NOVATEUR PUBLICATIONS
International Journal of Research Publications in Engineering and Technology [IJRPET]
ISSN: 2454-7875
VOLUME 2, ISSUE 9, September -2016

REAL TIME HAND GESTURE RECOGNITION AND VOICE CONVERSION SYSTEM FOR DEAF AND DUMB PERSON BASED ON IMAGE PROCESSING

SHWETA SONAJIRAO SHINDE
M.E Student. Electronics and Telecommunication Department, Deogiri Institute of Engineering and Management Studies, Aurangabad (MS), India

Dr. R.M. AUTEE
H.O.D. Electronics and Telecommunication Department, Deogiri Institute of Engineering and Management Studies, Aurangabad (MS), India

ABSTRACT:
Communication between normal and handicapped person such as deaf people, dumb people, and blind people has always been a challenging task. It has been observed that they find it really difficult at times to interact with normal people with their gestures, as only a very few of those are recognized by most people. Since people with hearing impairment or deaf people cannot talk like normal people so they have to depend on some sort of visual communication in most of the time. Sign Language is the primary means of communication in the deaf and dumb community. As like any other language it has also got grammar and vocabulary but uses visual modality for exchanging information. The importance of sign language is emphasized by the growing public approval and funds for international project. Interesting technologies are being developed for speech recognition but no real commercial product for sign recognition is actually there in the current market. So, to take this field of research to another higher level this project was studied and carried out. The basic objective of this research is to develop MATLAB based real time system for hand gesture recognition which recognize hand gestures, features of hands such as centroid, peak calculation, angle calculation and convert gesture images into voice and vice versa. To implement this system we used simple night vision web-cam with 20 megapixel intensity. The idea consisted of designing and building up an intelligent system using image processing, data mining and artificial intelligence concepts to take visual inputs of sign languages hand gestures and generate easily recognizable form of outputs in the form of text and voice with 82% accuracy.

KEYWORDS: hand gesture recognition, voice conversion, gesture to speech, speech to gesture conversion.

I. INTRODUCTION:

One of the important problems that our society faces is that people with disabilities are finding it hard to come up with the fast growing technology. In the recent years, there has been a rapid increase in the number of hearing impaired and speech disabled victims due to birth defects, oral diseases and accidents. When a deaf-dumb person speaks to a normal person, the normal person seldom understands and asks the deaf-dumb person to show gestures for his/her needs. Dumb persons have their own language to communicate with us; the only thing is that we need to understand their language. Generally dumb people use sign language for communication but they find difficulty in communicating with others who don't understand sign language. For better communication between deaf and normal people we proposed a system which converts gesture images into speech and vice versa.

The sign language translation system translates the normal sign language to speech and hence makes the communication between normal person and dumb people easier. Many research works related to Sign languages have been done as for example the American Sign Language, the British Sign Language, the Japanese Sign Language, and so on. Finding an experienced and qualified interpreters every time is a very difficult task and also unaffordable. Automated speech recognition system which aims to convert the speech signals into text form. Hence the two way communication is possible between deaf-mute person and normal person. Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. Gestures can originate from any bodily motion or state but commonly originate

from the face or hand. Current focuses in the field include emotion recognition from the face and hand gesture recognition. The approaches present can be mainly divided into "Data-Glove based" and "Vision Based" approaches. The Data-Glove based methods use sensor devices for digitizing hand and finger motions into multi-parametric data. The extra sensors make it easy to collect hand configuration and movement. However, the devices are quite expensive and bring much cumbersome experience to the users. In contrast, the Vision Based methods require only a camera, thus realizing a natural interaction between humans and computers without the use of any extra devices.

Hand gesture recognition is of great importance for human-computer interaction (HCI), because of its extensive applications in virtual reality and sign language recognition. Speech recognition is a broad subject, the commercial and personal use implementations are rare. Although some promising solutions are available for speech synthesis and recognition, most of them are tuned to English. There are several problems that need to be solved before speech recognition can become more useful. The amount of pattern matching and feature extraction techniques is large and the decision on which ones to use is debatable. After studying the history of speech recognition we found that the very popular feature extraction technique Mel frequency cepstral coefficients (MFCC) is used in many speech recognition applications and one of the most popular pattern matching techniques in speaker dependent speech recognition is Dynamic time warping (DTW). The signal processing techniques, MFCC and DTW are explained and discussed in detail and these techniques have been implemented in MATLAB.

II. RELATED WORK:

Some recent reviews explained number of applications on gesture recognition and its growing importance in our daily life especially for Human computer Interaction HCI, games, Robot control and surveillance, using different tools and algorithms. This work demonstrates the advancement of the gesture recognition systems, with the discussion of many stages are essential to build a complete system with less erroneous using different algorithms

Table 1: Comparison between existing systems

Author(s)	Name of research paper	Feature extraction and classification methods	Remark
Aditi Kalsh and N.S. Garewal[16]	Sign Language Recognition for Deaf & Dumb	gray scaling, edge detection and peak calculation	After the number of peaks is detected simple if-else rule is applied for gesture recognition.
Deepika Tewari and Sanjay Kumar Srivastava[17]	A Visual Recognition of Static Hand Gestures in Indian Sign Language based on Kohonen Self-Organizing Map Algorithm	Self-Organizing Feature Map and Artificial neural network	DCT-based feature vectors are classified to check whether sign mentioned in the input image is present or not present.
Bhupinder Singh[18]	Speech recognition with Hidden Markov model	Hidden Markov Model	Develop a voice based user machine interface system
Kavita Sharma[19]	Speech Denoising using Different Types of filters	FIR,IIR,WAVE-LETS,FILTER	Use of filter shows that estimation of clean speech and noise for speech enhancement in speech recognition.
Ibrahim Patel[20]	Speech Recognition Using HMM with MFCC-an analysis using Frequency Spectral Decomposition	Resolution decomposition with separating Frequency is the mapping approach	It shows an improvement in the quality metrics of Speech Recognition with respect to computational time, learning accuracy for a speech recognition system.

III. PROPOSED SYSTEM:

Figure drawn below shows the basic block diagram for hand gestures to speech conversion System

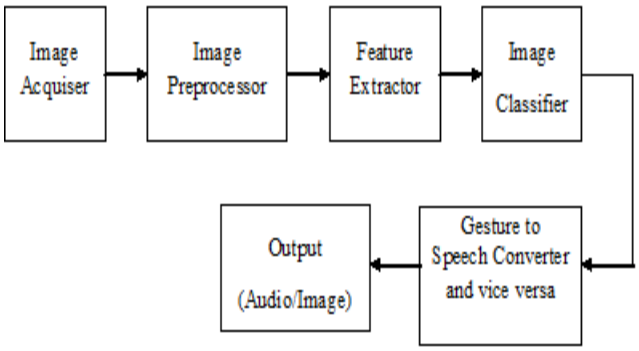

Figure1: Block diagram of hand gestures to Speech conversion System

Image Acquiser: There are a number of input devices for image acquisition. Some of them are hand images, data gloves, markers and drawings. In this system image is acquired by using 20 mega pixel web cam using MATLAB inbuilt command.

Image Preprocessor: Image preprocessing is very important step for image enhancement and for getting good results. The RGB images are captured using a 20 MP webcam. The input sequence of RGB images are converted into gray Images. Background segmentation is used to separate the hand object in the image from its background. Noise elimination steps are applied to remove connected components or insignificant smudges in the image that have fewer than P pixels, where P is has a variable value.

Feature Extractor: For feature Extraction different features like binary area, centroid, peak calculation, angle calculation, thumb detection, finger region detection of the hand region is calculated.

Image Classifier: 12 bit binary sequence is generated for each hand gesture which classifies the different hand gestures.

SPEECH TO GESTURE CONVERSION SYSTEM:

Speech recognition or Automatic Speech Recognition (ASR) is an essential and integral part of the human computer interaction and for a natural and ubiquitous computing speech plays an important part. It is the process of converting a speech signal to a set of words, by means of an algorithm implemented as a computer program. Voice Recognition based on the speaker can be classified into two types namely: Speaker-dependent and Speaker-independent. Figure 2 shows the speech to gesture conversion system.

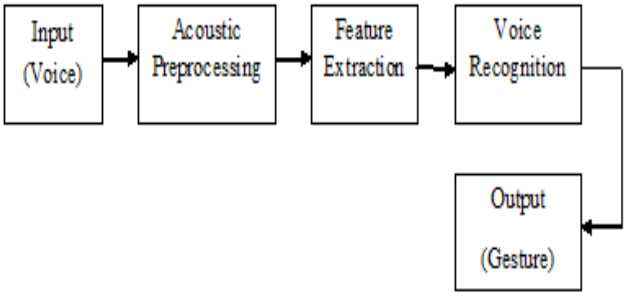

Figure 2: Block Diagram of Speech to Gesture Conversion

IV. RESULT:

RESULTS FOR GESTURES TO SPEECH CONVERSION SYSTEM:

Table 2: Percentage Accuracy of different Gestures

Different Gestures	No. of test	No .of correct test	error	Percentage of accuracy
A	5	4	1	80%
B	5	5	0	100%
C	5	5	0	100%
D	5	4	1	80%
E	5	3	2	60%
F	5	5	0	100%
G	5	4	1	80%
H	5	2	3	40%
I	5	4	1	80%

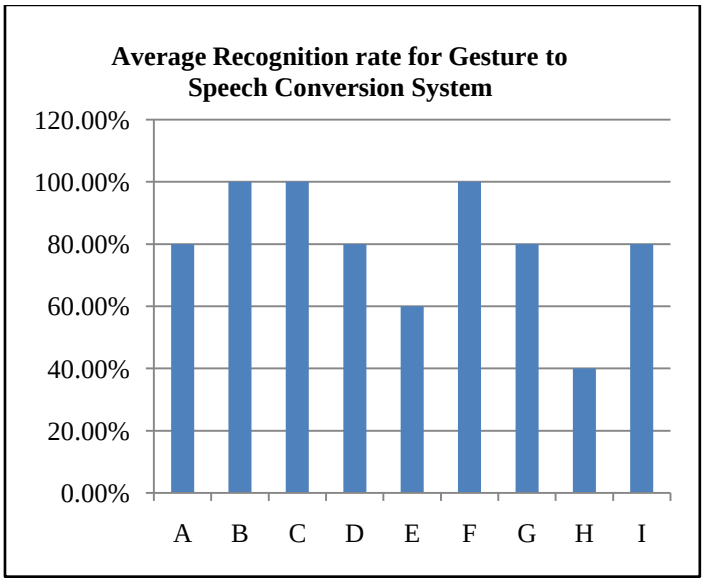

Figure 3: Average recognition rate for Gesture to Speech Conversion System

RESULTS FOR GESTURES TO SPEECH CONVERSION SYSTEM:

Table 3: Percentage Accuracy of different sound samples

Different Speech samples	No. of test Conduct	No. of correct matching test	Error	Percentage of accuracy
A	5	5	0	100%
B	5	4	1	80%
C	5	4	1	80%
D	5	5	0	100%
E	5	3	2	60%
F	5	4	1	80%
G	5	4	1	80%
H	5	4	1	80%
I	5	5	0	100%

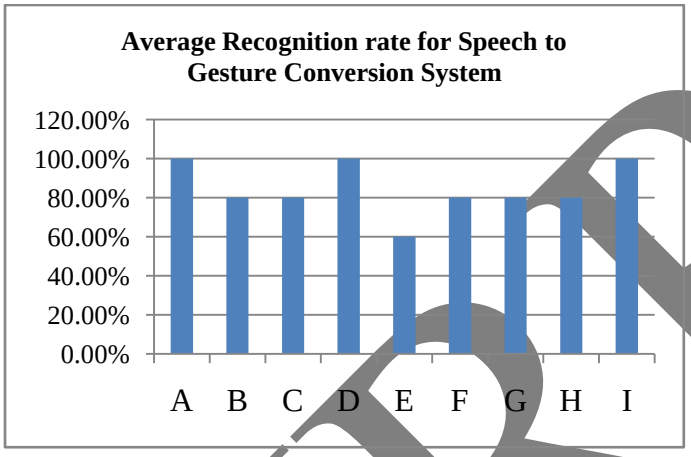

Figure 4: Average Recognition rate for Speech to Gesture Conversion System

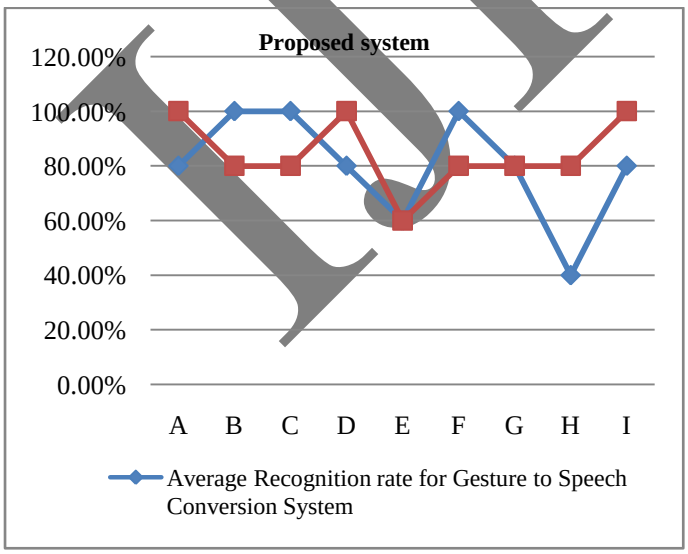

Figure 5: Recognition Result of Proposed System

V. CONCLUSION

The proposed system is very simple and easy to implement as there is no complex feature calculation, no significant amount of training or post- processing required. This system provides us with high gesture recognition rate with accuracy 80% within minimal computation time. Sign language is a useful tool to ease the communication between the deaf person and normal person. The system aims to lower the communication gap between deaf people and normal world, since it facilitates two way communications. The projected methodology interprets language into speech. The system overcomes the necessary time difficulties of dumb people and improves their manner. This system converts the language in associate passing voice that's well explicable by deaf people. With this project the deaf-mute people can use the hand gestures to perform sign language and it will be converted into speech with accuracy 84%; and the speech of normal person is converted into text and corresponding hand gesture, so the communication between them can take place easily, so the proposed system gives 82% accuracy. There is need of research in the area feature extraction and illumination so the system becomes more reliable. This system that enables impaired people to further connect with their society and aids them in overcoming communication obstacles created by the society's incapability of understanding and expressing sign language.

REFERENCES:

[1] Sakshi Goyal, Ishita Sharma, Shanu Sharma , "Sign Language Recognition System For Deaf And Dumb People", International Journal of Engineering Research & Technology (IJERT), Vol. 2, Issue 4,April 2013, pp. 382-387.

[2] 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.

[3] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, "Vision Based Hand Gesture Recognition",World Academy of Science, Engineering and Technology 49, 2009, pp. 972- 977.

[4] Mitra, S., & Acharya, T. Gesture recognition: A survey. IEEE Transactions on systems. Man and Cybernetics,Part

C: Applications and reviews, 37(3), 311-324. doi:10.1109/TSMCC.2007.893280,2007.

[5] Dipietro, L., Sabatini, A. M., & Dario, P. "Survey of glove-based systems and their applications". IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, 2008, 38(4), 461-482. http://dx.doi.org/10.1109/TSMCC.2008.923862.

[6] Lamberti, L., & Camastra, F. "Real-time hand gesture recognition using a color glove", 2011Springer-Verlag Berlin Heidelberg. ICIAP 2011, Part I, LNCS 6978, pp. 365-373.Retrieved from www.springerlink.com/index/02774HP03V0U4587.pdf.

[7] J. Joseph LaViola Jr., "A Survey of Hand Posture and Gesture Recognition Techniques and Technology", Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization, USA1999..

[8] Vajjarapu Lavanya, M.S. Akulapravin and Madhan Mohan, "Hand Gesture Recognition and Voice Conversion System Using Sign Language Transcription System", IJECT Vol. 5, Issue 4, Oct - Dec 2014.

[9] K. Sangeetha, and L. Barathi Krishna, "Gesture Detection For Deaf and Dumb People ", International Journal of Development Research Vol. 4, Issue, 3, pp. 749-752, March, 2014.

[10] S. Padmavathi , M.S.Saipreethy and V.Valliammai, "Indian Sign Language Character Recognition using Neural Networks ", IJCA Special Issue on "Recent Trends in Pattern Recognition and Image Analysis", RTPRIA 2013 .

[11] Sheng-Yu Peng, Kanoksak Wattanachote, Hwei-Jen Lin, Kuan-Ching Li, ―A Real-Time Hand Gesture Recognition System for Daily Information Retrieval from Internet‖, IEEE Fourth International Conference on Ubi-Media Computing, 978-0-7695-4493-9/11 © 2011.

[12] C.Maggioni and B. Kämmerer , "Gesture Computer – history, design and applications", Computer Vision for Human-Machine Interaction, Cambridge University Press,1998.

[13] R.Cutler and M. Turk,"View-based interpretation of real-time optical flow for gesture recognition", Proc. Third International Conference on Automatic Face and Gesture Recognition. Nara, Japan, 1998.

[14] W.Freeman,K. Tanaka ,J. Ohta , and K. Kyuma "Computer vision for computer games", Proc. Second International Conference on Automatic Face and Gesture Recognition. Killington, VT. 1996.

[15] M. Stark and M. Kohler, "Video based gesture recognition for human computer interaction", In W. D.1995.Fellner (Ed.), Modeling - Virtual Worlds - Distributed Graphics, 1995.

[16] Aditi Kalsh and N.S. Garewal ,"Sign Language Recognition for Deaf & Dumb", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 9, September 2013.

[17] Deepika Tewari, Sanjay Kumar Srivastava , "A Visual Recognition of Static Hand Gestures in Indian Sign Language based on Kohonen Self- Organizing Map Algorithm", International Journal of Engineering and Advanced Technology (IJEAT), Vol.2, Issue 2, pp. 165-170.Dec 2012.

[18] Bhupinder Singh, Neha Kapur, Puneet Kaur "Sppech Recognition with Hidden Markov Model:A Review" International Journal of Advanced Research in Computer and Software Engineering, Vol. 2, Issue 3, March 2012.

[19] Kavita Sharma, Prateek Hakar "Speech Denoising Using Different Types of Filters" International journal of Engineering Research and Applications Vol. 2, Issue 1, Jan-Feb 2012

[20] Ibrahim Patel, Dr. Y. Srinivas Rao, "Speech Recognition using HMM with MFCC-an analysis using Frequency Spectral Decomposition Technique" Signal and Image Processing:An International Journal(SIPIJ), Vol.1, Number.2, December 2010. [19] Pratt, W. K.,"Digital Image Processing", Fourth Edition, A John Wiley & Sons Inc. Publication,2007.

[21] Jingdong Chen, Member, Yiteng (Arden) Huang, Qi Li, Kuldip K. Paliwal, "Recognition of Noisy Speech using Dynamic Spectral Subband Centroids" IEEE SSIGNAL PROCESSING LETTERS, Vol. 11, Number 2, February 2004.

[22] Hakan Erdogan, Ruhi Sarikaya, Yuqing Gao, "Using semantic analysis to improve speech recognition performance" Computer Speech and Language, ELSEVIER 2005.

[23] Finlayson, G. D., Qiu, G., Qiu, M., Contrast Maximizing and Brightness Preserving Color to Grayscale Image Conversion,1999.

[24] Parveen, N. R.S., Sathik, M.M., Gray Scale Conversion Of X-Ray Image, ICISA 2010, and Chennai, 2010, India.

[25] Y. Benezeth; B. Emile; H. Laurent; C. Rosenberger ,"Review and evaluation of commonly-implemented background subtraction algorithms",. International Conference on Pattern Recognition, Dec 2008, pp. 1–4. Doi:10.1109/ICPR.2008.4760998