International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 2 Issue 4, April 2015

ISSN: 2394-4404

Sign Language Recognition System

Rohan Sam Kuruvilla

Loisy PL

Treasa Vellanikaran

Mar Athanasius College of Engineering,
Department of Electrical and Electronics,
Kothamangalam, Ernakulam, India

Abstract: ***In the present era where technology has thrived so much, allowing people around the world to communicate instantly through the border-less World Wide Web, it is highly deplorable to perceive the fact that an effective communication means is still deprived for the deaf and mute. Common people find it tiresome communicating with them. Relying on interpreters is a solution but even then finding experienced and qualified interpreters for their routine day to day life is impractical. Hence a Dumb/mute aid system with voice transmission is a requisite for their proper communication. The basic idea of this project is to make a system using which the mute can significantly communicate with their gestures. This system uses a process for recognizing hand gestures captured using video camera and a standard consumer personal computer, developed and implemented using the MATLAB mathematical environment. A pattern recognition system will be using a transform that converts an image into a feature vector, which will then be compared with the feature vectors of a training set of gestures and translates them into audio output so that its comprehensible for others. The idea aims to bridge the gap between the speech/hearing impaired and the normal people.***

Keywords: co-variance, sign language recognition, matlab, feature vector, image acquisition.

I. INTRODUCTION

We live in an era where communication knows no boundary but it is not so for the differently abled- those with a disability to speak. Their gestures often fail to convey their thoughts and emotions to the untrained eyes of ours. It is here where our system comes effective enabling the differently abled to feel at ease and communicate to the world around them.

Body language is an important way of communication among humans. Normal people can communicate their thoughts and ideas to others through speech. The only means of communication method for the hearing impaired community is the use of sign language. The hearing impaired community has developed their own culture and methods to communicate among themselves and with ordinary person by using sign gestures. Instead of conveying their thoughts and ideas acoustically they convey it by means of sign patterns. Sign gestures are a non-verbal visual language, different from the spoken language, but serving the same function. It is often very difficult for the hearing impaired community to communicate their ideas and creativity to the normal humans. This system was inspired by the special group of people who have difficulties to communicate in verbal form. It is designed with the ease of use for the hearing/speech impaired people. The objective of this project is to develop a system prototype that automatically helps to recognize sign languages of the signer and translate them into voice in real time. Hand gesture is one of the typical methods used in sign language for non-verbal communication. Sign gestures are a non-verbal visual language, different from the spoken language, but serving the same function. This project presents a system that will not only automatically recognize the hand gestures but also convert it into corresponding speech output so that speaking impaired person can easily communicate with normal people. The gesture to speech system, has been developed using the skin colour segmentation. The system consists of camera attached to computer that will take images of hand gestures. Image segmentation and feature extraction algorithm is used to recognize the hand gestures of the signer. According to recognized hand gestures, corresponding sound track will be played.

II. WORKING

This system involves two phases: training and running phase. In the training phase, the user shows the system one or more examples of hand gestures. The system stores the Eigen image of the hand shape and in the run phase the computer compares the current hand shape with each of stored shapes by coefficients. The best match gesture is selected by the covariance method.

A. IMAGE DATABASE

First of all an image database is created. These images are taken from American sign language (ASL).ASL consists of about 6000 gestures of common words out of which 8 gestures are used. These images were captured using webcam.

B. PRE-PROCESSING

Preprocessing is done so to make the image easier to analyze. It is applied to images before we can extract features from hand images. In the preprocessing stage the RGB image is converted into grayscale image and the resizing of the entire set of image is done.

Variance is another measure of the spread out of data in a set. In fact it is the square of the standard deviation, whereas Covariance is a statistical measure of the extent to which two random features vary together. Covariance can be a negative, positive or zero number, depending on what is the relation between two features. If the features increase together, the covariance is positive, one feature increases and other decreases, the covariance is negative, and if the two features are independent, the covariance is zero. The covariance matrix provides a natural way of fusing multiple features which might be correlated the covariance is a 2-dimensional measurement. The covariance will allow us to see if there is any relationship between the different dimensions of the data set. The definition for the covariance matrix for a set of data with n dimensions is:

$C_{nxn} = (C_{I,J} , C_{I,J}= \text{cov} (D_I,D_J))$ ....(1)

Where $C_{n \times n}$ is a matrix with n rows and n columns, and D_x is the xth dimension. If you have an n dimensional data set, then the matrix has n rows and columns (so is square) and each entry in the matrix is the result of calculating the covariance between two separate dimensions. For example the entry on row 2, column 3, is the covariance value calculate between the 2nd dimension and the 3rd dimension.

C. EIGEN VECTORS AND EIGEN VALUES

Eigenvectors of a linear operator are non-zero vectors which, when operated upon by the operator, result in a scalar multiple of them. Eigen value is a value which is associated with each Eigen vector. The Eigen value gives us some information about the importance of the eigenvector. The Eigen values are really important in the this method, because they will permit to realize some threshold to filter the nonsignificant eigenvectors, so that we can keep just the principal ones. Eigen Vectors can only be found for square matrices. Not every square matrix has Eigen vectors. An N x N matrix that has Eigen vectors will have N of them. If we scale the vector by some amount before multiplying it, we will still get the still get the same multiple of it as a result. All the Eigen vectors are perpendicular i.e. at right angle regardless of their dimensions. The concept of is to reduce the size of the images to be recognized from a high to a lower dimension. While lowering the dimensionality in the set, using the eigenvector method also highlights the variance within the set. By performing singular value decomposition on an average covariance matrix we can find the Eigen values and eigenvectors.

$Cu = \lambda k$............................... (2)

Where C is the average covariance image matrix, k corresponds to the eigenvectors and λ corresponds to the Eigen values of C.

D. GESTURE RECOGNITION

The main idea is based on the eigenvectors of the covariance matrix of the group of hand posture training images. These eigenvectors can be thought of as a group of features which characterize the variation between hand posture training images. Hand posture training images are projected into the subspace spanned by a hand posture space. The hand posture space is defined by the eigenspace which are the eigenvectors of the set of hand postures. Each training image corresponds to each Eigenvector. An eigenvector can be shown as an eigenspace.

Then in the testing stage, every detected hand posture is classified by comparing its position in hand posture space with the positions of every hand posture training images. Each hand posture image in the training set is represented a linear combination of the eigenspace, which indicate a feature space that spans the variations among the known hand posture training images. The number of Eigen space is equal to the number of hand posture images in the training set. The hand postures can also be approximated using only the best Eigen space, which have the largest Eigen values.

First we read all the data set pictures, to treat the contrast (to normalize the pictures), to apply a Gaussian filter and to resize them in a vector. Then we resize all images the database to 200 x 180. This is achieved in MATLAB using imresize command. Then we reshape the images to get elements along rows as we take Input Image. The next step is to calculate the mean of each direction. It is a fast step. We just had to take the first matrix of all the images, and then ask to MATLAB to calculate the mean of the matrix.

Then, we subtracted it to the first matrix. Cov(x) or Cov(x,y) normalizes by N – 1, if N > 1, where N is the number of observations. This makes Cov(X) the best unbiased estimate of the covariance matrix if the observations are from a normal distribution. For N = 1, Cov normalizes by N. It is really important to choose the good eigenvectors to express the data set with the best base. The number of eigenvectors chooses will be in direct relations with the results that we get. To recognize an unknown hand posture image in the testing stage, its weight (wi) is calculated by multiplying the eigenvector (ui) of the covariance matrix (C_v) with difference image $(\Gamma - \Psi) = \Phi$ $w_i = u_i^T(\Gamma - \Psi)$. Now the weight vector of the unknown image becomes

$\Omega = [w_1\ w_2\ w_3 \ldots\ldots w_m]^T$................ (3)

Fig.1 shows the flowchart of the system. The minimum Euclidean distance between test image and each training image is defined by $\in$ = minimum $\| \Omega - \Omega \|$....................... (4) k = 1, 2…..... M Where M is the number of images trained. Then, Γ is recognized as hand posture k from the training images set.

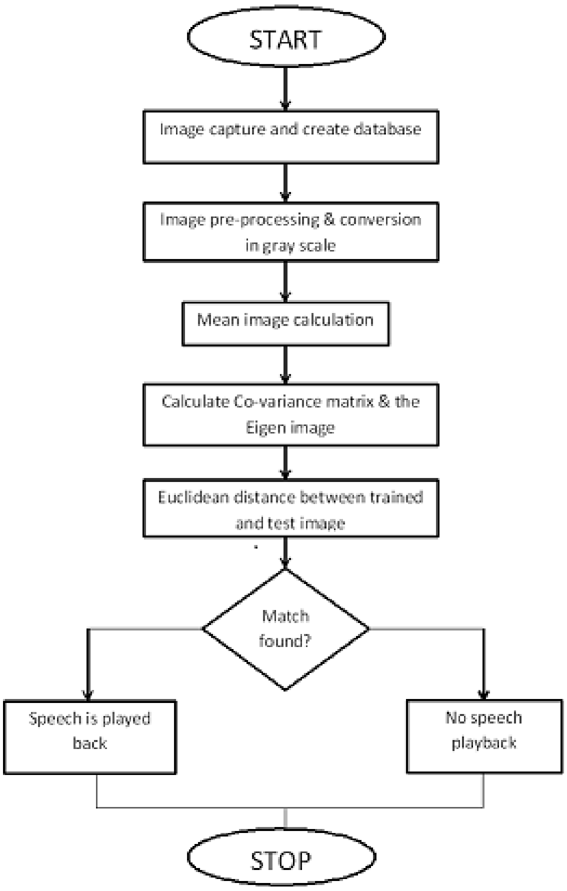

Figure 1: Flowchart of the system

These are processed and their corresponding values are generated for matching with the reference values. So once a reference value is matched its corresponding speech is played back. Here to avoid speech playback unnecessarily, we have set the threshold value at 0.5. This is to say that if no gesture is been showed no speech should be played back.

E. BLOCK DIAGRAM

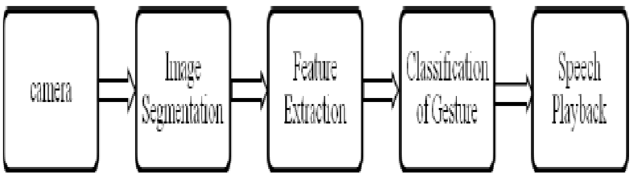

Figure 2: Block Diagram

The system has several stages of processing namely, Image Segmentation, Feature Extraction, Gesture Classification, speech playback. Fig.2 shows the block diagram of the proposed algorithm of the system. Firstly, the gesture sign images are captured with a resolution of 320X240 pixels.

The gesture is captured with a resolution of minimum 320X240 pixels. During the image segmentation stage basically the RGB colour Space is converted to grayscale by MATLAB functions.

Eigen value of each image is found. Then the standard deviation and mean of the images are calculated and stored as a single value against the corresponding gesture. And these values act as the reference for each gesture.

Now to recognize the signs images are captured. These are processed and their corresponding values are generated for matching with the reference values.

So once a reference value is matched its corresponding speech is played back. Here to avoid speech playback unnecessarily, we have set the threshold value at 0.5. This is to say that if no gesture is been showed no speech should be played back.

II. RESULT

The various tests were conducted and the results were found to be good. The various tests we did are:

- ✓ The system was tested on 20 people. They were made to do the same sign. For 70 percent it worked correctly and for the rest it did not. This happened mainly because they failed to reproduce the similar gesture. Fig.3 shows the pie chart of this test.

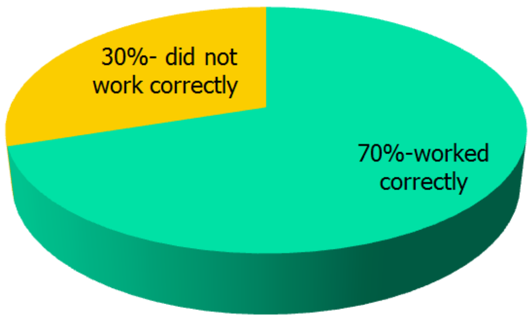

Figure 3: Pie Chart of the test conducted on 20 people

- ✓ We also did a test where the same person did the same sign around 20 times. For 95 percent it worked correctly and the rest it did not and this was mainly due to carelessness. Fig.4 shows the pie chart on a test done on the same person.

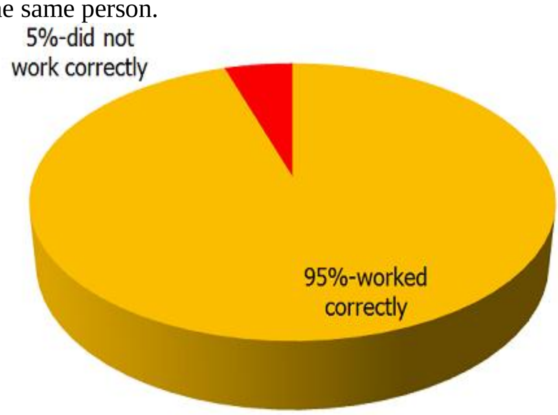

Figure 4: Pie chart on the test conducted on the same person

ISSN: 2394-4404

- ✓ Another test that we conducted was in different light conditions. The result was, for very low luminance the accuracy was nearly zero. As the light improved the accuracy was also seen to be improving. But beyond a certain light intensity the images became too lit up and gestures couldnt be distinguished. Fig.5 shows the graph in which Accuracy v/s Luminance has been plotted.

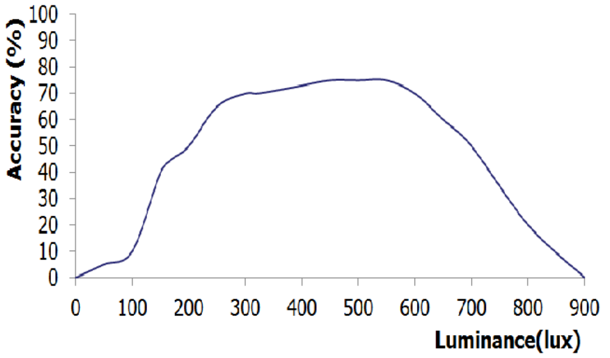

Figure 5: Graph of Accuracy v/s Luminance

III. CONCLUSION AND FUTURE SCOPE

Our work was done using static gestures, but with advances in recent technology hand gesture will have to be realized in real time i.e. dynamic gestures.

Some of the applications of this project can be in the field of robotics where the robots can be socially assistive by interpreting the meaning of the hand gesture recognized. Strong gesture recognition can be used an alternative computer interface, which can allow users to accomplish common tasks like opening up files etc. Another use of gesture recognition is in the field of security where authorization can be provided only on recognizing only any particular or combination of hand gestures.

Further enhancement of the techniques proposed is possible with better segmentation of the image. Experiments will have to be done on a larger scale so that results can be more accurate. Edge detection technology keeps on changing so constant improvement is possible, also line detection can also be used to solve some problems.

IV. ACKNOWLEDGEMENT

The authors would like to thank Mr. Benny Cherian, Professor Electrical and Electronics Engineering department, Mar Athanasius College of Engineering for his spirited Guidance and moral support. We thank all faculty members and staff of Electrical and Electronics Engineering Department.

REFERENCES

[1] Mayuresh Keni, Shireen Meher, Aniket Marathe,"Sign Language Recognition System" ,International Journal of Scientific and Engineering Research, Volume 4, Issue 12, December-2013 ISSN 2229–5518.

[2] Prof. R. R. Itkarkar, Dr. Anil V Nandi, "Hand Gesture to Speech Conversion using Matlab", Computing, Communications and Networking Technologies (ICCCNT),2013 Fourth International Conference, 4–6 July 2013, IEEE.

[3] Ginu Thomas, Rahul Vivek Purohit "Hand Gesture Recognition via Covariance Method",International Journal of Engineering and Advanced Technology (IJEAT)ISSN: 2249 8958, Volume-2, Issue-5, June 2013

[4] P. Subha Rajam , Dr. G Balakrishnan," Real time Indian sign language recognition system to aid deaf-dumb people", Communication Technology (ICCT), 2011 IEEE 13th International Conference.

[5] Oi Mean Foong, Tan Jung Low, and Satrio Wibowo,"Hand Gesture Recognition: Sign to Voice System (S2V)" , World Academy of Science, Engineering and Technology Vol:2,pp. 26–30, 2008.

[6] Qutaishat Munib , Moussa Habeeb, Bayan Takruri, Hiba Abed Al-Malik, "American sign language (ASL) recognition based on Hough transform and neural networks", Expert Systems with Applications 32 (2007) 24-37

[7] Aleem Khalid Alvi, M. Yousuf Bin Azhar, Mehmood Usman, Suleman Mumtaz, Sameer Rafiq, Razi Ur Rehman, Israr Ahmed T , "Pakistan Sign Language Recognition Using Statistical Template Matching" , World Academy of Science, Engineering and Technology, 2005.

[8] E.J. Ong, and R. Bowden, "A Boosted Classifier Tree for Hand Shape Detection", Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2004), IEEE Computer Society, 2004, pp. 889-894.

[9] W. Gao, G. Fang, D. Zhao, and Y. Chen," Transition Movement Models for Large Vocabulary Continuous Sign Language Recognition Proc. Sixth IEEE Intl Conf. Automatic Face and Gesture Recognition, pp. 553–558, May 2004.

[10] C. Vogler and D. Metaxas,"A Framework for Recognizing the Simultaneous Aspects of American Sign Language", Computer Vision and Image Understanding, vol. 81, no. 3, pp. 358–384, 2001.

[11] M. J. Jones, and J. M. Rehg," Statistical Color Models with Application to skin Detection", International Journal of Computer Vision, Jan. 2002, vol. 46, no.1, pp. 81–96.

[12] Byung-woo min, Ho-sub yoon, Jung soh, Takeshi ohashi and Toshiaki jima, "Visual Recognition of Static/Dynamic Gesture: Gesture-Driven Editing System", Journal of Visual Languages & Computing Volume10,Issue3, June 1999, Pages 291–309.