

Recognition of Two Hand Gestures of word in British Sign Language (BSL)

Pingale Prerna Rambhau

Department of Computer Engineering(University of Pune), Sir Vishaveshwaray Institute Technology, Chincholi, Tal-Sinner,Dist- Nashik ,India

Abstract- This paper presents the implementation and analysis of a real time stereo vision hand tracking system that can be used for interaction purpose. In real time, the system can track the 3D position and 2D orientation of thumb and index finger of each hand without the use of special markers and gloves. For someone who can hear, it is difficult to imagine what it is like to be deaf. Many people associate 'deaf with handicapped', but this is not really correct. A handicapped person, like someone who is blind, has trouble functioning independently in everyday life. Deaf people do not have this problem: they can ride a bike, drive a car, go to work, shop, and go on vacation like everybody else. The greatest problem for a deaf person is "Communication". The main motive of the paper is to simplify the communication between normal and deaf people and bring them in normal flow of society. To study the first phase of the finger spelling recognition task involves the segmentation of finger spelling hand gestures from image sequences. The aim of this paper to convert finger spelled word to text (Character by Character) as well as speech. Second phase includes Text to Image conversion process. In this sentences are scanned entered by user and then alphabets separation module separate the alphabets from sentences and corresponding image is display to viewers. Here images are nothing but finger spelling alphabets.

Index Terms- Finger spelled Segmentation, Sign language, 2D Orientation, 3D.

I. INTRODUCTION

Many commercial are being burdened by the problem of recognizing words from video, finger spelled using the British sign language (BSL) [1] finger spelling alphabets. BSL is a visual language used by deaf people, which uses word level sign (gestures), non manual features. E.g. Finger spelling (letter by letter signing) to convey meaning. This is challenging task since the BSL alphabet involves both hand occluding each other, and contains signs which are ambiguous from the observer's viewpoints.

Their intelligence can be used in the various development areas effectively with the help of this paper. Although sign language is the main communication medium of the hearing impaired, in term of automatic recognition, finger spelling has the advantage of using limited number of finger signs, corresponding to the letters/sounds in the alphabets.

The aim of this paper that translate the sign language to speech. Considering the current state of art and project duration, focusing on finger spelling is reasonable choice and will provide insight to next coming project to develop advanced system

Moreover as finger spelling is used in sign out of vocabulary words, the outcome of this project will provide modules that can be reused in a sign language to speech translator. The practical problems of recognizing finger spelling are very different from those of recognizing signs, in finger spelling there is only a fixed number of hand shapes that need to be recognized. Orientation and motion are only marginally important and location is always the same. Non-manual components can be ignored completely. This makes the problem much simpler than that of sign recognition.

In this paper work, there is two phases:

Sign Processing:

The first phase of the finger spelling recognition task involves the segmentation of finger spelling hand gestures from image sequences The aim of this paper to convert finger spelled word to text[1] (Character by Character) as well as speech[3].

Text to Sign Conversion:

Second phase includes Text to Image conversion process. In this sentences are scanned entered by user and then alphabets separation module separate the alphabets from sentences and corresponding image is display to viewers. Here images are nothing but finger spelling alphabets.

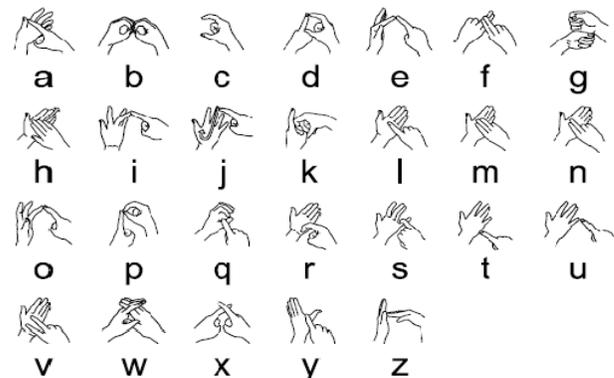


Fig. 1 Standard Alphabets in BSL

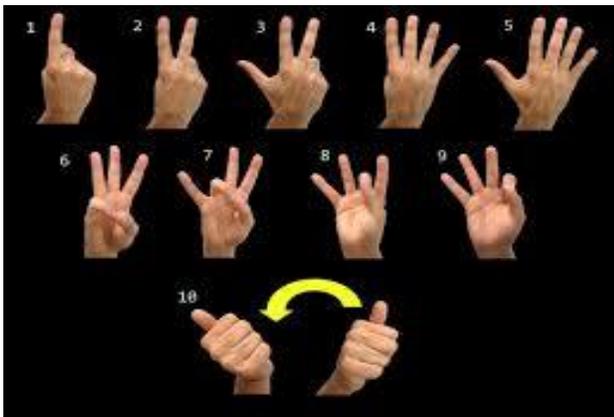


Fig. 2 Standard Numbers in BSL

II. FLOW DIAGRAM OF SYSTEM

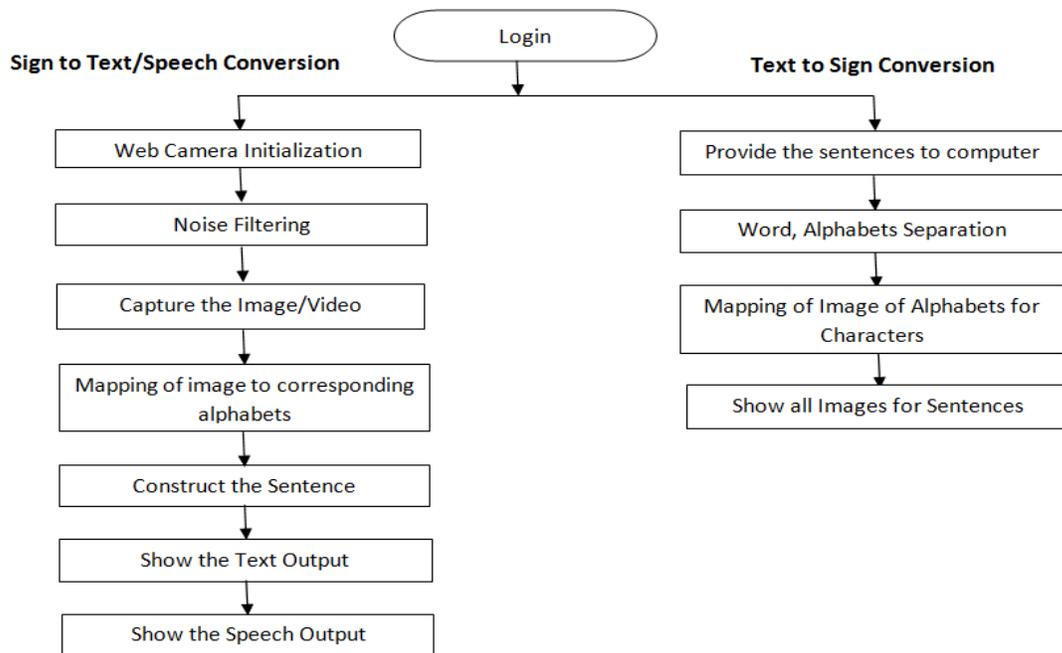


Fig. 3 Flow Diagram of system

The main purpose or goal of the project is to implement completely new computer based system for simplify the communication between deaf and dumb and normal people. This system will make the substantial changes in traditional way of communication and remove the barriers in communication. The basic working of the system is shown in above figure.

III. IMPLEMENTATION DETAILS

In this section, system works in different modules as describe below:

Modules:

Phase I: Sign Processing

- A. Finger Spelling Recognition

- Background Subtraction
- Skin Segmentation
- Region Extraction
- Feature Extraction
- Pose Estimation

B. Finger Spelling Synthesis

C. Comparison

D. Text Generation and Speech Generation

Phase II: Text to Sign Conversion

- A. Text Identification and Image Generation

Phase I: Sign Processing

- A. Module-1 Finger Spelling Recognition

The finger spelling recognition module involves the segmentation of finger spelling hand gestures from image sequences. The classification of features extracted from these images, make use of different methodologies:

- **Background Subtraction**

All subsequent image processing operations are then performing on this scaled and filtered image.

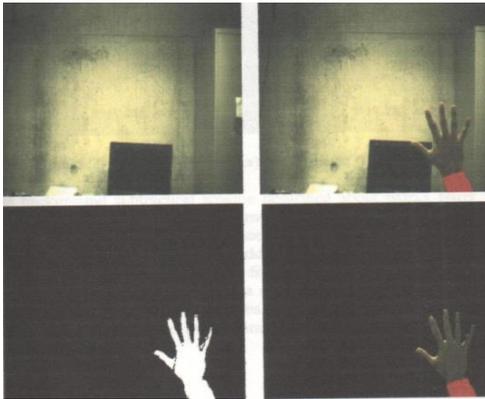


Fig. 4 Background image., Fig.5 Captured image, Fig 6 Foreground Mask, Fig. 7 Foreground image.

Background subtraction method is used to segment any foreground hand information from the non changing background image. At system startup, a background images IB, L and IB, R are captured to static workspace from each camera view. Background image to segment out moving foreground information using foreground mask image IF as:

$$IF = \begin{cases} 255 & \text{if } |I_i - IB| > \sigma B \\ 0 & \text{otherwise} \end{cases}$$

Where B is a threshold values from background information. The result of foreground mask is a binary image with a single 8 bit channel.

- **Skin Segmentation**

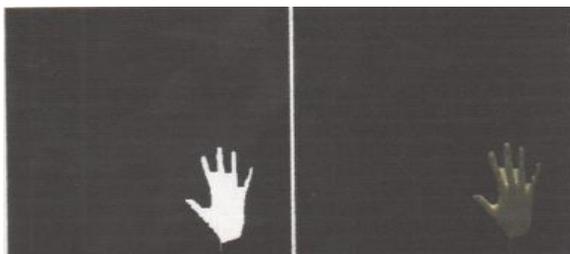


Fig.8 Skin Mask, Fig.9 Foreground skin image.

For each camera captured the small snapshots are taken of various hands with the range of different skin tones and poses. Thus this module captured image segmented into binary mask where white pixels represent skin area and black pixels represent non skin area [5].

- **Region Extraction**



Fig.10 Hand Contours

In binary image processing operations and connected component analysis, first extract the contours of all the detected skin regions in IS and for each region 'i', get a set of counter-clockwise perimeter coordinates $C_i(j) = \{(x_j, y_j)\}$ that trace the outline of each region.

- **Feature Extraction**



Fig. 11 Peak and Valley detection

1. Using frame Gabor filtering of the intensity and depth image.
2. Classify the points as either peaks or vally, convert the vector into 3D vectors in the XY plane.
3. Compute cross product.
4. If image [area of sign component of the cross product == positive]
 - Label the point as a peak.
 - Else
 - Label the point as a vally.

- **Pose Estimation**

A simple black and white planer checkerboard pattern is captured at four different poses using the stereo camera. These images are then passes to the further processing where corresponding corner features are manually selected in each image. Pose Estimation estimate the 2D and 3D poses of the Fingers and thumb.

B. Module-2 Finger Spelling Synthesis

The finger spelling synthesis can be seen as a part of sign language synthesis. Sign language synthesis can be used in two

forms. The first is real time generated avatar animation shown on computer screen that provide real time feedback. The second form is pre-generated short movie clips inserted into graphical user interface.

C. Module-3 Comparison

1. Contour matching and sign mapping
2. If image [Total_Hull_point== red && Total_defect_point==yellow] then
3. Display character whatever mapping with sign.

D. Module-4 Text Generation and Speech Generation

When above process is completed then text is generated (Character by Character). By combining the letters sentence will be formed and audio for that will get generated.

Phase II: Text to Sign Conversion

A. Module-5 Text Identification and Image Generation

In this module, Sentences are scanned entered by user and then alphabets separation module separate the alphabets from sentences and corresponding image is display to viewers.

IV. HIDDEN MARKOV MODEL

In this paper, a hand gesture recognition system to recognize continuous gesture before stationary background. The system consists of three modules: segmentation and preprocessing [6], feature extraction [6], hidden Markov model (HMM) [6] training, and gesture recognition [1] [7]. First, we apply a skin color detection and hand tracking algorithm and feature extraction algorithm(Gesture path and Orientation Estimation and vector quantization)[6] to trace the moving hand and extract the hand region, then we use the gesture database to combine the spatial and temporal features of the input image sequence as our feature vector. After having extracted the feature vectors, we apply HMMs to recognize the input gesture. The gesture to be recognized is separately scored against different HMMs.

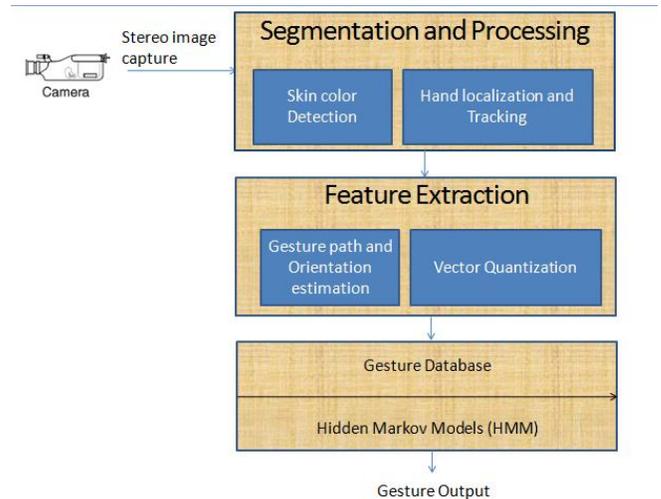


Fig. 12 Gesture recognition system using Hidden Markov Model.

V. CONCLUSION

In this paper, present a simple model of static gestures of British Sign Language alphabet recognition system, does not require any special markers or gloves and can operate in real-time on a commodity PC with low cost cameras. Specifically, the system can track the tip positions of thumb and index finger for each hand, assuming that a calibrated pair of cameras is viewing the hands. The motivation for this hand tracker was a desktop-based two-handed interaction system in which a user can select and manipulate 3D geometric in real time using natural hand motions. And finally with the help of this software communication barrier and dumb people and normal people get reduced.

ACKNOWLEDGEMENT

Many people have contributed to the success of this. Although a single sentence hardly suffices, I would like to thank Almighty God for blessing me with His grace. I extend my sincere and heartfelt thanks to **Mr.M.T.Jagtap Sir**, Head of the Computer Engineering Department (Pune Vidyarthi Griha's College of Engineering, Nashik, India) for providing me the right ambience for carrying out this work, encouragement and I sincerely express my gratitude to him. I express my immense pleasure and thankfulness to all staff of the Department of Computer Engineering for their co-operation support. Last but not the least, I thank all others, and especially my friends who in one way or another helped me in the successful completion of this work.

REFERENCES

- [1] Stephan Liwicki, Mark Everingham, "Automatic Recognition of Finger spelled Words in BritishSign Language," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 50-57, Miami, FL, USA, June, 2009.
- [2] R. Feris, M. Turk, R. Raskar, K. Tan, and G. Ohashi. Exploiting depth discontinuities for vision-based fingerspelling recognition.

- In In IEEE Workshop on Real-time Vision for Human-Computer Interaction, 2004.
- [3] X.D. Huang, Y. Ariki, and M. A. Jack. Hidden Markov Models for Speech Recognition. Edinburgh University Press, 1990.
- [4] L. R. Rabiner. A tutorial on hidden Markov models And selected applications in speech recognition. Proc. IEEE, 77:257–286, 1989.
- [5] M. Jones and J. Rehg. Statistical color models with application to skin detection. IJCV, 46(1):81–96, 2002.
- [6] J. Schlenzig, E. Hunter, and R. Jain. Recursive identification of gesture inputs using hidden Markov models. Proc. Second Annual Conference on Applications of Computer Vision, pages 187–194, December 1994.
- [7] D. B. Nguyen, S. Enokida, and E. Toshiaki, Real-Time Hand Tracking and Gesture Recognition System, IGVIP05 Conference, CICC, pp. 362-368, 2005.

AUTHORS

First Author – Pingale Prerna Rambhau, Department of Computer Engineering(University of Pune), Sir Vishaveshwaray Institute Technology, Chincholi, Tal-Sinner,Dist- Nashik ,India, Email- pingale.pruna795@gmail.com