

Static Image-Based Sign Language Recognition Using Eigen Object Recognizer

¹Shobha S. Raskar, ²Prutha Gandhi, ³Dhanashri Dalvi, ⁴Pallavi Gaikwad, ⁵Shubham Khode

¹Assistant Professor

Department of Computer Engineering,
Modern Education Society's College of Engineering, Pune, India

Abstract— A mediator person is required for communication between deaf person and a second person. But a mediator should know the sign language used by deaf person. But this is also not possible always since there are multiple sign languages for multiple languages. It is difficult for a deaf person to understand what a second person speaks. And therefore deaf person should keep track of lip movements of second person in order to know what he is speaking. But the lip movements does not give proper efficiency and accuracy. To overcome these problems we have proposed a system, i.e. an application for recognizing sign language using hand gestures with the facility for user to define and upload their own sign language into the system. The features of this system are the real time conversion of gesture to text and speech. For two-way communication between deaf person and second person, the speech of second person is converted into text. The processing steps include: gesture extraction, gesture matching and conversion of text to speech and vice-versa. The system can also be used by other people who migrate to different regions and do not know the regional/local language.

Index Terms— Digitization and Image Capture, Segmentation, Object recognition, Tracking

I. INTRODUCTION

We have implemented a system for recognizing sign language hand gestures. It will additionally provide the facility to each individual to define and his own sign language into the system. To interpret gestures performed by deaf person by converting gestures into text and speech for communication. For two-way communication the speech will be converted into text.

The previously implemented system had a predefined database with a limited scope. Thus, we are facilitating an application that will allow a user to define their own database i.e. to define his own sign language in the system. This feature will help deaf people to communicate with other people varying from different countries or regions.

II. SYSTEM DESIGN

The block diagram of the proposed system is given in Figure 1 where input image is converted from RGB to YCrCb [2], [5]. Then we get the filtered image which is morphed and converted into binary image in grey scale. Morphing consists of erosion and dilation of an image. This grey scale image is represented as the biggest blob. And hence text associated with recognized image is retrieved from the text file, where the text labelled to the images added by user are stored.

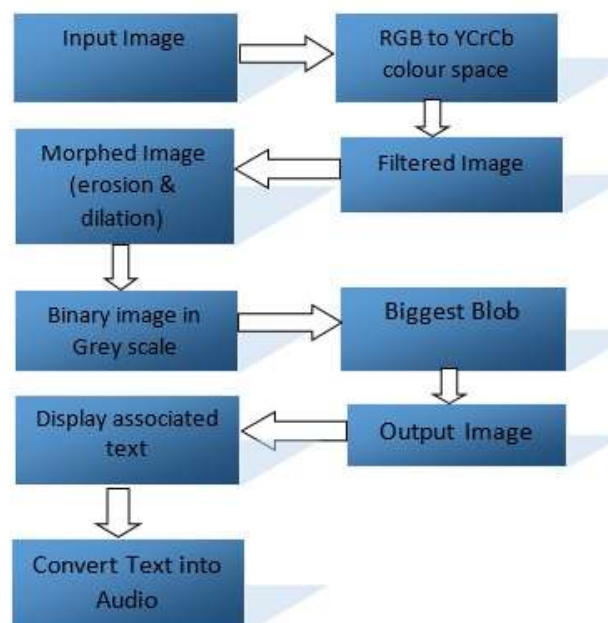


Fig. 1: Block Diagram of implemented System

III. SYSTEM IMPLEMENTATION

The proposed system is implemented using Microsoft .NET Framework, Microsoft Visual Studio Ultimate 2012 IDE, Emgu CV for Windows Desktop 2.4.10 Library [6], [7]. Microsoft’s Speech Synthesizer API is used to convert text into speech. We have to train the database before using the system. Screenshots of the implemented system are given below:



Fig. 2 GUI of the System

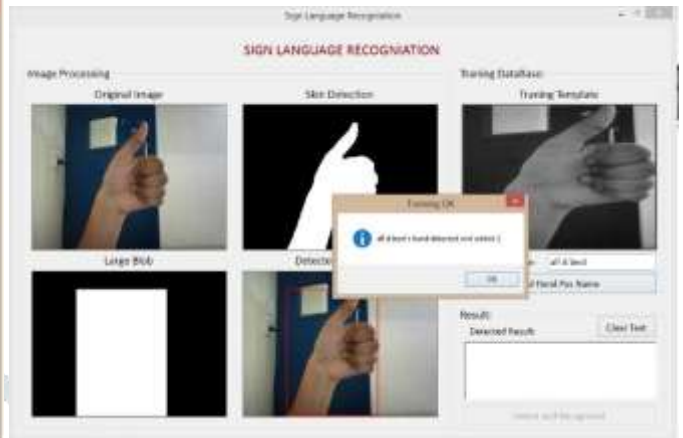


Fig. 3 Adding gestures in the database with its text



Fig. 4 Trained images are converted into Grey images and TrainedLables.txt file to store text



Fig. 5 Format of file in which text is stored

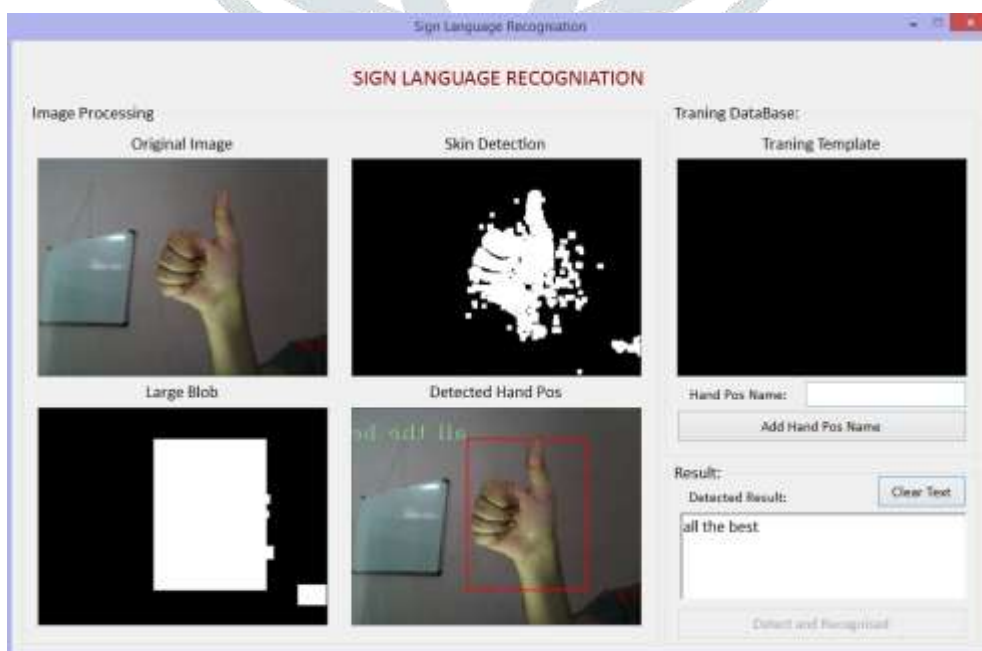


Fig. 6 Recognition of Image and displaying its associated text

IV. ALGORITHMS

We have implemented the following algorithms in the system.

Eigen value and Eigen vector

Eigen values and Eigen vectors are a part of linear transformations. Eigen vectors are the directions along which the linear transformation acts by stretching, compressing or flipping and Eigen values gives the factor by which the compression or stretching occurs. In case of analysis of data, the Eigen vectors of the covariance are being found out. Eigenvectors are set of basis function which describes variability of data. The more the Eigen vectors the better the information obtained from the linear transformation. Eigen values measures the variance of data of new coordinate system. For 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.

Eigen Object Recognition using Principal Component Analysis (PCA)

PCA looks for the correlation between the templates. It summarizes the variation in a correlated multi variables to a set of uncorrelated variables, each of which is a particular linear combination of the original variable. The extracted uncorrelated variables are called principal components (PC).

The number of Principal components is less than or equal to the number of original variables. Principal components are found by calculating the Eigen vectors and Eigen values of the data covariance matrix [8].

Step 1: Get the Eigen vectors that form the Eigen space. Let A be an nxn matrix. The Eigen values of A are defined as the roots of

$$\text{Determinant } (A-\lambda I) = | (A-\lambda I) | = 0 \quad (1)$$

Where I is an identity matrix. This Equation 1 is called characteristic equation. Let λ be an Eigen value of A. Then there exists a vector x such that

$$Ax = \lambda x \quad (2)$$

This vector x in Equation 2 is called an Eigen vector of A associated with the Eigen value λ . The covariance values are computed using following Equation 3:

$$S_{xy} = \frac{1}{n-1} \left(\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \right) \quad (3)$$

Here n in Equation 3 represents sample size. The images are all of the same resolution and are all equivalently framed.

Step 2: Get or set the Eigen distance threshold. The smaller the number, the more likely an examined image will be treated as unrecognized object. Set it to a huge number (e.g. 5000) and the recognizer will always treated the examined image as one of the known object.

Step 3: In image recognition an input image with n pixels can be treated as a point in an n-dimensional space called the image space. Most of the image pixels will be highly correlated. Thus we need to consider how to achieve a reduction in the number of variables. PCA helps in reducing the dimensions of the variables.

We can compress the data by using Eigen vectors.

Step 4: PCA can supply the user with a lower dimensional picture, then a projection or shadow of this object viewed from its most informative viewpoint. This is done by using only the first few principal components, so that the dimensionality of the transformed data is reduced.

Step 5: Then the recognition of the image is done with this reconstructed image.

Skin detection using YCbCr color space.

This method is purely based on the threshold value of three different components:

1. A RGB color image is transformed into YCbCr color image.
2. The result image is comprised of intensity component (Y) and chrominance components (Cb and Cr).
3. The threshold is applied for chrominance components only as $Cr > 150$ && $Cr < 200$ && $Cb > 100$ && $Cb < 150$.

Contour Functions

Contours are useful in creating masks for images and to segment and extract features from an image.

In order to use contour functions in a program we have to add reference to required dlls.

The two contour functions are: FindContours and DrawContours.

Syntax:

```
public Contour<Point> FindContours (
    CHAIN_APPROX_METHOD method,
    RETR_TYPE type,
    MemStorage store
)
```

Parameters:

- method

Type:Emgu.CV.CvEnum.CHAIN_APPROX_METHOD

The type of approximation method.

- type

Type:Emgu.CV.CvEnum.RETR_TYPE

The retrieval type.

- CV_RETR_EXTERNAL:
Retrieve only the extreme outer contours
- CV_RETR_LIST:
Retrieve all the contours and puts them in the list
- CV_RETR_CCOMP:
Retrieve all the contours and organizes them into two-level hierarchy: top level are external boundaries of the components, second level are bound a boundaries of the holes
- CV_RETR_TREE:
Retrieve all the contours and reconstructs the full hierarchy of nested contours

- store

Type:Emgu.CV.MemStorage

The storage used by the sequences

The contours are not indexed, but each Contour contains a pointer HNext that points to the next contour. If HNext is null, there are no more contours.

If you want to go through them again, you can use the HPrev pointer, which points to the previous contour. If HPrev is null, you have reached the first contour.

Morphing

Morphing is a familiar technology to produce special effects in image or videos. The Word "Morphing" comes from the word "metamorphosis" which means change shape, appearance or form. Morphing is done by coupling image warping with color interpolation. Morphing is the process in which the source image is gradually distorted and vanished while producing the target image. So earlier images in the sequence are similar to source image and last images are similar to target image. Middle image of the sequence is the average of the source image and the target image.

Morphing techniques may be classified into two based on the ways to specify their features. They are mesh based methods and feature based methods. In mesh based methods, features of the image specified by a non-uniform mesh. Feature based methods, features of the image specifies as line segment or a set of points. Feature based methods are popular.

Image erosion and dilation are implementations of morphological filters, a subset of Mathematical Morphology. It is typically applied to binary images, but there are versions that work on grayscale images. They have wide array of uses [4]:

- Removing noise.
- Isolation of individual elements and joining disparate elements in an image.
- Finding of intensity bumps or holes in an image

Dilation: The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels grow in size while holes within those regions become smaller.

Erosion: The basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels shrink in size, and holes within those areas become larger.

Therefore, Image Dilation increases the size of edges contained in an image. In contrast Image Erosion decreases or shrinks the size of an Image's edges.

Tips to remember for Morphing process-

1. Choose pictures or images to morph carefully. For example, if you wish to morph two animals, it should have same size and outline.
2. Selection of Background. For best results you should use same background for both pictures.

Morphing is common in entertainment industry. Morphing is widely used in movies, animation games etc. In addition to the usage of entertainment industry, morphing can be used in computer based trainings, electronic book illustrations, presentations, education purposes etc. morphing software is widely available in internet.

V. CONCLUSION AND FUTURE SCOPE

Sign languages are one of the main communication methods used by deaf people, but opposed to common thought, there is no universal sign language: every country or even regional group uses its own set of signs. The ultimate goal of this project is to interpret sign language defined by user in the system into its equivalent text or speech bases on input images.

This system will ease and encourage the interaction of common people with the handicapped people since the common people would no longer be required to learn the various sign languages in order to communicate with them. Handicapped people can

benefit from this system in their day to day life whenever they need to easily convey their message through their sign language to common people. Also speech of second person is converted into text for two-way communication.

Current issue in project is related to brightness which needs to be resolved means good accuracy is obtained in full brightness. The application is built for targeted Windows OS and hence requires porting onto different OS. Accuracy for conversion of speech to text can be increased by considering factors like age group, pitch, accent of speaker. The application can be ported on Android OS which will be an Android application that will be handy for use. The Edge Detection Algorithm can be used to achieve more accuracy.

VI. ACKNOWLEDGMENT

We would like to take this opportunity to thank our internal guide Prof. Shobha S. Raskar for giving us all the help and guidance we needed. We are really grateful to her for her kind support. Her valuable suggestions were very helpful. We are also grateful to Dr. N. F. Shaikh, Head of Computer Engineering Department, Modern Education Society's College of Engineering for her indispensable support and suggestions.

REFERENCES

- [1] M. Mohandes, L. Mohandes, Deriche, "Image-based and sensor-based approaches to Arabic sign language recognition," IEEE Transactions on Human-Machine Systems, 2014.
- [2] <https://en.wikipedia.org/wiki/YCbCr>
- [3] Joyeeta Singha, Karen Das Dept. of Electronics and Communication Engineering Assam Don Bosco University Guwahati, India, "Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique" (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, Vol. 4, No. 2, 2013.
- [4] Eroding and Dilating — OpenCV 2.4.13.0 documentation.
- [5] Khamar Basha Shaika, Ganesan P, V.Kalist, B.S.Sathish, J.Merlin Mary Jenitha, "Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Colour Space." 3rd International Conference on Recent Trends in Computing 2015
- [6] http://www.emgu.com/wiki/index.php/Main_Page.
- [7] <http://www.emgu.com/wiki/index.php/Tutorial>
- [8] Megha Sahu, Toran Verma, Department of Computer Sci. & Engg. Rungta College of Engineering & Technology Bhilai, Chhattisgarh-India, "Recognition of Object Using Principle Component Analysis", *International Journal of Advance Research in Computer Science and Management Studies*, volume 2, Issue 9, September 2014.
- [9] https://en.wikipedia.org/wiki/Principal_component_analysis.
- [10] https://www.youtube.com/watch?v=g5_tonFnfaQ