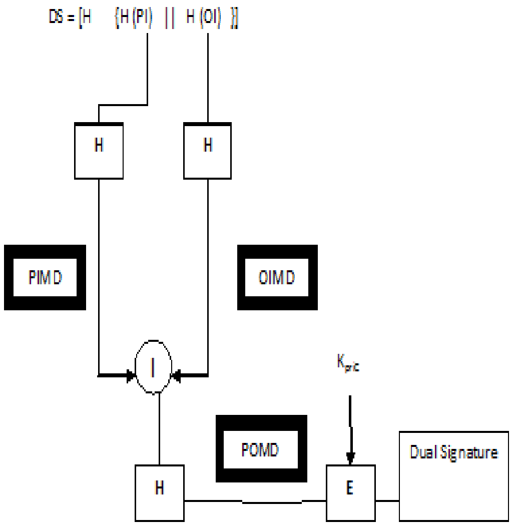


Figure 8: Construction of Dual Signatures in SET

Figure 8 illustrates the spatial circuit drawn for the dual signature generation. The digital envelope combines the speed of DES and efficient key-management of RSA. The envelope and the encrypted message is sent to the recipient who decrypts the digital envelope using his private key to generate the symmetric key and then uses this symmetric key to regenerate the original message.

Similarly, encryption and decryption algorithms can be easily transformed into spatial circuits. An algorithm to hardware transformation is an important concept to introduce in system security courses. Students learn cryptographic algorithms faster if they know how to transform equations and high level programming language constructs, such as arithmetic expressions, for loops and algorithms into spatial circuits or special purpose hardware. Figure 9 shows the *for* loop and the final round of the Blow Fish encryption algorithm

```
For i = 1 to 16 do
   REi = LEi-1 Ex-OR  Pi
   LEi = REi-1 Ex-OR    F (REi)
Final Round
LE17 = RE16  Ex-OR P16
RE17 = LE16  Ex-OR P17 .
```

V. Other Related Work

There are other electronic cashless payment protocols such as credit card, e-cash, e-check, smartcard and micropayment used over the Internet. In credit card based platforms, the consumer uses a card containing card holder's financial information issued by a bank. This credit card is used to purchase items over the Internet. E-cash is a digital form of money provided by a certified financial institution. Consumers need to install software on their machine called e-wallet. The e-wallet contains consumer's financial information that can be accessed using an ID and password. Consumers can use this account to transfer funds online and withdraw from or deposit to banks.

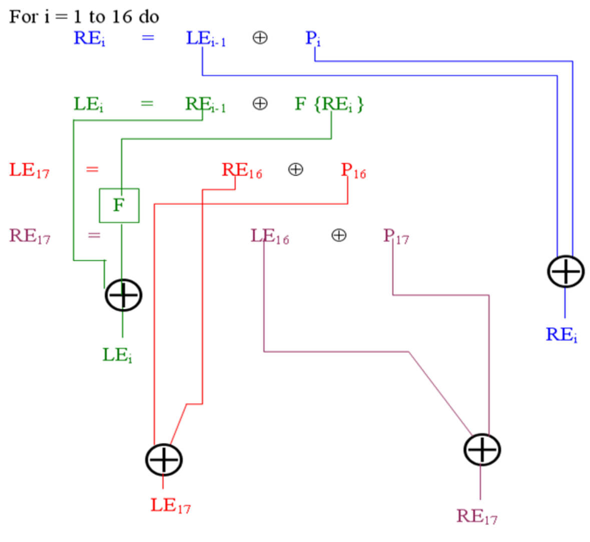

For i = 1 to 16 do
$RE_i = LE_{i-1} \oplus P_i$
$LE_i = RE_{i-1} \oplus F\{RE_i\}$
$LE_{17} = RE_{16} \oplus P_{16}$
$RE_{17} = LE_{16} \oplus P_{17}$
F
LE_i
RE_i
LE_{17}
RE_{17}

Figure 9 Algorithm to Spatial Circuit – BlowFish Encryption

PayPal is the most successful e-wallet application used in the industry today. It operates in many countries, manages millions of accounts and allows consumers to send, receive and hold funds in different currencies worldwide.

E-check is similar to e-cash except that it uses a check instead of digital money. E-check contains consumer's bank information such as account number, bank's routing number, check number, amount paid and the date of authorization. This information is used by the merchant to authenticate the consumer and the consumer's bank uses this information to authorize the payment. One advantage of e-check is that they can clear much faster than conventional check.

Micropayment systems are more practical for environments with low-cost transactions. Several platforms available in the industry today include CyberCoin, NetBill, PayWord and MicroMint. The biggest difference between micropayment and other payment systems is their operating costs. In order to make the payment system profitable, various payment approaches are used such as service prepayment, reduction of computational load, offline authorization and grouping of micropayments before financial clearance.

One common way of achieving this computational reduction is by using symmetric encryption algorithms over public key algorithms whenever possible. Using one key for both encryption and decryption will reduce the number of keys generated for the total transactions in a given day. Offline authorization can reduce computational load. This can be done by not doing any online verification with the verification center until each individual transaction is grouped offline. Another advantage is that it gives consumers partial anonymity for individual transactions.

VI. Conclusion

This paper summarized mathematical representations used in security as well as spatial circuits to represent cryptographic algorithms, providing examples related to confidentiality and integrity and their combinations. The active learning module developed can be easily adapted and effectively used in a classroom with senior undergraduate or graduate students in Computer Science, Engineering and Information Systems to teach other symmetric key algorithms and help students understand quickly. Both reading and interpreting equations are important in Computer Security classes. To survive in a highly competitive internet world the service provider need to be able to offer fast, reliable and secure service to their customers. In addition, providing trustworthiness among the merchant, the consumer, and the credit or economic institution is always required. We can assume that these e-commerce transactions are safe and trusted, but it is not easy to find out the degree of safeness and trustworthiness in the electronic world.

References & Bibliography

[1] Stallings, Williams. "Cryptography and Network Security Principles and Practice Second Edition" Prentice Hall, 2012

[2] "IBM International Technical Support." Secure Electronic Transaction: Credit Card Payment on the Web in Theory and Practice. Jun 1997. http://www.redbooks.ibm.com/redbooks/pdfs/sg244978.pdf

[3] Ganesh, Ramakrishnan. "Secure Electronic Transaction (SET) Protocol." Information Systems Control Journal, Volume 6. 2000

[4] Bella, G. Massacci, F. and Paulson, L. "Verifying the SET Registration Protocols." Proceedings of IEEE Journal of Selected Areas in Communications 2003. http://www.cl.cam.ac.uk/~lp15/papers/Bella/registration.pdf

[5] NIST's SHA-3 Contest: http://csrc.nist.gov/groups/ST/hash/sha-3/index.html

[6] Yousif Albastaki Ajantha Herath E-Learning of Security and Information Assurance with Sequence Diagrams, ACM Gulf Region special Interest Group of research, HCCE-2012, The Joint International Conference on Human-Centered Computer Environments (HCCE) Aizu Japan

[7] A. Herath S. Herath et al, Learning Digital Cashless Applications with Consolidation of Authenticity, Confidentiality and Integrity using SequenceDiagrams, Conference on Computer Science, Engineering and Applications ICCSEA-2011) Dubai May 2011

[8] VISA Partner Network." Visa Authenticated Payment Program. Apr 2007. https://partnernetwork.visa.com/vpn/global/retrieve_document.do?documentRetrievalId=118

[9] Andam, Z. (2003). e-commerce and e-business. Technical report, e-ASEAN Task Force and UNDPAPDIP.

[10] Androutsellis-Theotokis, Stephanos. Spinellis, Diomidis. "A survey of peer-to-peer content distribution technologies" ACM Computing Surveys (CSUR.). Dec 2004.

[11] Becker, A. (2008). Mobile commerce security and payment methods. In Electronic Commerce: Concepts, Methodologies, Tools and Applications, page 295. IGI Global.

[12] Bella, G. Massacci, F. and Paulson, L. "The Verification of an Industrial Payment Protocol: The SET purchase phase." Proceedings of 9th ACM Conference on Computer and Communications Security 2002

[13] Bella, G. Massacci, F. and Paulson, L. "Verifying the SET Registration Protocols." Proceedings of IEEE Journal of Selected Areas in Communications 2003. http://www.cl.cam.ac.uk/~lp15/papers/Bella/registration.pdf

[14] Bidgoli, H. (2002). Security issues and measures: Protecting electronic commerce resources. In Electronic Commerce: Principles and Practice, pages 363–394. Academic Press.

[15] Bollin, Sherrie. "E-commerce: a market analysis and prognostication." Communication Partners International, Carmel, CA. 1998

[16] Brustoloni, Jose. "Advertising and Security for E-Commerce: Protecting electronic commerce from distributed denial-of-service attacks." Proceedings of the 11th international conference on World Wide Web WWW '02. May 2002.

[17] Castelluccia, Claude. Mykletun, Einar. Tsudik, Gene "Improving secure server performance by re-balancing SSL/TLS handshakes" Proceedings of the 2006 ACM Symposium on Information, computer and communications security. Mar 2006.

[18] Chou, Jerry. Lin, Bill. Subhabrata, Sen. Oliver, Spatscheck "Minimizing collateral damage by proactive surge protection" Proceedings of the 2007 workshop on Large scale attack defense. Aug 2007.

[19] Cox, B.Tygar, J.D. and Sirbu, M., "NetBill security and transaction protocol." Proceedings of 1st USENIX Workshop on Electronic Commerce, New York. 1995

[20] Dai, Q. Kauffman, R. "Business Models for Internet Based E-Procurement Systems and B2B Electronic Markets: An Exploratory Assessment" 34th Annual Hawaii International Conference on System Sciences (HICSS-3))-Volume 7. Jan 2001.

[21] Wei, Kai. Chen, Yih-Farn and Smith, Alan. "WhoPay: A Scalable and Anonymous Payment System for Peer-to-Peer Environments." Department of Electrical Engineering and Computer Sciences Univeristy of California, Berkley, CA. May 2005.

[22] Hong, T. and Yuanzhi, Q. (2001). Legal guarantee of the security of electronic commerce. pages 167–170.

[23] D. Chaum, A. Fiat, M. Naor, Untraceable Electronic cash, Advances in Cryptology - CRYPTO '88, 1990.

[24] Yang, S., Su, S., and Lam, H. (2003). A nonrepudiation message transfer protocol for ecommerce. Proceedings of the IEEE International Conference on E-Commerce (CEC03), page 1.4

[25] Tanenbaum, Andrew. Steen, Maarten. "Distributed Systems Principles and Paradigms" Pearson Education. Oct 2006.

[26] Sherif, Mostafa H. "Protocols for Secure Electronic Commerce." CRC PRESS Advanced and Emerging Communications Technologies SERIES Second Edition. Nov 2003.

[27] Sirbu , M. Tygar, J.D. "NetBill: An Internet commerce system optimized for network delivered services" 40th IEEE Computer Society International Conference (COMPCON'95). Mar 1995.

[28] Shaw, M., Blanning, R., Strader, T., and Whinston,A. (2000). Virtual organization and e-commerce. In Handbook on Electronic Commerce, page 491. Spriner.

[29] Sherif , M.H. Serhrouchni , A. Gaid , A.Y. Farazmandnia, F. "SET and SSL: Electronic Payments on the Iner 1. D. Chaum, Blind signature for untraceable payments, Advances in Cryptology – Crypto - 82, Springer Verlag, p.p. 199-203.

[30] D. Chaum, A. Fiat, M. Naor, Untraceable Electronic cash, Advances in Cryptology - CRYPTO '88, 1990.

[31] T. Okamoto, "An Efficient Divisible Electronic Cash Scheme", Advances in Cryptology--Crypto 95, LNCS 963, Springer, pp.438-451, 1995.net" Third IEEE Symposium on Computers & Communications. Jun 1998.

[32] Makoto Matsumoto, Takuji Nishimura., () "Mersenne Twister: A 623-Dimensional Equidistributed Uniform Pseudo-Random Number Generator", ACM Transaction on Modeling and Computer Simulation 1998.

[33] Andrew Rukhin, Juan Soto, James Nechvatal, Miles Smid, Elaine Barker, Stefan Leigh, Mark Levenson, Mark Vangel, David Banks, Alan Heckert, James Dray, San Vo., () "Statistical Test Suite for Random and Pseudo Random Number Generators for Cryptographic Applications", National Institute of Standards and Technology Publication. 2001

[34] Lih-Yuan Deng, Hongquan Xu., () "A system of high-dimensional, efficient, long-cycle and portable uniform random number generator", ACM Transaction on Modeling and Computer Simulation, 2003.

[35] Jean-Philippe Aumasson, Luca Henzen,Willi Meier, and Raphael C.-W. Phan. SHA-3 proposal BLAKE, version 1.3. http://131002.net/blake/blake.pdf

[36] Niels Ferguson, Stefan Lucks, Bruce Schneier, Doug Whiting, Mihir Bellare, Tadayoshi Kohno, Jon Callas, and Jesse Walker. The Skein Hash Function Family. http://www.skein-hash.info/sites/default/files/skein1.1.pdf

[37] Praveen Gauravaram, Lars R. Knudsen, Krystian Matusiewicz, Florian Mendel, Christian Rechberger, and Søren S. Thomsen Martin Schläffer.
Grøstl – a SHA-3 candidate. http://www.groestl.info/Groestl.pdf

Authors Profile

Dr. Yousif AL-Bastaki is the Information Technology Advisor to the Hon. Deputy Prime Minster in the Kingdom of Bahrain. He has held several top rank administrative positions at the University of Bahrain and national level. He has edited, published research papers in international journals and conferences. He has written several books in e-government and well recognized as an excellent teacher, researcher and a very good speaker. Dr. Yousif earned his PhD from University of Nottingham, UK in 1996 and MSc from the University of Leeds, Currently he is an associate professor at the University of Bahrain. His research interests includes information assurance, neural networks, genetic algorithms E-Learning, e-commerce protocols secure network protocols, green computing and e-government strategies and implementation.

Dr. Ajantha Herath earned his PhD from the Gifu University, Japan, in 1997. His research interests include e-commerce protocols; secure network protocols, computer forensics and algorithm transformations to cryptographic hardware. He worked as the Professor at the University of Fiji's Department of Computer Science and Information Technology in 2011. At present he is teaching at the University of Bahrain. In 1988 he received the Monbusho research scholarship award. In 2007 he received the Outstanding Research Award for Commitment to Excellence in Computer Forensics and Development of Student Leaders and Researchers from the IEEE–Region 2 AIAA USA. He is a senior member of the IEEE. In 1986, Herath brothers established the Herath Foundation to help financially needy but talented students and awarded more than 7000 scholarships to continue their higher education.

Comparison of 2D and 3D Local Binary Pattern in Lung Cancer Diagnosis

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Yeni Herdiyeni
Department of Computer Science
Bogor Agricultural University
West Java, Indonesia

Hiroshi Okumura
Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*— Comparative study between 2D and 3D Local Binary Patter (LBP) methods for extraction from Computed Tomography (CT) imagery data in lung cancer diagnosis is conducted. The lung image classification is performed using probabilistic neural network (PNN) with histogram similarity as distance measure. The technique is evaluated on a set of CT lung images from Japan Society of Computer Aided Diagnosis of Medical Images. Experimental results show that 3D LBP has superior performance in accuracy compare to 2D LBP. The 2D LBP and 3D LBP achieved a classification accuracy of 43% and 78% respectively.**

Keywords- lung cancer detection; local binary pattern; probabilistic neural network.

I. Introduction

Lung cancer is one of the main cause deaths in the world among both men and woman, with an impressive rate about five million deadly cases per year [1]. Lung cancer is a disease of abnormal cells multiplying and growing into a tumor. Cancer cell can be carried away from the lungs in the blood, or lymph fluid that surrounds lung tissue. Lymph flows through lymphatic vessels, which drain into lymph nodes located in the lungs and in the center of the chest. Lung cancer often spreads toward the center of the chest because the natural flow of lymph out of the lungs is toward the center of the chest.

There are two major groups of lung cancer, Small Cell Lung Cancer (SCLC) and Non-Small Cell Lung Cancer (NSCLC). SCLC is a disease in which malignant (cancer) cells form in the tissues of the lung. Also known as oat cell cancer, NSCLC is the most common form of lung cancer that incorporates a variety of cancer sub types such as Squamous cell carcinoma, adenocarcinoma, and large cell carcinoma. This type of cancer also spreads to other parts of the body but tends to progress slowly to different parts of the body slower than SCLC.

One of the main imaging modality for this kind of disease diagnostic is computerized tomography (CT)[1] of the patient chest.

The advantage of the X-ray CT is pulmonary nodules that are typical shadow of pathological changes of the lung cancer can be detected more clearly compared to the chest X-ray examination even if they are at early stages. A number of methods have been explored for lung cancer detection using CT images. [2,3,4] uses measures on co-occurrence matrices, measures on run-length matrices, moments of the attenuation or intensity histogram, and in some cases fractal dimension as features. Sluimer et al. [5] used a filter bank of Gaussians and Gaussian derivatives.

The system using standard local binary patterns also has been developed and tested to 2D lung CT images. Sorensen et al [6] used joint local binary patterns (LBP) and intensity histogram for discriminating different texture pattern in 2D lung CT images. The extension of original LBP from 2D images to 3D volume data has been applied to facial expression [7].

[8] has been compared 3D texture representation for 3D brain MR images. The experimental results have shown that 3D LBP perform best with precision recall of 65% compare to 3D Grey Level Co-occurrence Matrices (3D GLCM), 3D Wavelet Transform (3D WT) [9], [10] and 3D Gabor Transform (3D GT).

In this research, we compare 2D LBP [11] and 3D LBP [7] as a texture features for lung cancer detection. We use probabilistic neural network (PNN) classifier [12] with histogram similarity as distance measure. To the best of our knowledge, this is the first application to use 3D LBP on lung CT image.

II. Proposed Method

A. Framework

In the following section the framework system is describe. The system is divided into two stages: training and test. In the training-stage, well-segmented lung are loaded into a feature extraction module.

In this research, we used 2D LBP and 3D LBP to extract texture feature of segmented lung. Based on the extracted feature vector from LBP, the trainer classifier will classify the lung cancer image. We used probabilistic neural network (PNN) as features classifier. The system flow is illustrated in Figure 1.

[1] http://en.wikipedia.org/wiki/X-ray_computed_tomography

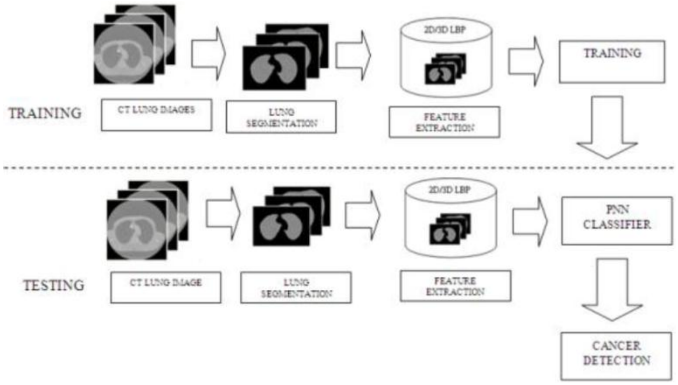


Figure 1. The framework system

B. Lung Segmentation

This section describes segmentation of the lung cancer. The goal of segmentation is to simplify the representation of an image into something that is more meaningful and easier to analyze.

In medical imaging, segmentation is important for feature extraction, image measurement and image display. In this research lung segmentation is done to eliminate of all artifacts external to patient's body and removal of thorax remaining just parenchyma. We used image enhancement and morphological operation for lung segmentation in the CT image to achieve a better orientation in the image.

Left and right lung are segmented using morphological operation and adaptive threshold based on statistical criteria. Morphology is a technique of image processing based on shape. The value of each pixel in the output image is based on comparison of the corresponding pixel in the input image with its neighbors.

C. 2D Local Binary Pattern (2DLBP)

After lung segmentation, the next step is feature extraction. Feature extraction is the process of defining a set of features, or image characteristic, which will most efficiently or meaningfully represent information that is important for analysis and classification. For 2D LBP, the features extraction that we use based on the local binary patterns (LBP) proposed by Ojala *et al.* [9].

In general, LBP measures the local structures at a given pixel using P samples on a circle of radius R around the pixel and summarizes this information with a unique code for each local structure or pattern (Figure 2).

The operator is highly non-linear and detects microstructures in the image at different resolutions governed by the parameter R, for example spots, edges, corners, etc., exemplified in the right part of Figure 3.

To obtain LBP value, thresholding performed on the neighborhood circular pixels using the central pixel, and then multiply by binary weighting. As an example for the sampling points P = 8 and radius R = 1, the calculation of LBP value is illustrated in Figure 4.

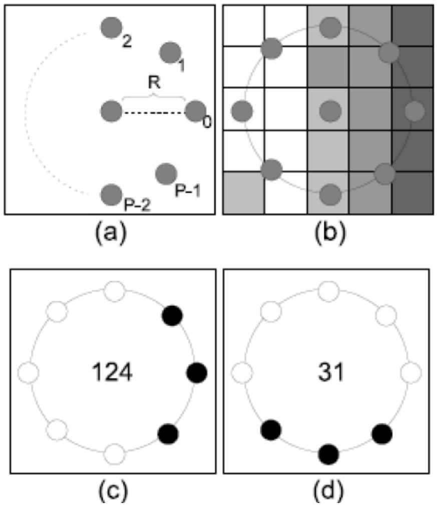


Figure 2. Illustration of LBP. (a) The filter is defined by two parameters; the circle radius R and the number of samples P on the circle. (b) Local structure is measured with reference to a given pixel by placing the center of the circle in the position of that pixel. (c) The samples on the circle are binarized by thresholding with the intensity in the center pixel as threshold value. Black is zero and white is one. (d) Rotating the example image in (b) ninety degrees clock wise reduce the LBP code to 31 which is smallest possible code for this binary pattern. This principle is used to achieve rotation invariant.

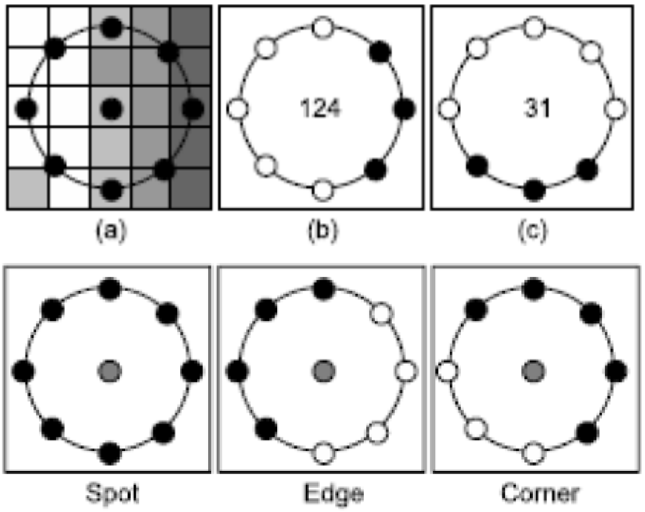


Figure 3. Some of the microstructures that LBP are measuring

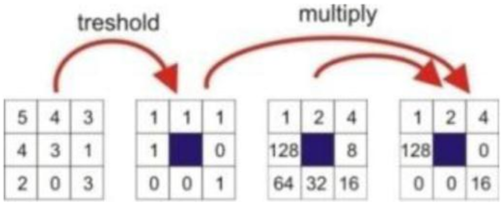


Figure 4. Calculation of LBP value.

LBP can be formulated as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{1}$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad (2)$$

where x_c and y_c are the coordinate of center pixel, p is circular sampling points, P is number of sampling points or neighborhood pixels, g_p is gray scale value of p, g_c is center pixel, and s or sign is threshold function. For classification purpose, the LBP values are represented as a histogram.

D. 3D Local Binary Pattern (3DLBP)

3D local binary pattern (3D LBP) or volume local binary pattern (VLBP) is an extension of LBP operator that combining the motion and appearance [7]. The features extracted in a small local neighborhood of the volume are not insensitive with respect to translation and rotation, but also robust with respect to monotonic gray-scale changes. In terms of 3D LBP form, Zhao, et. al [7] have proposed a 3D dynamic texture recognition by concatenating three histogram obtained from LBP on three orthogonal planes (LBP-TOP), i.e., XY, XZ and Y planes, the idea that has been adopted in this investigation.

VLBP is defined as the joint distribution V of gray level of *3P+3(P>1)* image pixel [7]. P is the number of local neighboring points around the central pixel in one frame. Both are shown in equation (3).

$$V = v(g_{t_c-L,c}, g_{t_c-L,0}, \cdots\cdots g_{t_c-L,P-1}, g_{t_c,0}, \cdots,$$
$$g_{t_c,P-1}, g_{t_c+L,0}, \cdots g_{t_c+L,P-1}, g_{t_c+L,c}). \qquad (3)$$

where the gray value $g_{t_c,c}$ corresponds to the gray value of the center pixel of the local volume neighborhood, $g_{t_c-L,c}$ and $g_{t_c+L,c}$ correspond to the gray value of the center pixel in the previous and posterior neighboring frames with time interval L; $g_{t,p}(t = t_c - L, t_c, t_c + L; p = 0, \ldots, P-1)$ correspond to the gray values of P equally spaced pixel on a circle of radius R *(R>0)* in image t, which form a circularly symmetric neighbor set.

Volume local binary pattern (*VLBP$_{L,P,R}$*) characterizes the spatial structure of the local volume dynamic textures shown in equation (4).

$$VLBP_{L,P,R} = \sum_{q=0}^{3P+1} v_q 2^q \qquad (4)$$

Figure 6 shows the whole computing procedures for *VLBP$_{1,4,1}$*. We begin by sampling neighboring points in the volume, and then thresholding every point in the neighborhood with the value of the center pixel to get a binary value. Finally we produce the VLBP code by multiplying the threshold binary values with weights given to the corresponding pixel and sum up the results.

The basic VLBP code is calculated for each pixel in the cropped portion of the DT, and the distribution of the code is used as a feature vector denoted by D shown in equation (5):

$$D = v(VLBP_{L,P,R}(x, y, t)), x \in \{\lceil R \rceil, \ldots, X - 1 - \lceil R \rceil\},$$
$$y \in \{\lceil R \rceil, \ldots Y - 1 - \lceil R \rceil\}, t \in \{\lceil L \rceil, \ldots T - 1 - \lceil L \rceil\} \qquad (5)$$

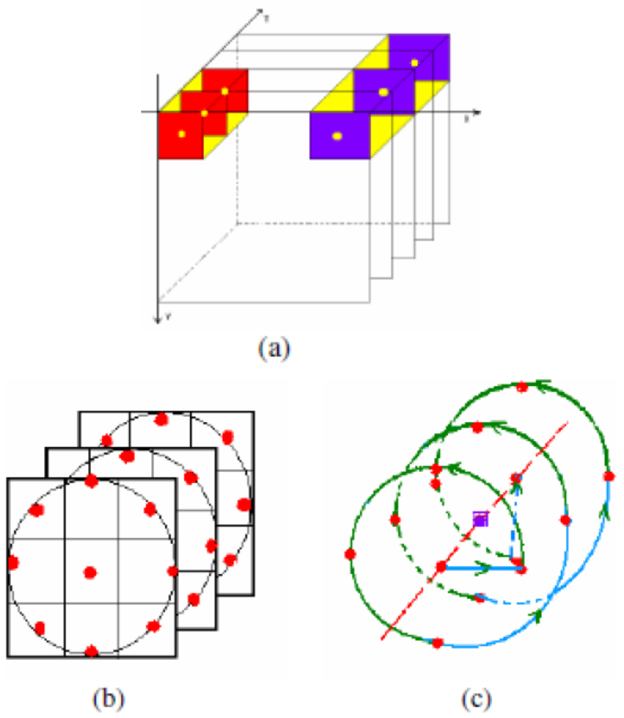


Figure 5. (a) Volume in dynamic texture (Right volume with L=1, while left volume with L=2. (b) Circularly symmetric neighbor sets in volume (R=1 and P=8), (c) Neighboring points along the helix surface of the cylinder (P=4).

The histograms are normalized with respect to volume size variations by setting the sum of their bins to unity. Because the dynamic texture is viewed as sets of volumes and their features are extracted on the basis of those volume textons[2], VLBP combines the motion and appearance to describe dynamic texture [7].

To make the VLBP computationally simple and easy to extend, only the co-occurrences on three separated planes are then considered. The textures are modeled with concatenated Local Binary Pattern histograms from Three Orthogonal Planes (LBP-TOP). The circular neighborhoods are generalized to elliptical sampling to fit the space-time statistics. A block-based approach combining pixel-level, region-level and volume-level features is proposed for dealing with such nontraditional dynamic textures in which local information and its spatial locations should also be taken into account.

The LBP-TOP (LBP on three orthogonal planes) only considers the co-occurrences statistics in three direction, XY, XT, and YT. LBP-TOP code is extracted from the XY, XT and YT planes, which are denoted as XY-LBP, XT-LBP, and YT-LBP, for all pixels and statistic of three different planes are obtained, and then concatenated into a single histogram. The procedure is demonstrated in Figure 7.

There are two differences between VLBP and LBP-TOP. Firstly, the VLBP uses three parallel planes of which only the middle one contains the center pixel. The LBP-TOP, on the other hand, uses three orthogonal planes which intersect in the center pixel.

[2] http://www.cis.upenn.edu/~jshi/papers/ICCV99b_final.pdf

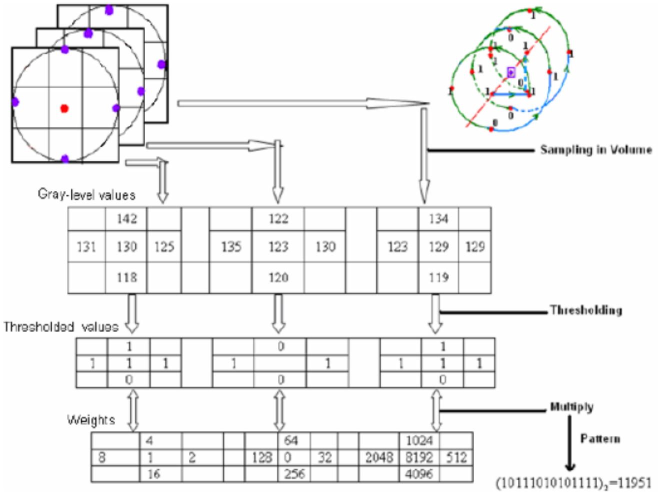


Figure 6. Procedure of VLBP1,4,1

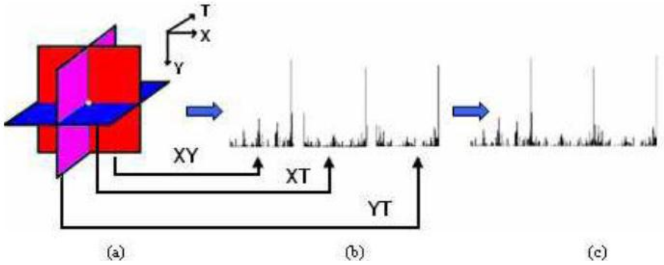


Figure 7. (a) Three planes in 3D-LBP (b) LBP histogram from each plane (c) Concatenated feature histogram

Secondly, VLBP considers the co-occurrences of all neighboring points from three parallel frames, which tends to make the feature vector too long. LBP-TOP considers the features distributions from each separate plane and then concatenates them together, making the feature vector much shorter when the number of neighboring points increase.

E. Probabilistic Neural Network(PNN)

Probabilistic Neural Network (PNN) proposed by Donald Specht in 1990 as an alternative back-propagation neural network. PNN has several advantages i.e. training requires only one iteration, and general solution is obtained by using a Bayesian approach [13]. PNN is a neural network that uses radial basis function (RBF). RBF is a function that is shaped like a bell that scales a nonlinear variable [14].

PNN consists of four layers, input layer, pattern layer, summation layer and output layer. PNN structure is shown in Figure 8. The layers that make up the PNN are as:

1) Input layer

Input layer is input *x* consisting of *k* value to be classified in one class of *n* classes.

2) Pattern layer

Pattern layer performs dot product between input and weight x_{ij} , or $Z_i = x \cdot x_{ij}$, then divided by a certain bias σ then inserted into the radial basis functions, that is $radbas(n) = exp(-n)^2$.

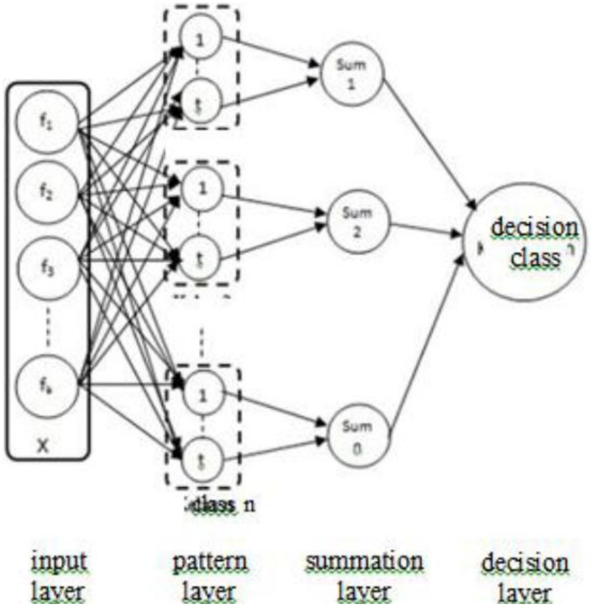


Figure 8 Structure of PNN.

Thus, the equation used in pattern layer is computed as follows:

$$f(x) = exp\left(-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right)^2 \qquad (6)$$

with x_{ij} express training vector class *i* order *j*.

3) Summation layer

In this layer, each pattern in each class is summed to produce a population density function for each class. The equation used at this layer is as follows:

$$p(x) = \frac{1}{(2\pi)^{\frac{k}{2}}\sigma^k t}\sum_{i=1}^{t}\exp(-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2})^2 \qquad (7)$$

4) Output layer

At the decision layer input *x* will be classified into class *I* if the value is larger than any other class.

F. Cross Validation

Cross-validation is widely used for model assessment and comparison. Cross-validation is a method of measuring generalization error through the use of holdout data. In this research we used *K*-fold cross validation. In this procedure, dataset *D* is divided into *K* partitions of roughly equal size, $D = \bigcup_{k=1}^{k} D_k$ and each partition is termed a “fold” of the dataset (thus there are *K*-folds). The model is trained on *K-1* folds and the K-th fold is used for testing [10]. This repeated K times such that each fold is used for testing exactly once.

The *K*-fold cross validation procedure is often applied to choose a model specific parameter. Suppose that the models are indexed by parameter $\lambda \in \Lambda$ with corresponding estimated model function $f_{\lambda}^{-k(i)}$ to be evaluated on the *k*-th fold. The optimal $\hat{\lambda}$ is chosen as follows:

$$\lambda = \arg\min_{\lambda} KCV(\lambda) = \arg\min_{\lambda} \frac{1}{n}\sum_{i=1}^{n}(y_i - y_{\lambda}{}^{-k(i)})^2 \qquad (8)$$

The final model is trained with the optimal parameter over the entire data, with the KCV statistic reported as the cross validation prediction error.

G. Performance Measures

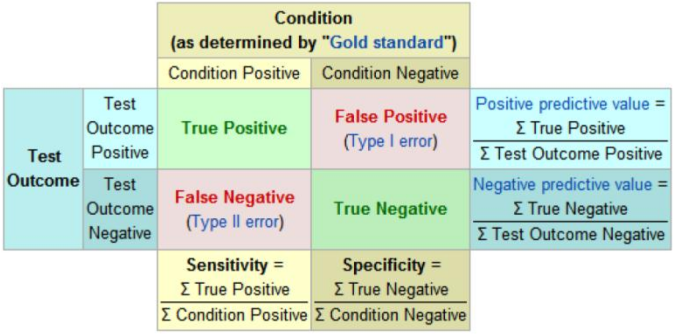
To evaluate performance of the system, we used receiver operating characteristic curve (ROC)[3] analysis. Four basic measures are defined from the set of true conditions and observed information as shown in Figure 9.

These basic measures are true positive, false positive, false negative and true negative rates or fraction. A “positive” observation in an image means that the object was observed in the test. A ”negative” observation means that the object was not observed in the test. A “true condition” is the actual truth, while an observation is the outcome of the test.

A graph between TP and FP is called receiver operating characteristic (ROC) curve for a specific medical imaging or diagnostic test for detection of an object. A true positive fraction is also called the sensitivity while the true negative fraction is known as specificity of the test for detection of an object. In each test, we measured the following values:

$$Sensitivity = \frac{TP}{TP + FN} \qquad (9)$$

$$Specificity = \frac{TN}{TN + FP} \qquad (10)$$



		Condition (as determined by "Gold standard")		
		Condition Positive	Condition Negative	
Test Outcome	Test Outcome Positive	True Positive	False Positive (Type I error)	Positive predictive value = Σ True Positive Σ Test Outcome Positive
	Test Outcome Negative	False Negative (Type II error)	True Negative	Negative predictive value = Σ True Negative Σ Test Outcome Negative
		Sensitivity = Σ True Positive Σ Condition Positive	Specificity = Σ True Negative Σ Condition Negative	

Figure 9. A conditional matrix for defining four basic performance measures of receiver operating characteristics curve (ROC) analysis.
- True positive fraction (TP) is the ratio of the number of positive observations to the number of positive true condition cases.
- False negative fraction (FN) is the ratio of the number of negative observations to the number of positive true condition cases.
- False positive fraction (FP) is the ratio of the number of positive observations to the number of negative true condition cases.
- True negative fraction (TN) is the ratio of the number of negative observations to the number of negative true condition cases.

Accuracy of the test is given by a ratio of correct observation to the total number of examination cases. Thus, the accuracy is expressed with equation (11).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (11)$$

III. Experimental Results

A. Data Collections

The input of our system is a set of lung CT images with doses from 150*mAs* to 200*mAs*, in plane resolution 0.63mm/pixel, slice thickness of 5mm. The number of data is 14 patients' data. The image sizes are all 512 by 512 pixel within cross section, with 43 to 81 slices in Z direction. Data set are collected from Japan Society of Computer Aided Diagnosis of Medical Images. There are 14 patients and five classes of lung cancer that we used in the experiments:

1) *Small cell carcinoma (6 patients)*
2) *Tubercoluma of the lung (2 patients)*
3) *Inflammantory pseudotumor (2 patients)*
4) *Adecarcinoma (1 patient)*
5) *Squamous cell carcinoma (3 patients)*

Due to the fact that we only have a relatively smaller size of database, then we divide data into two types cancer, benign (non-cancerous) and malignant (cancerous). So, in these experiments, tubercoluma of the lung and inflammantory pseudotumor we classify into benign and the remaining data we classify into malignant. Before feature extraction stage, lung CT images are segmented. The process is started with thorax extraction. This stage comprises the removal of all artifacts external to the patient’s body. There are mainly consists of five steps for lung segmentation, i.e.:

1) *Crop the original image from 512 by 512 pixels to 466 by 351 pixels to get the lung region.*

2) *Compute the original image histogram. Apply adaptive threshold based on statistical criteria using Otsu Algorithm*[4] *from original image histogram to make binary image.*

3) *Remove object which touching the border.*

4) *Apply morphology operation by apply Structuring Element (SE) to an input image. In this step we use erosion and dilation to remove pixel on the object boundaries. The number of pixel removed from the object in an image depends on the size of structuring element used to process the image. In this research we used SE (3 by 3).*

Process flow is illustrated in Figure 10. After lung segmentation, the next stage is feature extraction. In our study, we use original LBP (2D LBP) [9] and volume LBP (3D LBP) [7] to extract lung cancer features. There three experiments we conducted to compare performance in accuracy of 2D LBP and 3D LBP.

In this study, the 3D LBP-TOP uses three orthogonal planes that intersect in the center voxel[5] as shown in Figure 11.

[3] http://en.wikipedia.org/wiki/Receiver_operating_characteristic

[4] http://en.wikipedia.org/wiki/Otsu's_method

[5] http://www.webopedia.com/TERM/V/voxel.html

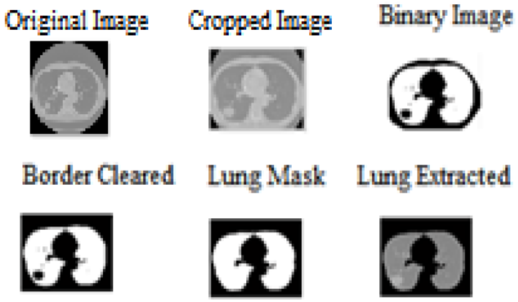


Figure 10. Lung segmentation

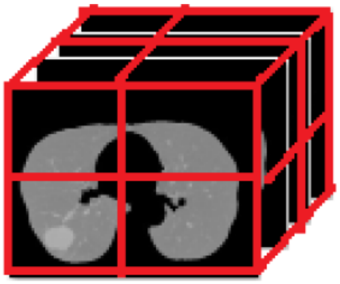


Figure 11. Block for 3D lung cancer.

In the experiment we select 4 and 8 neighbors with 1 voxel radius as a local neighborhood. Fifty-nine uniform LBP code is then extracted from XY, XT, YT planes respectively, producing a 59 bin histogram for each plane by accumulating the 59 binary patterns. Finally, three histogram are concatenated to generate a 3D texture representation, giving the size of a feature vector being 177 (=59 by 3).

In order to describe local features parts of a lung, in this experiment 3D volume is divided into four non-overlapping equality sized blocks, which are 2 block along each of x, y, z axes respectively as shown in Figure 8. The method of 3D LBP is applied to extract local features from each block respectively. Therefore, the dimension of the feature vector for 3D lung cancer is the size of local features multiplied by 4.

B. Results

In this section, the performance of 2D LBP and 3D LBP in lung cancer detection is evaluated. In the experiment we select 8 neighbors with 1 voxel radius as a local neighborhood. Table 1 presents the overall classification rates of 2D LBP and Table 4 presents the overall classification rates of 3D LBP. Table 2,3 shows ROC analyzed results for original 2D LBR and 2D block LBR, respectively. Meanwhile, Table 5,6 shows ROC analyzed results for original 3D LBR and 3D block LBR, respectively.

As shown in Table 1 and Table 4, in 2D LBP and 3D LBP, the block LBP method has better recognition rates compare to original LBP. Experimental results show that 3D block LBP is superior to 2D block LBP in accuracy. The 2D Block LBP and 3D block LBP achieved a classification accuracy of 43% and 78% respectively. We only use a relatively small size of database in the experiments so further study will be needed to verify this finding.

TABLE I. THE CLASSIFICATION RATES OF 2D LBR (8,1)

Sample	Accuracy (Original)	Accuracy (Block)
1	0.30	0.41
2	0.38	0.42
3	0.43	0.50
4	0.39	0.43
5	0.34	0.41
6	0.29	0.44
7	0.36	0.42
Average	0.36	0.43

Note : (Neighbor points, Radius)

TABLE II. ROC ANANALYSIS FOR ORIGINAL 2D LBR (8,1)

	0 = benign	1 = malignant
0 = benign	0	6
1 = malignant	3	5
	Sensitivity =0	Specificity =0.46

TABLE III. ROC ANALYSIS FOR 2D BLOCK LBR (8,1)

	0 = benign	1 = malignant
0 = benign	0	5
1 = malignant	3	6
	Sensitivity =0	Specificity =0.54

TABLE IV. THE CLASSIFICATION RATES OF 3D LBP (8,8,8,1,1,1,1)

Sample	Accuracy (Original)	Accuracy (Block)
1	0.54	1.00
2	0.75	0.75
3	0.68	0.75
4	0.75	0.73
5	0.78	0.73
6	0.73	0.69
7	0.72	0.78
Average	0.71	0.78

Note : (Neighbor point (XY, XT, YT), FxRadius, FyRadius, Interval)

TABLE V. ROC ANALYSIS OF ORIGINAL 3D LBP

	0 = benign	1 = malignant
0 = benign	0	0
1 = malignant	4	10
	Sensitivity =0	Specificity =1

TABLE VI. ROC ANALYSIS OF 3D BLOCK LBP

	0 = benign	1 = malignant
0 = benign	1	0
1 = malignant	3	10
	Sensitivity =0.25	Specificity =1

IV. CONCLUSION AND FUTURE DIRECTION

We conducted a comparative study between two techniques of feature extraction from lung computed tomography images for lung cancer diagnosis. According to experiments results, the block LBP method has better recognition rates compared to the original LBP.

As overall, it may conclude that 3D LBP is superior to 2D LBP in accuracy.

The 2D block LBP and 3D block LBP achieved a classification accuracy of 43% and 78%, respectively. Due to the fact that we only have a relatively smaller size of database, further study is needed to justify these results.

ACKNOWLEDGMENT

This work is financially funded by Directorate of Higher Education, the ministry of National Education, Republic of Indonesia for Program Academic Recharging (PAR-C).

The authors would like thank to Department of Information Science, Saga University for research collaboration. Their support is gratefully acknowledged.

REFERENCES

[1] Nunzio, G. D, *et al*. Automatic Lung Segmentation in CT Images with Accurate Handling of Hilar Region. Journal of Digital Imaging. 24, 1,. 11-27, 2011.

[2] Uppalari, et al. Quantification of pulmonary emphysema from lung computed tomography images. Am. J. Respair. Crit. Care. Med. 156(1), 248-254, 1997.

[3] Chabat, et. al. Obstructive lung diseases: texture classification for differentiation at CT. Radiology 228(3), 871-877, 2003.

[4] Xu, et. al. MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007 Lecture Notes in Computer Science, 2007, 4791/2007, 825-833, 2007.

[5] Sluimer, et. al. Computer Aided diagnosis in high resolution CT of the lungs. Med. Phys. 30(12), 3081-3090, 2003.

[6] Sorensen, L., Shaker, S.B., Bruijne, M. L. Texture Classification in the Lung CT using Local Binary Patterns. Medical Image Computing and Computer- Assisted Intervention - MICCAI 2008. LNCS 5241, 925-933, 2008.

[7] Zhao, G and Pietikainen, M. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expression. IEEE Transaction on Pattern Analysis and Machine Intelligence. 2007.

[8] Gao, X.,Qian, Y., Hui, R., Loomes, M., Comley, R., Barn, B., Chapman, A., and Rix, J. Texture-Based 3D Image Retrieval for Medical Application. IADIS International Conference e-Health 2010,.101-108, 2010.

[9] Arai K, Fundamental theory on wavelet and wavelet analysis, Morikita-Shuppan Publishing Co. Ltd., 2000.

[10] Arai K., Self learning for wavelet and wavelet analysis, Kindaikagakusha Publishing Co. Ltd., 2006.

[11] M¨aenp¨a¨a T. and M. Pietik¨ainen. In C. Chen and P. Wang, editors, Handbook of Pattern Recognition and Computer Vision, chapter Texture analysis with local bi-nary patterns., pages 197–216. World Scientific, 2005

[12] Specht D.F., Probabilistic Neural Networks for Classification, Mapping, or Associative Memory, IEEE International Conference on Neural Networks, I, 525-532, July 1998

[13] Holder M. and P.O.Lewis, Phylogeny estimation: traditional Baysian approaches, Nature Reviews, 275-284 , 2003

[14] Yee, Paul V. and Haykin, Simon, Regularized Radial Basis Function Networks: Theory and Applications. John Wiley, 2001.

AUTHORS PROFILE

Kohei Arai received a PhD from Nihon University in 1982. He was subsequently appointed to the University of Tokyo, CCRS, and the Japan Aerospace Exploration Agency. He was appointed professor at Saga University in 1990. He is also an adjunct professor at the University of Arizona and is Vice Chairman of ICSU/COSPAR Commission A

Balancing a Sphere in a Linear Oscillatory Movement through Fuzzy Control

Gustavo Ozuna, German Figueroa
Department of Industrial Engineering and Systems
University of Sonora
Hermosillo, Mexico

Marek Wosniak
Politechnika Lodzka,
Poland

Abstract— **The following paper describes an intelligent control problem, which depends on the balance of a metallic sphere on a beam, that oscillates in only one point, localized in the middle of the beam, using it for the balance of this fuzzy control system.**

Keywords— ball beam system; balancing a metallic sphere; inclination; equilibrium forces.

I. INTRODUCTION

There exist several problems with the control of electromechanical and mechatronics systems, for example the inverted pendulum [1], which solution has been approached from various methodologies, another example is the braking of a vehicle at high speed, the direction and the detection of obstacles, this is why the following control problem is proposed, which consist in balancing a metallic sphere on a beam with linear oscillatory movement, at a supporting point located in the center as shown in figure 1, in which the sphere will experiment several positions changes in an unstable form losing balance, for the control system handle the stabilizing of the sphere in the center, the fuzzy control is used.

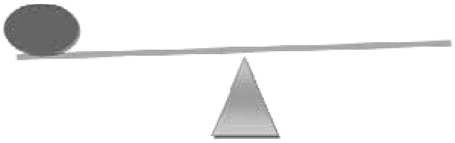


Figure 1. Balancing a sphere over the beam

II. PROBLEM

The following problem is to success the balance of a sphere on a beam which has the point of support in the center, when the sphere travels form right to left and left to right, makes the beam lose the balance, currently there are several similar problems called ball beam system [2].

This system consists in slope a beam, in which a metal sphere travels, searching the right inclination to achieve the same position of another sphere that is in constant movement outside the system.

The system uses a distance sensor which provides the information to take the right inclination of the beam to place the sphere at the same position as the reference sphere that is in constant movement, unlike the previously cited this only includes one sphere.

III. PROTOTYPE

To control the inclination of this system a direct current motor is connected to the center of the beam, which rotates the beam form left to right and right to left, as appropriate to achieve the balance of the sphere, taking the signal of a sensor of inclination as shown in figure 2.

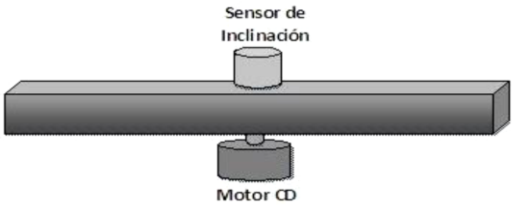


Figure 2. Direct current motor and

The leaver system in which the sphere moves freely is represented by the next equations, such as the Torque Moment and the equilibrium forces on the beam:

$$F_N - F_A - F_B = 0 \qquad (1)$$

$$M=Fd \qquad (2)$$

$$F=mg \qquad (3)$$

In which F_N is the upward force of the supporting point and F_A and F_B are the forces that the sphere exercises when it moves and M is the Torque that are changing according to the distance of the sphere from the pivot [4] as shown in figure 3.

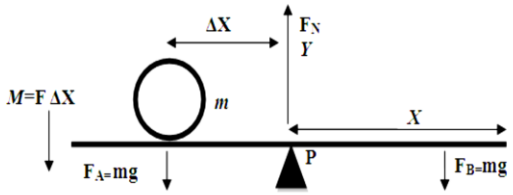


Figure 3. Forces that interact with the system

IV. CONTROL

The fuzzy control system is proposed for the system, which will emulate the human reasoning, in the system structure is identified the logic flow fuzzy inference of the input variable inclination with respect of the output variables direction and speed, the fuzzy inferences [6], the figure 4 shows the control system structure showing the analogical inputs, the block of fuzzy rules as the analogical outputs, the lines that connect symbolizes the flux of data simulated in the simulator FuzzyTech.

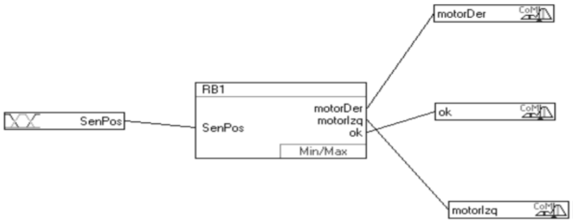


Figure 4. System simulated in FuzzyTech

The linguistic variables used in the system translates the real values of the position sensor into linguistic values and also into the outputs of the system shown in the table 1 and 2.

TABLE 1. INPUT VARIABLE

#	Variable Name	Type	Unit	Min	Max	Default	Term Names
1	SenPosicion		Units	0	1	0.5	extremoI medioI Bajoi ok bajoD meidioD extremoD

Table 2. Output Variables

#	Variable Name	Type	Unit	Min	Max	Default	Term Names
1	motorDer		Units	0	1	0	MDlow MDmedium MDhigh
2	motorIzq		Units	0	1	0	MIlow MImedium MIhigh
3	Ok		Units	0	1	0	ok

The fuzzy conjunct shown in the figure 5 shows the position sensor input using linguistic variables related to the right or left extreme of the leave like this until arriving the central position, place that is search to place the sphere accomplishing balance.

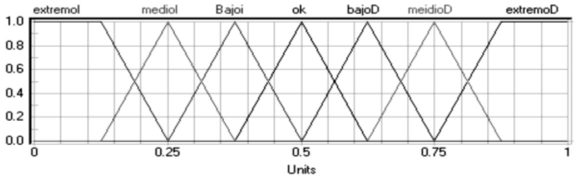


Figure 5. Fuzzy conjunct of input

The outputs of the system represented in the fuzzy conjunct of the figures 6, 7, 8 shows how the direct wire motor should behave, the spin direction, as well as the signal that presents when the system is balance.

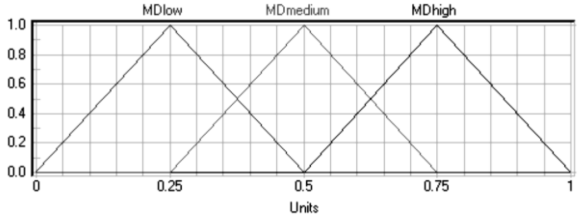


Figure 6. Fuzzy conjunct of output right to left motor

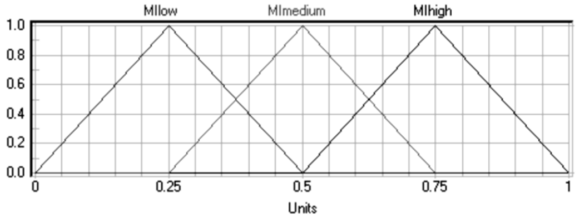


Figure 8. Fuzzy conjunct of output left to right motor

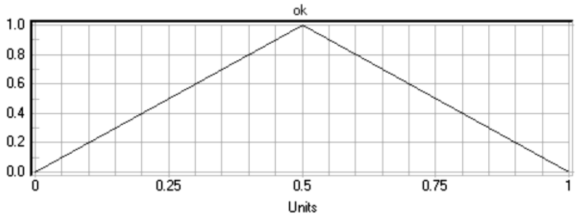


Figure 9. Fuzzy conjunct of output of the motor ok

The block rules shows the strategy for the control of the fuzzy system, the context is defined by the same variables of input and output operated by the maximum and minimum method shown in the table 3.

TABLE 3. FUZZY RULES

SenPos	DoS	motorDer	DoS	motorIzq	DoS	ok
extremoI	1.00	MDhigh				
medioI	1.00	MDmedium				
Bajoi	1.00	MDlow				
Ok					1.00	ok
bajoD			1.00	MIlow		
meidioD			1.00	MImedium		
extremoD			1.00	MIhigh		

The behavior of the dc motor and sensor following the linguistic strategies to balance the sphere on the beam are shown in the figures 10 and 11.

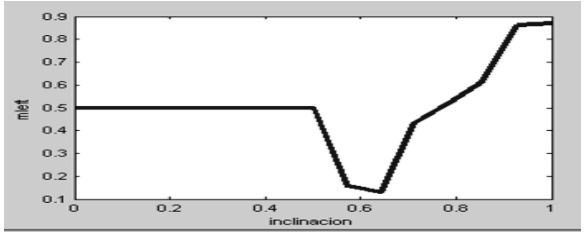


Figure 10. Fuzzy conjunct of output of the motor ok

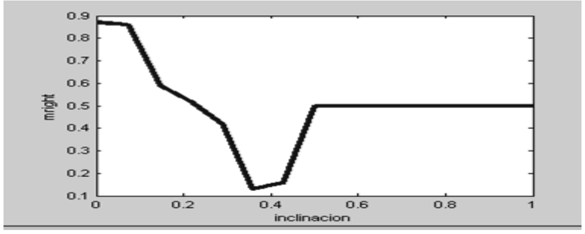


Figure 11. Fuzzy conjunct of output of the motor ok

V. ELECTRONIC PROPOSAL

The position sensor, that it is proposed for the following system, consist of a variable resistant which is connected to two operational amplifiers configured as comparators, the circuit will provide the sense of inclination by the differential of voltage, shown in figure 12.

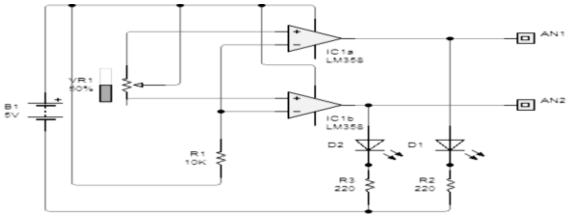


Figure 12. Circuit of the inclination sensor

The proposed system uses a microcontroller, 16F684A in which the fuzzy rules are programmed, this possess four analog inputs which connects the inclination sensor and the outputs for the direct wire motor in an array of transistors (H Bridge) to control the spin of it, to stabilize the beam.

VI. CONCLUSIONS

The research of new methods of solution for electromechanical problems, help to put into practice the knowledge of intelligent control into the solution of Mechatronic systems, it is expected that this type of control problems may be helpful for other investigators to put in practice the different types of intelligence control.

REFERENCES

[1] Ogata K. *"Ingeniería de Control Moderna"*, Pearson, España, 4ta Ed. 2003.

[2] Bazañes L. *"Control Digital",* UPC, España, 1ra Ed. 1994.

Hewit P. *"Física Conceptual"*, Pearson, 9na Ed, 2004.

[3] Giancoli D. *"Física"*, Pearson, 6ta Ed. México, 2006.

[4] Alvarez L. "Fundamentos de Inteligencia *Artificial"*, Universidad de Murcia, 1ra. España, 1994

[5] Xafha F. *"Programación en C++ para Ingenieros",* Paraninfo, 1ra Ed. España, 2006.

[6] Spartacus C. *"Teoria de Control"*, Alfa omega, 1ra Ed. España, 1999

[7] Alciatore D. "Introducción a la Mecatrónica", McGraw-Hill, 3ra Ed. México, 2008.

AUTHORS PROFILE

Gustavo Ozuna he is a robotics and electronics professor, in University of Sonora Mexico. gozuna@industrial.uson.mx

Marek Wozniak he is a professor in Lodz University of Technology, Department of Vehicles and Fundamentals of Machine Design, Poland.

German Figueroa, he is a industrial engiener student in the University of Sonora Mexico.

E-learning Document Search Method with Supplemental Keywords Derived from Keywords in Meta-Tag and Descriptions which are Included in the Header of the First Search Result

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Herman Tolle
Software Engineering Department
Brawijaya University
Malang, Indonesia

Abstract— **Optimization method for e-learning document search with keywords which are derived from the keywords and descriptions in the meta-tag of web search results together with thesaurus engine is proposed. 15 to 20% of improvement on hit rate of search performance is confirmed with the proposed search engine.**

Keywords- Search engine; e-learning content; thesaurus engine.

I. INTRODUCTION

When a word or words are typed in search engines, a list of web sites that contain those words is displayed. The words you enter are known as a query [1]. Baeza-Yates and Ribeiro-Neto linked Information Retrieval to the user information needs which can be expressed as a query submitted to a search engine [2]. Search engines were also known as some of the brightest stars in the Internet investing frenzy that occurred in the late 1990s [3]. Although search engines are programmed to rank websites based on their popularity and relevancy, empirical studies indicate various political, economic, and social biases in the information they provide [4],[5].

Based on our previous experiment [6], our system can detect client mobile browser and provide proper format for document and mobile markup language. We propose a new method and approach for developing a new search engine for helping people search E-Learning document on the Internet. We develop a new system based on the open search engine API (Application Protocol Interface) like Google and Yahoo. We create an e-learning specific search engine with improvement in the efficiency and effectiveness in searching document file format comparing with just using original Google or Yahoo. The new system is also accessible through mobile browser on mobile devices for support recent and future technology in the mobile area [7] Method for e-learning contents search engine of ELDOXEA is proposed already. ELDOXEA allows search e-learning contents with a single keyword. Hit rate of search performance of the ELDOXEA is not good enough. In order to improve hit rate, supplemental keywords are required in addition to the firstly input keyword which is referred to primary keyword hereafter.

In order to choose appropriate supplemental keywords for improvement of hit rate, several attempts are performed. First one is to use keywords and descriptions in meta-tag of the header of the first web search result. Keywords are used to be in the meta-tag. Also, there are some keywords in the descriptions in the header of the first web search result. Therefore, these supplemental keywords in the meta-tag and descriptions in the header are applicable to add to the primary keyword.

Second one is to use keywords which are derived from thesaurus engines. Thesaurus engine provides similar words to the primary keyword with their priority. Therefore, these are used for supplemental keywords.

Third one is to use twitter for gathering suggestions of supplemental keywords from twitters. Reliability of the twitters can be evaluated with the previously proposed method.

Fourth one is to use Bulletin Board System: BBS for gathering suggestions of supplemental keywords from the community members. Firstly, the user has to send a message which is requesting supplemental keywords to BBS system with the primary keyword. Then the user has reply message with supplemental keywords. These are to be candidates of supplemental keywords.

Experiments with some queries, “Linear Algebra” and “Hazardous Materials Handler examination” are conducted for searching e-learning contents. By adding supplemental keywords to the primary keyword based on the aforementioned four methods, we confirm their efficiency, hit rate improvements.

The second section describes the proposed methods for choosing supplemental keywords while the third section describes some experimental results followed by some concluding remarks and some discussions.

II. PROPOSED METHODS

A. Search Engine

There are three types of search engines, (1) Directory type,

(2) Information Collection Robot type, and (3) Hybrid type.

(1) Directory type

Content in the Web directories is searched and examined by operators so that it is reliable. Information content is not so much. There is some response delay after updating web sites.

(2) Information Collection Robot: ICR type

ICR is collecting from the web sites so that information contents are rich. On the other hands, classification items are not so many as appropriate for a wide variety of search purposes.

Ex) Google[1], Inforseek[2], etc.

(3) Hybrid type

It has the aforementioned both benefits.

Ex) Yahoo[3], MSN[4], ELDOXEA[5], etc.

The proposed search engine for e-learning document search is based on the ELDOXEA of Hybrid type.

B. *Efficiency Improvement on Searching Process of E-Learning Document on the Internet*

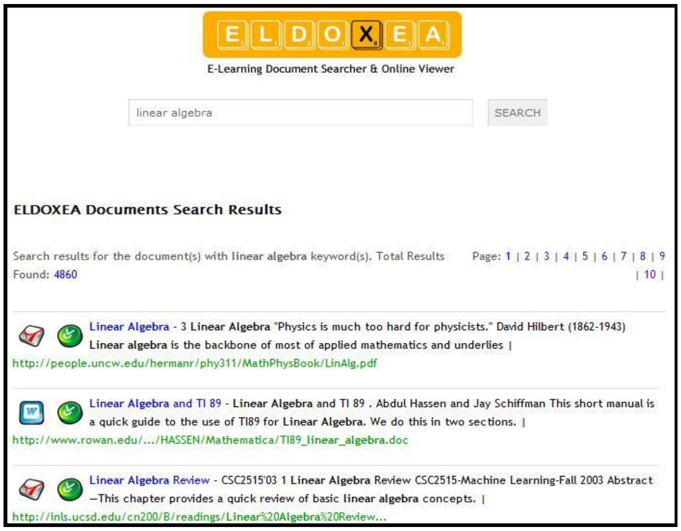
To improve the efficiency of the searching process of E-Learning document, we design a new process for searching and display the document. One of the most problems on efficiency while we search a document file is how to get the appropriate files in a fast way. People usually have to check on each document files of the results set, start from the first results.

Most of document files format is not able to display directly on browser without additional plug-in or application. So, in the conventional way, we should download the document file first, and then open it in our PC, for example, we open PPT files using *Microsoft PowerPoint*. After check the content then we can decide to keep this file or not. This process takes time if the file size is large and we should wait for download process. Another problem with this process is storage problem for download too many unrelated files.

We design a new process that help user to preview document, before they decide to save (download) it. The preview processes displaying the document file in the same results area, so users still stay in the same page while checking each result. This will lead the user to control which document is related or not related and decide to save or skip it. The proposed algorithm as follows:

1. Get *SearchKey* from user input
2. Get *RelatedKey* and *NotRelatedKey* from a Database based on user's *SearchKey*
3. Create the *CompleteKeyword* (1)
4. Search using the *CompleteKeyword*
5. Get results and display it
6. If user click one of the results, preview the document files in same page
7. Preview display, user can choose Download/Save or close button
8. If user click Download, then save the files

Example can be seen through http://www.eldoxea.com as shown in Figure 1 ((a) for Internet terminals and (b) for mobile phone).



linear algebra

ELDOXEA Documents Search Results

Search results for the document(s) with linear algebra keyword(s). Total Results Found: 4860

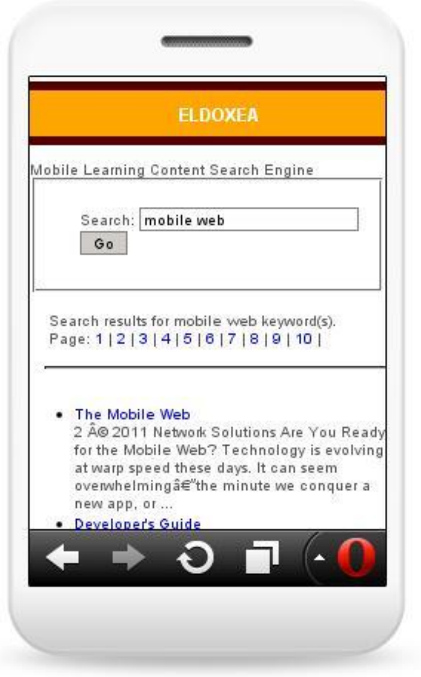
Page: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Linear Algebra - 3 Linear Algebra "Physics is much too hard for physicists." David Hilbert (1862-1943) Linear algebra is the backbone of most of applied mathematics and underlies |
http://people.uncw.edu/hermanr/phy311/MathPhysBook/LinAlg.pdf

Linear Algebra and TI 89 - Linear Algebra and TI 89 . Abdul Hassen and Jay Schiffman This short manual is a quick guide to the use of TI89 for Linear Algebra. We do this in two sections. |
http://www.rowan.edu/.../HASSEN/Mathematica/TI89_linear_algebra.doc

Linear Algebra Review - CSC2515'03 1 Linear Algebra Review CSC2515-Machine Learning-Fall 2003 Abstract —This chapter provides a quick review of basic linear algebra concepts. |
http://inls.ucsd.edu/cn200/B/readings/Linear%20Algebra%20Review...

(a)ELDOXEA for internet terminals



Mobile Learning Content Search Engine

Search: mobile web

Search results for mobile web keyword(s).
Page: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

- The Mobile Web
  2 Â© 2011 Network Solutions Are You Ready for the Mobile Web? Technology is evolving at warp speed these days. It can seem overwhelmingâ€"the minute we conquer a new app, or ...
- Developer's Guide

(b) ELDOXEA for mobile phone

Figure 1. Top page of ELDOXEA

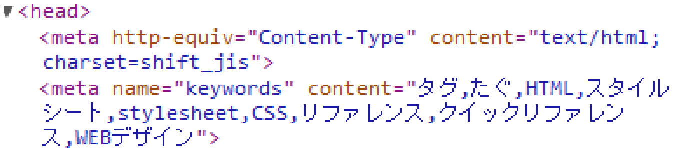
C. *Supplemental Keyword Selection with the Keywords in the Meta-Tag and the Descriptions in the First Search Result*

In order to choose appropriate supplemental keywords for improvement of hit rate, several attempts are performed. First one is to use *keywords* and *descriptions* in meta-tag of the first web search result. Keywords are used to be in the meta-tag.

[1] http://en.wikipedia.org/wiki/Google_Search
[2] http://en.wikipedia.org/wiki/Infoseek
[3] http://en.wikipedia.org/wiki/Yahoo!_Search
[4] http://en.wikipedia.org/wiki/Bing
[5] http://b.hatena.ne.jp/entry/www.eldoxea.com/

Also, there are some keywords in the descriptions in the first web search result. Therefore, these supplemental keywords in the meta-tag and descriptions are applicable to add to the primary keyword. This approach using the assumption that related website should contain the similar keyword, while not related website containing another not related keyword. So, we try to find the intersection of keyword in meta-tag of page header between webs in the search results set.

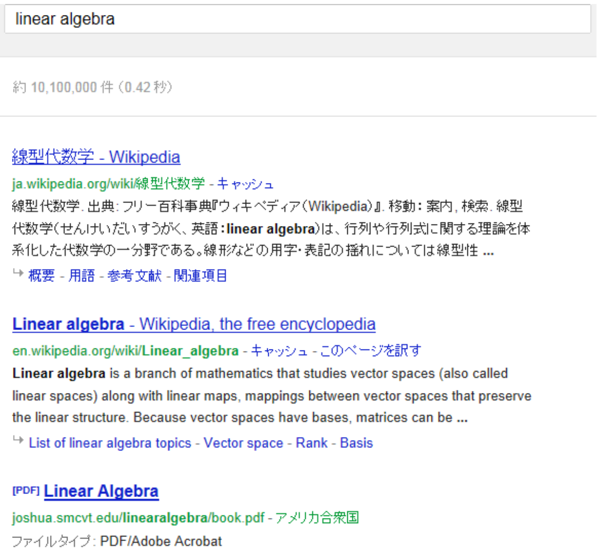
Figure 2 shows an example of header. In the header, there are meta-tag and descriptions. In these meta-tags and the descriptions, there are some keywords. We could use these keywords as supplemental keywords for search.



```
▼<head>
   <meta http-equiv="Content-Type" content="text/html;
   charset=shift_jis">
   <meta name="keywords" content="タグ,たぐ,HTML,スタイル
   シート,stylesheet,CSS,リファレンス,クイックリファレン
   ス,WEBデザイン">
```

Figure 2. shows an example of header

Figure 3 shows the first three search results based on Google search with the keyword "Linear Algebra". When I visit the first URL of Wikipedia in Japanese, then Figure 4 appears. Then source code can be displayed as shown in Figure 5. Although we cannot get any keyword in the meta-tag sometime or description in the header, it used to be appeared in the header. If we repeat the same keyword twice as keyword for search, then we get the other search results as shown in Figure 6



linear algebra

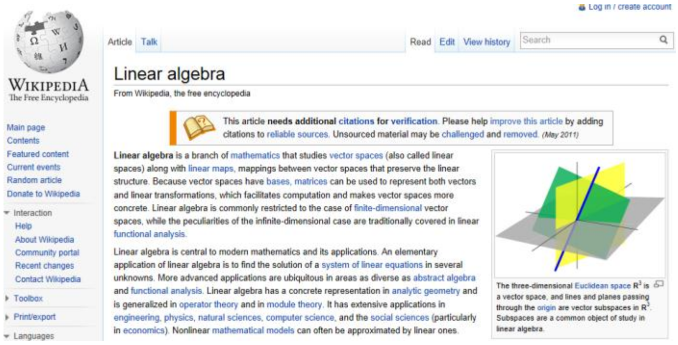
約 10,100,000 件 (0.42 秒)

線型代数学 - Wikipedia
ja.wikipedia.org/wiki/線型代数学 - キャッシュ
線型代数学. 出典: フリー百科事典『ウィキペディア（Wikipedia）』. 移動: 案内, 検索. 線型代数学（せんけいだいすうがく、英語：**linear algebra**）は、行列や行列式に関する理論を体系化した代数学の一分野である。線形などの用字･表記の揺れについては線型性 ...
↳ 概要 - 用語 - 参考文献 - 関連項目

Linear algebra - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/**Linear_algebra** - キャッシュ - このページを訳す
Linear algebra is a branch of mathematics that studies vector spaces (also called linear spaces) along with linear maps, mappings between vector spaces that preserve the linear structure. Because vector spaces have bases, matrices can be ...
↳ List of linear algebra topics - Vector space - Rank - Basis

[PDF] **Linear Algebra**
joshua.smcvt.edu/**linearalgebra**/book.pdf - アメリカ合衆国
ファイルタイプ: PDF/Adobe Acrobat

Figure 3. First three search results based on Google search with the keyword "Linear Algebra".



Linear algebra

From Wikipedia, the free encyclopedia

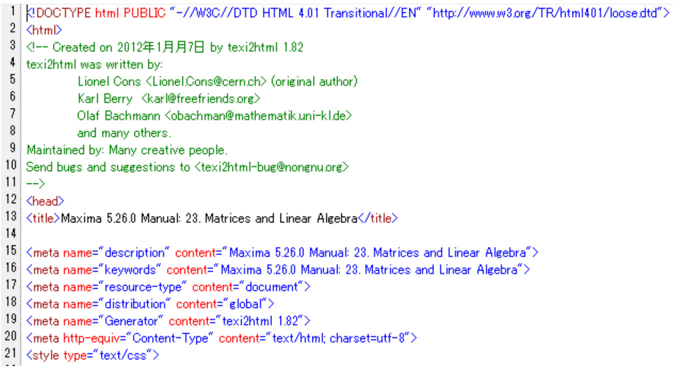
This article **needs additional citations for verification**. Please help improve this article by adding citations to reliable sources. Unsourced material may be challenged and removed. *(May 2011)*

Linear algebra is a branch of mathematics that studies vector spaces (also called linear spaces) along with linear maps, mappings between vector spaces that preserve the linear structure. Because vector spaces have bases, matrices can be used to represent both vectors and linear transformations, which facilitates computation and makes vector spaces more concrete. Linear algebra is commonly restricted to the case of finite-dimensional vector spaces, while the peculiarities of the infinite-dimensional case are traditionally covered in linear functional analysis.

Linear algebra is central to modern mathematics and its applications. An elementary application of linear algebra is to find the solution of a system of linear equations in several unknowns. More advanced applications are ubiquitous in areas as diverse as abstract algebra and functional analysis. Linear algebra has a concrete representation in analytic geometry and is generalized in operator theory and in module theory. It has extensive applications in engineering, physics, natural sciences, computer science, and the social sciences (particularly in economics). Nonlinear mathematical models can often be approximated by linear ones.

The three-dimensional Euclidean space R^3 is a vector space, and lines and planes passing through the origin are vector subspaces in R^3. Subspaces are a common object of study in linear algebra.

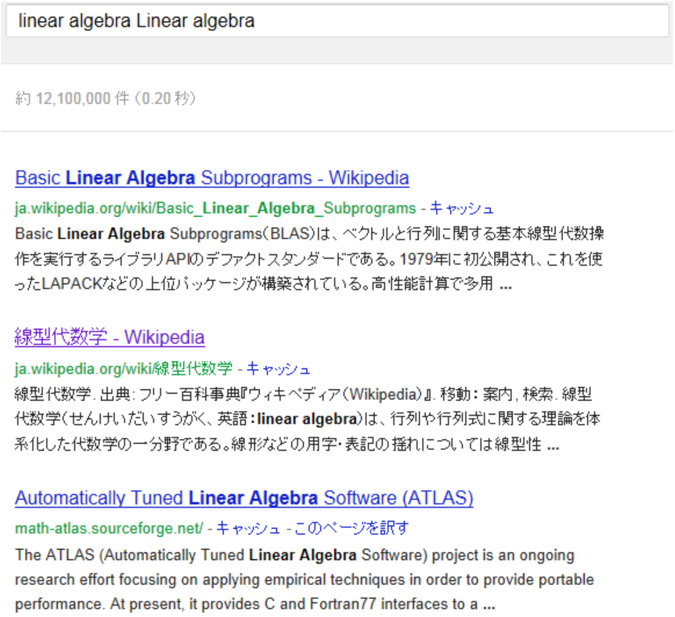
Figure 4. Top page of the first URL of Wikipedia of Linear Algebra in Japanese



```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html401/loose.dtd">
<html>
<!-- Created on 2012年1月月7日 by texi2html 1.82
texi2html was written by:
            Lionel Cons <Lionel.Cons@cern.ch> (original author)
            Karl Berry  <karl@freefriends.org>
            Olaf Bachmann <obachman@mathematik.uni-kl.de>
            and many others.
Maintained by: Many creative people.
Send bugs and suggestions to <texi2html-bug@nongnu.org>
-->
<head>
<title>Maxima 5.26.0 Manual: 23. Matrices and Linear Algebra</title>

<meta name="description" content="Maxima 5.26.0 Manual: 23. Matrices and Linear Algebra">
<meta name="keywords" content="Maxima 5.26.0 Manual: 23. Matrices and Linear Algebra">
<meta name="resource-type" content="document">
<meta name="distribution" content="global">
<meta name="Generator" content="texi2html 1.82">
<meta http-equiv="Content-Type" content="text/html; charset=utf-8">
<style type="text/css">
```

Figure 5. The source code of the first URL of Linear Algebra.



linear algebra Linear algebra

約 12,100,000 件 (0.20 秒)

Basic **Linear Algebra** Subprograms - Wikipedia
ja.wikipedia.org/wiki/Basic_**Linear**_**Algebra**_Subprograms - キャッシュ
Basic **Linear Algebra** Subprograms（BLAS）は、ベクトルと行列に関する基本線型代数操作を実行するライブラリAPIのデファクトスタンダードである。1979年に初公開され、これを使ったLAPACKなどの上位パッケージが構築されている。高性能計算で多用 ...

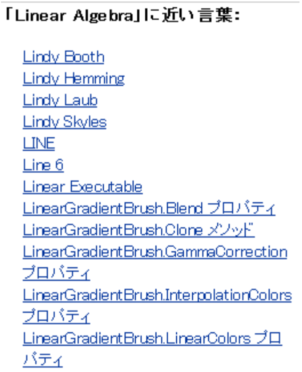
線型代数学 - Wikipedia
ja.wikipedia.org/wiki/線型代数学 - キャッシュ
線型代数学. 出典: フリー百科事典『ウィキペディア（Wikipedia）』. 移動: 案内, 検索. 線型代数学（せんけいだいすうがく、英語：**linear algebra**）は、行列や行列式に関する理論を体系化した代数学の一分野である。線形などの用字･表記の揺れについては線型性 ...

Automatically Tuned **Linear Algebra** Software (ATLAS)
math-atlas.sourceforge.net/ - キャッシュ - このページを訳す
The ATLAS (Automatically Tuned **Linear Algebra** Software) project is an ongoing research effort focusing on applying empirical techniques in order to provide portable performance. At present, it provides C and Fortran77 interfaces to a ...

Figure 6. Other search results we used to get with double keyword for search (in this case "Linear Algebra is refrained twice)

D. Supplemental Keyword Selection from Thesaurus Engine

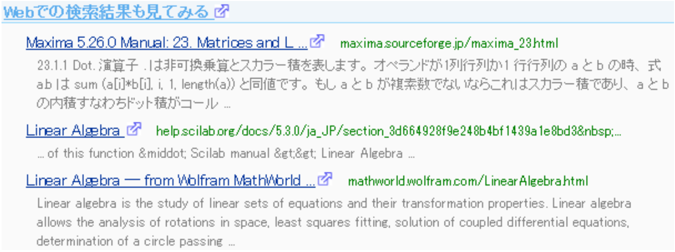
Second one is to use *keywords* which are derived from thesaurus engines. Thesaurus engine provides similar words to the primary keyword with their priority. Therefore, these are used for supplemental keywords. An example of the search results of thesaurus engine with "Linear Algebra" is shown in Figure 7. As shown in Figure 7, similar words to "Linear Algebra" are listed in the order of priority. Also, you can refer to the URLs as results of thesaurus engine as shown in Figure 8. Figure 9 shows an example of top page of the first priority of URLs of the search results of thesaurus engine, *Weblio*[6]. Then we can get the keywords in the meta-tag and the descriptions in the header when you check the source word of the web pages as shown in Figure 10.



「Linear Algebra」に近い言葉:

Lindy Booth
Lindy Hemming
Lindy Laub
Lindy Skyles
LINE
Line 6
Linear Executable
LinearGradientBrush.Blend プロパティ
LinearGradientBrush.Clone メソッド
LinearGradientBrush.GammaCorrection プロパティ
LinearGradientBrush.InterpolationColors プロパティ
LinearGradientBrush.LinearColors プロパティ

Figure 7. An example of search results of thesaurus engine of Weblio with the keyword "Linear Algebra" in Japanese.



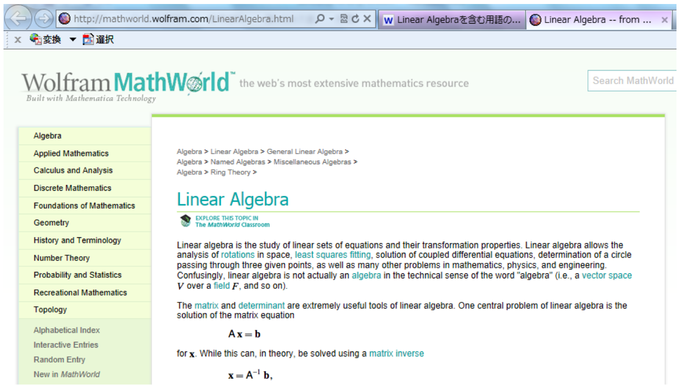
Webでの検索結果も見てみる

Maxima 5.26.0 Manual: 23. Matrices and L ... maxima.sourceforge.jp/maxima_23.html
23.1.1 Dot. 演算子 . は非可換乗算とスカラー積を表します。オペランドが1列行列か1 行行列の a と b の時、式 a.b は sum (a[i]*b[i], i, 1, length(a)) と同値です。もし a と b が複素数でないならこれはスカラー積であり、a と b の内積すなわちドット積がコール ...

Linear Algebra help.scilab.org/docs/5.3.0/ja_JP/section_3d664928f9e248b4bf1439a1e8bd3&nbsp;...
... of this function &middot; Scilab manual &gt;&gt; Linear Algebra ...

Linear Algebra — from Wolfram MathWorld ... mathworld.wolfram.com/LinearAlgebra.html
Linear algebra is the study of linear sets of equations and their transformation properties. Linear algebra allows the analysis of rotations in space, least squares fitting, solution of coupled differential equations, determination of a circle passing ...

Figure 8. URLs can be referred as results of thesaurus engine (in this case with the keyword of "Linear Algebra").



Algebra > Linear Algebra > General Linear Algebra >
Algebra > Named Algebras > Miscellaneous Algebras >
Algebra > Ring Theory >

Linear Algebra

Linear algebra is the study of linear sets of equations and their transformation properties. Linear algebra allows the analysis of rotations in space, least squares fitting, solution of coupled differential equations, determination of a circle passing through three given points, as well as many other problems in mathematics, physics, and engineering. Confusingly, linear algebra is not actually an algebra in the technical sense of the word "algebra" (i.e., a vector space V over a field F, and so on).

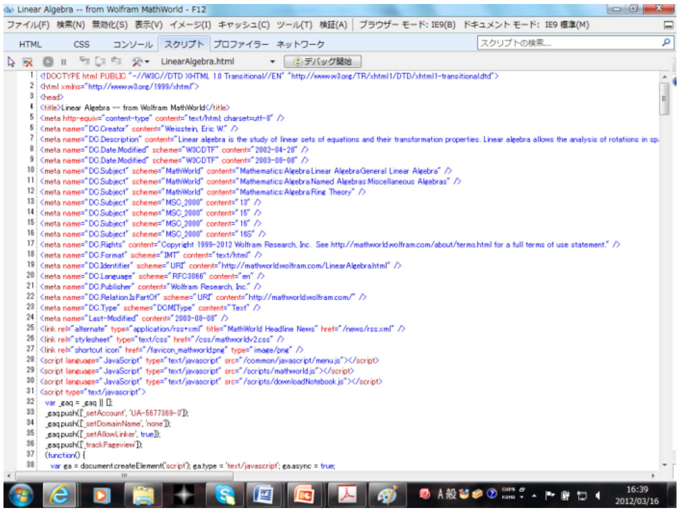
The matrix and determinant are extremely useful tools of linear algebra. One central problem of linear algebra is the solution of the matrix equation

$$\mathsf{A}\mathbf{x} = \mathbf{b}$$

for $\mathbf{x}$. While this can, in theory, be solved using a matrix inverse

$$\mathbf{x} = \mathsf{A}^{-1}\mathbf{b},$$

Figure 9. Example of top page of the first priority of URLs of the search results of thesaurus engine, *Weblio*.


Linear Algebra -- from Wolfram MathWorld - F12
ファイル(F) 検索(N) 無効化(S) 表示(V) イメージ(I) キャッシュ(C) ツール(T) 検証(A) ブラウザー モード: IE9(B) ドキュメント モード: IE9 標準(M)
HTML
CSS
コンソール
スクリプト
プロファイラー
ネットワーク
スクリプトの検索...
LinearAlgebra.html
デバッグ開始
16:39
2012/03/16

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>Linear Algebra -- from Wolfram MathWorld</title>
<meta http-equiv="content-type" content="text/html; charset=utf-8" />
<meta name="DC.Creator" content="Weisstein, Eric W." />
<meta name="DC.Description" content="Linear algebra is the study of linear sets of equations and their transformation properties. Linear algebra allows the analysis of rotations in sp
<meta name="DC.Date.Modified" scheme="W3CDTF" content="2002-04-28" />
<meta name="DC.Date.Modified" scheme="W3CDTF" content="2003-08-08" />
<meta name="DC.Subject" scheme="MathWorld" content="Mathematics:Algebra:Linear Algebra:General Linear Algebra" />
<meta name="DC.Subject" scheme="MathWorld" content="Mathematics:Algebra:Named Algebras:Miscellaneous Algebras" />
<meta name="DC.Subject" scheme="MathWorld" content="Mathematics:Algebra:Ring Theory" />
<meta name="DC.Subject" scheme="MSC_2000" content="13" />
<meta name="DC.Subject" scheme="MSC_2000" content="15" />
<meta name="DC.Subject" scheme="MSC_2000" content="16" />
<meta name="DC.Subject" scheme="MSC_2000" content="16S" />
<meta name="DC.Rights" content="Copyright 1999-2012 Wolfram Research, Inc.  See http://mathworld.wolfram.com/about/terms.html for a full terms of use statement." />
<meta name="DC.Format" scheme="IMT" content="text/html" />
<meta name="DC.Identifier" scheme="URI" content="http://mathworld.wolfram.com/LinearAlgebra.html" />
<meta name="DC.Language" scheme="RFC3066" content="en" />
<meta name="DC.Publisher" content="Wolfram Research, Inc." />
<meta name="DC.Relation.IsPartOf" scheme="URI" content="http://mathworld.wolfram.com/" />
<meta name="DC.Type" scheme="DCMIType" content="Text" />
<meta name="Last-Modified" content="2003-08-08" />
<link rel="alternate" type="application/rss+xml" title="MathWorld Headline News" href="/news/rss.xml" />
<link rel="stylesheet" type="text/css" href="/css/mathworldv2.css" />
<link rel="shortcut icon" href="/favicon_mathworld.png" type="image/png" />
<script language="JavaScript" type="text/javascript" src="/common/javascript/menu.js"></script>
<script language="JavaScript" type="text/javascript" src="/scripts/mathworld.js"></script>
<script language="JavaScript" type="text/javascript" src="/scripts/downloadNotebook.js"></script>
<script type="text/javascript">
  var _gaq = _gaq || [];
  _gaq.push(['_setAccount', 'UA-5677369-8']);
  _gaq.push(['_setDomainName', 'none']);
  _gaq.push(['_setAllowLinker', true]);
  _gaq.push(['_trackPageview']);
  (function() {
    var ga = document.createElement('script'); ga.type = 'text/javascript'; ga.async = true;
```

Figure 10. The keywords in the meta-tag and the descriptions in the header when you check the source word of the web pages

After that, we could determine the supplemental keywords in accordance with their priority.

E. Supplemental Keyword Selection from Twitters

Third one is to use twitter for gathering suggestions of supplemental keywords from twitters. Reliability of the twitters can be evaluated with the previously proposed method[7]. The *Weblio* of thesaurus engine provides the gate for twitter. From this gate, we can bet valuable information of URLs related to the primary keyword. Reliability of the twitter has to be checked though.

F. Supplemental Keyword Selection from BBS

Fourth one is to use Bulletin Board System: BBS and chatting for gathering suggestions of supplemental keywords from the community members. Firstly, the user has to send a message which is requesting supplemental keywords to BBS system as well as chat with the primary keyword. Then the user has reply message with supplemental keywords. These are to be candidates of supplemental keywords. Most of Learning Management System like *Moodle*[8] provides BBS and chatting capabilities. Using these functions, we can get valuable information relating to the primary keyword.

III. EXPERIMENTS

A. Method for Experiments Conducted

We conduct the experiment for our propose methods in two ways. Subjectively search through Yahoo search engine and objectively creating a new search engine using Yahoo API. We check the results of a search engine that only using primary keyword, and then comparing the results while using combination of primary key and supplementary keyword. First, correct answer of Yahoo search with queries of "Linear Algebra" and "Hazardous Material Handler test" is determined. Yahoo search is a kind of hybrid type of search engine that is

[6] http://thesaurus.weblio.jp/content/エンジン

[7] http://www.readwriteweb.com/archives/twazzup_a_better_twitter_search_engine.php
http://www.govloop.com/profiles/blogs/twitters-reliability-an-issue

[8] http://moodle.org/

same as ELDOXEA. The first 50 candidates of URLs are selected. All of the 50 sites is visited and evaluated subjectively. Then these sites are divided into "*appropriate*" or "*not appropriate*". The term "*appropriate*" means that the web content is containing e-learning materials that related to the input keyword. The term "*not appropriate*" means that the web content may not related to the input keyword in the area of e-learning.

The first four keywords for the primary keywords, "Linear Algebra" and "Hazardous Material Handler Examination" are as follows,

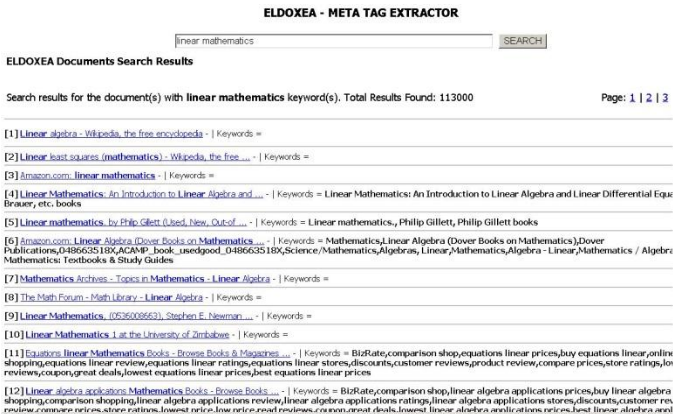
(1) Linear Algebra: Linear Algebra, Mathematics, Matrix Algebra/Geometry

(2) Hazardous Material Handler Examination: Hazardous Material Handler Examination, Gasoline Handler, Hazardous, National Certificate

Figure 12 shows hit rate of the proposed method and of the search results with same primary keywords repeatedly. Square denotes the hit rate with the proposed method while the upside down triangle denotes search result with using the same primary keyword "Linear Algebra" repeatedly. In the case, of usage of the same primary keyword of "Linear Algebra" repeatedly, search result shows that hit rate is saturated at the number of supplemental keyword is 1, which is corresponding to the search with two same primary keyword results in maximum hit rate.

On the other hands, the first two supplemental keywords show the maximum hit rate for the "Hazardous Material Handler Examination" case while repeated usage of the same primary keyword does work for the "Hazardous Material Handler Examination" case.

Second, keywords in the meta-tag and the descriptions are extracted from the header of these sites. Then the keywords are sorted with priority depending on their frequency. These processes are automatically done by our search engine system. Figure 11 shows the screenshot of our meta-tag keyword extractor for automatically extract *keywords* from the meta-tag of a search results.


ELDOXEA - META TAG EXTRACTOR

linear mathematics SEARCH

ELDOXEA Documents Search Results

Search results for the document(s) with **linear mathematics** keyword(s). Total Results Found: 113000 Page: 1 | 2 | 3

[1] Linear algebra - Wikipedia, the free encyclopedia - | Keywords =

[2] Linear least squares (mathematics) - Wikipedia, the free ... - | Keywords =

[3] Amazon.com: linear mathematics - | Keywords =

[4] Linear Mathematics: An Introduction to Linear Algebra and ... - | Keywords = Linear Mathematics: An Introduction to Linear Algebra and Linear Differential Equa
Brauer, etc. books

[5] Linear mathematics, by Philip Gillett (Used, New, Out-of ... - | Keywords = Linear mathematics., Philip Gillett, Philip Gillett books

[6] Amazon.com: Linear Algebra (Dover Books on Mathematics ... - | Keywords = Mathematics,Linear Algebra (Dover Books on Mathematics),Dover
Publications,048663518X,ACAMP_book_usedgood_048663518X,Science/Mathematics,Algebras, Linear,Mathematics,Algebra - Linear,Mathematics / Algebra
Mathematics: Textbooks & Study Guides

[7] Mathematics Archives - Topics in Mathematics - Linear Algebra - | Keywords =

[8] The Math Forum - Math Library - Linear Algebra - | Keywords =

[9] Linear Mathematics, (0536008663), Stephen E. Newman ... - | Keywords =

[10] Linear Mathematics 1 at the University of Zimbabwe - | Keywords =

[11] Equations linear Mathematics Books - Browse Books & Magazines ... - | Keywords = BizRate,comparison shop,equations linear prices,buy equations linear,online
shopping,equations linear review,equations linear ratings,equations linear stores,discounts,customer reviews,product review,compare prices,store ratings,lov
reviews,coupon,great deals,lowest equations linear prices,best equations linear prices

[12] Linear algebra applications Mathematics Books - Browse Books ... - | Keywords = BizRate,comparison shop,linear algebra applications prices,buy linear algebra
shopping,comparison shopping,linear algebra applications review,linear algebra applications ratings,linear algebra applications stores,discounts,customer rev
review,compare prices,store ratings,lowest price,low price,read reviews,coupon,great deals,lowest linear algebra applications prices,best linear algebra appl

Figure 11. Our automatic Meta Tag Keywords extractor system

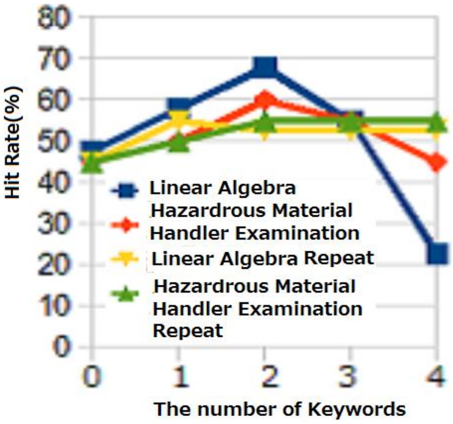


Figure 12. Hit rate of the proposed method and of the search results with same primary keywords repeatedly.

B. Supplemental Keyword Selection with Thesaurus Engine, Weblio

As shown in Figure 7, similar words to "Linear Algebra" are listed in the order of priority. There, however, is less related keyword from the Weblio of thesaurus engine. Therefore, it would not be worked for finding supplemental keyword at all.

C. Applicability of the Proposed Search Engine

The proposed search engine is applicable to the other primary keywords. Other than "Linear Algebra", the proposed search engine is applied to the chemistry, mechanics, etc. Improvements of hit rate for these primary keywords are evaluated. Figure 13 shows the improvements of hit rate for searching the primary keyword with the primary and supplemental keywords.

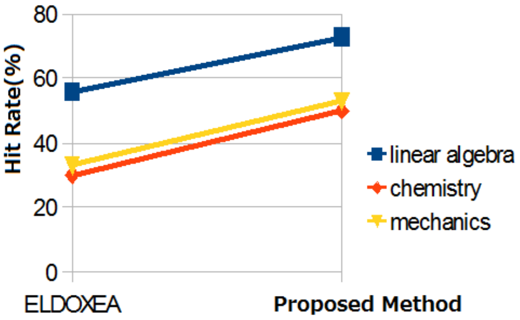


Figure 13. Improvement of hit rate of the proposed method in comparison to the previously proposed ELDOXEA of e-learning content search engine.

17 to 20% of improvements are confirmed for the primary keywords, Linear Algebra", Chemistry" and "Mechanics".

Other two methods by utilizing twitter and BBS system will be discussed in the other paper in the near future.

IV. CONCLUSIONS

The proposed search method uses not only one single primary keyword, but also supplemental keywords which are derived from the keywords in the meta-tag and descriptions in the page header of a website appeared in the first search result.

Hit rate is defined as matching accuracy between subjectively determined success search and the current search results of URLs. The hit rate of the proposed method is compared to that of the search method with the same primary keyword with repeatedly used supplemental keyword. Depending on the number of supplemental keyword, hit rate is increasing. Improvement of the hit rate of the proposed search method is 15 to 20 % while that of the search method with repeatedly used supplemental keyword which is same as primary keyword is around 10 %. It is also found that the proposed search method is applicable to the other primary keywords.

ACKNOWLEDGMENT

The authors would like to thank Miss Makoto Hara for her effort to conduct the experiments.

REFERENCES

[1] S. Brin and L. Page. The anatomy of a large-scale hypertextual Web search engine. In WWW, 1998.

[2] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. Modern Information Retrieval. Addison Wesley, May 1999.

[3] Gandal, Neil (2001). "The dynamics of competition in the internet search engine market". *International Journal of Industrial Organization* **19** (7): 1103–1117

[4] Segev, Elad (2010). Google and the Digital Divide: The Biases of Online Knowledge, Oxford: Chandos Publishing

[5] Vaughan, L. & Thelwall, M. (2004). Search engine coverage bias: evidence and possible causes, Information Processing & Management, 40(4), 693-707

[6] Kohei Arai, Herman Tolle “Module Based Content Adaptation of Composite E-Learning Content for Delivering to Mobile Learners”, International Journal of Computer Theory and Engineering (IJCTE), Vol 3, No. 3, pp. 381-386, June 2011

[7] Kohei Arai, Herman Tolle, Efficiency improvements of e-learning document search engine for mobile browser, International Journal of Research and Reviews on Computer Science, 2, 6, 1287-1291, 2011.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science, and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission “A” of ICSU/COSPAR since 2008. He wrote 30 books and published 322 journal papers.

Herman Tolle, He graduated Bachelor degree in Electrical Engineering from Brawijaya University, Malang in 1998, also graduated Master degree in Telecommunication Information System from Bandung Institute of Technology (ITB), Bandung in 2002. He is with Engineering Faculty of Brawijaya University from 2002 to present. He is now a Doctoral student in Department of Information Science, Faculty of Science and Engineering, Saga University Japan. He has a major concern of research in image analysis, multimedia, content adaptation and web engineering.

The Construction of a Web-Based Learning Platform from the Perspective of Computer Support for Collaborative Design

Hsu, Cheng Mei
Department of Visual Communication and Design
China University of Technology
Taipei City, Taiwan (R.O.C.)

Abstract—**The purpose of this study is to construct a web-based learning platform of Computer Support for Collaborative Design (CSCD) based on theories related to a constructivist learning environment model, mind mapping and computer-supported collaborative learning. The platform conforms to the needs of design students and provides effective tools for interaction and collaborative learning by integrating the tools of mind mapping into a learning environment that utilizes CSCD, a computer-assisted support system that can support and enhance group collaboration. The establishment of the CSCD learning platform represents a significant advance from the fixed functions and existing models of current online learning platforms and is the only learning platform in the world that focuses on learners in design departments. The platform is outstanding for its excellence, user-friendly functions, and innovative technology. In terms of funding, technical ability, human resources, organizational strategies, and risk analysis and evaluations, the learning platform is also worthy of expansion and implementation.**

Keywords- computer support for collaborative design; computer support for collaborative learning; design education; mind-map; web-based learning platform.

I. Introduction

The advancement of information media and the convenience of the Internet have fostered the growth of web-based instruction, interactive functions, and resources available for learners to engage in interaction, learning, discussion, and access without time and geographic constraints. In design practice, the integration of computer technology has allowed designers to practice with consistency and innovation design approaches that were long believed to be possible only through the use of traditional tools or simulated hand-made physical objects, and ways in which humans communicate have been improved [20], [44]. Hence, web-based instruction is essential for design education.

However, because of how quickly departments have implemented network-based teaching, design departments fall short in comparison. The reasons for this delay include the many unique features of design education, such as difficulties with the design materials or works into digital formats [8], [22], [56], [61], [62], and the functions of current general web-based learning platforms are insufficient for teaching design, thus resulting in slow progress of web-based design instruction [10]. Based on the special knowledge and techniques of design, design students also possess unique learning styles and specialties, and they need to integrate related knowledge and coordinate professionals from different fields during the design process. For these reasons, the establishment of an individualized and adaptive web-based instructional platform that can enhance the effectiveness of design education remains a significant issue [11].

Creative thinking is one of the most important abilities in business and academic circles ([32], [52], [59]. Design students are highly sensitive to graphics and colors, and they are skilled at using visual-thinking models and the diffusion of images related to creative thinking methods to cultivate their creativity [30], [53], [55], [58]. Moreover, mind mapping is a learning method used to develop the potential of the left and right hemispheres of the human brain through graphic techniques. The design concept of mind mapping is associated with radiant thinking, which is a graphic type of organizational skill that incorporates graphics, colors, spaces, and imagination and effectively utilizes a whole-brain thinking method. The use of both graphics and text, in turn, enhances creativity, thereby completing the associative process of brainstorming [27], [49]. With advances in computer technology, computer-based mind mapping allows designers to create new concepts by freely linking concepts and integrating graphics, text, voice, video, and other multimedia through methods related to spatial and visual organization.

A designer must possess visual thinking capabilities, and mind mapping is a learning tool for graphic, visual, and spatial thinking. When applied to the creative thinking process in the early stages of a design, this tool may enhance a designer’s creative thinking ability. To establish a web-based instructional platform that meets the user’s needs and enhances learning effectiveness, the researcher proposed the CSCD platform, a design model based on design teachers’ and students’ needs for web-based instructional platform functions and learning tools. This model was developed by considering previously conducted research [63] applying the analytic hierarchy process (AHP) and by using learning theory as the basis for integrating and analyzing the design of the constructivist learning environment model, the Blackboard platform, the National Sun Yat-Sen Cyber University platform, and e-learning tools of technology companies for designing. Then, the researcher put forward the "Computer Support for Collaborative Design (CSCD) platform" design model based on the design department teachers' and students' needs for the web-based instruction platform functions and learning tools. Subsequently, the technical capabilities of many technology experts were integrated to establish a CSCD platform prototype, and modifications were conducted based on expert validity and user tests. After completing the CSCD platform, the experts and users assessed its usability and the overall design results through formative evaluations. The evaluation results show that the experts and users provided very positive feedback and believed that the platform not only possessed strengths lacking in general online learning platforms but also that the completion of the setup marked the beginning of a customized web-based learning platforms. The platform combines several web-based technologies and teaching resources and displays works through multimedia graphics and animation, with user-friendly and image-style characteristics that are useful for learners with an expertise in imaging.

A. *The Purposes of the Research*

The research purposes include the following: 1) to design and set up a teaching website with the constructivist learning environment model, computer-supported collaborative learning theory, and mind mapping-related theories that serve as the basis; 2) to design and set up a teaching website that aids in understanding the influence of the CSCD and collaborative mind mapping on design department students' learning attitudes, learning effectiveness, and creativity.

II. THEORETICAL FRAMEWORK

With the constructivist learning environment, mind mapping-related theories, and computer-supported collaborative learning as its bases, the CSCD learning platform setup in this study can be described in the following section.

A. *The Construction of the Constructivist Learning Environment Model*

Jonassen [23] proposed the constructivist learning environment (CLE) model (Figure 1) with six elements. Based on the design teaching, design purposes, design methods, and other characteristics and from the perspective of the constructivist learning environment design, the relationship between the model's application and design education is proposed as follows:

1) *The Six Components of the Constructivist Learning Environment Model*

a) *Problems/Cases/Projects*

The CLE drills on a problem or issue, which are poorly structured, the answers are uncertain, and they are real-life problems that elicit motivation to train the learner to attempt to resolve them, complete projects, and apply the results to real-life situations.

a) *Relevant Cases*

In the CLE, relevant cases are provided. Through the introduction of cases and the demonstration of model works, novices have the opportunity to observe and learn. Moreover, through case-based reasoning, necessary referential experiences and scaffolding aids are provided for the learner to understand and resolve problems [45].

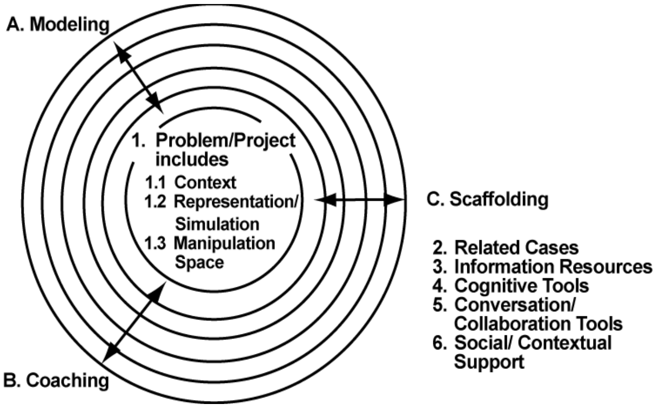


Figure 1. *Constructivist Learning Environment Model (Source: Jonassen, 1999, p.218)*

b) *Information Resources*

The CLE provides appropriate text files, pictures, sound resources, videos, animation, among other support information for building students' mental models and for forming a space for manipulating and thus solving problems.

c) *Cognitive (Knowledge Construction) Tools*

The CLE should provide cognitive (knowledge construction) tools such as databases, mind mapping, expert systems, and hypermedia to help solve problems and aid in extended thinking.

d) *Dialogue and Cooperation Tools*

The CLE should provide a variety of computer-mediated communication tools, such as various types of computer conferences, listservs, email, bulletin board, and NetNews services to support the learner in community cooperation, agreement, and decision-making, thus achieving a common goal (Scardamalia, 1994 cited from [23], p. 228;).

e) *Social or Situational Support*

The CLE's social or situational support includes completing and perfecting hardware, instruments, equipment, and other physical environments in the learning environment. Additionally, it includes the user's familiarity with the system tools of the learning platform and the operating method and the teachers' explanation or clarification of their opinions on students' questions.

2) *The Three Teaching Strategies that Support the Implementation of a CLE*

In a CLE, the teaching strategies should support the goals of the learning activities where appropriate. The implementation of the CLE is supported by three teaching strategies: modeling, guidance, and scaffolding aids.

a) *Modeling*

In a CLE, modeling provided by teachers can be divided into behavioral modeling and cognitive modeling; the former refers to the modeling of how activities are conducted, while the latter clearly expresses to students the need for reasoning during activities.

b) Guidance

Guidance is provided when the learner seeks assistance. The guidance coach analyzes the learner's performance and provides feedback, reflections, suggestions, and skills for adjusting the learner's development, with emphasis on the learner's performance.

c) Scaffolding Aid

Unlike modeling and guidance, the scaffolding aid generally compensates for a learner's lack of prior knowledge. When a problem is encountered while executing a task, the teacher should employ a scaffolding-aided learning approach, such as adjusting the degree of operational difficulty, accommodating the learner's ability, restructuring the task to compensate for the lack of prior knowledge, or providing an alternative evaluation method to complete the task.

The three teaching strategies that support the CLE implementation coincide with the design teaching concept. In the design of teaching activities, the teacher's modeling and individual guidance enhance students' interest in learning through actual operation and the application of theories in practice [15].

B. Mind Mapping

1) The Definition of Mind Mapping

Mind mapping is a thinking-integration skill that is constantly linked to external concepts through an illustrated technique of association based on imagination. This technique transforms a central theme into a concrete one. Mind mapping associates objects through brainstorming to assist thinking and decision-making and to encourage problem solving by innovative means. By changing traditional logical and convergent thinking patterns through the use of lines, colors, text, numbers, symbols, graphics, keywords, and other methods of radiant thinking, the brain can freely radiate and associate, which encourages infinite creativity. Mind mapping not only characterizes the structure of knowledge but also aids in its absorption, compilation, and comprehension while enabling students to fully employ whole-brain creative thinking and problem solving ([18], [21], [36], [37], [60].Mind mapping presents systematic knowledge through visual diagrams and text-based themes. The advantages of mind mapping include visualization, streamlining, integration, and focus, which all help the learner to fully apply brainstorming and free association through radiant thinking and composition methods. In education, mind mapping is not only an effective learning strategy but is also an essential tool for enhancing creativity. The computer-based mind-mapping technology uses a visualized thinking model, combines the Internet and multimedia (hypermedia) and thus allows users to arbitrarily organize and develop systematic knowledge, develop creativity, and express personal style. For these reasons, the technology greatly aids in teaching and learning.

2) The Theoretical Foundation of Mind Mapping

The process of mind mapping is compatible with the constructivist theory, meaningful learning, the radiant thinking model, and theories of whole-brain thinking.

a) The Constructivist Theory

In constructivism, knowledge is constructed, and the significance of the knowledge lies in the learner's integration of new ideas with past knowledge and experience, thus giving the phenomenon meaning and constructing his understanding about knowledge. This process of knowledge building is known as "meaningful learning". Initially, it is easy for the learner to assign meaning to things and phenomena, but with the accumulation of experience, the meanings assigned become more complex [24]. In summary, Constructivism emphasizes the following points: 1) the learning should be learner-centered, and knowledge is voluntarily obtained from the learner's past experiences, not directly from the teacher's instruction; 2) the learner combines new knowledge with old cognitive framework in a non-arbitrary manner; 3) learning cannot be accomplished by memorizing books; instead, the learner should integrate new knowledge with accumulated knowledge and assign it new meanings; 4) knowledge is complex and situational; therefore, the learner's understanding of knowledge should be characterized by complex knowledge [25].

b) Meaningful Learning

Meaningful learning refers to the combination of the learner's new knowledge and old cognitive framework in a nonarbitrary manner. Many scholars [1], [42] have suggested that learning is only effective if it is learner-oriented and if the learner understands the significance of the concepts gained. For learning to be meaningful, the teacher should first understand the learner's existing cognitive framework and then transfer new knowledge based on this cognitive structure, thereby enabling the learner to link the new and old cognitive structures.

c) The Radiant Thinking Model

Research has found that human brain waves do not travel in straight lines. Instead, the brain delivers messages in a radial manner. When a concept appears, the brain may generate a series of related ideas, and the relationship between the ideas and theme concepts is expressed in a radiant pattern. From each idea branched out (in text or images), more ideas can be generated. Mind mapping adopts such radiant thinking to simulate the brain's thinking patterns. Compared with traditional linear thinking, radiant thinking more effectively sparks creative ideas because it is characterized by lateral and vertical thinking, resulting in the free production of ideas, free association, and the generation of infinite creativity [3], [6],[33] , [35].

d) The Whole-Brain Learning Model

In the late 1960s, Professor Roger Sperry, from California, U.S.A., published his research results on the brain cortex, which showed that the left and right cortexes are responsible for different mental functions. The left brain specializes in learning-related functions, including linguistic and logical thinking; the right brain focuses on creative aspects, such as spatial perception, the Gestalt concept, color sensitivity, and range concepts [2], [6],[4], [7], [35], [46], [48]. However, the whole-brain thinking model involves the presentation of visually outlined views, messages turned into images, and memory associations, thus allowing the free display of thoughts and ideas and generating an integrated whole-brain operation, which further enhances the ability to learn and solve problems [6].

Mind mapping helps teachers to design meaningful learning environments. While a learning theory-based teaching environment encourages students to learn actively and enhances learning effectiveness, it also supports students' ability to solve complex problems and complete innovative tasks.

C. Computer Support for Collaborative Design

1) The Meaning of Computer Support for Collaborative Design

The CSCL applied to the field of design is known as the Computer Support for Collaborative Design (CSCD). Kvan [28] defines collaborative design as team members' negotiation, agreement, compromise, satisfying, and completion of stages in the design process and the achievement of goals that cannot be achieved individually. Kalay [26] suggested that a good design must be continuous and must integrate different specialties and shared knowledge. Lee [31] also found that collaborative design refers to the synchronized interaction, communication, and discussion of team members during the product development process. Moreover, to implement a design mission, the communication and collaborative tools must meet the needs of all the team members. CSCD means designing activities that support and strengthen teamwork through computer aids and support systems so that such activities can be integrated into computer technology-related applications, incorporate professional knowledge from different fields, share interdisciplinary expertise, and achieve perfect design results [47]. In practice, the designer must conduct complex communication and coordination among proprietors, material providers, and even partners and competitors in large projects [16]. Cross and Clayburn [14] also pointed out that in the design process, most problems are complex and difficult to solve, and groups are better able than individuals to acquire more information, which aids in the decision-making process. Teamwork with specialized personnel from other fields helps designers to complete projects. Hence, teamwork is essential in design [17], [64].

2) The CSCD Applications

Collaborative design has numerous benefits. Empirical studies have found that by incorporating the CSCL design in the planning of courses, the interactions among the learners via the computer platform could be enhanced, which contributed to learning effectiveness. The cross-functional team approach helps the designer to resolve design-related problems, shorten the time for product design and development, and reduce costs [9], [12], [54]. Teamwork is emphasized in practical design projects. The introduction of the CSCL in the design field and its further development into the CSCD enables team members to interact more frequently through collaborative discussions, helps to solve problems and enhances learning motivation and results. More importantly, with this kind of interaction, team members will elicit more creativity in one another. Thus, compared with other fields, the importance and value of the CSCD for design are critical. Tang, Lin and Chen [50] also found that concepts produced during the collaborative process can be applied to creative performance, and different knowledge and design experiences result in clearer and more complete design concepts, thereby producing better results. The division of tasks through teamwork can integrate the members' different skills and knowledge to enhance the diversity and integrity of design [43], [51], [57]. In particular, with the progress and popularization of network technologies, traditional face-to-face and synchronous forms of communication, in view of their time and cost considerations, have gradually been supplanted by other forms of interaction and collaborative learning tools. These changes, while enabling students to engage in networked learning activities through constructive and collaborative learning methods, have reduced costs and enhanced communication and work efficiencies [13], [19], [23], [29].

The integration of network technology and teamwork has led to major changes in the design environment. The team members create interactive relationships through collaborative discussion, which helps to resolve problems, elicit creativity, and enhance learning motivation and results. Moreover, digital media have overcome the time and space constraints of traditional media, thus substantially improving production and operation techniques in the design industry. In other words, the support of computer technologies provides learners with useful resources and channels for collaborative learning so that learning can occur at any time and in any place. In addition, teachers' effectiveness can also be enhanced. The importance of design education is therefore clear.

D. The Characteristics of Design Students

Unlike other types of learners, design students possess characteristics such as graphics ability, creativity, and learning style. As Chen and You [12] stated, design departments greatly emphasize communication, discussion, and cooperation.

1) The Personality Traits of Design Students

Through in-depth interviews, Lin [34] found that design students possess the following characteristics: (1) they are more confident and demand perfection more than other students; (2) their interests in life are often related to what they have learned, such as computer graphics for design and painting, indicating their ability to integrate learning into life; (3) information gathering and sketching are the most common learning activities for finding inspiration and creative expression; (4) to enhance their professional capabilities, the students need to continually absorb information; (5) accumulated learning experiences provide design students with superior creative thinking, orderly analysis, and precise hand-eye coordination; (6) their preferred courses have stimulating designs, rather than recitation or theory-based instructional designs, to allow the expression of creativity.

2) The Graphics Ability of Design Students

Tseng [53] showed that compared with students from other departments, design students possess better creative thinking ability, specifically in dispersed graphics. Lee [30] also found that approximately 75% of college design students have a greater need for graphics data than for text data. Moreover, graphics, audio-visuals, and multimedia are important sources of inspiration and creativity for design students. In this study, computer-based mind mapping elicited students' creativity through constructive and organizational ideas and thus enhanced learning effectiveness. Wavering conducted research on high school students' logical and thinking abilities required for constructing linear diagrams. The results showed that the learners' operational and cognitive abilities were indeed associated with the learner's graphics capability.

3) The Creativity of Design Students

Design students possess not only an aesthetic sense but also the powers of observation, imagination, and creativity to develop design experiences and resolve design issues from different perspectives. Design students consider the ability to think creatively as most important, followed by the ability to engage in technical work [52]. Methods for cultivating design students' creativity should begin by improving teaching activities and learning environments. To foster students' creativity, teachers should design and implement innovative lessons based on students' characteristics [34].

Design students have greater sensitivity to images and colors, are skilled at visual thinking, know how to flexibly apply their learning, and are fond of dynamic learning methods and knowledge that can be applied to daily life. Hence, graphics and the teaching tools of text, voice, video, and multimedia should be applied to learning environment designs and teaching materials where appropriate. Additionally, open-ended creative activities should be designed, and learning aids and tools that meet the students' needs should be provided to enable design students to display their strengths, build their own knowledge structures, enhance their learning motivation, and develop good attitudes, thereby promoting academic achievement.

III. REASERCH METHOD OF THE PRIMARY STUDY

A. The Research Framework

This study is based on the ADDIE teaching design model, which consists of the following five processes: analysis, design, development, implementation, and evaluation underwent platform construction. First, the design teachers' and students' needs for the platform functions and learning tools on the platforms were analyzed; then, a CSCD learning platform model was designed. Next, feasible analysis was conducted, and the course content was planned to develop the CSCD platform. After integrating the information engineers' techniques, the CSCD platform was constructed. Finally, through expert opinions and user evaluations, the platform was corrected.

B. The Research Procedures

1) The Demand Phase

Based on previous studies analyzing theories and literatures related to web-based instructional platforms and domestic and foreign learning platform designs, this study constructed the functional design and integration framework of the web-based instructional platform. Then, the AHP hierarchical analysis was adopted to produce a questionnaire intended to reveal the design teachers' and students' (the main platform users) views on the platform functions and learning tools.

2) The Design Phase

In this study, the design model of the CSCD platform was established according to the users' needs. The CSCD model covers four functions: course information, teaching content, learning tools, and assistance (Q&A and help). The model also includes seven learning tools: mind mapping, work display, electronic whiteboard, course discussion, virtual classroom, audio-video media, and relative links.

3) The Develpoment Phase

To understand the feasibility of the CSCD platform model setup, five engineers from information technology companies and school e-learning centres were interviewed in this study. Analysis and evaluation were conducted from the perspectives of costs, technical capacity, human resources, organizational strategies, possible risks, and prior setup-related experiences.

4) The Setup Phase

a) The Construction of the Platform Prototype

The results of the feasibility evaluation showed the possibility of the construction. Thus, numerous information engineers' techniques were integrated to construct the platform prototype.

b) The Incorporation of Course Materials

After constructing the platform prototype, the course materials were incorporated. With the "graduation topic" course as the example in this study, relevant teaching materials and references were provided by teachers with related experience and students who had taken the course.

c) Expert Validity and User Testing

After completing the CSCD platform prototype, the expert and user testings were conducted, and corrections were made based on the comments.

5) The Evaluation Phase

After correcting the CSCD platform, the experts and users then engaged in formative evaluation of the usability and overall design results of the platform.

C. The Research Method

1) In-Depth Interviews

After constructing the study's previous design and development phase and completing the construction and evaluation phases, the opinions of engineers' and school network managers' were collected through interviews.

2) The Website Framework of the CSCD Platform

The completed CSCD learning platform website framework is shown in Figure 2.

IV. FINDINGS AND DISCUSSION

A. Findings

The CSCD platform functions are as follows:

1) The Homepage

After entering the platform, users register a new account, create a password on the right-hand side of the website and log in. The main page has nine course buttons. By clicking one of them, users are directed to the course profile page. The CSCD homepage is shown in Figure 3.

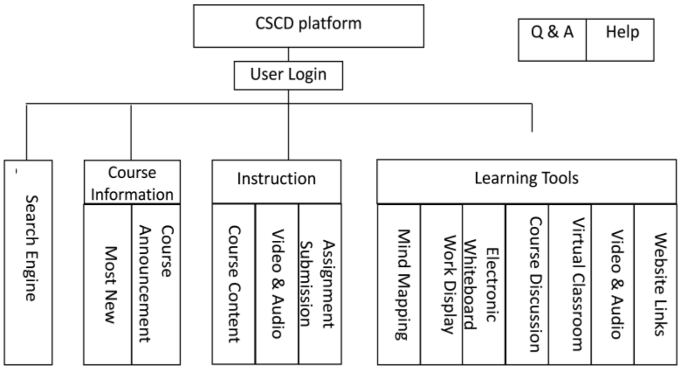


Figure 2. *Website Framework of the CSCD Learning Platform*

2) The Drop-Down Menu in the Top Middle Part of the Webpage

a) The Course information

The course information is divided into two parts: 1) News Updates. Users click this to enter the CSCD course discussion page. They should first read the instruction articles on the course discussion board to access the board; 2) Course Announcements. Users click this to enter and read the CSCD course-related news updates.

b) The Teaching Content

The items under this menu are managed by the course managers or teacher users and include course contents, audio-visual media, and assignment submissions. The maximum upload for assignment submissions is 20M per file, and the limit is subject to change as needed.

The capacity varies depending on the hard disc space available. The file extension formats of uploaded files are .zip, .rar, .pdf, .txt, .jpg, .tif, and .bmp, which are subject to change as needed, except ndex.htm, index.html, and index.php.

c) The Student's Work

This function in the menu classifies the students' work by topic. Users click on a work album to enter the browser screen and enjoy the various creative design projects.

3) The Left Side of the Website

a) The Search Engine

Users click the keyword to search pictures and data needed on the CSCD platform.

b) The Display of Students' Work

The students' work is automatically displayed randomly.

c) The Counter

The number of visitors to the CSCD platform is recorded.

4) The Function Menu on the Right-Hand Side

a) The Course Menu

The menu displays all nine courses provided on the CSCD platform: introduction to design, text style, digital editing and special effects, design creativity, composition, advertising design, photography, special topics design, and work collections.

b) Q & A

This section introduces the menu functions and describes the various links while accessing the CSCD platform. Descriptions and solutions are also provided for various common operating problems.

c) Help

The help menu includes the CSCD course discussion boards, course announcements, and bulletin boards. New course information is constantly updated, and space is provided for users to engage in exchanges and discussions.

d) The Website Homepage

Users click this option to be directed to the homepage screen of the CSCD platform.

5) The Learning Tool Area on the Right Side

According to the needs of users, the CSCD platform provides learning tools such as mind mapping, work display, electronic whiteboard, course discussion, virtual classroom, audio-visual media, and links to websites. The functions of the various learning tools are as follows:

a) Mind Mapping

Students sketch mind mapping through the use of the program X-Mind and upload the sketches to the platform. The rotate, zoom-in and zoom-out, or full-screen functions can be used to view mind mapping. The mind-mapping display window is shown in Figure 4.

b) The Work Display

This platform provides space for displaying work, thereby allowing students to learn from each other. The platform is presented in images to better suit design students' characteristics and preferences. The students can preview works via thumbnail. Clicking on a picture will enlarge it for a better view. The students' work display window is shown in Figure 5.

c) The Electronic Whiteboard

Once the teacher removes the authorization limit, the students can freely sketch on the electronic whiteboard to explain their abstract concepts, which are considered very conducive to interaction. The teacher's explanations can also be viewed on the whiteboard, which produces excellent interaction results.

d) Course Discussions

This discussion board provides students with a forum for discussing their work and exchanging information outside of classes. In particular, because most design projects require teamwork, this interaction mechanism is critical.

e) The Virtual Classroom

Network learning environments with video conferencing allow students to feel as though they are in an actual classroom. Students should bring their own video cameras and recording devices. First-time users should first install the related equipment and programs. The Co-Life system allows 2-29 webcams to operate simultaneously (depending on the bandwidth and the devices being used).

f) Audiovisual Media

Some audio and video courses, digital video editing or audio effects production courses, come with multimedia teaching materials and related resources. Course content presented via audiovisual media is more stimulating, and students can advance and rewind as they like, which enhances their learning interest and leads to more effective learning outcomes.

g) Links to Websites

Various multimedia information is provided, including photo gallery, illustration gallery, design gallery, audio and video gallery, and unclassified. Clicking the class directs the user to the corresponding website. These resources can better inspire designers.

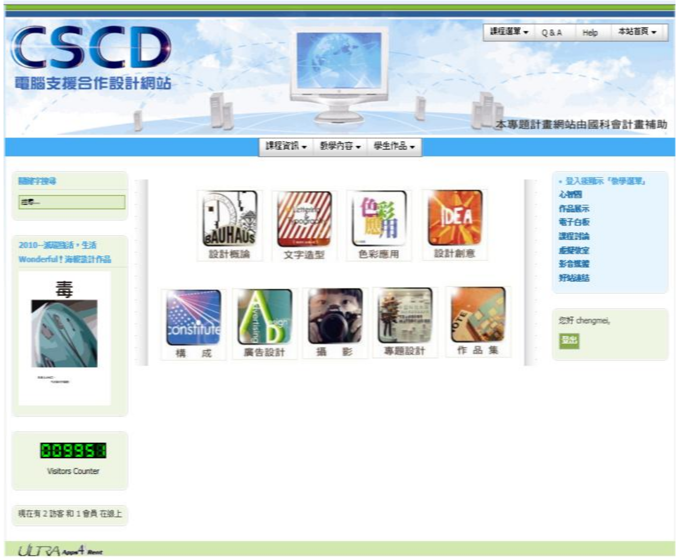


Figure 3. *The CSCD Homepage*

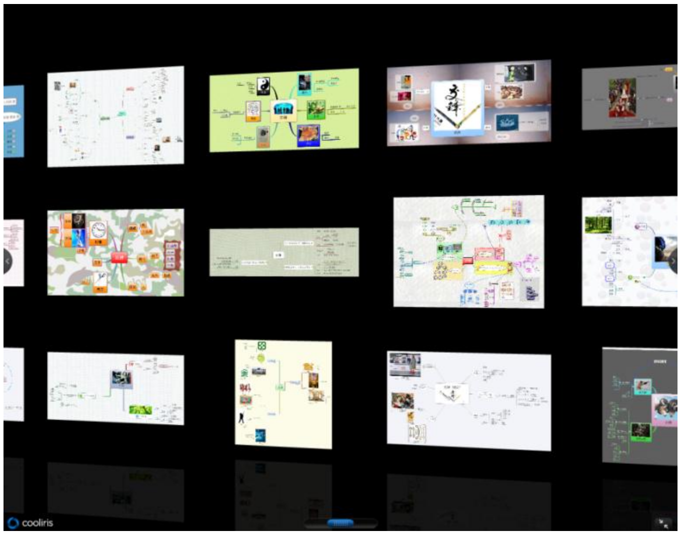


Figure 4. *The Mind Mapping Display Window*

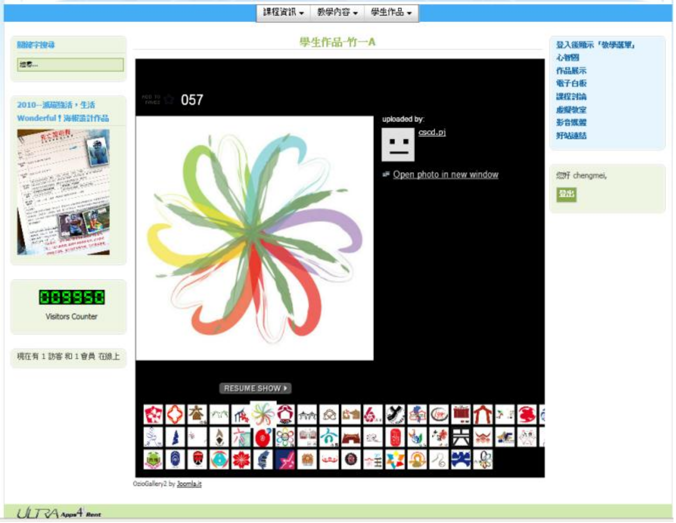


Figure 5. *The Students' Work Display Window*

B. Expert Opinions

Interviews were conducted with the information engineers before and after setting up the CSCD platform. Their opinions of the platform setup are summarized as follows:

1) Related Supplementary Measures for Maintaining the CSCD Operations

Teaching and learning must be based on sound network environments and course-related information. Hence, the operation of the learning platform relies on the coordination of three aspects to smoothly carry out the platform functions:

(1) hardware resources; because the host is usually open 24 hours a day, it is normally deployed in an air-conditioned network-controlled room to be uniformly configured and managed by the school; (2) the management manpower on the back-end learning platform is the monitoring and management unit that supports network learning. The unit's role is to process quickly to ensure the normal operation of network learning and to assist platform users in delivering system and course announcements and consultations; (3) teaching assistant; the main task of this assistant is to support the teacher with instructional operations and to communicate with students regarding the various website functions. In the coordination and operation of these three aspects, the teaching assistant is more likely to support the course operation through searches and training, while the procurement of the hardware host (including the placement of the equipment, the power supply, the network cable configuration, the operating system installation, and related software program settings and modifications) and the management of human resources require further budget and other considerations.

2) The Estimation of the CSCD Platform Setup Costs

The platform setup costs primarily include the software and hardware equipment and the staff.

The software and hardware equipment costs include the operating system and database software authorization fees and the budget for the software program of the learning platform management system. For the construction of the CSCD course functions and related technologies, a free content management system, the CMS Joomla website construction program, was used. Additionally, resources from the free database MySQL were adopted. To allow open source organizations to engage in development and support, Joomla applied various websites and progressive new technologies to enhance the responsiveness and performance of the websites. Furthermore, thousands of different website applications from around the globe, such as add-ons, graphic design and scenery developments, rapid deployment resources, and free and powerful extensions are performance strengths of the platform.

3) The Assessment of the Overall Technical Capacity of the CSCD Platform

First, with regard to the platform technology for constructing the CSCD platform, the operating system version of the web-based instructional environment as a whole should first be determined. Second, the various network and course function programs needed to set up the website are installed. Third, the various parameter settings are adjusted and modified. Fourth, art templates that match the visual style of the course program are edited and drawn. To maintain the normal operation of the platform, the host status should be regularly checked and maintained, and the course backup data should be downloaded.

4) The Construction of the Human Resource Assessment of the CSCD Platform

In addition to the school network managers, a teaching assistant is required to help students solve problems related to webpage browsing or platform use. The assistant also download assignments, edits course announcements and notices, contacts students, responds to the various discussions, explains the teacher-student contacts, and sends user comments, system problems and status updates to the system manager. These duties make the teaching assistant an important communication bridge between the system and the class.

5) The Assessment of the Organizational Strategies on the CSCD Platform Setup

The contribution level of the CSCD can be analyzed in two categories:

(1) platform usage, user satisfaction, and the ease of the platform operation; can the platform compensate for the inadequacies of classroom teaching? If the conducting of classes, teacher-student interactions, assignment assessments, and other teaching items can be fully applied to ensure CSCD performance, and the system can also review and modify the course implementation where appropriate, the CSCD setup will serve as a good model for similar implementations in the future, and the e-learning effectiveness of schools will also be enhanced; (2) analysis of the operating costs shows that the CSCD platform was constructed on the premises of low cost and high effectiveness, which will encourage the introduction of more courses and gradually increase the use of such courses. For example, there can be more synchronized online courses, more courses with simultaneous online access, and more online course activities can be used to observe the development of teaching.

6) The Assessment of Risks in the CSCD Platform Setup

In terms of a school's implementation of e-learning, the risks involved with this site are relatively low because of the minimal equipment and low system setup costs as well as the high expandability of the course functions. To maintain the smooth operation of the platform, the network technologies and related system developments should include periodic testing and updates.

7) The Assessment of the Innovative Technologies of the Learning Tools on the CSCD Platform

a) The Setup

The seven main learning tools of the web-based instructional platform (CSCD platform) include mind mapping, work display, electronic whiteboard, course discussion, virtual classroom, audio-visual media, and links to websites.

b) The Mind-Mapping Tools: Ozio Gallery 2: Cooliris Components

The course options were originally designed for connecting to the Xmind website for online functional uses, but due to many online user authorizations and nonsynchronized teaching considerations, the options were changed to the Joomla components Ozio Gallery: Ozio Cooliris skin menu functions, and the students' mind-mapping operating files were used as a display. The mind-mapping work samples are displayed through the component functions. Clicks visually emphasize the dynamic results, or they can be watched on the full screen.

c) The Work Display Tools: Flickr SliderShow Components

The work display function is one of the website's most characteristic features. In terms of the website's course design attributes, related assignments and work displays are provided. Hence, in consideration of the website's file upload management and performance, the Flickr SliderShow components were adopted. With these components, including the storage spaces and application features provided on the Flickr website, intervals can be placed seconds apart, and file displays can be selected.

The visual output is stimulating and quick, while the click responses are sound. In addition to the static picture files in the website's design category, there are also mp3 files of multimedia sources and e-books that undergo audiovisual editing. Through the component functions, these options are directly available for listening on the website.

d) The Electronic Whiteboard Tools: the Website Co-Life System

The electronic whiteboard connects to the Co-Life audiovisual conferencing system. The system was developed by the National Centre for High-Performance Computing (NCHC) and can synchronize with the electronic whiteboard, the desktop, text, and other instant messages, and it can provide a remote common processing platform during the project plan development period. The system is an essential tool for a virtual team.

e) The Course Discussion Board Tools: the Component Version Kunena Forum

The application of the discussion board comes in multiple usage operations, examples of which include the latest news, course announcements, and course information that uses the discussion board functions. With the privacy protection function, some discussion boards authorize only the teachers and students to use them by logging in, at which point they can participate in presentations and interactive discussions.

f) The Virtual Classroom Tools: the Website Co-Life System

The virtual classroom combines the Co-Life online conferencing system functions of the NCHC. Users should bring their own camera video communication devices, and first-time users must install the related equipment and programs. The Co-Life online conferencing system will also automatically detect and install the relevant software needed for the environment. The current computer automated functions, coupled with devices such as the USB, have been successfully completed for installation and use. The microphone audio input devices, however, can be divided into the sound card and the webcam. Users can decide which sound input to adopt for voice recording and play. In principle, with the academic network transmission conditions and the ADSL environment, the Co-Life online conferencing system is used for synchronized teaching in online conferencing, thereby achieving smooth two-way audiovisual communication quality. However, for users with lower network bandwidths, systems of this type may result in connection problems or poor video quality.

g) The Audiovisual Media Tools

The audiovisual media tools adopt the following Joomla expansion component: the Expose Flash Gallery 4.6.3 Alpha3c module, which can be classified or edited and is a powerful function for photographs and audiovisual displays.

h) The Tools for Links to Websites

The links to websites feature adopts the built-in components of the Joomla system: the version of the webpage link and classification lists will show all the lists for the webpage link classifications, which are managed by the background links for adding and editing.

8) The Superiority of the CSCD Platform

This website adopts the world-renowned Content Management System (CMS) Joomla website program setup, which features three advantages: (1) among the Joomla system expansion components, the multimedia graphics and animation performance are outstanding. These components not only display smooth and beautiful foreground results but also allow detailed adjustments of the background management functions, such as file uploading and parameter setting. When constructing and setting up website programs for various course tasks, the foreground and background performance and the operation of management-related configurations should be taken into consideration because these functions depend on the strengths and weaknesses of the website service program design, adjustments and optimized modifications of the relevant parameters to enable successful links to websites and normal launches. When designers consider setting up the Joomla version of the Content Management System, they should also consider numerous relevant support component functions, the technical community usage, and other related issues, as well as whether these functions meet the website requirements. From installing the host operating system to setting up the host system database and related components for convenient operating performance, the advantages of the CMS system for the course website are clear; (2) Joomla applies a wide range of new technologies that contribute to website progress, such as website caching technology, and can enhance responsiveness and performance of the website. Thousands of web-based add-ons and graphic design sets from around the world have been developed with this system as the basis. The expansion components with the features of quick setup, free resources, and powerful functions contribute to the website's outstanding performance; (3) when updates for the various essential component versions are available, the component version-related information can be quickly understood, and component function updates can be considered and selected.

The website combines various web-based technology and teaching resources, such as placing a large number of photographs and student image files on Flickr to save the website storage space and upload time. Additionally, the CMS system management function provides rapid and convenient browsing, and through classification links and overviews, the various themes are distinguished and presented. Furthermore, online submissions of work files, class announcements, and problem discussions can be directly conducted on this website. All the teaching resources and functions available on this website can be accessed only when teachers and students log in through their accounts and passwords, except for special authorizations for announcements and related course information. This feature thus protects information and allows convenient use. In addition, the recommended links on this website are mostly cited from the CC information network created by the Ministry of Education. Unless otherwise specified, the website contents are all labeled in CC names to promote the CC creation concept of network resource sharing.

V. Recommendations

A. The Promotion of the CSCD Learning Platform

The CSCD learning platform was completed based on the findings of this study, and the teaching theories include many network learning environment components that meet teaching requirements, thus offering a new opportunity for networking design teaching. In recent years, the Ministry of Education has actively promoted high-quality e-learning content applications [38], [39], [40], [41] to improve the breadth and depth of information network education in our country, thus providing a favorable social environment for implementing a "network design academy." Moreover, with the richer and more mature technologies and experiences of local companies' web-based instructional platform setups, sound external conditions for these setups have been provided. With the support of the above environments and conditions, the implementation of the "network design academy" model should be feasible. In this study, the environment of design network learning serves as the starting point for promoting the web-based design instruction, and it is hoped that this model will receive valuable feedback and encourage other professionals to join the R&D, thereby achieving the goal of implementing the web-based design instruction.

B. Making Good Use of the Network Technology and Multimedia Features in Teaching Design

To strengthen the learners' intrinsic cognitive processes and enhance learning results, the design teaching network should make good use of the flexibility of the technology and the multimedia features when engaging teaching material content and learning activity designs. For example, problem-solving strategies and online questions are techniques used to continually test learning effectiveness and produce the power of urge, and these techniques provide onsite video recording of visitation and speech activities, which enhance the students' learning motivation and understanding of the content. Relevant videos and case files may also be provided for students to download, thus achieving the goal of active learning. In addition, the convenience of the network interactions also facilitates collaborative learning and online discussion, thereby achieving the goals of the CSCD.

C. Establishing Responsible Units and Training Professionals to Develop Web-Based Design Instruction

Currently, network learning programs are mostly promoted by school units (such as computing centers) or teaching development centers funded through program subsidies (such as the Teaching Excellence Program). Due to restrictions on funding and subsidies, problems related to network learning platforms cannot be effectively improved, and better platforms cannot be researched and developed. Therefore, it is suggested that schools establish responsible units to promote web-based design instruction, including the design of web-based teaching materials, the arrangement of web-based instruction activities, the design of network learning tools and interfaces, and other teaching resources with the assistance of professionals and administrative units (such as computing centers) in schools to construct high-quality web-based instructional environments.

D. Promoting Mutual Growth and Complementarities between Academia and Industry through Industry-University Cooperation

Numerous domestic universities have developed their own teaching materials and web-based learning platforms, but few have successfully applied them to industry settings. Because design departments have special requirements for learning tools and hardware for the web-based instructional platform, the software, hardware and R&D expenditures of design departments are much higher than those of other departments. Hence, for future platform setups, mutual cooperation between academia and industry should be adopted to obtain funding assistance, fully engage the different professions, construct an attractive platform with substantial teaching results, and meet users' needs. In addition, it is suggested that academia/university cooperation be adopted in the future to share research results and relevant innovative applications and concepts with the industry, thereby continuing to develop competitive products to be marketed.

VI. Implications

A. The Planning of the Overall Learning Environment

This study started from the learning environment; through the web-based instructional platform, students' network learning motivation and performance can be enhanced. However, even with an appropriate web-based instructional platform that meets the requirements, there is no guarantee that the expected and ideal goal will be accomplished. Planning of the overall learning environment, including the coordination of teachers, course content designs, teaching activity designs, and comprehensive learning assessment methods that complement one another are required to achieve maximum network learning performance.

B. The Selection of Appropriate Courses to Implement Network Teaching

Although this study advocates design teaching and the implementation of networking, it is emphasized that network learning is not applicable to all design courses. For instance, traditional technical courses, including physical operation and other diverse contents, cannot and inappropriately be completely replaced by the web-based instruction. However, web-based instruction will have value for certain courses, such as digital technology classes (e.g., webpage production) and creative thinking classes (e.g., design creativity and creativity development) if the courses can be adapted to implement network teaching or serve as a learning aid with the help of advanced technology, network convenience, and effective learning tools. In other words, in design education, appropriate courses can be selected to implement network teaching and maximize the teaching value.

C. Expanding the Platform Functions and Tool Applications within the Scope of Design Teaching

Hindered by current web-based learning platforms that provide only general features, the network ranges and function-related applications of design teachers are focused mostly on discussion boards, data transmission, and assignment submissions, while other functions are rarely used. The CSCD platform constructed in this study provides mind mapping, work display, electronic whiteboards, and other learning tools, which are very helpful for web-based design instruction. Therefore, the researchers hope that the relevant design education units understand the value and uses of design teaching, that platform functions will be expanded, and that the relevant tools will be applied to the scope of design teaching to enhance the performance of assisted teaching.

Acknowledgment

This article adopted part of the research results of the 2009 National Science Council Research Projects (Project Number: NSC 99-2410-H-163 -004). We sincerely appreciate the funding of the government.

References

[1] Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A cognitive view* (2nd ed.). New York: John Wiley & Sons.

[2] Buzan, T. (2000). *The mind map book*. London: BBC.

[3] Buzan, T. (2005). Mind maps at work: How to be the best at your job and still have time to play. New York: Plume.

[4] Buzan, T. (2007). *Buzan center*. Retrieved December 7, 2009, from http://mind-map.com

[5] Buzan, T. (2007). *The speed reading book*. (Mickey, S., & Chen S.-Y. Trans.). Taipei: Yale International Cultural. (Original work published 2003).

[6] Buzan, T., & Buzan, B. (2007). *The Mind Map Book.* (Mickey ,S. Trans.). Taipei: Yale International Culture Publishing Co., Ltd. (Original work published 1993).

[7] Colin, R. (2004). *Accelerated learning for the 21th century.* (Pao-Lo Tai, Trans.). Taipei: Classic Communications Co. (Original work published 1997).

[8] Chang, M. -M. (2001). *A study on the business strategy and the CFs of e-learning Web sites*. Unpublished master thesis, National Changhua University of Education Institute of Business Education, Changhua, Taiwan.

[9] Chen, I. -W. (2004). *The influence of designers' cognition styles on their team communication and problem-solving.* Unpublished master thesis, Tatung University Graduate School of Industrial Design, Taipei, Taiwan.

[10] Chen, W. -Z., & You, M. -L. (2001a). Internet mediated design course: The pilot study on IMDC model. *The Journal of Design Research, 2*, 109-115. Retrieved December 14, 2011, from http://thinkdesign.cgu.edu.tw/File_uploads/wenzhi/2001_CID6.pdf

[11] Chen, W. -Z., & You, M. -L. (2001b). On application of internet in design education. *Industrial Design, 29*(2), 105-110. Retrieved December 14, 2011, from http://thinkdesign.cgu.edu.tw/File_uploads/wenzhi/2001_MIT.pdf

[12] Chen, W. -Z., & You, M. -L. (2001c). *On application of internet in design education.* Paper presented at the meeting of the 16th National Technical and Vocational Education Seminar, Hualien, Taiwan. Retrieved June 3, 2007, from http://thinkdesign.cgu.edu.tw/File_uploads/wenzhi/2001_TVE_16.pdf

[13] Chen, Y. -C., & Huang, H. -H. (2001). An analysis of the interactive process of network collaborative learning and the role of the learner - the "teaching system design" e-course case study. Paper presented at the meeting of 2001 The International Seminar on Information Literacy and Lifelong Learning, Taichung, Taiwan.

[14] Cross, N., & Clayburn-Cross, A. (1996). Observations of teamwork and social processes in design. In Cross N., Christiaans H. & Dorst K. (Eds.), *Analyzing design activity* (pp. 291–317). New York: John Wiley and Sons.

[15] Curriculum Research and Development Center for Commerce Education (2001). Curriculum Final Report/ Final report on the design field. Retrieved December 22, 2011, from http://bcc.yuntech.edu.tw/TVEC/一貫課程期末報告/設計群期末報告.pdf

[16] Department of Interior Design, Tainan university of technology (2010). *Future prospects of the Department of Interior Design*. Retrieved December 7, 2011, from http://203.68.182.102/idtut2/index.php?option=com_content&task=view&id=17&Itemid=32

[17] Dorta, T., Lesage, A., & Pérez, E. (2008). Point and Sketch: Collaboration in the hybrid ideation space, in C Bastien & N Carbonell (eds), IHM 2008, ACM Interaction humain machine: 20 ans d'interaction homme-machine francophone: De l'interaction à la fusion entre l'humain et la technologie, Association for Computing Machinery (ACM), Metz, pp. 129-136.

[18] Features-Xmind (2010). *Mind mapping and storming.* Retrieved December 7, 2010, from http://www.xmind.net/pro/features

[19] Geer, I. W., Brey, J. A., Weinbeck, R.S., Moran, J. M., Ficek, M. M., Hopkins, E. J., & Blair, B. A. (2000). Online weather studies: An introductory college level distance-learning course. *American Meteorological Society, 9th Symposium on Education*, AMS, Long Beach, CA, 144-147.

[20] Haymaker, J., Keel, P., Ackermann, E., & Porter, W. (2000). Filter mediated design: Generating coherence in collaborative design, *Design Studies,* 21(2), 205-220.

[21] Hobin, E., & Anderson, A. (2008). Middle-school students' concepts of health in Ontario and the British Virgin Islands and the implications for school health education. *Physical & Health Education Journal,* 74(2), 16-22.

[22] Hsu, C. -M., & Yen, J. (2006). The perspective analysis of the design teachers to the digitalized design courses –An application of WinMAX. *The Journal of Design Research*, 6, 189-199.

[23] Jonassen, D. H. (1999). Designing constructivist leaning environment. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (pp. 215-239). Mahwah, NJ: Lawrence Erlbaum Associates.

[24] Jonassen, D. H. (2000a). *Jonassen's class notes*, November 15, 2000.

[25] Jonassen, D. H. (Ed.) (2000b). *Computer as mindtools for schools: Engaging critical thinking.* Upper Saddle River, NJ: Prentice-Hall.

[26] Kalay, Y. E. (1999) The future of CAAD: From computer-aided design to computer-aided collaboration. In Augenbore, G. & Eastman, C. (Ed.), Proceedings of the CAAD Futures' 99, Kluwer Academic, 13 -30.

[27] Karen Bromley, Linda Irwin-DeVitis, & Marcia Modlo (2005). *Graphic organizers.* (Hsin-Jung Lee, Trans.). Taipei: Yuan-Liou. (Original work published 1995).

[28] Kvan, T. (2000). Collaborative design: what is it? *Automation in Construction*, 9, 409-415.

[29] Kwon, E. S. (2004). A new constructivist learning theory for web-based design learning with its implementation and interpretation for design education. Unpublished doctoral dissertation, Ohio State University, Ohio.

[30] Lee, E. (2000). *The study of information-seeking behavior: A case of the students of the design college of shih chien university.* Unpublished master thesis, Tamkang University, Institute of Information and Library, Taipei, Taiwan.

[31] Lee, L. -C. (2004). The concept and characteristics of collaborative design (I). Retrieved December 7, 2011, from http://webcache.googleusercontent.com/search?q=cache:ZpRO156EjaEJ:www.boco.com.tw/newsdetail.aspx%3Fbid%3DB20070117002924+%E6%9D%8E%E4%BE%86%E6%98%A5+2004+%E5%8D%94%E5%90%8C%E8%A8%AD%E8%A8%88+collaborative+design&cd=9&hl=zh-TW&ct=clnk&gl=tw

[32] Lee, H. -J. (2009). *Teaching and learning for excellence-introduction and plan.* Retrieved December 8, 2010, from http://www.tec.tcu.edu.tw/overview.aspx.

[33] Lee, I. -L. (2009). The impact of the mind mapping teaching plan on the visual art learning of sixth grade elementary school students - exemplified by Chi Wen elementary school in Miao-Li county. Unpublished master thesis, National Hsinchu University of Education Fine art and crafts education, Hsinchu, Taiwan.

[34] Lin, R. -T. (2000). *A study on the learning attitude of students of design department.* Unpublished master thesis, National Taiwan University of Science and Technology, Graduate Institute of Applied Science and Technology, Taipei, Taiwan.

[35] Liu, K. -W., & Hsieh, C. -C. (2009). The study of the effects of mind mapping program on the fifth-grade students' linguistic creativity in elementary school. *STUT Journal of Humanities and Social Sciences,* 1, 75-106.

[36] MindMapper Operation (2010). *Mindmapper basics plus important features.* Retrieved December 7, 2010, from http://www.mindmapperusa.com/features.htm

[37] Mind Mapping Software Helps Personal Development (2009). *Computers and technology: Software.* Retrieved December 7, 2010, from http://mind-map.com

[38] Ministry of Education (1999a). *The 4th annual first-session report of the Education Committee, Legislative Yuan.* Retrieved September 16, 2008, from http://www.edu.tw/content.aspx?site_content_sn=546

[39] Ministry of Education (1999b). *The 4th annual second-session report of the Education Committee, Legislative Yuan.* Retrieved September 16, 2008, from http://www.edu.tw/content.aspx?site_content_sn=580

[40] Ministry of Education (2000). *The 4th annual third-session report of the Education Committee, Legislative Yuan.* Retrieved September 16, 2008, from http://www.edu.tw/content.aspx?site_content_sn=627

[41] Ministry of Education (2005). *The action plan for education policy implementation focus.* Retrieved September 18, 2008, from http://www.edu.tw/content.aspx?site_content_sn=13535

[42] Novak, J. D., Gowin, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge via mapping with junior high school science students. *Science Education,* 67, 625-645.

[43] Paulus, P. B. (2000). Groups, teams, and creativity: The creative potential of idea-generating groups. *applied psychology: An International Review*, 49, 237-262.

[44] Phelan, A. (2006). Studio Art Education Today: the impact of Digital Media and Technology on the Pedagogical Structure. *The International Journal of Arts Education*, 4(1), 9-22.

[45] Schank, R. C., Berman, T. R., & Macpherson, K. A. (1999). Learning by doing. In C. M. Reigeluth (Ed.), *Instructional design theories and models.* Mahwah, NJ: Erlbaum, 161-181.

[46] Scott, D. M. (2005). Thinking right about content. *EContent,* 28(6), 48.

[47] Shen, W., Hao, Q., & Li, W. (2008). Computer supported collaborative design: Retrospective and perspective. *Computers in Industry,* 59, 855-862.

[48] Sousa, D. A. (2003). *How the gifted brain learns*. Thousand Oaks, CA: Corwin.

[49] Sun, M. (2003). *Mind map thinking method*. Taipei: Radiant Mind Co., Ltd.

[50] Tang, H. -H., Lin, C. -W., & Chen, W. -C. (2009). Exploring the relationship between personal creativity, creativity process, Concept Evolution, and Design Performance in a Collaborative Design Process. *Journal of Design,* 14(3), 51-71.

[51] Tidafi, T., & Dorta, T. (Eds) (2009).*Design tools and collaborative ideation.* Tomás Dorta, Annemarie Lesage, Edgar Pérez Hybridlab, School Of Industrial Design, Université De Montréal, Canada. Retrieved December 7, 2010, from http://www.hybridlab.umontreal.ca/documents/23-CAADFutures.pdf

[52] Tseng, C. -F. (1987). A study of the development of the design learning. *Design Education Proceeding,* 130, 121-138.

[53] Tseng, Y. (2002). *Divergent thinking detection and analysis for students of industrial design*. Unpublished master thesis, National Chiao Tung University, Hsinchu, Taiwan

[54] Tung, Y. -J. (2005). The influence of communication modes among cross-functional research and development teams regarding design knowledge integration. Unpublished master thesis, Tatung University Graduate School of Industrial Design, Taipei, Taiwan.

[55] University Campus Suffolk (2010). *BA (Hons) design*. Retrieved December 7, 2010, from http://www.ucs.ac.uk/study/SchoolsAndCentres/Lowestoft/graphicsspec.aspx

[56] Wang, C. -W. (2003). *Factors influencing teachers to involve in using web-based teaching platforms*. Unpublished master thesis, Shu Te university institute of information management, Kaohsiung, Taiwan.

[57] Warr, A., & O'Neill, E. (2005). Understanding design as a social creative process, Proceedings of the 5th Conference on Creativity & Cognition (pp. 629-642). London: ACM.

[58] Wavering, M. (1985). The logical reasoning necessary to make line graphs. *Journal of Research in Science Teaching,* 26, 373-379.

[59] Wu, M. -C. (2010). An analytic study of the effectiveness of thinking skills training courses of technical vocational high schools in enhancing students' problem solving abilities. Paper presented at the meeting of the 2009 Enhancement of Research Energy in National Science Council's Applied Science and Education and Result Presentation Seminar, Taipei, Taiwan.

[60] Wycoff, J. (1991). *Mind mapping*. New York: Berkley.

[61] Yang, L. -H. (1999). *Distance learning curriculum in colleges and universities-the current situation and future developments.* Unpublished master thesis, NTNU Institute of Industrial Technology Educational, Taipei, Taiwan.

[62] Yao, Y. -C. (2002). An exploration of critical success factors of the implementation of cyber university – A case study from "National Sun Yat-Sen cyber university". Unpublished master thesis, National Sun Yat-sen University, Institute of Information Management, Kaohsiung, Taiwan.

[63] Yen, J., & Hsu, C. -M. (2007). The functional designs of the web-based teaching platform for design field: An application of analytical hierarchy process. *Journal of Science and Technology,* 16(1), 61-80.

[64] You, M. -L., & Chen, W. -Z. (2009). The problems and implications of online collaborative design learning projects. *Research in Arts Education,* 15, 105-137.

Knowledge Discovery in Health Care Datasets Using Data Mining Tools

MD. Ezaz Ahmed
Department of Computer Science & Engineering
ITM University,
Gurgaon, India

Dr. Y.K. Mathur
183 First Floor, Vaishali, Delhi University
Teacher's Housing Society
Delhi, India

Dr Varun Kumar
Head of Department
Department of CSE
MVN,
Palwal, India

***Abstract*—Non communicable diseases (NCDs) are the biggest global killers today. Sixty-three percent of all deaths in 2008 – 36 million people – were caused by NCDs. Nearly 80% of these deaths occurred in low- and middle-income countries, where the highest proportion of deaths under the age of 70 from NCDs occur [1]. The commonness of NCDs, and the resulting number of related deaths, are expected to increase substantially in the future, particularly in low and middle-income countries, due to population growth and ageing, in conjunction with economic transition and resulting changes in behavioral, occupational and environmental risk factors. NCDs already disproportionately affect low and middle-income countries. Current projections indicate that by 2020 the largest increases in NCD mortality will occur in Africa, India and other low and middle-income countries [2].**

Computer-based support in healthcare is becoming ever more important. No other domain has so many innovative changes that have such a high social impact. There has already been a long standing tradition for computer-based decision support, dealing with complex problems in medicine such as diagnosing disease, managerial decisions and assisting in the prescription of appropriate treatment. As we know that "Research is for the people not for yourself" so we are pleased to work for the healthcare and hence for the society and ultimately the MANKIND.

Keywords: NCDs; Web; Web Data; Web Mining; data Mining Healthcare.

I. INTRODUCTION

Healthcare researchers as well as practitioners require a lot of information to make their healthcare related activities and practices either with drug prescriptions which can efficiently cure patients' illness or with correct and efficient medical/clinical procedures and services. Over the last decade, we have witnessed a likely to explode growth in the information available on the World Wide Web. Today, web browsers provide easy access to innumerable sources of text and multimedia data. More than 1000000000 pages are indexed by search engines, and finding the preferred information is not an easy task. This abundance of resources has provoked the need for developing automatic mining techniques on the World Wide Web, thereby giving rise to the term "web mining." Information technology has been playing a vital and critical role in this field for many years. Therefore being able to promptly and correctly access required medical databases and information system resources and effectively communicate across different medical Institutes or countries become necessary. To proceed toward web intelligence, requires the need for human intrusion, we need to integrate and embed knowledge discovery, and machine learning into web tools.

The Web has become a major vehicle in performing research and practice related activities for healthcare researchers and practitioners, because it has so many resources and potentials to offer in their specialized professional fields. There is tremendous amount of information and knowledge existing on the Web and waiting to be discovered, shared and utilized. The research in improving the quality of life through the Web has become attractive. This paper summarizes. The reason for considering web mining, a separate field from data mining, is explained. The limitations of some of the existing web mining methods and tools are enunciated, and the significance of proposed model in health care. Scope for future research in developing "Proposed model of web mining" systems is explained. We present an approach regarding Semantic Web and mining [3] in healthcare, which can be used to not only improve the quality of Web mining results but also enhances the functions and services and the interoperability of medical information systems and standards in the healthcare field.

The objective of this article is to present an outline of web mining, Knowledge web mining, its subtasks, and to give a perspective to the research community about the potential of applying knowledge discovery techniques to its different components. The article, gives emphasis on possible enhancements of these tools using "Knowledge Web Mining". It should be noted that the use of knowledge discovery in "web mining" is a field in its origins, and thus the worth of this paper at this point in time is evident.

The rest of this paper is organized as follows: Section II deals with the web and web mining, knowledge web mining discussed in next Section i.e. in section III. Section IV provides an introduction to knowledge discovery. Sections V cover, in detail, the possible healthcare application and future model a practical approach, Section VI provides the conclusion and scope of future research in the area of knowledge web mining by proposed model.

II. WEB AND WEB MINING

Web is a collection of inter-related files on one or more Web servers. The web is a immense collection of completely uncontrolled heterogeneous documents.

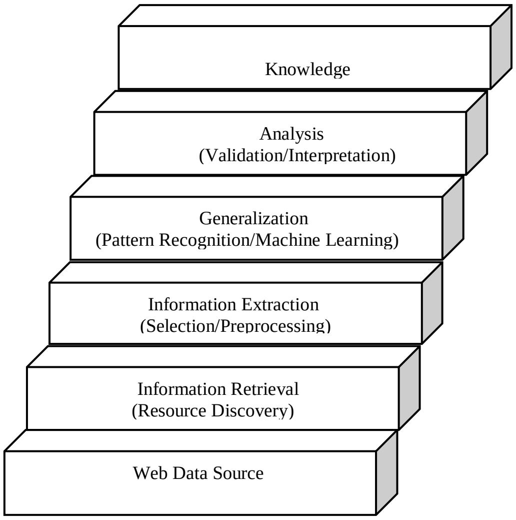


Figure1. The steps of extracting knowledge from data

Thus, it is huge, varied, and dynamic, and raises the issues of scalability, heterogeneity, and dynamism, respectively. Due to these characteristics, we are currently drowning in information, but famished for knowledge; thereby making the web a fertile area of data mining research with the vast amount of information available online. Data mining refers to the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data. Web mining can be mostly defined as the discovery and analysis of useful information from the World Wide Web.

Web mining is the application of data mining techniques to Web data [5]. Web mining helps to solve the problem of discovering how users are using Web sites. It involves mining logs (or log analysis) and the steps that typically have to be gone through to get meaningful data from Web logs - data collection, pre-processing, data enrichment and pattern analysis and discovery as given in figure 1.

Web mining is the application of data mining techniques to extract knowledge from Web data

Web data is

- Web content –text, image, records, etc.
- Web structure –hyperlinks, tags, etc.
- Web usage –http logs, app server logs, etc.

In web mining data can be collected at the server side, client side, proxy servers, or obtained from an organization's database. Depending on the location of the source, the type of collected data differs. It also has extreme variation both in its content (e.g., text, image, audio, symbolic etc.) and meta information, that might be available. This makes the techniques to be used for a particular task in web mining widely varying.

Some of the issues which have come to light, as a result, concern

- Need for handling context sensitive and imprecise
- Queries;
- Need for summarization and deduction;
- Need for personalization and learning.

Thus, web mining, though considered to be a particular application of data mining, warrants a separate field of research, mainly because of the aforesaid characteristics of the data and human related issues.

III. KNOWLEDGE WEB MINING

Relevance of knowledge web mining is extensively established in the literature recently, the application of knowledge web mining in health care problems has also drawn the attention of researchers.

Thus better healthcare related recommendations can be constructed with ontologies and with little human intervention. For an on-line healthcare web site, two important ontologies would need to be built: one of the ontology describing all the healthcare services provided, with the relation between each other, and the other ontology describing the web site. Thus Semantic Web ontology can help build better web mining analysis in healthcare, and web mining in-turn helps construct better, more powerful ontology in healthcare. Web personalization is to display and offer information to the healthcare web site users according to their interests and needs, which are already stored in the database. Personalization requires implicitly or explicitly collecting web site users information and leveraging that knowledge in the content delivery framework to choose what information to present to the users and how to present it with tailored pages according to information gathered about the particular health care web site user. Web mining is the application of data mining techniques to Web data.

Web mining helps to solve the problem of discovering how users are using Web sites. It involves mining logs (or log analysis) and the steps that typically have to be gone through to get meaningful data from Web logs - data collection, pre-processing, data enrichment and pattern analysis and discovery. We have proposed a new type of intelligent model in health care which is web based, In which we use web mining and data mining.

IV. KNOWLEDGE DISCOVERY

A. *Proposed Model*

Our work is for the architecture of a web based decision support system model. Means we have to work on basically three areas.

a) *Web based Model*
b) *DSS*
c) *Intelligent web based Model (IDSS)*

This will give us a Web based DSS model [6] which is Intelligent. Our work area will be health care, so we need data related to health care such as cardio, OSA, diabetes, breast cancer etc. Now for the development of the model we have to use Data Mining Tools such as Weka, Tanagra and PSW @12 Modeler, Statistica.

As per our research related literature survey we came to know what was the traditional model what is the current model and what model we are going to proposed. All this will be discussed below and we have to work accordingly.

V. Possible HealthCare Application And Future Model A Practical Approach

Now finally we are going to propose our model in which we have to use Web and Web mining. From web we fetch data and through web and web server we store data and after mining we make these data intelligent using our model. In this model we use data mining, web mining tools, web and web servers. Our proposed model will be as follows.

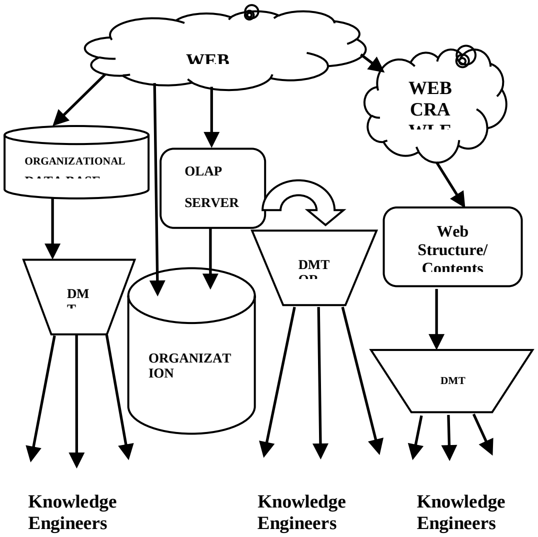


Figure2. Proposed Web based IDSS Model

As given in the above figure our research will base on three modes. In first mode we fetch data from the web and that organizational data base will be treated by data mining. By the help of data mining tools (DMT) or Knowledge mining Tools (KMT) we get knowledge data that will be used by the knowledge engineers or the users.

In second method our research will based on data which will be fetched by the web [6]. We fetch data and collect data as organizations' operational database. By the help of OLAP server we fetch data and on that data we apply DMT or KMT. This will result as knowledge data. Data will be further used by the knowledge engineers or the users. These data would be further send to web for the global use. Similarly in the third mode we fetch data from web by different web crawlers. In these data we have web structure, contents and web usage. On these data we use DMT or KMT and finally acquire knowledge data and that would be further used by knowledge engineers or end users given in figure 2.

A. *Diabetes and data mining*

When considering the healthcare business, we may find several interesting and demanding applications for DM [10]. Following our analytical formulation, we now present a real-life application for identifying diabetic patients in a small Indian town.

TABLE1: SHOWS THE PATIENT SUFFERING FROM THE DISEASE

Patient id	Disease 1	Disease 2	Disease 3
1	Blood Sugar	Blood Pressure	Heart Disease
2	Blood Sugar	Blood Pressure	Heart Disease
3	Blood Sugar	Blood Pressure	Heart Disease
4	Blood Sugar	Blood Pressure	Kidney Problem
5	Blood Sugar	Blood Pressure	Eye Problem

B. *Classification and Prediction*

Classification is the processing of finding a set of models (or functions) which describe and distinguish data classes or concepts, for the purposes of being able to use the model to predict the class of objects whose class label is unknown. The derived model may be represented in various forms, such as classification (IF-THEN) rules, decision trees, mathematical formulae, or neural networks.

Classification can be used for predicting the class label of data objects. However, in many applications, one may like to predict some missing or unavailable data values rather than class labels. This is usually the case when the predicted values are numerical data, and is often specifically referred to as prediction [7].

IF-THEN rules are specified as IF condition THEN conclusion

e.g. IF age=old and patient=diabetic then heart disease prone=yes

C. *Clustering Analysis*

Unlike classification and predication, which analyze class-labeled data objects, clustering analyzes data objects without consulting a known class label. In general, the class labels are not present in the training data simply because they are not known to begin with. Clustering can be used to generate such labels. The objects are clustered or grouped based on the principle of maximizing the intra class similarity and minimizing the interclass similarity.

That is, clusters of objects are formed so that objects within a cluster have high similarity in comparison to one another, but are very dissimilar to objects in other clusters. Each cluster that is formed can be viewed as a class of objects, from which rules can be derived. [4]

Application of clustering in medical can help medical institutes' group individual patient into classes of similar behavior [7]. Partition the patient into clusters, so those patients within a cluster (e.g. healthy) are similar to each other while dissimilar to patient in other clusters (e.g. disease prone or Weak). As given in figure 3.

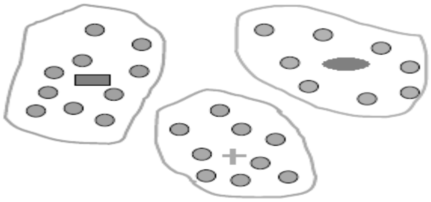


Figure 3. Picture showing the partition of patients in clusters

The main goal of this application is to be familiar with what causes diabetics. We were capable to obtain a patient database and conduct an analysis looking for to identify which patients have high probability of being diabetic.

Association Rules that can be derived from Table 1 are of the form:

$$(X, disease1) \Rightarrow (X, disease2)$$

$$(X, disease1)\hat{}(X, disease2) \Rightarrow (X, disease3)$$

$(X, "Bloodsugar") \Rightarrow (X, "Bloodpressure")$ [support=2% and confidence=60%]

(X,"Bloodsuger")^(X," Bloodpressure")

$\Longrightarrow$(X,"Heartdisease")

[Support=1% and confidence=50%]

Where support factor of the association rule shows that 1% of the patient suffering from the disease blood sugar and blood pressure, confidence factor shows that there is a chance that 50% of the patients who have "Blood sugar" will also have "Blood pressure".

This way we can find the strongly related disease and can optimize the database of a healthcare programme. As given in Table 1.

D. *Generating Knowledge*

The use of Information Theory is primarily interesting as this theory relates also to the Information Systems field. When integrates those concepts together we were capable to show that our method is relatively excellent compared to other traditional methods. Therefore, one outcome is establishing our method as a legitimate method for DM [8].

Second, we used to the DM procedure to gain knowledge about diabetes. We wrap up that the following variables can provide good indicators for identifying probable diabetic patients: family history, body weight (BMI Body Mass Index), pregnancy (in case of female patent), SFT (Skin Fold Thickness) and age. This may become a powerful predictive tool for any organization seeking to perform a more accurate and informed patient selection process to recognize diabetic patients.

E. *Working with Tanagra on Diabetes Dataset:*

Open Tanagra then open data file which is in txt, xls or arff format. We use tanagra for finding the class of particular disease and its various attributes such as Max, Min, Mean, std. deviation etc. then we find its class of diabetic and non diabetic patients then finally we are able to give description of both class of diabetic and non-diabetic patients [9] with there accuracy in percentage.

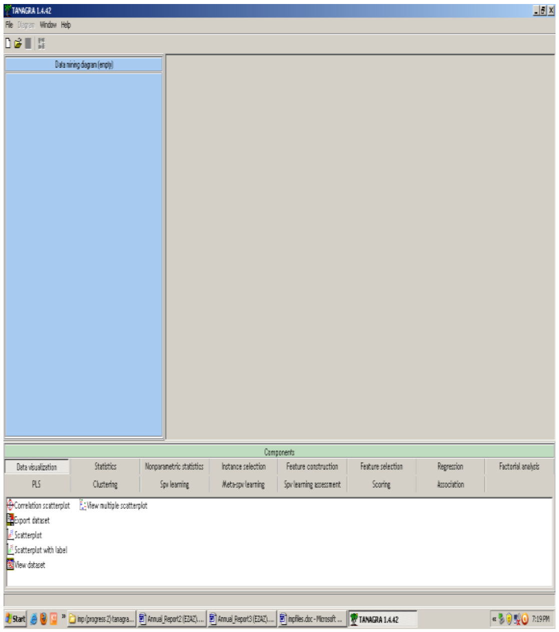
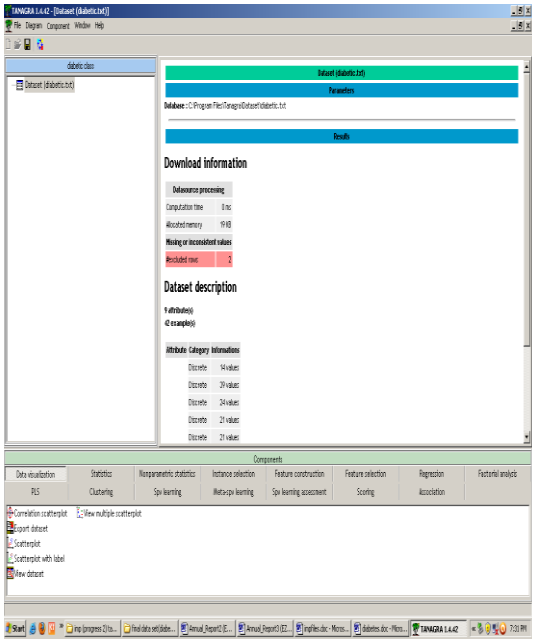


figure 4. Tanagra page

As we open the file in dataset it will appear as given below, the screen shot given below clearly indicate the open file name diabetic.txt in title diabetic class and also on the task bar of Tanagra. The down load information is on the right side of the screen given in figure 4 and 5.



Dataset (diabetic.txt)

Parameters

Database : C:\Program Files\Tanagra\Dataset\diabetic.txt

Results

Download information

Datasource processing	
Computation time	0 ms
Allocated memory	19 KB
Missing or inconsistent values	
#excluded rows	2

Dataset description

9 attribute(s)
42 example(s)

Attribute	Category	Informations
	Discrete	14 values
	Discrete	39 values
	Discrete	24 values
	Discrete	21 values
	Discrete	21 values

Figure 5. Open file information

First we open a new sheet of Tanagra and then open a dataset of diabetic patient in the format given in Tanagra. First look of Tanagra is given in figure 4. Dataset file name diabetic.txt. Now we right click on view dataset and choose view from pop-up menu which will appear after right click on view dataset. This gives the data on the Tanagra sheet given below in figure 6. Now we select view dataset from data visualization tab and drag it and drop to that on dataset. Now we select define status from feature selection tab and then drag it and drop to dataset then right click on define status and select parameters from pop up menu we get above attributes given in figure 7.

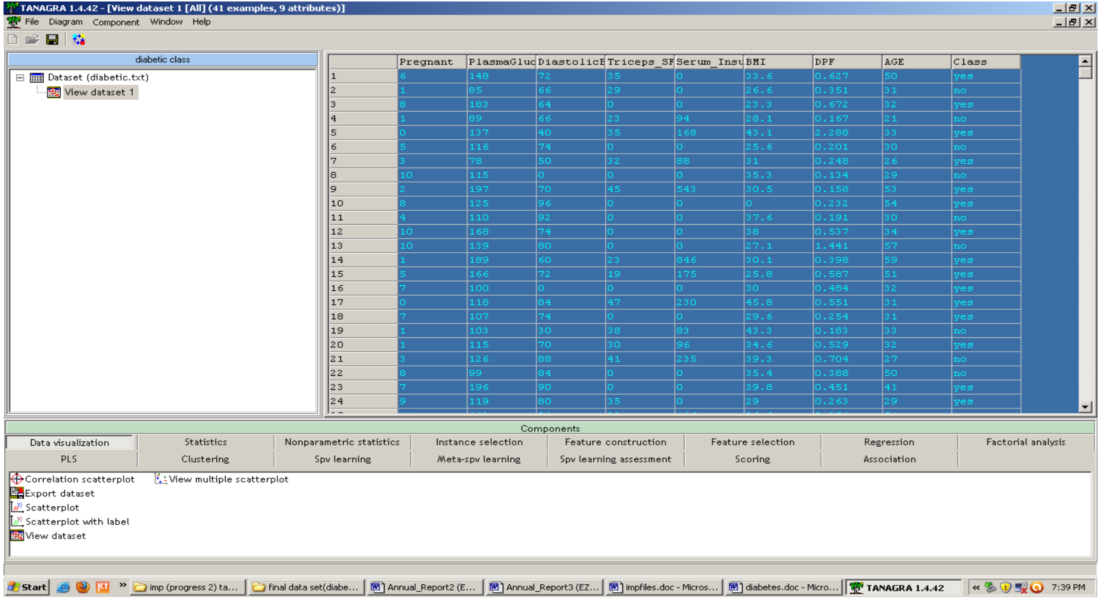
Now select four attributes as input as age, BMI, DPF and Plasma Glucose and press OK button. Now we get following result given in figure number 8.

From statistics tab we choose Univariate continuous stat, drag and drop it in define status1. Then we use view command from pop up menu we will get following figure 8. In above figure it is clear that we get result as Min, Max, Average, Std-dev and avg. Std-dev.

From example Plasma Glucose min value is 78 max values is 197 Average 126.7 Std dev is 32.11 and avg. std dev is 0.253 BMI (Body Mass Index) result is 0, 45.8, 31.42, 7.9643, and 0.2534.

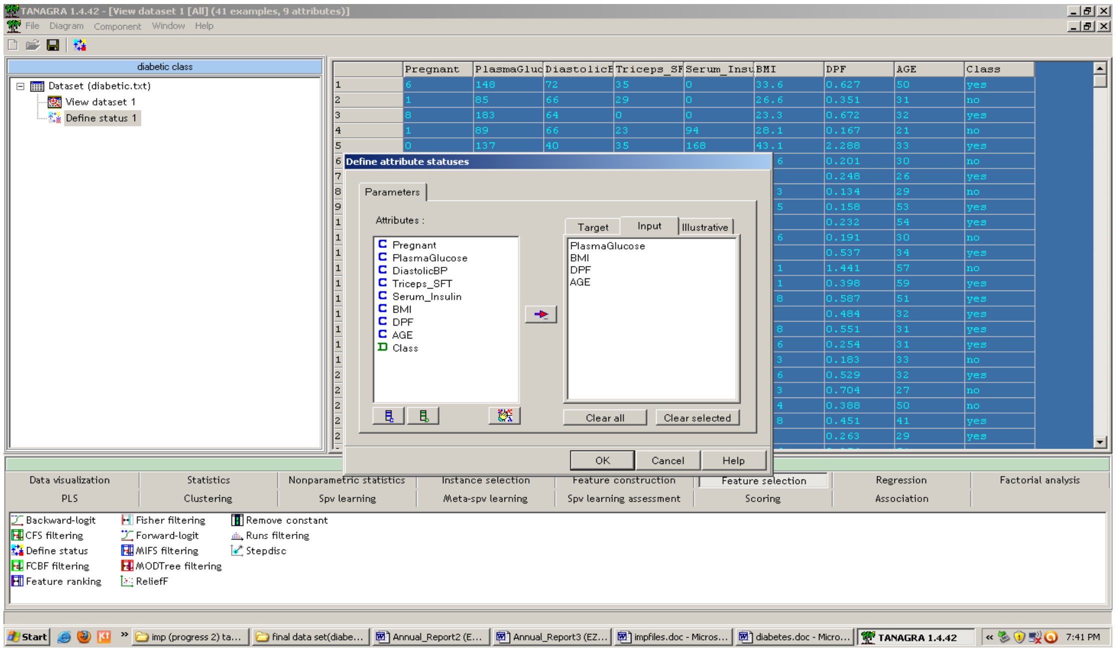
Similarly for age min age is 21 max ages in dataset is 60, average is 37.9024, std. dev is 11.6293 and avg. std dev is 0.3068 as given in above figure 8. Again we select define status3 and drag and drop group characterization figure 9 from statistics tab. Then press right click and choose view we will get following result in figure 9.

In the next section we have outcomes of the work explained.



	Pregnant	PlasmaGluc	DiastolicE	Triceps_SF	Serum_Insu	BMI	DPF	AGE	Class
1	6	148	72	35	0	33.6	0.627	50	yes
2	1	85	66	29	0	26.6	0.351	31	no
3	8	183	64	0	0	23.3	0.672	32	yes
4	1	89	66	23	94	28.1	0.167	21	no
5	0	137	40	35	168	43.1	2.288	33	yes
6	5	116	74	0	0	25.6	0.201	30	no
7	3	78	50	32	88	31	0.248	26	yes
8	10	115	0	0	0	35.3	0.134	29	no
9	2	197	70	45	543	30.5	0.158	53	yes
10	8	125	96	0	0	0	0.232	54	yes
11	4	110	92	0	0	37.6	0.191	30	no
12	10	168	74	0	0	38	0.537	34	yes
13	10	139	80	0	0	27.1	1.441	57	no
14	1	189	60	23	846	30.1	0.398	59	yes
15	5	166	72	19	175	25.8	0.587	51	yes
16	7	100	0	0	0	30	0.484	32	yes
17	0	118	84	47	230	45.8	0.551	31	yes
18	7	107	74	0	0	29.6	0.254	31	yes
19	1	103	30	38	83	43.3	0.183	33	no
20	1	115	70	30	96	34.6	0.529	32	yes
21	3	126	88	41	235	39.3	0.704	27	no
22	8	99	84	0	0	35.4	0.388	50	no
23	7	196	90	0	0	39.8	0.451	41	yes
24	9	119	80	35	0	29	0.263	29	yes

Figure 6. Diabetes dataset in Tanagra



	Pregnant	PlasmaGluc	DiastolicE	Triceps_SF	Serum_Insu	BMI	DPF	AGE	Class
1	6	148	72	35	0	33.6	0.627	50	yes
2	1	85	66	29	0	26.6	0.351	31	no
3	8	183	64	0	0	23.3	0.672	32	yes
4	1	89	66	23	94	28.1	0.167	21	no
5	0	137	40	35	168	43.1	2.288	33	yes

Define attribute statuses

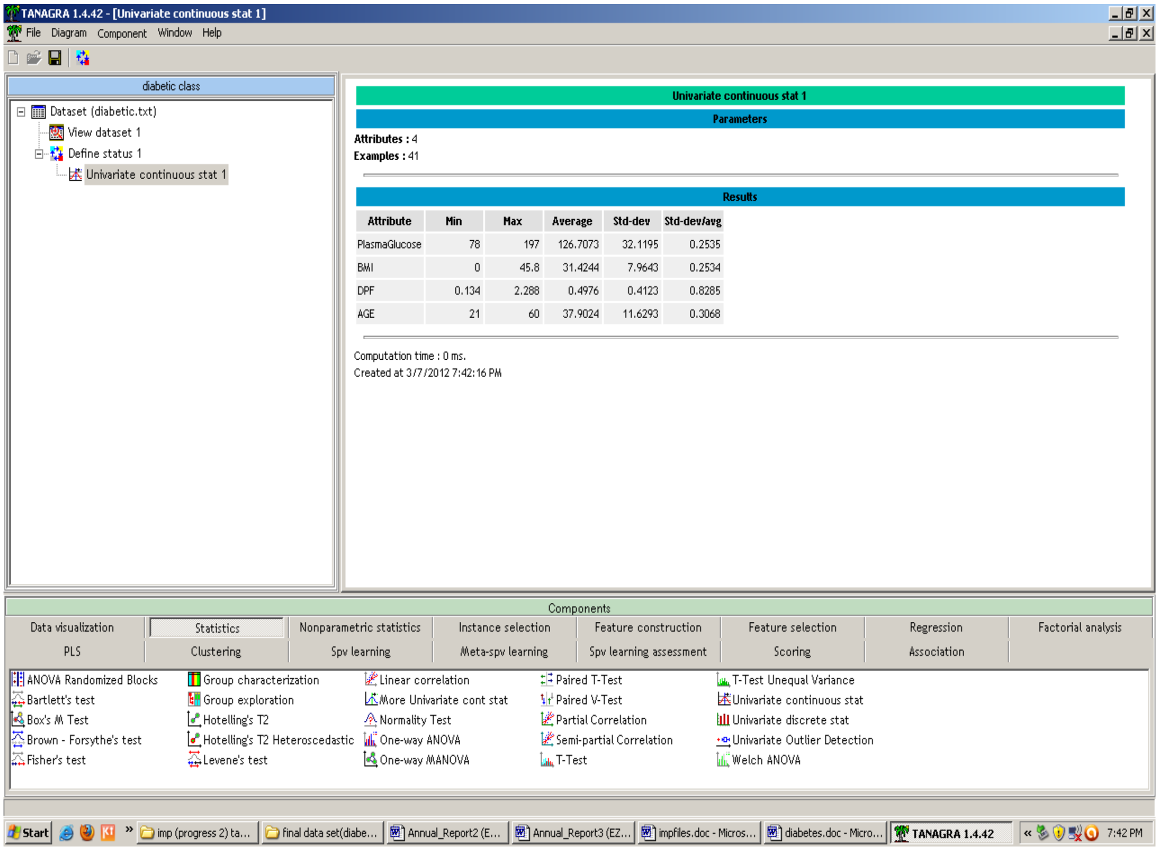
Parameters

Attributes :
C Pregnant
C PlasmaGlucose
C DiastolicBP
C Triceps_SFT
C Serum_Insulin
C BMI
C DPF
C AGE
D Class

Target | Input | Illustrative

Input:
PlasmaGlucose
BMI
DPF
AGE

Figure 7. Selection of different attributes



Univariate continuous stat 1

Parameters

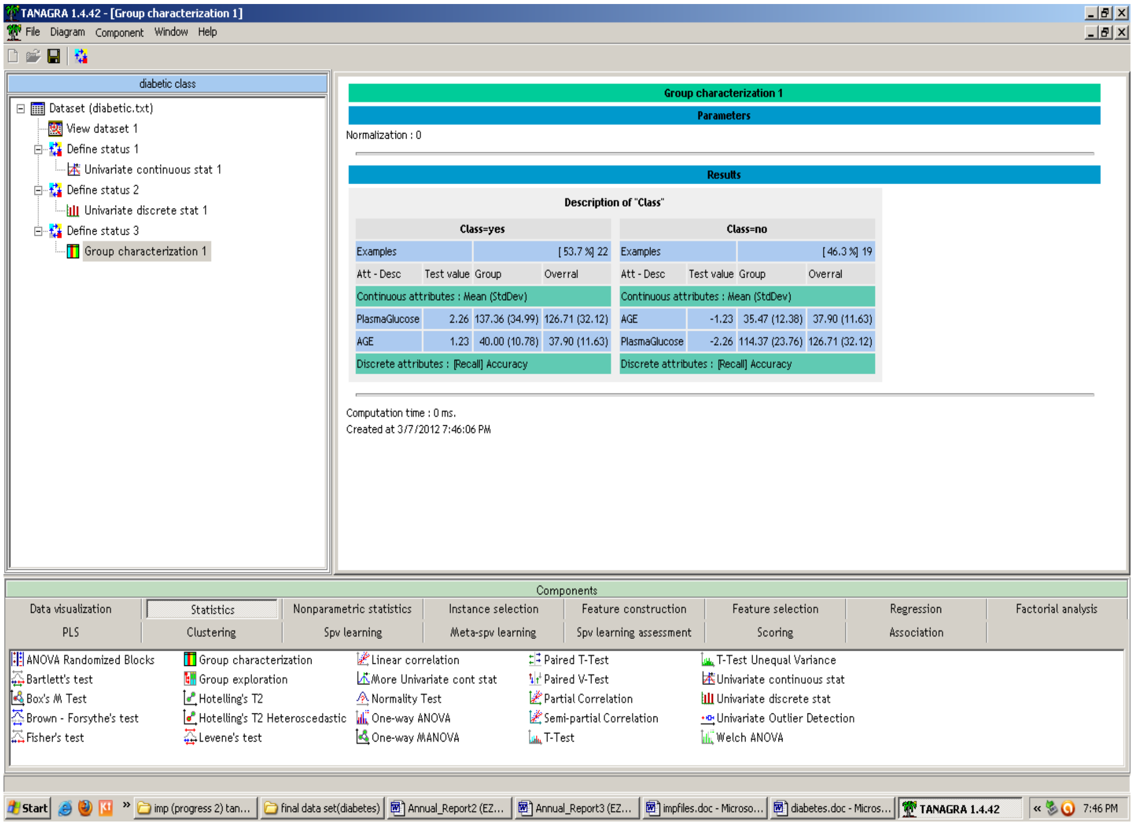
Attributes : 4
Examples : 41

Results

Attribute	Min	Max	Average	Std-dev	Std-dev/avg
PlasmaGlucose	78	197	126.7073	32.1195	0.2535
BMI	0	45.8	31.4244	7.9643	0.2534
DPF	0.134	2.288	0.4976	0.4123	0.8285
AGE	21	60	37.9024	11.6293	0.3068

Computation time : 0 ms.
Created at 3/7/2012 7:42:16 PM

Figure 8. Min, Max, average, std. deviation of different attributes



Group characterization 1

Parameters

Normalization : 0

Results

Description of "Class"

Class=yes				Class=no			
Examples			[53.7 %] 22	Examples			[46.3 %] 19
Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
PlasmaGlucose	2.26	137.36 (34.99)	126.71 (32.12)	AGE	-1.23	35.47 (12.38)	37.90 (11.63)
AGE	1.23	40.00 (10.78)	37.90 (11.63)	PlasmaGlucose	-2.26	114.37 (23.76)	126.71 (32.12)
Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy			

Computation time : 0 ms.
Created at 3/7/2012 7:46:06 PM

Figure 9. final outcomes on given diabetes dataset.

In this result we could conclude that the group 'Yes' means diabetic class having mean value of plasma glucose is 137.36 and standard deviation is 34.99 for the average mean age group is 40 years.

Whereas for group 'No' means non-diabetic class the mean plasma glucose value is 114.37 and standard deviation is 23.76 for average age group is 35.47. This is overall 41 patients clinical records test dataset. We could draw a conclusion that average age for non-diabetic is 35-36 years; increase of plasma glucose depends on after 40 years of age. So after 40 years of age a person is more prone to diabetic according to his/her plasma glucose value.

VI. Conclusion

As we continue our fight against diabetes, sharing and benchmarking diabetes care is essential to influence health policy and improve outcomes and quality of life for people with diabetes. As more and more data is collected, Diabetes measurement will become an even more powerful resource for inspiring and driving change in diabetes care. The fundamental goal of the Diabetes measurement is to measure, share, and improve diabetes outcomes. There are so many ways to get involved in reversing diabetes trends, from collecting data to sharing better practice models to improving public visibility and advocating for the quality of diabetes care at the global, national, clinic, and patient levels.

In this work we have presented an intelligent proposed model for healthcare which is related with healthcare. In this proposed model diabetic patient. This web based decision support system will helpful in health care management. Since the application of data mining brings a lot of advantages in higher well equipped hospitals, it is recommended to apply these techniques in the areas like optimization of resources, prediction of disease of a patient in the hospital.

As shown in the proposed model mentioned above, the main components of IDSS are intelligent techniques that generate knowledge which further helps health care planners to take more accurate decisions. Future work will be done on heart disease clinical dataset and find outcomes on heart disease. Work will be done on weka and will get knowledge data.

References

[1] The world health report 2002: Reducing risks, promoting healthy life. Geneva, World Health Organization, 2002.

[2] Global health risks: mortality and burden of disease attributable to selected major risks. Geneva, World Health Organization, 2009.

[3] Gerd Stumme, Andreas Hotho, Bettina Berendt, "Usage Mining for and on the Semantic Web," *Proc. Of the Joint Conference on Information Sciences*, pp. 200-204, 2003.

[4] A. Z. Broder, S. C. Glassman, M. S. Manasse, and G. Zweig, "Syntactic clustering of the web," in Proc. 6th Int.WWWConf., 1997, pp. 391–404.

[5] B. Mobasher, N. Jain, E.-H. Han, and J. Srivastava, "Web Mining: Patterns from WWW Transactions," Dept. Comput. Sci., Univ. Minnesota, Tech. Rep. TR96-050, Mar. 1997

[6] C. Romero, S. Ventura "Educational data Mining: A Survey from 1995 to 2005", Expert Systems with Applications (33), pp. 135-146, 2007

[7] R.Cooley, B. Mobasher, and J. Srivastava. "Web mining: Information and Pattern Discovery on the World Wide Web." *In Proceedings of Ninth IEEE International Conference on Tools with Artificial Intelligence" (ICTAI'97),* November 1997.

[8] Bourlas, P., Giakoumakis, E., and Papakonstantinou, G. (1999). A Knowledge Acquisition and management System for ECG Diagnosis. Machine Learning and Applications: Machine Learning in Medical Applications. Chania, Greece, pp. 27-29.

[9] Icks A et al. Incidence of lower-limb amputations in the diabetic compared to the non-diabetic population. Findings from nationwide insurance data, Germany, 2005-2007. Experimental and Clinical Endocrinology & Diabetes, 2009, 117:500–504.

[10] Han Jiawei, Micheline Kamber, Data Mining: Concepts and Technique. Morgan Kaufmann Publishers,2000

AUTHOR'S PROFILE

Md. Ezaz Ahmed

Pursuing Ph.D. under the supervision of Prof, (DR.) Yogesh K. Mathur. and co-supervision of DR. Varun kumar. Currently, he is working with itm University as Asst. Professor. He did his M.E (CSE) in first division with honors. He has more than 17 years of experience out of which 15 years teaching and 2 years industry experience. He has published 12 research papers, 2 in international Journal others in national conference and in departmental journal. His area of interest includes Web Development, Web Mining, Data Mining Software Engineering, Software verification validation and testing, and Basics of computer and C programming. He is a member of Indian Society of Technical Education (ISTE). Co-author of one project book published in 1998. He has 1 project book, 3 lab manuals to his credit.

Dr. Varun Kumar

Ph.D. (Computer Science), Head , CSE Deptt., MVN Engineering College, Palwal, Haryana ,India. Presently 3 Ph. D students are working under his supervision. Dr. Varun Kumar, completed his PhD in Computer Science. He received his M. Phil. in Computer Science and M. Tech. in Information Technology. He has 13 years of teaching experience. He is recipient of Gold Medal at his Master's degree. His area of interest includes Data Warehousing, Data Mining, and Object Oriented Languages like C++, JAVA, C# etc. He has published more than 35 research papers in Journals/Conferences/Seminars at international/national levels. He is working as an Editorial Board Member / Reviewer of various International Journals and Conferences. He has 3 books, 5 study materials and 3 lab manuals to his credit.

Dr. Y. K. Mathur

Ph.D. (Theoretical Physics) - Moscow University, Moscow, Russia (1982), M.Sc. Physics and Mathematics – Moscow University, Moscow, Russia (1978). Post Doctoral Positions held: Joint Institute for Nuclear Investigation, DUBNA, Russia (April 1982-Feb.1983), Department of Physics, University of Rochester, USA(Feb.1983-Dec.1983) and Department of Physics, University of Bielefeld, Germany (Dec.1983-Dec.1984). Academic Positions Held: CSIR Pool Officer, Department od Physics and Astrophysics, University of Delhi (as a member of Quark Physics Team headed by Prof. A.N.Mitra (FNA)) (Jan.1985-Jan. 1986), CSIR Research Scientist , Department of Physics and Astrophysics, University of Delhi.(Jan.1986-Aug.1988), Lecturer (Astt. Professor), Department of Physics and Astrophysics, University of Delhi (August 1988-May 1994), Reader (Associate Professor), Department of Physics and Astrophysics, University of Delhi(December1994-May 2005), Professor, Department of Applied Sciences and Humanities, ITM, Gurgaon Delhi(July 2005 –January 2011), Head ASH and School of Physics (April, 2010 to January 2011) and Professor, Department of Applied Sciences, PDM college of Engineering, Bahadurgarh (January 2011 till date).

Teaching Experience : Post Graduate: 21years, at the Department of Physics and Astrophysics, University of Delhi

Undergraduate : 6years (5 years and 6 months at ITM, Gurgaon and 6 months at PDM college of Engineering, Bahadurgarh) Haryana, India.

Mining Scientific Data from Pub-Med Database

G .Charles Babu
Professor, Dept. of CSE
HITS, Bogaram

Dr. A.GOVARDHAN
Professor and Director of Evaluation
JNTU Hyderabad

***Abstract*- The continuous, rapidly growing volume of scientific literature and increasing diversification of inter-disciplinary fields of science and their answers to unsolved problems in medical and allied fields of science present a major problem to scientists and librarians. It should be recalled in this aspect that today as many as 4800 scientific journals exist in the internet of which some are online only. The list of journals located in subject citation indexes in Thomson Reuters can be obtained from the website. From researchers' point of view, the problem is amplified when we consider today's competition where we may not be able to spend time on experimental work merely because of already published information. Therefore, considering these facts partly and the volume of serials on the other, a study has been initiated in evaluating the scientific literature published in various journal sources. The scope of the study does not permit inclusion of all periodicals in the extensive fields of biology and hence a text mining routine was employed to extract data based on keywords such as bioinformatics, algorithms, genomics and proteomics. The wide availability of genome sequence data has created abundant opportunities, most notably in the realm of functional genomics and proteomics. This quiet revolution in biological sciences has been enabled by our ability to collect, manage, analyze, and integrate large quantities of data.**

I. Introduction

Scientific discovery in genomics and related biomedical disciplines increased the amount of data and information [3] whereas text mining provide useful tools to assist in the curation process [4] in extracting relevant information using automatic techniques, text-mining and information-extraction approaches [5]. Text literature is playing an increasingly important role in biomedical discovery.

Most text mining applications require the ability to identify and classify words, or multi-word terms, that authors use in an article. Several strategies have been tried to recognize biological entity names in articles.

Some methods rely on protein and gene databases to assemble dictionaries of protein names. Most of these methods were developed for abstracts, because abstracts are readily available for millions of articles (e.g., PubMed)[6]. To support data interpretation, bioinformatics tools were utilized to identify relevant information from literature databases. On the other hand, success has been achieved in developing biomedical literature mining software using semantic analyses to automatically extract information [7]. This method uses a pattern discovery algorithm to identify relevant keywords in abstracts.

In this paper, we present segregated information of journals that contain or publish data on bioinformatics, proteomics and genomics. Keyword searches in PubMed database with a list of countries and their involvement in research publications have also been presented. Most of the articles in bioinformatics journals are often technology centred, focusing on rapidly evolved techniques for analysis of sequences, structures and phylogenies [8]. Some articles emphasized on data integration and analysis with data-driven data management for integrative bioinformatics systems [9]

For the purposes of investigation, the evaluation was confined to the scientific journals hosted in PubMed only [1]. It is obvious that in compiling the information on the volume of data published in journals and that even the most careful check could not exclude the possibility of errors; however it is understood that the influence of such errors is minimal considering the huge volume of information in PubMed database.

II. Materials and Methods

NCBI PubMed literature database was selected for the study. Initially a generalized search without any limits was employed to retrieve articles related to *bioinformatics* and *computational biology*. As search results indicated the presence of keyword anywhere in the article (title, abstract, address, keywords and text), a more stringent search criterion was employed to identify the number of articles appeared when a search performed either by individual or in combinations of keywords by limiting the search within Title and Abstract.

Title and abstract only search were considered in this study because the Title field in some articles refers to the most important keywords relative to the subject. Therefore, a validated disparity in information retrieved through text mining limited to Titles and Abstract terms only.

Articles belonging to bioinformatics, computational biology are explicitly reported in journals, some may have the term in Title/Abstract while some are representative of the field without keywords. Therefore, though a myriad of pertinent articles are located; preference is given to the two search techniques: Title and Abstract.

Title/Abstract is selected as limit to search the database in order to overcome false hits and to identify true positives. Therefore, an article is considered true positive only if the keyword is explicitly identified in Title/Abstract.

Records without abstracts are counted as true positives only if title contained the keywords [10]. Finally, year wise growth in number of articles in each field was carried to find out the enormous amount of data deposited in PubMed.

III. Results and Discussion

A generalized search in NCBI PubMed literature database, on March 28th 2012, using *bioinformatics* as keyword resulted in 97618 articles, of which 42.9 % are free full text and 15.9 % constitute review only articles. On the other hand, a search for *computational biology* articles in PubMed resulted in 79965 articles, of which 39.7 % and 17.9 % constitute free full text and review articles (see Table I).

TABLE I: DISTRIBUTION OF MAXIMUM NUMBER OF ARTICLES IN PUBMED DATABASE

Keyword	Total no. of articles	Free full text	Review
Bioinformatics	97618	41966	15551
Computational biology	79965	31820	14370

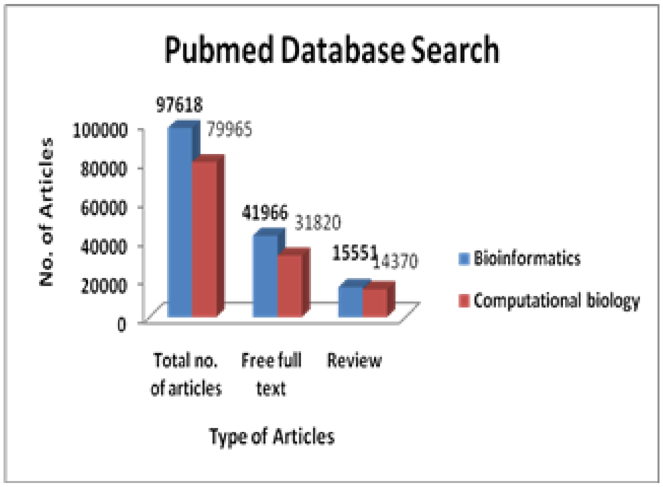


Fig. I PubMed database search with *bioinformatics* and *computational biology* as keywords.

Fig 1. illustrates the experimental results when the PubMed database was searched with *bioinformatics* and *computational* biology as keywords. The numbers over each bar represent the total number of articles from each field.

However, a more stringent search with Title/Abstract as key words revealed 11728 *bioinformatics* articles (11.9% of wild search as given in Table-1) and *computational biology* 2608 articles (3.6%) (See Table II).

TABLE II: DISTRIBUTION OF MAXIMUM NUMBER OF ARTICLES IN PUBMED WITH TITLE/ABSTRACT AS LIMIT

Keyword	Total no. of articles	Free full text	Review
Bioinformatics	11728	5664	1726
Computational biology	2934	1386	716

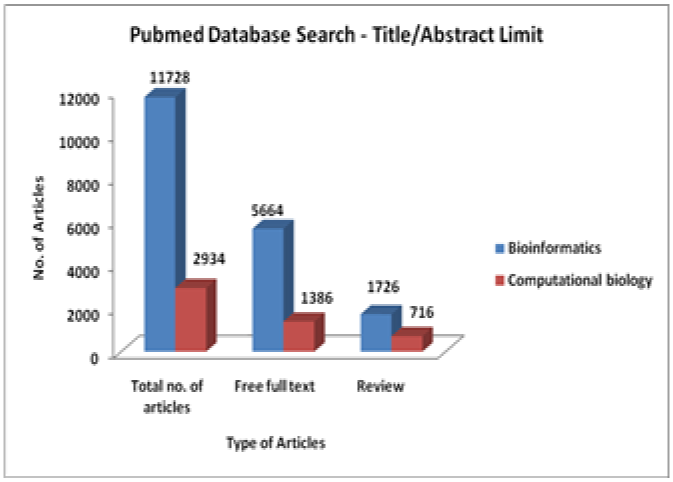


Fig. 2 PubMed database search with Title/Abstract limit for the two fields.

Fig 2. illustrates the experimental results when the PubMed database search with Title/Abstract limit for the two fields. The numbers over each bar represent the total number of articles from each field.

Boolean operator search enabled in PubMed database was used to extract combined keywords (See Table III). This shows the impact of these two ever-growing areas in sharing information and influencing the research publications.

TABLE III: NUMBER OF ARTICLES RETRIEVED IN A BOOLEAN SEARCH FROM PUBMED DATABASE

Boolean Operator	Total no. of articles	Free full text	Review
AND	80679	31873	14381
OR	97736	42018	15562

TABLE IV, Fig. 3 illustrate the annual data of the articles published on Bio-Informatics in the PubMed database.

TABLE IV : THE ANNUAL DATA OF THE ARTICLES PUBLISHED ON BIO-INFORMATICS IN THE PUBMED DATABASE.

Year	Total no. of articles	Free full text	Review
2000	1506	512	278
2001	2245	716	498
2002	3295	1209	826
2003	4674	1840	953
2004	6745	2989	1248
2005	8568	3802	1438
2006	9526	4106	1634
2007	10582	4753	1716
2008	11035	5371	1757
2009	13229	6782	1833
2010	14687	7202	2059
2011	14151	4908	1707

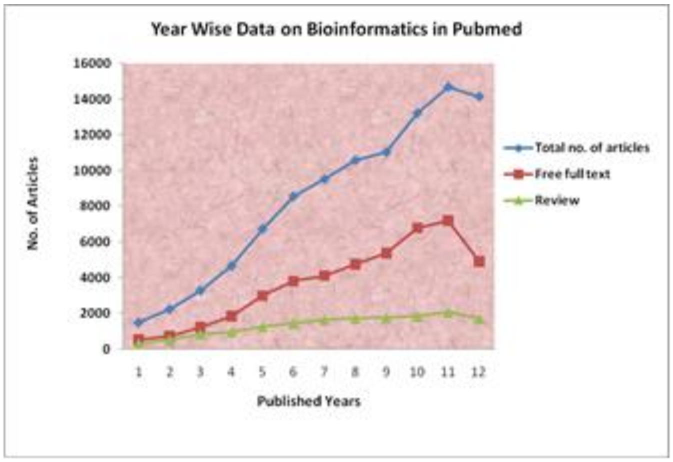


Fig. 3 Annual data on bioinformatics articles published in PubMed

IV. Conclusion

From the report, it can be emphasized that text mining is a useful alternative considering the enormous amount of data present in literature database such as PubMed. From an informatics perspective, integrated literature database like PubMed provides new insights for research in areas such as bioinformatics and computational biology. Though many research and review papers aimed at these two fields and as keywords are limited to Title/Abstract only, data suggests the phenomenal rise in number of papers in their respective fields. Therefore, from the work reported here, it can be suggested that scientific literature and approaches towards text mining have greater impact on data integration that support research for potential gains in life sciences and enable to understand the literature database applications.

References

[1] http://science.thomsonreuters.com/mjl/

[2] JohnQuackenbush *"Open-Source Software Accelerates Bioinformatics."* Genome Biology 4: p 336, 2003.

[3] Hersh W, Bhupatiraju RT, Corley S *"Enhancing Access To The Bibliome: The TREC Genomics Track."* Medinfo p 773–777, 2004.

[4] Yeh AS, Hirschman L, Morgan AA *"Evaluation Of Text Data Mining For Database Curation"* Lessons Learned From The KDD Challenge Cup. Bioinformatics 19: Suppl 1:i331–339, 2003.

[5] Krallinger M, Erhardt RA, Valencia A *"Text-Mining Approaches In Molecular Biology And Biomedicine"* Drug Discovery Today 10 p 439-445, 2005 .

[6] Shi L, Campagne F. *"Building A Protein Name Dictionary From Full Text: A Machine Learning Term Extraction Approach."* BMC Bioinformatics 6 p 88, 2005.

[7] Chaussabel D.*"Biomedical Literature Mining: Challenges And Solutions In The 'Omics' Era"*. American Journal of Pharmacogenomics 4 p383-393, 2004.

[8] David B. Searls *"Using Bioinformatics In Gene And Drug Discovery".* Drug Discovery Today 5 p135-143, 2000.

[9] Jain, E. and Jain, K. *"Integrated Bioinformatics – High Throughput Interpretation Of Pathways And Biology".* Trends Biotechnology 19,p 157–158, 2001.

[10] Kaveh G. Shojania, Lisa A. Bero,*"Taking Advantage Of The Explosion Of Systematics Reviews: An Efficient MEDLINE Search Strategy".* Effective Clinical Practice,July/August 2001.

Authors Profile

He received his M.Tech from JNTU in 1997 and B.Tech from KLCE in 1997.He is working as a Professor & Head in Dept. of CSE , Holy Mary Institutte of Technology & Science, Bogaram, Hyedarabad, India.

Received Ph.D in Computer Science & Engg from JNTU in 2003. M.Tech from JNTU in 1994.B.E from Osmaina University in 1992. He is working as a Director of Evaluation in JNTU Hyderabad.

He has published around 120 papers in various national and international Journals/conferences. His research of interest includes Data Mining, Information Retrieval & Search Engines.