Automatic Sign Language Recognition: Performance Comparison of Word based Approach with Spelling based Approach

Shazia Saqib

Department of Computer Science, GC University,
Lahore, Pakistan,
Works at Lahore Garrison University, Pakistan

Syed Asad Raza Kazmi

Department of Computer Science, GC University,
Lahore, Pakistan

Dr. Khalid Masood

Lahore Garrison University,
Pakistan

Saleh Alrashed

University of Dammam,
Dammam, KS

Abstract--**Evolution of computer based interaction has been through a number of phases. From command line interface to menu driven environment to Graphics User Interface, the communication has evolved to a better user friendly environment. A new form of communication is on the rise and that is Gesture Based Communication, which is a touch free environment basically. Although its applications are mainly for deaf community but smart mobiles, laptops and other similar devices are encouraging this new kind of communication. Sign languages all over the world have a dictionary of signs of several thousand words. Mostly these signs are word based which means that these signs do not make use of basic alphabet signs, rather a new sign has to be designed for every new word added to the dictionary. This paper suggests use of spelling-based gestures especially while communicating with smart phones and laptops.**

Keywords—Feature extraction; human computer interaction; image segmentation; object recognition

I. INTRODUCTION

A "sign" means a meaningful unit element of communication in a sign language. A sign may be a simple sign or it may represent a word. Signs may represent hand gestures or other signs. Whole sign recognition system works around to identify signs conveyed in a gesture. Sign languages pose the challenge that there is unfortunately no international sign language. All sign languages suffer from lack of rules and regulations [1], [2].

A gesture conveys an information. When we press a key, although it creates a piece of information but it is not a gesture, so we need a different class of input devices to understand gestures. Human hand and fingers have larger degree of freedom as compared to mouse or other pointing devices. The human hand has roughly 27 degrees of freedom. While dealing with hand gestures we need to learn two concepts:

- Hand posture: A hand posture is a simple hand sign that has no hand movement.

- Hand gesture: A hand gesture is collection of one or more continuous hand postures. When we hold our fist in a particular way it is a hand posture. A hand gesture is a dynamic movement, such as moving hand to wish farewell [3].

Many fiction writers like in "Star Trek" had foreseen this future development and used these ideas in the movies. They used the idea of shouting in space and let the computer process the command [3]. Another way has been Gaze Detection, use of Bolt added gaze detection to improve the interface. There were up to thirty different moving images at a time on wall with all sound tracks as one track. Gaze detection helped to resize moving images.

The gesture-based input gives us a natural and interactive interface. When we use touch free interface we can be more focused on our problem rather than the input [4]. The following are the application areas of hand gestures system:

A. Evolution of Sign Language

Gestures have gained popularity all over the world for communication with smart devices. Another major area is helping deaf community to make smooth interaction with mobile and computers. American Sign Language (ASL) and British Sign Language (BSL) are based on English language [5]-[7] whereas, German Sign Language (GSL) [8], Argentinian Sign Language (ArSL) [9], Indian Sign Language (ISL) [10], Persian sign language[11], Arabian Sign Language [12], [13] and Chinese Sign Language (CSL) [14] are also among the well-known sign languages.

B. Robot Control

Robot control is another area where gestures are highly useful. One interesting application of robotics, i.e. all finger gestures are assigned special meaning e.g. 1 may mean come forward. These systems used one hand fingers only.

C. Gestures in Graphic Editing

In this hand gesture movement is tracked and depending on the movement of hand a particular action is taken care of.

Min et al. has involved 12 dynamic gestures to make tools available for creating graphic systems [15]. It uses building blocks like triangle, rectangular, circle, arc, lines for creating the environment for various graphics functions.

D. Virtual Environments (VEs)

Another popular application of gestures is virtual environments VEs for 3D pointing gesture recognition.

E. Numbers Identification

Gestures are used to identify numbers. Elmezain et al. has proposed algorithm for recognition of Arabic numbers in real time using HMM [16].

F. 3D Modeling

Sign languages are more powerful in building 3D models, sometimes even hand shadows are used to build 2D and 3D objects [17].

The following are the reasons, why users love gesture-based interface:

Reason 1: It is the safest method to communicate

Gestures are most ancient way of communication. Even for a foreign language communication, where we fail to use words, we speak our mind by using our hands, and you can be sure that you'll have a good chance of being understood. Moreover regardless of nationality, No matter where you come from, you have the same methods of indicating that something is big, that the meal you've just eaten was tasty or that something smells awful.

Reason 2: Trend of adopting gesture-based design in applications

In last few years, the trend of adopting gesture-based design in applications (mobile applications) has become more common thanks to growing usage of devices with touchscreen interfaces. In the not-so-distant past, everyone mainly used mobile devices such as smartphones and tablets. This new segment of devices requires a completely new approach to the problem, because the interface that we will be working with is different. It's worth pointing out that we can use these new devices on our lap or on the table, so both hands are free**.** This wipes out a whole load of restrictions so designers are able to discover and develop new solutions in the area of interfaces and interactions.

Reason 3: Touch Free interface is fun and invokes Powerful Response

The popularity of applications such as Snap Chat or Tinder has influenced users' awareness and feel of interfaces based on gestures. This has opened up the option of honing app design down to simplicity by getting rid of some visual elements of navigation. A key trend to notice in this field is the gradual disappearance of differences between popular platforms (iOS, Android) and increasing unification in terms of possible interactions based on gestures.

Simpler design and unified solutions allow the user to concentrate fully on the content and absorb it in a more intuitive way. This type of application is way more fun for the user and elicits a more positive and powerful emotional response. A more intuitive approach to usage that allows us to get into the app's content more directly and simply makes our apps more attractive and, therefore, more highly rated.

Reason 4: Gesture-based approach is constantly evolving and is a game changer

The gesture-based approach is constantly evolving and, in the future when combined with voice recognition, it may well turn out to be a game changer.

Usually gestures are a part of a continuous sequence of signs. This is very active area of research in gesture recognition as it is very hard to locate start and end of a gesture. Different signers have different gesture boundaries thus making it quite hard to enumerate all gestures. A signer gives a gesture starting from a pause state and ending in a pause state even when gesturing continuously [3].

The following performance parameters were kept in mind during the design of the research:

Recognition time: A good recognition time varies from 0.25 to 0.50 seconds.

Continuous and automatic recognition: System accuracy should be high enough to recognize many dynamic gestures in one go.

Recognition Accuracy: A good accuracy rate varies in between 80-90 percent [18], [19].

Gesture based systems are affected by image transition, scaling and illumination affects. Response time and the cost of interpretation of gesture are other important factors that measure the effectiveness of the system [20]. Fig. 1 shows few such factors. Our goal is to help this disabled population and side by side providing a touch free environment for our smart devices [21].
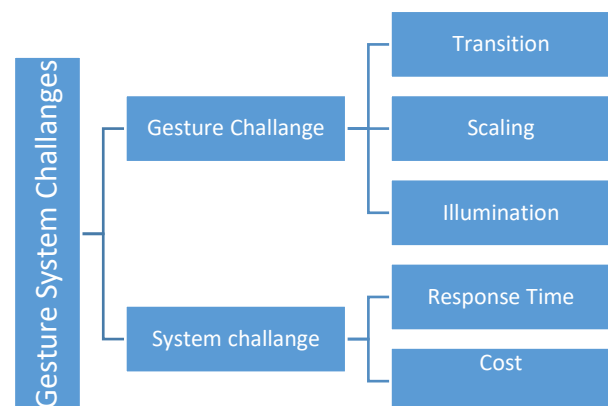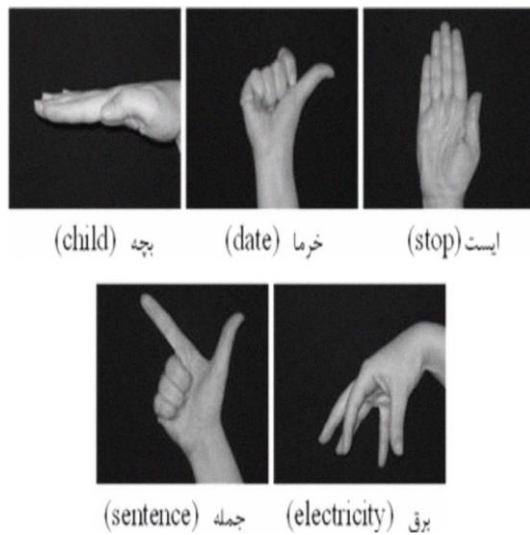

Fig. 1.   Gesture system challenges.

Fig. 2. Few words in Persian Sign language.

Different words have different gestures in different sign languages, e.g. the word "date" in Persian Sign Language has same symbol as that of Seen (س) in Pakistan Sign Language and the gesture for the word "sentence" in Persian Sign Language resembles the symbol laam(ل) in Urdu language. That means there are no universal signs [11], [22], [23]. Fig. 2 shows few word based signs in Persian sign language.

As compared to verbal languages, sign languages are regional languages. Even in Arab countries lots of efforts have been made to establish same standard sign language used in individual countries [7]. The significance of using hand gestures for communication becomes clearer when sign language is considered. Sign language uses gestures, postures, and facial expressions to identify the signer's input [3], [24].

A sign language uses manual communication and body language to convey meaning rather than using sound to communicate [25]. Every gesture represents a letter or a word and it may hold different meaning in different languages for the word "What" the sign in different Sign languages is shown in Fig. 3.

Fig. 3. Sign of "What" in different sign languages.

We can see here some sign languages use two hand gestures while some use one hand gesture.

II. PROPOSED APPROACH

Most of the sign languages used are using word based gestures, which means suppose there are 5000 words in the sign language and we need to add one more word to our gesture database, we will need a new gesture which does not base on any algorithm or automata structure. So whenever a new word becomes part of sign language, the word gesture has to be learned.

The proposed **spelling based gestures** are based on the fact that alphabets make words. So static gestures can be used to form gestures for dynamic gestures. In this way whenever new words are added to dictionary of sign language, there will be no need to learn a new sign. New words will be actually sequence of static signs.

Moreover, if word gestures are spelling based, we can use any technique like regular expression, expression tree or finite automata for gesture recognition. The gesture based input are very powerful means for communication. Their advantages are: naturalness, more expressive, direct interaction and freedom of expression through signs. Fig. 4 shows word-based gestures for shoe while Fig. 5 shows same when we use spelling-based gestures.

Recognition and interpretation of a sign holds biggest challenge in gesture recognition. When a gesture starts or ends, this is another research question at the moment for image processing experts. However, when we deal with word-based gestures the intensity of problem is reduced considerably. In spelling based gestures a sequence of inputs is given however boundaries are blurred [4].
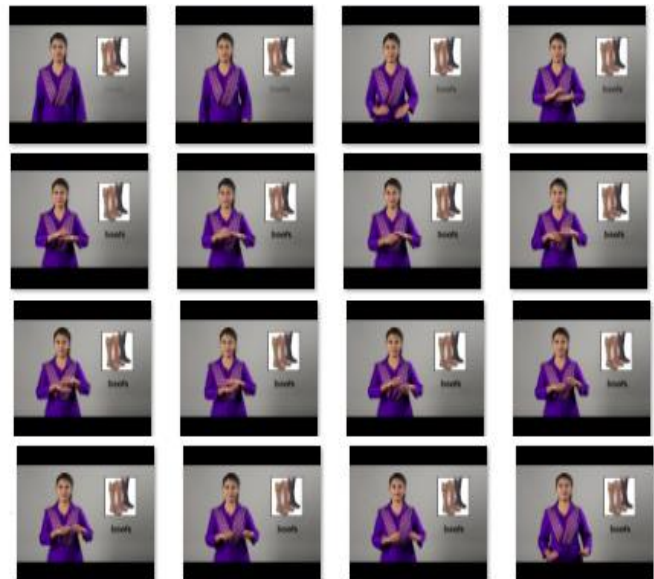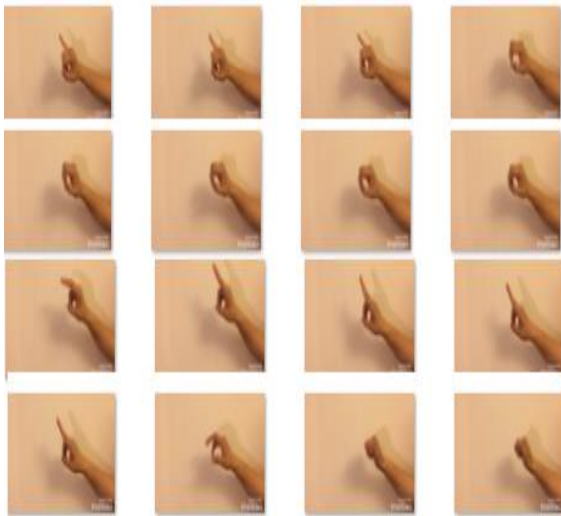

Fig. 4. Word based gesture for shoe.

Fig. 5. Spelling based gesture for shoe.

III. EXPERIMENTAL SETUP

Spelling based gestures can be one hand or two hand, they are usually one hand. That is another added benefit, we are usually holding something in our one hand. So, one handed interaction will give us more freedom.

Videos for both have been converted to frames. It was found that the average time taken by word-based gesture shoe(joota) took 4 seconds while the spelling-based gesture for shoe(joota) took 3 seconds.

To compare different word based and spelling based gestures, we used 30 words with five different signers. The camera for capturing images can be any webcam, mobile camera or ordinary laptop camera. The images are captured at a resolution of 8 mega pixels. The images are taken from an average distance of 4 to 5 feet.
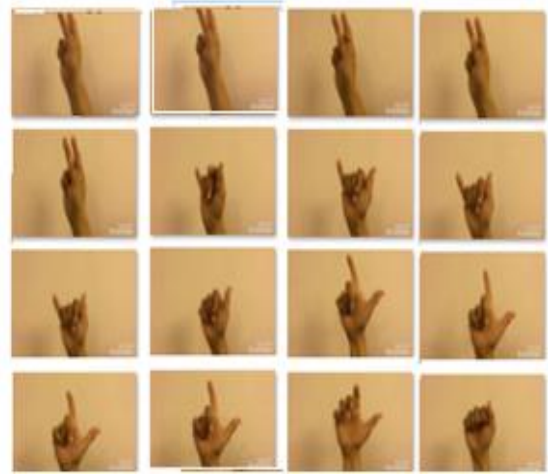

Fig. 6. Word-based gesture for banana.

Fig. 7. Spelling based gesture for banana.

Fig. 6and 7 shows gestures for Banana in both word based and spelling-based gestures. Fig. 8 shows a chart of comparison of word based and spelling-based gestures of few words.

Using 5 signers, different videos for different words were made by ordinary mobile and laptop camera. The average distance between the signer and image capturing device was 5 feet appr. The experiment was repeated time and again for different signers. Table I gives measure of average time taken by each gesture.

TABLE. I. TIME COMPARISON OF SPELLING BASED AND WORD BASED GESTURES (TIME IN SECONDS)

	Spelling based gestures	Word based gestures
Apple	6	4.5
Banana	7	6.5
Raisins	7	6.5
Lychee	7	7
Mango	4	5.5
Shoe	3.5	5

IV. RESULTS AND DISCUSSION

Pakistan sign language has been taken as a case study for experimentation however any sign language can be chosen. By looking at the video comparison result of both types, we can draw the following conclusions:

1) Both types of gestures take almost the same time. Difference of time in both is almost negligible.

2) Spelling based gestures are very easy to learn and generate. A signer needs to learn only 26 alphabets and 10 numeric signs.

3) If we can devise an algorithm that can separate character boundaries, we can use a construct like finite automata or fuzzy logic to identify the sign.
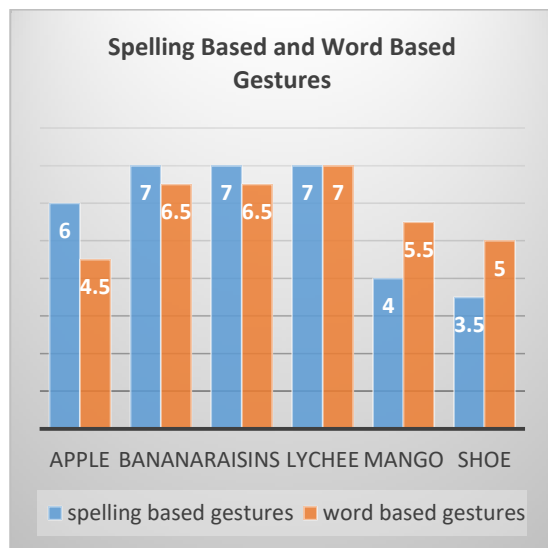
Fig. 8.   Graph of word based gestures and spelling based gestures time measured in second.

4) If we can devise an algorithm that can identify first alphabet of the gesture, searching the gesture in data set becomes very easy.

V.   CONCLUSION AND FUTURE WORK

Although word-based gestures are in use all over the world, adopting spelling-based approach can bring a revolution for deaf as well as for an effective touch free interface. Every region has its own sign language. Although the whole research in this area started with an idea to facilitate the deaf community but it actually formed the basis of touch free interface for smart devices. This leads to a very strong need for **universal sign language**. Moreover, estimating start of a gesture and end of a gesture can be a very attractive area of research. At the moment it is very hard to separate symbols however in next few years advancements in technology will be able to easily handle this issue. Present research can recognize static gestures with great accuracy. Very soon spelling based gestures will be adopted as this is most easy way to expand the sign language.

ACKNOWLEDGMENTS

REFERENCES

[1]   S. Tehsin et al., "Text Localization and Detection Method for Born-digital Images Text Localization and Detection Method for Born-digital Images," vol. 2063, no. February, 2016.

[2]   M. Melnyk, V. Shadrova, and B. Karwatsky, "Towards Computer Assisted International Sign Language Recognition System : A Systematic Survey," Int. J. Comput. Appl., vol. 89, no. 17, pp. 44–51, 2014.

[3]   Q. Chen, M. D. Cordea, E. M. Petriu, T. E. Whalen, I. J. Rudas, and A. Varkonyi-koczy, "Hand-Gesture and Facial-Expression Human-Computer Interfaces for Intelligent Space Applications," 2008.

[4]   M. Van Beurden, W. Ijsselsteijn, and Y. De Kort, "User experience of gesture-based interfaces : A comparison with traditional interaction methods on pragmatic and hedonic qualities," pp. 121–124.

[5]   R. Varga and Z. Prekopcsák, "Creating a Database for Objective Comparison of Gesture Recognition Systems," pp. 1–6, 2011.

[6]   J. R. Pansare, S. H. Gawande, and M. Ingle, "Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background," vol. 2012, no. August, pp. 364–367, 2012.

[7]   P. Pandey and V. Jain, "Hand Gesture Recognition for Sign Language Recognition : A Review," vol. 4, no. 3, 2015.

[8]   J. Forster, C. Oberdꙅɑrfer, O. Koller, and H. Ney, "Modality Combination Techniques for Continuous Sign Language Recognition," Pattern Recognit. Image Anal., pp. 89–99, 2013.

[9]   F. Ronchetti, F. Quiroga, and L. Lanzarini, "LSA64 : An Argentinian Sign Language Dataset," pp. 794–803.

[10]   J. Singha and K. Das, "Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique," vol. 4, no. 2, pp. 188–195, 2013.

[11]   A. K. Sarkalehl, F. Poorahangaryan, B. Zan, and A. Karami, "A neural network based system for persian sign language recognition," ICSIPA09 - 2009 IEEE Int. Conf. Signal Image Process. Appl. Conf. Proc., pp. 145–149, 2009.

[12]   N. R. Albelwi, "Real-Time Arabic Sign Language (ArSL) Recognition," pp. 497–501, 2012.

[13]   B. Jalilian and A. Chalechale, "Face and Hand Shape Segmentation Using Statistical Skin Detection for Sign Language Recognition," vol. 1, no. 3, pp. 196–201, 2013.

[14]   C. Wang, X. Chen, and W. Gao, "Expanding Training Set for Chinese Sign Language Recognition," pp. 0–5, 2006.

[15]   B. Min, H. Yoon, J. Soh, T. Ejimau, and I. Engineering, "Recogrution Using," pp. 4232–4235, 1997.

[16]   M. Elmezain, A. Al-Hamadi, J. Appenrodt, and B. Michaelis, "A Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition," Proc. World Acad. Sci. Eng. Technol., vol. 43, no. July, p. pages 394-401, 2009.

[17]   R. Zaman Khan and N. A. Ibraheem, "Comparative Study of Hand Gesture Recognition System," Comput. Sci. Inf. Technol. (CS IT), vol. 6, no. 4, pp. 203–213, 2012.

[18]   A. K. Alvi, A. Muzzaffar, and M. Usman, "Project Mentor : Team Members :"

[19]   A. K. Alvi et al., "Pakistan Sign Language Recognition Using Statistical Template Matching," vol. 1, no. 3, pp. 1–4, 2007.

[20]   S. M. Darwish, M. M. Madbouly, and M. B. Khorsheed, "Hand Gesture Recognition for Sign Language : A New Higher Order Fuzzy HMM Approach," vol. 8, no. 3, 2016.

[21]   M. S. Abdalla and E. E. Hemayed, "Dynamic Hand Gesture Recognition of Arabic," vol. 13, no. 5, 2013.

[22]   M. Moghaddam, M. Nahvi, and R. H. Pak, "Static Persian Sign Language Recognition using Kernel-based Feature Extraction," 2011.

[23]   S. N. Sawant, "Sign Language Recognition System to aid Deaf-dumb People Using PCA," vol. 5, no. 5, pp. 570–574, 2014.

[24]   N. El-bendary, H. M. Zawbaa, M. S. Daoud, A. E. Hassanien, and K. Nakamatsu, "ArSLAT : Arabic Sign Language Alphabets Translator," vol. 3, pp. 498–506, 2011.

[25]   N. S. Khan, A. Abid, K. Abid, U. Farooq, M. S. Farooq, and H. Jameel, "Speak Pakistan : Challenges in Developing Pakistan Sign," vol. 30, no. 2, pp. 367–379, 2015.