

Different Techniques for Hand Gesture Recognition and Speech Conversion: A Survey

Shruti S. Mohrir¹, Prof. Jaya H.Dewan²

¹Department of Information Technology, Pimpri Chinchwad College of Engineering

²Department of Information Technology, Pimpri Chinchwad College of Engineering

1mohrirshruti@gmail.com, 2 jaya.h.dewan@gmail.com

Abstract

In the era of digital communication, people have various ways of communicating with each other. However, this is not the case with hearing impaired people. Hearing impaired people have no other way except for hand gestures to communicate with hearing-abled people. Digital communication is also possible if the way of communication between the communicators is same. In many parts of the world, Sign languages have different vocabulary as compared to normal language. American Sign Language is still a better-standardized version over Indian Sign Language. In emergencies such as accidents, hospitalization, it becomes very tedious for a people to communicate with a hearing-impaired person. Hearing impaired people face difficulties in every day situation due to this communication barrier. Providing a device or a way of communication that bridges this gap between a normal person and a hearing-impaired person would make their life even simpler. Any emergency can be efficiently tackled, if we have an interpreter that could understand the sign language and make us understand the meaning behind the gestures. This survey paper represents a study of some techniques for recognition of Hand Gesture to precisely interpret the hand gestures and covert them to speech.

Keywords: *Gesture Recognition, ASL, Motion synthesis, vision-based gesture, sensor-based gesture.*

1. Introduction

People communicate with each other on regular basis. Normal people communicate verbally and through gestures or a combination of the two. Whereas in case of a hearing impaired person the communication is solely done through gestures. This communication is also possible only if the two people communicating are trained under a common language. A person can communicate using a Sign Language with the help of hand gestures and some other non-verbal behaviors or actions simultaneously in order to convey their intended speech. A sign language can be effectively communicated when we combine the finger movements, the hand movements, orientation, arms and movement of the body along with facial expressions. In a Recognized Sign Language like American Sign Language, every gesture or hand movement or the combination of them is uniquely associated with a meaning. Hand gesture can be recognized using vision-based approach or sensor-based approach. Sign languages also differ on basis of the region where it is used. The Gestures used in communication are not only used for Human to Human interaction but are also found to be effective for Human to Machine as well as Machine to machine communication. Sign languages for countries differ from each other. These hand gestures used for communication can be still frames i.e. static in nature or can be of a continuous nature. Static hand gestures have a less robust application as compared to dynamic hand gestures. Any language will be a combination of sentences and a sentence will be a combination of words. Words can be interpreted through static hand gestures. However, a sentence will have dynamic or continuous hand gesture stream. Hand and finger movement in gesture recognition when given as a dynamic input will make the use of human-computer Interaction as a major part in the implementation. This will be same for vision-based approach. Whereas in case of sensor based techniques, sensors are used to capture the hand movement through gloves. This paper provides a study of various techniques used in hand gesture recognition that provide apt

results and are robust, cost-effective and easy to use.

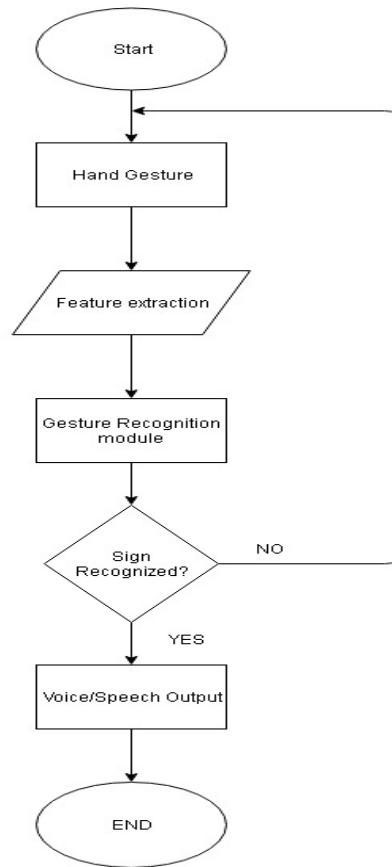


Figure 1. Flowchart for General methodology of Gesture Recognition System.

2. Literature Survey

Aarthi M, Vijayalakshmi P et al; in their paper [1] presented where the authors have studied various gesture recognition system and have categorized them broadly into system that are based on the vision as an input and systems that make the use of sensors to the input. They have emphasized on the study of systems that make the use of sensors. The basic presentation in this paper is of two modules i.e. Finger recognition system and voice synthesis module.

The author has proposed a system that uses a combination of flex sensors and tactile sensors that are embedded onto a glove to get the input for the system. The data received from these sensors along with the accelerometer are further processed with the help of a microcontroller. The hand movement and orientation while performing a gesture is captured by the accelerometer which senses the tilt. When there is a hand movement, a certain amount of bend is enforced on the sensors. The resistance value in the Flex sensor would vary according to this amount of bend. A tactile sensor is a robust polymer thick film device whose resistance changes when a force is applied[1].The resistance of the tactile sensor will vary when the pressure applied on it is variable.

The microcontroller would be given the input which is the output received from the Flex and the Tactile sensors and the accelerometer. Atmega 328 is a microcontroller unit present in the Arduino board[1].

This microcontroller compares the fed-in readings with the standard readings it has in the database/memory, which are the threshold values, and recognizes the corresponding gesture to the value given as input.

The Text-to-Speech module is a statistical parametric approach for speech synthesis based on hidden Markov models (HMMs)[10]. This module will provide the voice output for the recognized gesture.

Jingjing Meng, Junsong Yuan, Zhengyou and Zhang Zhou Ren et al; in their paper[2] have studied and worked on a robust way for finger movements and gesture recognition through Kinect sensors. The paper is built on a novel method for addressing the dissimilarities between the hand gestures. The method is entitled as Finger-Earthmover's Distance (FEMD)[2].

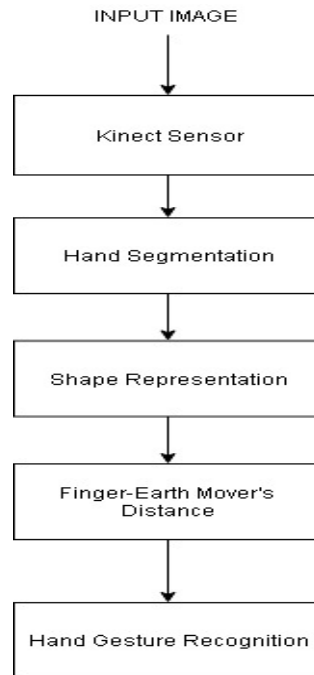


Figure 2. Architecture diagram for hand movement recognition system using FEMD.

The input from the hand movements are captured by the Kinect sensors. These Kinect sensors are not sensitive to distinguish between two fingers if they are very close to one another. These inputs are captured by embedding the sensor on a glove. The input is then evaluated to eliminate the dissimilarities through FEMD for gesture recognition. The authors have tested the module on two different frameworks namely, detection of hand movement and gesture recognition. The system provides a robust hand recognition system over traditional methods.

Swapnil Chandel, Mrs. Akansha Awasthi et al; in their paper[3] have presented a review for Sign language recognition using gestures. They have studied three approaches considering the input parameter: Vision as an input, Color marker as an input and a glove system with sensors as an input.

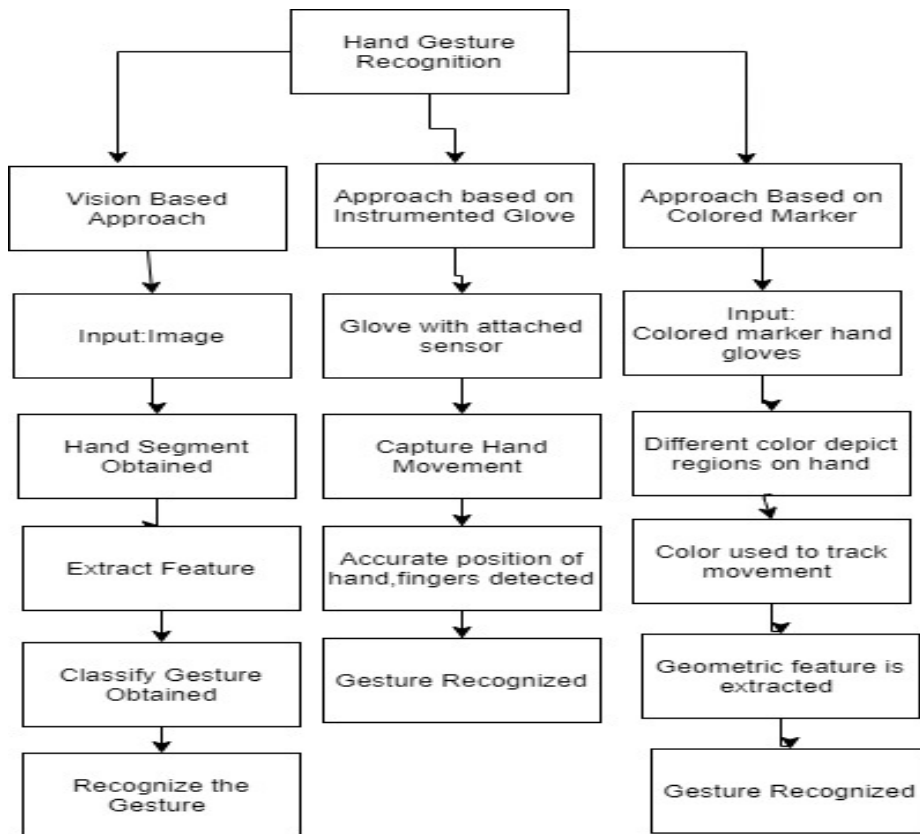


Figure 3. Block diagram for different ways to recognize the gestures

In Vision-based approach a camera is used to feed-in to the system for further processing. The camera captures the gestures of the hand or body part. These gestures captured by the input device are provided for further processing in the form of frames or a series of frames i.e. a video input. This application is most suited when the issue is faced in real time. The vision-based approach also faces many issues. In the appearance-based approach the appearance of hand image is worked-upon visually with the features extracted from the set of features of the images obtained from the previous step.

The Instrumented glove based approach helps to get the accurate orientation and position of the hand and fingers. The disadvantage being that it requires the person who uses it to connect with the computer which makes it not a very practically comfortable technique. The use of sensory gloves makes it more costly than the other methods.

When a color glove is used, various colors are used to depict various areas of the palm. The implementation of the color based method is simple and also cost-effective. Its limitation is that it is also not a natural method of computer human interaction[3]. Computer vision or Human-Computer interaction or Computer-interaction are all efficient ways of dealing with Hand gesture recognition and then processing it to obtain the results as a speech or the equivalent text accurately. Therefore, vision-based approach is more progressed over other methods for hand recognition.

Sh. Saqib Munawwar Qureshi, Syed Muhammad Baber Ali and Syed Faiz Ahmed, et el in their paper[4] have implemented an electronic glove for hand gesture detection for understanding the sign language. The authors have made the use of flex sensors to understand the hand gestures. These flex sensors can sense the movement when they are mounted on the glove. The glove is then worn on the hand, which is performing the gestures. Flex sensor will sense the bend in the hand when a gesture is performed. The bend in the hand will fluctuate the resistance in the flex sensor.

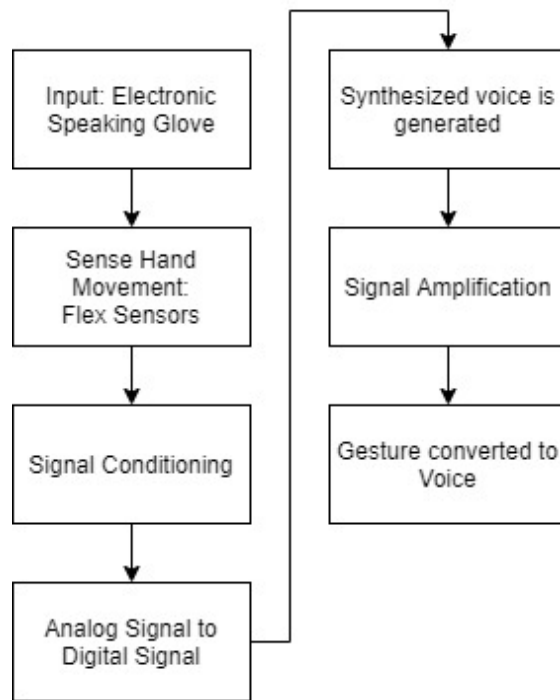


Figure 4. Block Diagram for Electronic Speaking Glove

A signal-conditioning unit is fed the output of the flex sensors. This unit will store the values of the resistance and will give it as an input to the module that uses the micro-controller. This module is responsible for analog to digital signal. The output from the microcontroller is fed to the Speak Jet that has MSA (Mathematical Sound Architecture), predefined allophones that are spoken according to the data received from AVR microcontroller[4]. The synthesized voice output from SpeakJet is not much audible to human ears therefore; it is fed to an amplifier that enhances its volume[4]. Hence the authors have provided a low power consuming portable device for Hand Movement recognition.

3. Comparative Analysis

Parameters	Arthi M et al[1]	Zhou Ren et el[2]	Swapnil chandel[3]	Syed Faiz Ahmed et el[4]
Year	2016	2013	2016	2010
Applications	Vision based system	Vision based system	Review on various systems	Vision based system
Sensor	Flex and tactile are used	Kinect sensors are used	Flex, tactile and Kinect sensors are studied	Flex sensor are used
Methodology	The study of sensor-based system is emphasized	A novel method for addressing the dissimilarities between hand gestures. The method used is	The methodology used considers the parameters : Vision as an input, Color marker as an input and a glove system	A glove-based system is implemented that uses an electronic glove where flex

		known as Finger-Earthmover's Distance.	with sensors as an input	sensors are used for gesture recognition in Sign Language communication.
Algorithms	Hidden Markov model	Finger-Earthmover's Distance.	Comparative study of various techniques	Mathematical Sound Architecture for audio output.
Focused Factors/ Criteria	Full degree of freedom for hand movement	Kinect sensors are widely used to recognize gestures in a part-based system.	HCI for hand gesture recognition based on vision inputs.	Cost effective, Portable, Less weight, Less hardware, Robust, Fully embedded.
Accuracy	87.5%	93.2%	Vision-based system more likely to be used the systems	Found more efficient over traditional glove based system

Table 1. Comparative analysis of various methods surveyed in the Literature Survey

4. Conclusion

Gesture recognition and interpretation provides a better alternative over traditional human interpreters. It provides a medium for deaf-dumb to communicate with the world in an independent manner. Emergencies will be well addressed with the help of this electronic interpreter.

References

1. Aarthi M ,Vijayalakshmi P “Sign Language to speech conversion” 2016 fifth international conference on recent trends in information technology.
2. Zhou Ren, Junsong Yuan, Jingjing Meng, Zhengyou Zhang, “Robust Part-Based Hand Gesture Recognition Using Kinect Sensor” IEEE TRANSACTIONS ON MULTIMEDIA, VOL.15 ,NO.5, AUGUST 2013.
3. Swapnil Chandel and Mrs. Akanksha Awasthi “Hand Gesture Recognition for Sign Language Recognition: A Review” IJSTE - International Journal of Science Technology & Engineering | Volume 2 | Issue 10 | April 2016.
4. Syed Faiz Ahmed, Syed Muhammad Baber Ali, Sh. Saqib Munawwar Qureshi “ Electronic Speaking Glove for Speechless Patients A Tongue to a Dumb” Proceedings of the 2010 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology Universiti Tunku Abdul Rahman November 2010
5. Anbarasi Rajamohan, Hemavathy R., Dhanalakshmi M. “DeafMute Communication Interpreter” International Journal of Scientific Engineering and Technology, May 2013.
6. Priyanka Lokhande, Riya Prajapati and Sandeep Pansare, “Data Gloves for Sign Language Recognition System” IJCA Proceedings on National Conference on Emerging Trends in Advanced Communication Technologies NCETACT, pp. 11-14, Jun 2015

7. Rajam, P. Subha and Dr G Balakrishnan, "Real Time Indian Sign Language Recognition System to aid Deaf and Dumb people", 13th International Conference on Communication Technology (ICCT), 2011, pp 737-742.
8. Mokhtar M. Hasan, Pramoud K. Misra, (2011). "Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization", International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2).
9. Vinod J Thomas, Diljo Thomas, "A Motion based Sign Language Recognition Method", International Journal of Engineering Technology, vol 3 Issue 4, ISSN 2349-4476, Apr 2015.
10. K. Tokuda, H. Zen, A.W. Black, An HMM-based speech synthesis system applied to English, Proc. of 2002 IEEE Speech Synthesis Workshop, Sep. 2002, pp 227-230.
11. Noor Adnan Ibraheem, RafiqulZaman Khan, "Survey on Various Gesture Recognition Technologies and Techniques", International Journal of Computer Applications, Vol. 50, No.7, July 2012, pp. 38-44.