International Journal of Engineering Research and General Science Volume 2, Issue 3, April-May 2014
ISSN 2091-2730

An Approach to Global Gesture Recognition Translator

Apeksha Agarwal [1], Parul Yadav[2]

[1]Amity school of Engieering Technology, lucknow Uttar Pradesh, india

[2]Department of Computer Science and Engineering, Amity Lucknow, Lucknow, India

E-mail- pyadav@lko.amity.edu

Abstract— Hand gestures play a vital role in communication between people during their daily lives. The major use of hand gestures as a mean of communication can be found mostly in the form of sign languages. Sign language is a popular communication method that can be used between deaf and dumb people. A translator is definitely needed when a person wants to communicate with a deaf one. Sign Language is the only mode of communication between deaf/dumb and normal human beings. The major difficulty of sign language recognition is determined by the fact that there exist a variety of sign language sets in the world even for a single language such as English. No Global gesture recognition translator is proposed to overcome this difficulty and therefore it is impossible for users of different sign language groups to understand each other. In this research paper, we have proposed an approach to Global sign language recognition translator.

Keywords— Master gesture(MG), Canonical frame, Sign Language(SL), Kohonen, Translator, Eigen, Degree of freedom(DOF).

INTRODUCTION

Human–Computer Interaction (HCI) is getting increasingly important as computer's Influence on our lives is becoming more and more significant[1]. With the advancement in the world of computers, the already-existing HCI devices (the mouse and the keyboard for example) are not satisfying the increasing demands anymore. Designers are working to make HCI faster, easier, and to look more natural[2]. To achieve this, *Human-to-Human Interaction* techniques are being introduced into the field of Human-Computer Interaction. One of the most fertile Human-to-Human Interaction fields is the use of hand gestures. People use hand gestures mainly to communicate and to express ideas[3].The importance of using hand gestures for communication becomes clearer when *sign language* is considered. The sign language is the fundamental communication method between people who suffer from hearing imperfections. Sign language is a collection of gestures, movements, postures, and facial expressions corresponding to letters and words in natural languages[4]. In order for an ordinary person to communicate with deaf people,an interpreter is usually needed to translate sign language into natural language and vice-versa. In the recent years, the idea of designing a GSL translator has become an attractive research area.

We can express our feelings and thoughts through gestures, gestures can go beyond this point, hostility and enmity can be expressed as well during speech, approval and emotion are also expressed by gestures[5]. The development of user interface requires a good understanding of the structure of human hands to specify the kinds of postures and gestures. To clarify the difference between hand postures and gestures, hand posture is considered to be a static form of hand poses an example of posture is the hand posture like 'stop' hand sign , it's called also static gesture, or Static Recognition[6]. On the other hand; a hand gesture is a comprised of a sequence static postures that form one single gesture and presented within a specific time period, example for such gesture the orchestra conductor that applies many gestures to coordinate the concert, also called dynamic recognition, or dynamic gesture. Some gestures might have both static and dynamic characteristics as in sign languages[7]. We can define gesture as a meaningful physical movement of the fingers, hands, arms, or other parts of the body, with the purpose to convey information or meaning for the environment interaction . Gesture recognition, needs a good interpretation of the hand movement as effectively meaningful commands[15]. For human computer interaction (HCI)interpretation system there are two commonly approaches:

a. ***Data Gloves Approaches:*** These methods employs mechanical or optical sensors attached to a glove that transforms finger flexions into electrical signals to determine the hand posture [16]. Further in this method the data is collected by one or more data- glove instruments which have different measures for the joint angles of the hand and degree of freedom (DOF) that contain data position and orientation of the hand used for tracking the hand. However, this method requires a wearisome device with a load of cables connected to the computer, which will hampers the naturalness of user-computer interaction.

b. ***Vision Based Approaches:*** These techniques based on the how person realize information about the environment. These methods are usually followed by capturing the input image using camera(s).In this ,the architecture is generally divided into two parts i.e feature extraction and recognition[17]. The recognizer uses and implements machine learning algorithms. Artificial Neural Network (ANN) and Hidden Markov Model (HMM) are the most common tool to be used and implemented. In order to create the database for gesture system, the gestures should be selected with their relevant meaning and each gesture may contain multi samples for increasing the accuracy of the system.

2. RELATED WORK

Work for automatic recognition of sign language became visible in the 90s. Researches on hand gestures and two major classes have been identified. First category relies on electromechanical devices (sensor) that are used to observe different gesture parameters such as hand's position, angle, and the location of fingertips and gives numerical values[1]. Systems that use such devices are called wearable or glove computation-based systems. So to avoid this inconvenience ,the second class exploits image processing techniques to create visual based hand gesture recognition systems. Visual gesture recognition systems are again classified into two categories: The first one depends on specially designed gloves with visual markers called "vision based glove-marker gestures (VBGMG)" that helps in determining the hand position and postures[2]. But using gloves and markers leads to the limitation and do not provide the natural feeling required in human computer interaction systems. Besides, colored gloves increases, the processing complexity. The next one that is an alternative to the second kind of visual based gesture recognition systems can be called "natural visual-based gesture (NVBG)" means visual based gesture without gloves & markers[3]. And this type tries to provide the ultimate convenience and naturalness by using images of bare hands for gesture recognition. Also many researchers have been trying very hard to introduce hand gestures to Human-Machine Interaction field. Year 1992: Charayaphan and Marble developed a way to understand American Sign Language using image processing. Previous work on sign language recognition focuses primarily on finger spelling recognition and isolated sign recognition. Some work uses neural networks ,

For the work to apply to continuous ASL recognition, the problem of explicit temporal segmentation must be solved. HMM-based approaches take care of this problem implicitly. Mohammed Waleed Kadous uses Power Gloves to recognize a set of 95 isolated Auslan signs with 80% accuracy, with an emphasis on computationally inexpensive methods There is very little previous work on continualus ASL recognition. Thad Starner and Alex Pentland use a view-based approach to extract two-dimensional features as input to HMMs with a 40-word vocabulary and a strongly constrained sentence structure consisting of a pronoun, verb, noun, adjective, and pronoun in sequence. Annelies Braf€ort describes ARGO, an architecture for LSF recognition based on linguistic principles and HMMs , but provides limited experimentation results. Yanghee Nam and KwangYoenWohn use three-dimensional data as input to HMMs for continuous recognition of a very small set of gestures.

Year1996: Grobel and Assan used HMMs to recognize isolated signs and Braffort also presented a recognition system for sentences of French Sign language. Year 1997: Vogler and Metaxas used computer vision methods and HMMs. Year 1998: Yoshinori, Kang-Hyun, Nobutaka, and Yoshiaki used colored gloves and have proven that using colored gloves faster & easier hand features extraction can be done than simply wearing no gloves at all. Year 1998: Liang and Ouhyoung developed a continuous recognition of Taiwan sign language using HMMs with a vocabulary between 71 and 250 signs using data glove as an input device. However their system required that gestures performed by the signer to be slow to so that word boundary can be detected. Year 1999: Yang and Ahuja developed dynamic gestures recognition as they used skin color detection and transforms of skin regions which are in motion to figure out the motion path of ASL signs. Using a neural network, they recognized 40 ASL gestures with a success rate of around 96%. But their technique has very high computational cost whenever wrong skin regions are detected. Year 2000: Halawani used subtractive clustering algorithm and least- squares estimator are used to identify fuzzy inference system[7]. Taniguchi, Arita, and Igi used Eigen method to detect hand shapes which was appearance based. Using a clustering technique, they generate clusters of hands on an eigen space. They have achieved accuracy of around 93%. Year 2000: Symeoinidis used orientation histograms to detect static hand gestures, specifically, a subset of American Sign Language. . Year 2003: Nielsen and Tezera[9] have developed hand posture detection and recognition approach using vision based.

Later a project in Oxford University using a red color wrist band was used to recognize hand from image and fast template matching was used for pattern recognition with an accuracy of 99.1%[10].

3. PROPOSED METHOD

This proposed method uses 2-D image captured from normal digital camera, webcam as input. An example of an input image is shown in Fig.1.Using Watch detection method, the watch is detected in the image and image is preprocessed. If watch is not detected then a different methods are used. If watch is detected then our proposed method can be applied on the preprocessed image. Then image is matched using the technique pattern recognition. After that quantization of image is done. Tree technique described in detail below is used to reduce time complexity and unnecessary matching's done and accordingly winning neuron will be chosen. The output can be audio or textual[10].

3.1 WATCH DETECTION

For detecting watch in the image three sets namely A,B,C are taken.

Set A {all pixels of image}
Set B {all skin color pixels}
Set C {wrist watch color pixels}
Then,
Preprocessed image = $A \cap (B \cup C)$
Wrists watch pixels can be detected as between the line passing through the adjacent pixels of skin color and a different color (watch) will be nearly parallel to each other as shown below[10].

FIG. 1 Wrist Watch

.3.2 TREE TECHNIQUE

Many symbols are similar and have a bit difference and some are totally different from others. In this technique we group visually similar symbols in groups and the input image is matched with the group to which it is most likely to be matched. This increases the possibility and chances of match to be increased.

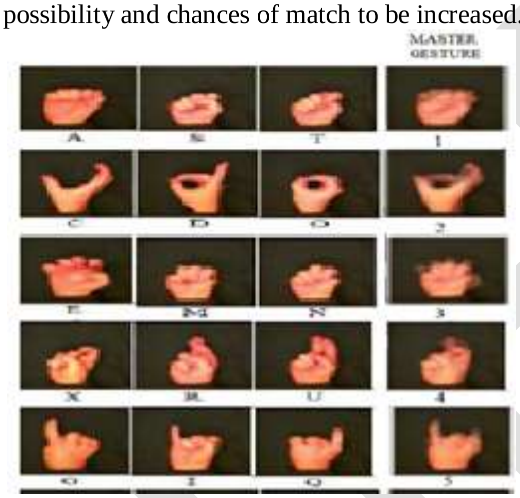


Fig.2 Master Gesture

Here we have divided 26 alphabets into 9 sets with 3 and less than three gestures per set. All the 3 gestures within a set are superimposed over each other to gives a **master gesture** representing that set. Then the input is matched with 9 master gestures of all sets. The closest match with a master gesture provides a security that the gesture lies in that particular set. Best case will have 2 matches for output and in worst case there will be 11 to 12 comparisons for 26 alphabets, average case will be 6-7 comparisons[10].
We took four countries namely India, Australia, Irish and South Africa. The original difficulty of SL recognition is aggravated by the fact that there are many variety of SL sets exist in the world for even a single language such as English. Each country has its own symbols and gestures. Thus, it is impossible for users of different SL groups to understand each other. So by defining a Global SL translator system that is capable of providing recognition and interpreting of signs from different regions/countries will be useful for deaf communities. So if we not make this we have to do 104 comparisons for 4 countries gestures and it is obviously time consuming process and would not be applicable for real time system. So we have to make use of this proposed method so as to reduce no. of comparisons and to make this process real time.

3.5 ARCHITECTURAL DESIGN

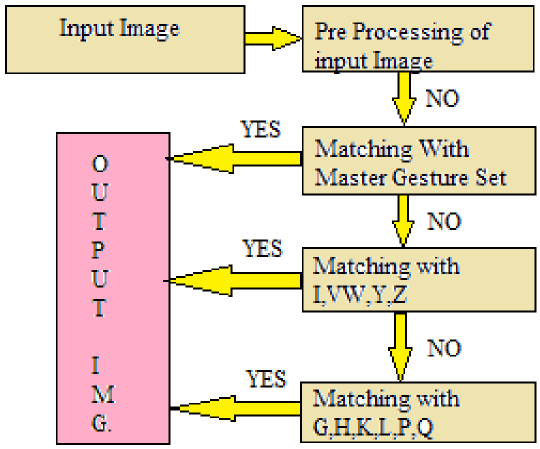


Fig.4 Architectural Design

For every country we know that SL gestures are different. So we are developing a language platform recognizer. We consider here four languages .In 3 languages all alphabets are same except 6 alphabets which are different. So that means 20 alphabets are same and remaining 6 alphabets which are different are given below-

G, H, K, L, P, Q

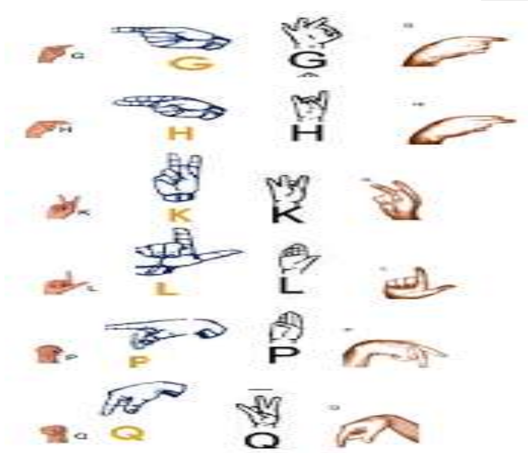


Fig.3 Signs Language Gestures

In three countries there are same gestures for same alphabets so we will take one image in our data set for matching that and in one country there are few different gestures for various alphabets so we will take two images for those alphabets one to represent those three countries in which it is same an one of that different one. Now we have to compare for six alphabets with two images as discussed above. Further we will take the concept of Master Gesture. Five Master Gesture we can make for fifteen alphabets and five alphabets remained that different that they can't form a MG set (I, V, W, Y, Z). **MG** set are

1. A S T
2. C D O
3. E M N
4. R U X

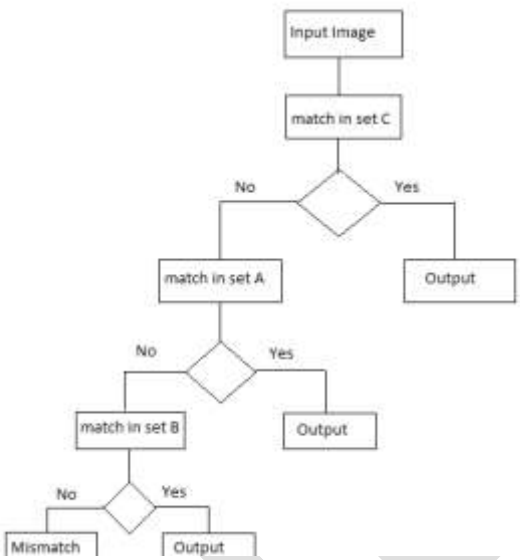
5. J F B

So now totally we have:-
Five alphabets which are same in all four countries but no master gesture exists for them.
Six alphabets which have two images in dataset for their recognition. Fifteen alphabets in five different master gesture sets. So total one hundred four gestures of four different countries can now be compared in:-
Best case: - 2 comparisons.
Worst case: - 22 comparisons.
Average case: - 12 comparisons.

3.4 WINNING NEURON

After the selection of group to which the input lies comes the choice of winning neuron. In winning neuron technique there come only one output which is most similar in nature using kohonen algorithm. The output can be in audio and textual manner whatever desired.

4. MATCHING PROCESS



CONCLUSION

This research has a new idea for the purpose of recognition of global sign languages with an attempt of minimizing computational complexity. This proposed approach overcomes various limitations of previously used techniques like gloves, multicolored gloves, and sensor gloves. Method summarized in this paper, is helpful in human-machine interface. By using Tree technique, time complexity gets reduced that results in reduction in number of comparisons.

FUTURE SCOPE

Methodology proposed in this paper will be implemented in our next research paper. Vision based control and access can lead to overcome several basic limitations of computers like mouse & keyboard.

REFERENCES:

[1] Syed Atif Mehdi & Yasir Niaz Khan, “SIGN LANGUAGE RECOGNITION USING SENSOR GLOVES”, Proceedings of the 9th International Conference on Neural Information Processing (ICONIP‘02), Vol. 5
[2] M. B. Waldron and S. Kim, “Isolated ASL sign recognition system for deaf persons”. IEEE transactions on Rehabilitation Engineering, 3(3):261-71, September 1995

[3] T. Starner. and A. Pentland, " Visual Recognition of American Sign Language Using Hidden Markov Models", Technical Report TR306, Media Lab, MIT, 1995.

[4] Kim, W. Jang and Z. Bien, "A Dynamic Gesture Recognition System for the Korean Sign Language (KSL)", IEEE Trans. On Systems, Man, and Cybernetics, Vol. 26, No. 2, pp. 354-359, 1996.

[5] R.-H. Liang, "Continuous Gesture Recognition System for Taiwanese Sign Language", Ph.D Thesis, National Taiwan University, Taiwan, R.O.C., 1997.

[6] Shin –Han Yu,Chung-Lin Huang,Shih-Chung Hsu & Hung-Wei Lin,andHau- Wei Wang "Vision based continuous sign language reconition using product HMM",IEEE, 2011.

[7] Omar Al-Jarrah ,AlaaHalawani "recognition of gstures in Arabic sign language", Elsevier Science , 2001

[8] B.Lekhashri & A.ArunPratap (2011), "Use of Motion-Print in Sign Language Recognition", Proceedings of the National Conference on Innovations in Emerging Technology-2011,Kongu Engineering College, Perundurai, Erode, Tamilnadu, India.17 & 18 February, 2011.pp.99-102.

[9] Ray Lockton & Balliol College Oxford University "Hand Gesture Recognition Using Computer Vision".

[10] Sonal Singh & Alpika Tripathi "gesture recognition using wrist watch"International journal of scientific & engineering research".

[11] V.I. Pavlovic, R. Sharma, T.S. Huang, "Visual interpretation of hand gestures for human–computer interaction": A review, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (1997) 677–695.

[12] Iwan Njoto Sandjaja & Nelson Marcos (2009) "Sign Language Number Recognition**",** 2009 Fifth International Joint Conference on INC, IMS and IDC.

[13] B.Lekhashri & A.Arun Pratap(2011), "Use of Motion-Print in Sign Language Recognition", Proceedings of the National Conference on Innovations in Emerging Technology-2011,Kongu Engineering College, Perundurai, Erode, Tamilnadu, India.17 & 18 February, 2011.pp.99-102.

[14] Ray Lockton & Balliol College Oxford University "Hand Gesture Recognition Using Computer Vision"

[15] A. BrafTort. ARGO: An architecture for sign language recognition and interpretation. In P. A. Hading, A. D. N. Edwards, editors, Progress in gestural interaction. Proceedings of Gesture Workshop '96, pp. 17-30. Springer, Berlin, New York 1997.

[16] Ankit Chaudhary, J. L. Raheja, Karen Das, Sonia Raheja, " Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way", International Journal of Computer Science & Engineering Survey (IJCSES) Vol.2, No.1, Feb 2011

[17] Rafiqul Zaman Khan and Noor Adnan Ibraheem" HAND GESTURE RECOGNITION", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012