A Survey On Gesture Recognition in Sign Language Recognition for Mute and Dumb People

Mr. Sumit R. Dhobale¹, Prof. Vaishali B. Bhagat²

¹ME (CSE) Scholar, Department of CSE, P R Patil College of Engg. & Tech., Amravati-444602, India
²Assistant Professor, Department of CSE, P R Patil College of Engg. & Tech., Amravati -444602, India

ABSTRACT
Gesture recognition is a field in computer science and sign-language technology with the aim of interpreting human gestures via mathematical algorithms. Current focuses in the field include emotion recognition from the face and hand gesture recognition. Many approaches have been made using cameras and computer vision algorithms to interpret sign language.

The Sign language is very important for people having hearing and speaking deficiency generally called Dumb and Mute. Sign language is the only way of communication for such peoples to convey their messages and it becomes important for peoples to understand their language. But as many peoples are not aware with the sign language and thus it is difficult for mute and dumb people to properly communicate with the society. So this gesture recognition provides a solution for this problem which will convert a sequence of hand gestures in text and then into the spoken words. Every hand gesture recognition is depend on character in American Sign Language (ASL). The related research area of sign gesture recognition is Human Computer Interaction (HCI) and image processing which help in the solution to this problem.

Keywords: Sign –language technology, gesture recognition, image processing, Human Computer Interaction.

1. INTRODUCTION
Gesture recognition is interface with computers using the human body, typically hand movements. In gesture recognition technology, a camera reads the movements of the human body and provides the data to a computer which makes use of the gestures as input to control devices or applications. For example zooming system in smart phones that zoom in and zoom out particular object which typically uses hand gestures. Thus sign language gesture recognition provides valuable assistance for physically impaired to easily and efficiently interact with computers.
This technology having the potential to change the traditional way in which user interact with computers by eliminating input devices such as joysticks, mice and keyboards and allowing the human body to give signals to the computer through gestures such as finger pointing. The gestures of the body are read by a camera instead of sensors attached to a device such as a data glove. Gesture recognition technology also can be used to read facial and speech expressions (i.e., lip reading), and eye movements.

1.1 2D Vision: (x, y) coordinate-based Limitations
2D vision co-ordinate having some limitations to recognize the object. 2D vision co-ordinate having lack of understanding sense include segmentation, object representation and recognition. It is not possible to get more accurate and valuable result from (x, y) co-ordinate based vision. While the possibility include whole body tracking and other technique that combine multiple cues, it is difficult to scenes using only 2D representation.

1.2 3D vision: “Z” (depth) innovation
3 D vision enables our capabilities to recognize an object via a 3 dimensional view having a detailed structural info as compared to traditional 2 dimensional techniques. “Z” co-ordinate gives us depth information, enables capabilities well beyond gesture recognition. The human eye naturally registers x, y and z coordinates for everything it sees, and the brain then interprets those coordinates into a 3D image. By using 3D image output increase correctness and efficiency which is more valuable.

1.3 Are there different types of gestures?
1.3.1 Pose gestures:
The pose gesture can be captured by static single snapshot .The sign of snapshot provide the sense of gesture but it is stable.
1.3.2 Single-motion gestures:
In this case, the body or part of the body performs a specific motion, over a finite, usually time restriction.

![Figure 1 Pose gestures.]

1.3.3 Continuous-motion gestures:
In continue motion refers to a repetitive action with no time limit, for example: a person running.

2. Materials and Methods
In recent years there has been a lot of research on hand gesture recognition. Several techniques have been reported on gesture recognition which includes skin segmentation using colour pixel classification [1], Parametric Hidden Markov models for gesture recognition [3], statistical database comparison method [4], orientation histograms for gesture recognition [6], Finger Detection for Sign Language Recognition[7] etc.

2.1 Eigen value and Eigen vector
The Eigen value and Eigen vector are the functions of linear transformations. Eigen vector provide the direction along which the linear transformation acts by stretching, compressing or flipping. Eigen values gives the factor by which the compression or stretching occurs Eigen vector is the coordinate system for the new coordinate system is uncorrelated. If there are more the Eigen vector then more better information obtain from linear transformation. The Eigen vector calculates variance of data of new coordinate system. The compression of the data only few significant Eigen values are being selected which reduces the dimension of the data allowing the data to get compressed. Mathematically, it is explained in (1).

If \( x \) is a one column vector with \( n \) rows and \( A \) is a square matrix with \( n \) rows and columns, then the matrix product \( Ax \) will result in vector \( y \). When these two vectors are parallel, \( Ax = \lambda x \), (\( \lambda \) being any real number) then \( x \) is an eigenvector of \( A \) and the scaling factor \( \lambda \) is the respective Eigen value.

\[
\begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}
\iff
\begin{bmatrix}
    A_{11} & A_{12} & \cdots & A_{1n} \\
    A_{21} & A_{22} & \cdots & A_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    A_{n1} & A_{n2} & \cdots & A_{nn}
\end{bmatrix}
\begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}
\iff
\begin{bmatrix}
    z_1 \\
    z_2 \\
    \vdots \\
    z_n
\end{bmatrix}
\]

2.2 PROPOSED SYSTEM
The proposed method consists of following phases:

2.2.1 Image capture & Image Processing
The high resolution web camera is used to capture the snapshots and movements of the gestured objects. It is essential to specify the format of the videos while creating any video. If the video format is not specified, the video input function uses the default format. The video function is used to determine which video format support to the device.

![Figure 2 Single motion gesture]

![Figure 3 Image capturing.]
2.1.1 Skin Filtering
The second phase for proposed system is the skin filtering which differentiate human body objects on the basis of skin colour pixels from the non-skin colours pixels. This system is efficiently used for detection of hand movements, face gestures etc. Now a day, this filtering method can be used for video processing. This skin filtering steps can be viewed in given Figure

![Diagram of Skin Filtering Process](image)

**Figure 4** Steps for carrying out Skin Filtering.

The input captured image in the form of RBG. That RBG image is first converted into the HSV image. The conversion is taken because of there is some limitation in RBG image. In HSV image space there are three attributes as: H stands for Hue means set of pure colour from colour space, S stands for saturation that declare the grade of purity of colour image and V stands for value used to provide relative lightness and darkness of colour. In this system RBG convert to HSV use following methodology.

\[
H = \begin{cases} 
60 \left( \frac{G - B}{\delta} \right) & \text{if } \text{MAX} = R \\
60 \left( \frac{B - R}{\delta} + 2 \right) & \text{if } \text{MAX} = G \\
60 \left( \frac{R - G}{\delta} + 4 \right) & \text{if } \text{MAX} = B \\
\text{not defined} & \text{if } \text{MAX} = 0 
\end{cases} \tag{2}
\]

\[
S = \begin{cases} 
\frac{S}{\text{MAX}} & \text{if } \text{MAX} \neq 0 \\
0 & \text{if } \text{MAX} = 0 
\end{cases} \tag{3}
\]

Where \( \delta = (\text{MAX} - \text{MIN}) \), \( \text{MAX} = \max(R,B,G) \) and \( \text{MIN} = \min(R,B,G) \)

The HSV image is filter and provides the output image, which compare with skin colour pixels. It is converted to smoothened image and gray scale colour image. Finally to eliminate these, it take the biggest binary linked object (BLOB) which considers only the region comprising of biggest linked skin-colour pixels. The steps of conversion as follows.

![Diagram of RBG to HSV Conversion](image)

**Figure 5** Steps of conversion RBG to HSV
2.1.2 Feature Extraction

The desired output image is then cropped in this phase. The Eigen value and Eigen vector is used to crop the image. The steps of calculating Eigen vector and Eigen value as follows:

Step1: Assume that input image is X.
Step2: The mean of above vector X is
\[ \mathbf{M} = E\{X\} \]  \hspace{1cm} (4)
Step3: The covariance matrix \( \mathbf{C} \) of the above input vector X was found out and it given by
\[ \mathbf{C} = E\{(X - \mathbf{M})(X - \mathbf{M})'\} \]  \hspace{1cm} (5)
Step4: By using covariance matrix \( \mathbf{C} \) Eigen value and Eigen vector are calculated.
Step5: In that last step the Eigen vector can arrange corresponding to decreasing order of Eigen value.

2.1.3 Gesture Matching

The output image matching with the stored database image. It gives required output correctly with the help of matching probability of database stored images. That is each sigh stored 10 image and matching with nearest possible matching image that gives correct output. That text converts into the voice which is also store to database for particular symbols. Each sign indirectly produce voice for communication.

3. CONCLUSIONS

The proposed system was implemented with Eigen vector approach for gesture to text conversion. The method of gesture recognition, the approach of skin colour segmentation method for gesture extraction and correlation approach for the matching is analysed. The output text then converts into the voice which stored on the database. The future scope of gesture recognition is new interacting method communicating and controlling with computer. The future enhancement for sign gesture technology implement on the android platform running on smart mobile for its broader use.

REFERENCES


AUTHOR

Mr. Sumit R. Dhobale
Received Bachelor degree in computer science and Engg from Amravati University in 2012 and pursuing master degree in C.S.E from P.R. Patil college of Engg Amravati -444602

Prof. Vaishali B. Bhagat
Received the Master degree in Computer Science from Amravati University in 2011. Working as Assistant Professor in department of C.S.E at P.R. Patil College of Engg Amravati -444602