VISION-BASED SIGN LANGUAGE TRANSLATION DEVICE

Yellapu Madhuri[1], Anitha.G[2], Anburajan.M[3]

[1, 2, 3] *Department of Biomedical Engineering, SRM University,kattankulathur,Tamilnadu,India*

ymadhury@rediffmail.com[1],mail2anitha@gmail.com[2],anbufelix@yahoo.com[3]

***Abstract*__This report presents a mobile VISION-BASED SIGN LANGUAGE TRANSLATION DEVICE for automatic translation of Indian sign language into speech in English to assist the hearing and/or speech impaired people to communicate with hearing people. It could be used as a translator for people that do not understand sign language, avoiding by this way the intervention of an intermediate person and allow communication using their natural way of speaking. The proposed system is an interactive application program developed using LABVIEW software and incorporated into a mobile phone. The sign language gesture images are acquired using the inbuilt camera of the mobile phone; vision analysis functions are performed in the operating system and provide speech output through the inbuilt audio device thereby minimizing hardware requirements and expense. The experienced lag time between the sign language and the translation is little because of parallel processing. This allows for almost instantaneous recognition from finger and hand movements to translation. This is able to recognize one handed sign representations of alphabets (A-Z) and numbers (0-9).The results are found to be highly consistent, reproducible, with fairly high precision and accuracy**.

Keywords:
Image processing, LabVIEW, pattern matching, sign language.

I. INTRODUCTION

Sign Language (SL) is the natural way of communication of hearing and/or speech -impaired people. A sign is a movement of one or both hands, accompanied with facial expression, which corresponds to a specific meaning. Although the deaf, hard of hearing and hearing signers can communicate without problems amongst themselves, there is a serious challenge for the deaf community trying to integrate into educational, social and work environments. The overall goal of this project is to develop a new vision based technology for recognizing and translating continuous sign language to text. To paper presents a mobile interactive application program for automatic translation of Indian sign language into speech in English to assist the hearing and/or speech impaired people to communicate with hearing people. This sign language translator should be able to translate alphabets (A-Z) and numbers (0-9).

Although facial expressions add important information to the emotional aspect of the sign; but in this project work they are excluded from the area of interest, since its analysis complicates the already difficult problem. Our system aims at listening to deaf, which means that it could be used as a translator between deaf and people that do not understand sign language, avoiding by this way the intervention of an intermediate person. Both deaf and people that do not have hearing problems would communicate using their natural way of speaking.

II.METHODOLOGY

We present a real-time vision-based system for recognizing finger spelling continuous Sign Language (ASL) using a single camera to track the user's unadorned hands. This system is broken down into three main parts starting with the image acquisition followed by image processing to extract features for recognition and last comes the recognition stage where signs are identified and audio output is given. This section explains the details of each process. The program starts with image acquisition, i.e. sign images capturing by the camera. The acquired images are pre-processed to differentiate static and dynamic signs, and also the start and end of a sign. The images are processed to identify the region of interest. The unique features of each sign in the region of interest are extracted to be used in the recognition stage. In the recognition stage, the features extracted are compared with the available database of pattern matching templates. A threshold value is set for the maximum difference between the input sign and the database, if the difference is below the maximum limit, a match is found and the sign is recognized. Corresponding audio file is played on audio device. The program can be implemented in a laptop, desktop or an IOS mobile phone to operate with its inbuilt camera, processor and audio device. The experiments use a 36 word lexicon.

2.1 IMAGE ACQUISITION

The gestures of sign language are captured by the inbuilt camera to detect the movement of the hand. Capturing thirty frames per second (fps) is found to be sufficient. Higher fps will only lead to higher computation time of the computer as more input data to be processed. As the acquisition process runs at real time, this part of the process has to be efficient. Thus, previous frame that has been processed will be automatically deleted to free the limited memory space in the buffer.

2.2 ENVIRONMENTAL SET-UP:

Image acquisition process is subjected to many environmental concerns such as the position of the camera, lighting sensitivity and background condition. The camera is placed to focus on an area that can capture the maximum possible movement of the hand and take into account the difference in height of individual signers. Sufficient lighting is required to ensure that it is bright enough to be seen and analyzed.

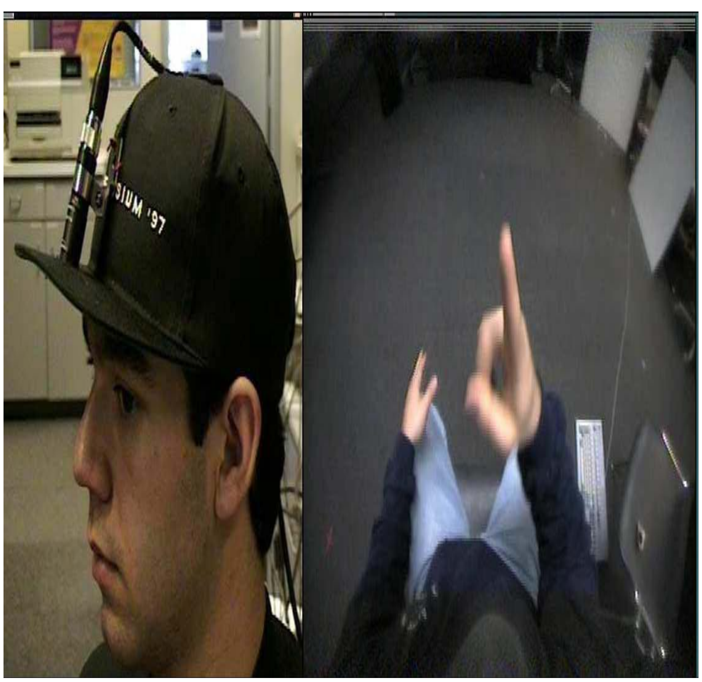

Fig 2.1 Environmental Setup

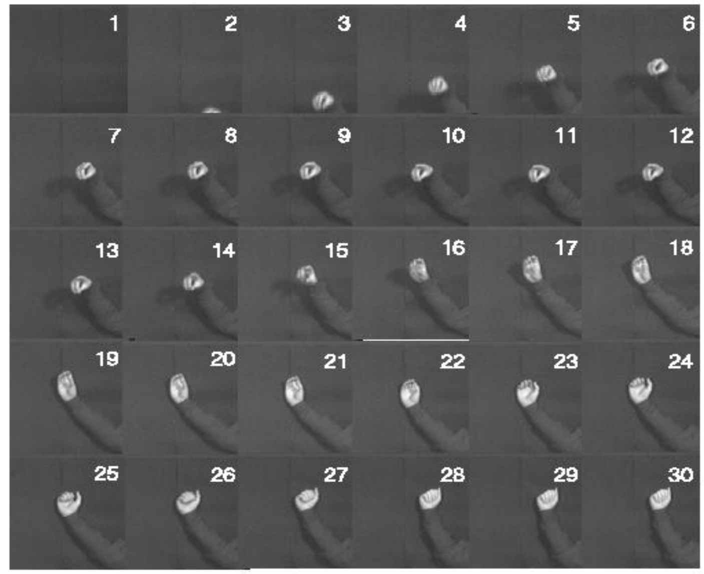

Figure 2.2 Image Acquisition

2.3 IMAGE PROCESSING

Captured images are processed using Vision Assistant to identify the unique features of each sign. Image processing eliminates any object in the background that will interfere in the recognition of the sign. The camera captures images at 30 frames per second. At this rate, the difference between subsequent images will be too small. Hence, the images are sampled at 5 frames per second. In the program, one frame is saved and numbered sequentially every 200 milliseconds so that the image classifying and processing can be done systematically. The position of the hand is monitored. To track the position of the hand, the color region will be segmented using the color thresholding method. Using the color thresholding method, the position of the region of interest can be determined. The image acquisition runs continuously until the signer indicates a sign to stop acquisition.

2.4 FEATURE EXTRACTION

After color thresholding, the segmented images are analyzed to obtain the unique features of each sign. Each sign will have a unique combination of finger tip position. Hence, this will be the only feature needed for recognition. As the frames from acquisition are read one by one, the centroid of each frame are computed and inserted into a $2 \times n$ array, which represents a set *X*, and *Y* coordinates. The $2 \times n$ array is then transposed into an $n \times 2$ array so that it can be passed on to the recognition stage which requires an $n \times 2$ array input.

2.5 SIGN RECOGNITION

The last stage of this project is the recognition stage and providing audio output. The recognition of static sign is done based on the position of the finger in the bounding box. This method of recognition is chosen because each alphabet and number has a unique combination of hand postures. A threshold value is set for the maximum difference between the input sign and the database, if the difference is below the maximum limit, a match is found and the sign is recognized.

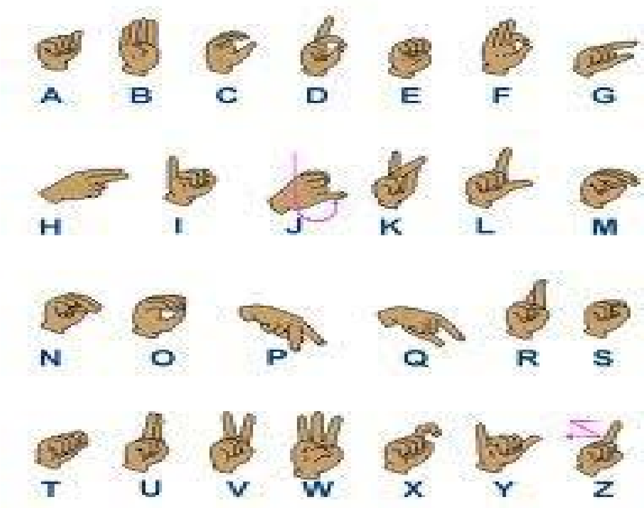

Fig 2.3 Database Of One-Handed Alphabet Signs

Fig 2.4 Database Of One-Handed Number Sign

3. RESULT

This sign language translator is able to translate alphabets (A-Z) and numbers (0-9). All the signs can be translated real-time. But signs that are similar in posture and gesture to another sign can be misinterpreted, resulting in a decrease in accuracy of the system. The current system has only been trained on a very small database. Since there will always be variation in either the signers hand posture or motion trajectory, a larger database accommodating a larger variety of hand posture for each sign is required.

4. CONCLUSION

In this work, a vision based sign language recognition system using LABVIEW for automatic sign language translation has been presented. This approach uses the feature vectors which include whole image frames containing all the aspects of the sign. This project has investigated the different issues of this new approach to sign language recognition to recognize on the hand sign language alphabets and numbers using appearance based features which are extracted directly from a video stream recorded with a conventional camera making recognition system more practical. Although sign language contains many different aspects from manual and non-manual cues, the position, the orientation and the configuration or shape of the dominant hand of the signer conveys a large portion of the information of the signs. Therefore, the geometric features which are extracted from the signers' dominant hand, improve the accuracy of the system to a great degree. Facial expressions are not focused, although it is well known that facial expressions convey important part of sign-languages. A wearable IOS phone system provides the greatest utility for automatic Sign Language to spoken English translator. It can be worn by the signer whenever communication with a non-signer might be necessary, such as for business or on vacation. Providing the signer with a self-contained and unobtrusive first-person view translation system is more feasible than trying to provide second-person translation systems for everyone whom the signer might encounter during the day.

5. FUTURE WORK

To increase the performance and accuracy of the Automatic Sign Language Translator (ASLT), the quality of the training database used should be enhanced to ensure that the ASLT picks up correct and significant characteristics in each individual sign and further improve the performance more efficiently. A larger dataset will also allow experimenting further on performance in different environments. Such a comparison will allow to tangibly measuring the robustness of the system in changing environments and provide training examples for a wider variety of situations. Adaptive color models and improved tracking could boost performance of the vision system.

Current collaboration with Assistive Technology researchers and members of the Deaf community for continued design work is under progress. The gesture recognition technology is only one component of a larger system that we hope to one day be an active tool for the Deaf community. This project did not focus on facial expressions although it is well known that facial expressions convey important part of sign-languages. The facial expressions can e.g. be extracted by tracking the signers' face. Then, the most discriminative features can be selected by employing a dimensionality reduction method and this cue could also be fused into the recognition system. This system can be implemented in many application areas examples include accessing government offices for filling out forms whereby no interpreter may be present to help.

6. REFERENCES

1. Andreas Domingo, Rini Akmeliawati, Kuang Ye Chow 'Pattern Matching for Automatic Sign Language Translation System using LabVIEW', International Conference on Intelligent and Advanced Systems 2007.

2. Beifang Yi Dr. Frederick C. Harris 'A Framework for a Sign Language Interfacing System', A dissertation submitted in partial fulllment of the requirements for the degree of Doctor of Philosophy in Computer Science and Engineering May 2006 University of Nevada, Reno.

3. Helene Brashear & Thad Starner 'Using Multiple Sensors for Mobile Sign Language Recognition', ETH - Swiss Federal Institute of Technology Wearable Computing Laboratory 8092 Zurich, Switzerland flukowicz, junker g@ife.ee.ethz.ch

4. Jose L. Hernandez-Rebollar1, Nicholas Kyriakopoulos1, Robert W. Lindeman2 'A New Instrumented Approach For Translating American Sign Language Into Sound And Text', Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04) 0-7695-2122-3/04 $ 20.00 © 2004 IEEE.

5. K. Abe, H. Saito, S. Ozawa: Virtual 3D Interface System via Hand Motion Recognition From Two Cameras. IEEE Trans. Systems, Man, and Cybernetics, Vol. 32, No. 4, pp. 536–540, July 2002.

6. Paschaloudi N. Vassilia, Margaritis G. Konstantinos "Listening to deaf': A Greek sign language translator', 0-7803-9521-2/06/$20.00 §2006IEEE.