

Hand Gesture Recognition and Voice Conversion for Deaf and Dumb

Rupesh Prajapati¹, Vedant Pandey², Nupur Jamindar³, Neeraj Yadav⁴, Prof. Neelam Phadnis⁵

^{1,2,3,4,5} Department of Computer Engineering, Shree L.R. Tiwari College of Engineering, Maharashtra, India

Abstract - Communication is the main channel between people to communicate with each other. In the recent years, there has been rapid increase in the number of deaf and dumb victims due to birth defects, accidents and oral diseases. Since deaf and dumb people cannot communicate with normal person so they have to depend on some sort of visual communication. There are many languages spoken all around the world and interpreted. "Special people", that is people who have difficulty in speaking and hearing "The dumb" and "The deaf" people respectively find it difficult to understand what exactly the other person is trying to express and so with the deaf people. Sometimes people interpret these messages wrongly either through sign language or through lip reading or lip sync. This project is made in such a way to help these specially challenged people hold equal par in the society.

Key Words: Sign language, Hand gesture, Feature extraction, Gesture recognition, Principal Component Analysis, Dumb and Deaf.

1. INTRODUCTION

Nowadays we always hear about new technology that improves our lifestyle, that makes our life easier. Technology has revolutionized the human mankind. Human race has put a gear in technology and they are not in a mood to move the pedals away from this gear. There is huge research on various technology sector such as Artificial Intelligence, Smart phones and many more. This research lead to new inventions and making one's life easier. But there has been a very less research for Deaf and Dumb people. This topic has get less attention as compared to other sectors. The Main challenges that this special person facing is the communication gap between -special person and normal person. Deaf and Dumb people always find difficulties to communicate with normal person. This huge challenge makes them uncomfortable and they feel discriminated in society. Because of miss communication Deaf and Dumb people feel not to communicate and hence they never able to express their feelings. HGRVC (Hand Gesture Recognition and Voice Conversion) system localizes and track the hand gestures of the dumb and deaf people in order to maintain a communication channel with the other people. The detection of hand gestures can be done using web camera. The pictures are then converted into standard size with the help of pre-processing. The aim of this project is to develop a system that can convert the hand gestures into text. The focus of this project is to place the pictures in the database and with database matching the image is converted into text. The detection involves observation of hand movement. The method gives output in text format that helps to reduce the communication gap between deaf-mute and people.

The paper is divided into 5 sections. The first section consists of all the researches related to topic. Second section is about the hand gesture recognition technique. Third section is all about how the system works. Fourth section is about the result of the system and the last section tells about the conclusion and accuracy of our system.

2. LITERATURE REVIEW

In Literature Review, we studied about existing project related to this topic and try to understand about the existing system behavior.

Shweta S. Shinde, Rajesh M. Autee and Vitthal K. Bhosale [1] have proposed a method in which the angle and peak calculation approach is used to extract the features of hand gestures by using MATLAB and then they convert the recognized gesture into speech using MATLAB inbuilt command.

Sangeetha .R.K, Valliammai .V and Padmavathi .S [2] have proposed a system based on the Indian hand sign language which contains both hands to create a gesture unlike the American sign language in which one hand is used. Their system is implemented using MATLAB without using any other external hardware for the user, here the runtime live image is captured after which image frames are extracted and image processing is applied using HIS model and then the feature extraction is done by distance transform method. The results obtained by this model is found to be satisfactorily good for most of the hand signs.

Anchal Sood and Anju Mishra [3] have proposed a sign recognition system based on Harris algorithm for extraction of feature in which after the image pre-processing part, the feature is extracted and stored in the Nx2 matrix. This matrix is further used to match the image from the database. There are some limitations to the system. The very light brown to somewhat dark brown background gives error as they are considered in the range value for skin segmentation. But the results are efficient.

Prashant G. Ahire, Kshitija B. Tilekar, Tejaswini A. Jawake, Pramod B. Warale [4] system works on MATLAB for hand gesture recognition, here they have taken a real time video as an input and after applying the image processing stage they have used the correlation based approach for mapping purpose and then at last the audio is generated using google TTS API. The system provides an efficient result as per the system is proposed.

Mrs. Neela Harish and Dr. S. Poonguzhali [5] have proposed a system based on the hardware approach, the system consist of hardware called as data glove, the glove consist of sensory part, that consist of flex sensors, accelerometer and PIC-microcontroller which provides all the input and output for the system. The results obtained are efficient and satisfactory.

Sonal Kumari and Suman K. Mitra [6] have purposed a system based on hand action recognition using background subtraction technique for image processing and they use Direct Fourier transform (DTE) algorithm for image extraction based on the MATLAB.

3. Hand Gesture Recognition Technique

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too.

There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction.

3.1 Image Matching Process

Image matching is a process in which the captured image is compared with the images stored in database.

3.1.1 Feature Extraction

Principal component analysis (PCA) is one of the statistical techniques frequently used in signal processing to the data dimension reduction or to the data decorrelation. Presented paper deals with two distinct applications of PCA in image processing. The first application consists in the image color reduction while the three color components are reduced into one containing a major part of information. The second use of PCA takes advantage of eigenvectors properties for determination of selected object orientation. Various methods can be used for previous object detection. Quality of image segmentation implies to results of the following process of object orientation evaluation based on PCA as well. Presented paper briefly introduces the PCA theory and Results are documented for the selected real pictures.

Principal component analysis (PCA) belongs to linear transforms based on the statistical techniques. This method

provides a powerful tool for data analysis and pattern recognition which is often used in signal and image processing as a technique for data compression, data dimension reduction or their decorrelation as well. There are various algorithms based on multivariate analysis or neural network that can perform PCA on a given data set. Presented paper introduces PCA as a possible tool in image enhancement and analysis.

ASL Fisher is a feature extraction algorithm which is an in-built function in MATLAB. It basically used to extract the Euclidian distance between two images.

3.1.2. Classification Method

K-Nearest Neighbor is a Learning algorithm that Defer in the decision to generalize beyond the training examples till a new query is encountered.

Whenever we have a new point to classify, classify, we find its K nearest neighbors from the training data. The distance is calculated using n Euclidean Distance.

“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well. Support Vectors are simply the co-ordinates of individual observation. Support Vector Machine is a frontier which best segregates the two classes.

4. SYSTEM ARCHITECTURE

The figure below represents the System Architecture of our system that basically show each component of the system, how the system works, and the flow of the system and so on. Images that are taken from the web camera goes under pre-processing stages to enhance the feature of an image. Then there is a removal of object and background from the images which later convert into binary form. Feature extraction and reorganization helps to match the images that is stored in database and we get the desired output in the form of text and converts that text to speech.

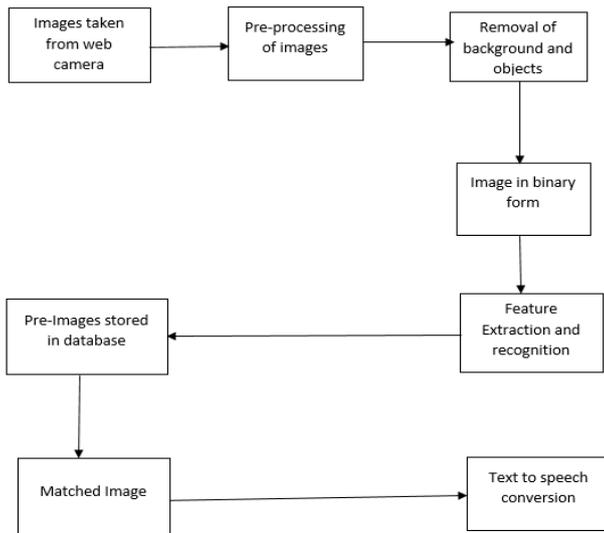


Fig -1: System Architecture

5. IMPLEMENTATION OF THE SYSTEM

Here we will discuss about how we implemented our system and is represented in a flowchart manner in Figure 1.

5.1 Training of System

User have to enter number of samples to store in the database. The number of samples should be more than 5 in order to get better accuracy. User have to select the folder where the images will get saved.

Click on start video to open the web camera in order to start the process of database creation. Click capture image to store the number of images in the training folder as per the no. of sample specified. When the number of images will be equal to number of captured image then done storing will get displayed that means database creation is done successfully.

5.2 Image Pre-Processing

The captured images go under pre-processing stage in order to enhance the property of an image. Pre-Processing is basically done to remove the object and background of an image and focus on the hand gestures only. The pre-processed image is then represented in the form of black and white pixels which basically means binarized image.

5.3 Feature Extraction and Recognition

PCA algorithm is used in order to extract the feature of an image. PCA algorithm is applied on the captured images in order to extract the best featured image from the database. PCA converts the images into some independent linear set of variables which refers to the information in the original data which is referred as principal components.

Following are the steps that is used to extract the feature of an image using PCA:

Step 1: Convert all images into the column matrix.

Step 2: Evaluate the Mean column matrix of column matrix.

Step 3: Calculate the difference for each vector set.

Step 4: Calculate a Covariance matrix.

Step 5: Calculate Eigen value and mean Eigen value for Covariance matrix.

Step 6: Sort the Eigen value.

Step 7: Calculate mapping eigenvectors and for matching, project the data.

After applying all the above steps, we get the reduced features of PCA and then calculate the result.

After feature Extraction, Recognition of an image is done by KNN and SVM algorithm. Recognized hand gesture is converted into text which converts into speech.

6. RESULTS

6.1 Database Creation

Database can be creating efficiently by user which is used for training the system, the only limitation while creating database is that background should white and clear. The user has to specify the number of samples; number of sample basically means number of images in database.



Fig -2: Number of sample



Fig -3: Database of images.

6.2 Image Capturing

Image Capturing is done by clicking on capture image and that image get stored in respective directory that we chose to save.

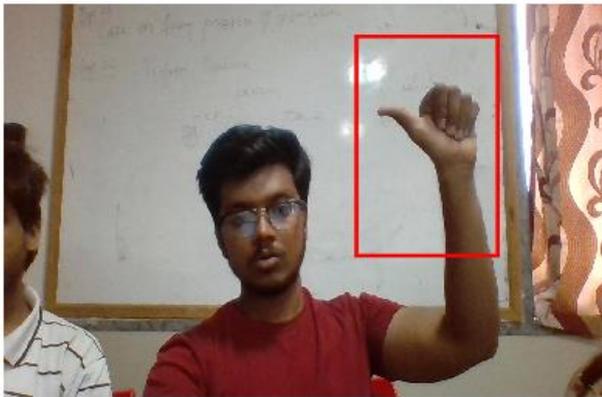


Fig -4: Image Capturing

6.3 Recognition of Image

For recognition of an image, click on recognized image, in order to recognize an image, capturing of an image should be done, after image capturing the image get recognized using PCA algorithm and classification of an images is done through KNN and SVM algorithm.



Fig -5: Recognition of Image

6.4 Final Result

Final result gets displayed in static text when captured image get extracted from image database. Hand gesture converts into through feature extraction and classification of an image.



Fig -6: Final Result

The resultant text gets convert into speech as a final output.

7. CONCLUSION

Hand Gesture recognition and voice conversion for dumb and deaf person was successfully executed using image processing. The method takes image as input and gives text and speech as an output. Implementation of this system gives up to 90% accuracy and works successfully in most of the test cases.

REFERENCES

- [1] Shinde, Shweta S., Rajesh M. Autee, and Vitthal K. Bhosale. "Real time two way communication approach for hearing impaired and dumb person based on image processing." Computational Intelligence and Computing Research (ICCIC), 2016 IEEE International Conference on. IEEE, 2016.
- [2] Shangeetha, R. K., V. Valliammai, and S. Padmavathi. "Computer vision based approach for Indian Sign Language character recognition." Machine Vision and Image Processing (MVIP), 2012 International Conference on. IEEE, 2012.
- [3] Sood, Anchal, and Anju Mishra. "AAWAAZ: A communication system for deaf and dumb." Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), 2016 5th International Conference on. IEEE, 2016.
- [4] Ahire, Prashant G., et al. "Two Way Communicator between Deaf and Dumb People and Normal People." Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on. IEEE, 2015.
- [5] Ms R. Vinitha and Ms A. Theerthana. "Design And Development Of Hand Gesture Recognition System For Speech Impaired People."
- [6] Kumari, Sonal, and Suman K. Mitra. "Human action recognition using DFT." Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2011 Third National Conference on. IEEE, 2011.