International Research Journal of Emerging Trends in Multidisciplinary
ISSN 2395-4434
Volume 1, Issue 5 July 2015
www.irjetm.com

Two Way Communicator between Deaf and Dumb People and Normal People

Prof. Prashant G. Ahire*, Pramod B. Warale†, Tejaswini A. Jawake‡, Kshitija B. Tilekar§

Department of Computer Engineering
Pimpri Chinchwad College of Engineering, Nigdi, Pune 411044, India
*ksprashantahire@gmail.com
†pramodwarale18@gmail.com
‡tajawake@gmail.com
§ktilekar263@gmail.com

Abstract—One of the most precious gift of nature to human beings is the ability to express himself by responding to the events occurring in his surroundings. Every normal human being sees, listens and then reacts to the situations by speaking himself out. But there are some unfortunate ones who are deprived of this valuable gift. This creates a gap between the normal human beings and the deprived ones. This application will help for both of them to communicate with each other. The system is mainly consists of two modules, first module is drawing out Indian Sign Language(ISL) gestures from real-time video and mapping it with human-understandable speech. Accordingly, second module will take natural language as input and map it with equivalent Indian Sign Language animated gestures. Processing from video to speech will include frame formation from videos, finding region of interest (ROI) and mapping of images with language knowledge base using Correlational based approach then relevant audio generation using Google Text-to-Speech (TTS) API. The other way round, natural language is mapped with equivalent Indian Sign Language gestures by conversion of speech to text using Google Speech-to-Text (STT) API, further mapping the text to relevant animated gestures from the database.

Keywords- Correlational based approach, Region of Intrest, Region growing,STT, TTS, ISL.

I. INTRODUCTION

The current era is a zoom of technology. Each and every field has an impact of the technological advancements onto it. Once such rapidly growing technical advancement is the increasing impact of mobile phones on human life. The enormous and ever increasing Internet along smart phones has proven a boon to mankind. The smart phone is a most important electronic gadget in our life, because it is with the smart phone that we stay connected. But, despite of these advancements, there is a certain part of the society which is deprived of these benefits.

SponserOranization : Persistent Systems Pvt. Ltd.

The hearing disable and mute people cannot mingle into the social world because of their physical disabilities. Unintentionally, they are treated in an unusual manner by the rest of the society. They cannot be a part of the social events, say students cannot study in schools with normal students, elderly persons cannot work at work places, and much more. Simple activities like going and buying a commodity from the grocery shop is very complicated task for the deaf and dumb person. The gap between normal human beings and deaf and dumb is wide and ever increasing day-by-day. Today, the national count of hearing disabled and mute persons throughout India is approximately 17 lakh. Despite of this large number, very less research is done in order to bridge the communication barrier.

An attempt to bridge the communication barrier, we propose an Application which helps normal and deaf and dumb people to effectively communicate with each other. Human Computer Interaction and Image Processing are the related areas of research which will help us build a solution to this problem. The rest of the paper is organized as follows.

International Research Journal of Emerging Trends in Multidisciplinary
ISSN 2395-4434
Volume 1, Issue 5 July 2015
www.irjetm.com

Section II we will introduce with modules included in the application and the architectural diagrams of the module. In the section III it mainly consists of the processing description of the first module and section IV contain the module description of the second module of project, following acknowledge and conclusion.

II. PROPOSED METHODOLOGY

In this section, we present an overall description of how we develop the solution. Our solution consists of two main modules, as depicted below:

A. Conversion of real time video(ISL) to equivalent human natural language speech.[1]

As shown by fig. 1, the process of drawing out gestures from real time video and converting it to natural language comprises of the following phases:

- Frame formation from real time video
- Pre-processing and noise removal
- Finding ROI using Region growing
- Mapping using Correlation based approach
- Relevant audio generation using Google TTS API

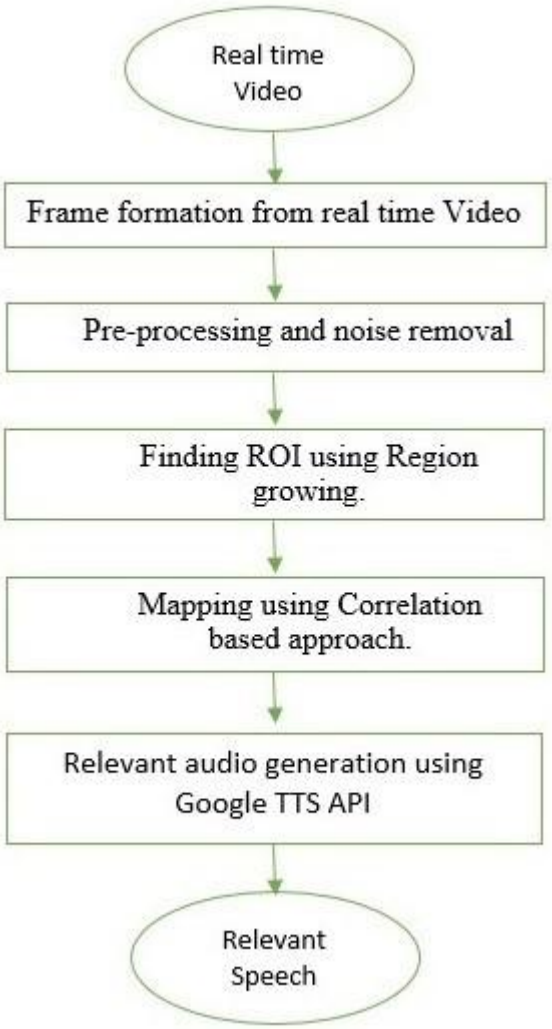

Fig. 1. Overall view of real time video to speech

B. Mapping of human speech (natural language) to animated videos corresponding to ISL gestures.

As shown in fig.2, the process of converting natural language to corresponding gesture animation equivalents comprises of following phases:

- Generation of text from real time audio speech using Google STT API
- Mapping text to gesture images in database.
- Sequencing the images to form video
- Display animated gesture video.

International Research Journal of Emerging Trends in Multidisciplinary
ISSN 2395-4434
Volume 1, Issue 5 July 2015
www.irjetm.com

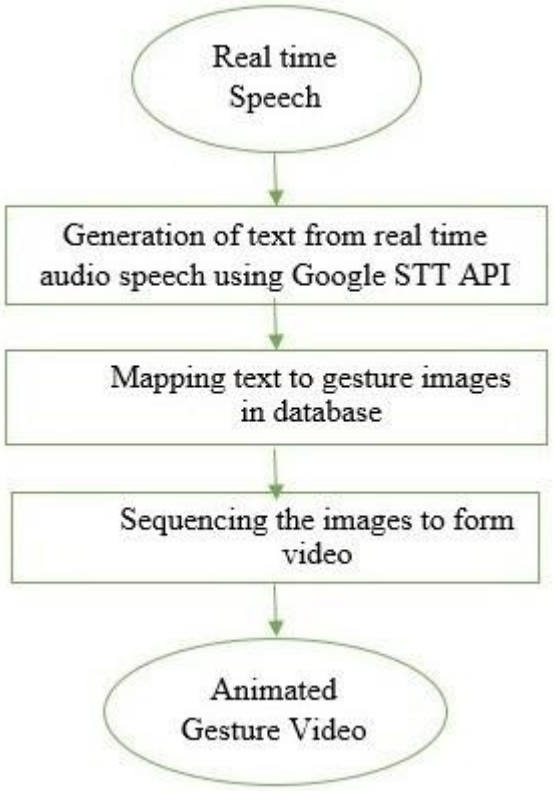

Fig. 2. Overall view of speech to animated video

III. MODULE I

A. Frame formation from real time video

Video sequences for sign language gestures are captured using the camera device. When we talk about processing real time videos with computer, it is complicated and tedious task as compared to processing of images. Thus, we simplify this process by breaking the video into image frames. Basically, a video is nothing but huge sequence of images captured at high data rate. So, formation of frames form video is breaking the video into its original base form that is images [2].

B. Pre-processing and noise removal

Nosie removal in image processing deals with improving the image in terms of brightness, accuracy and other similar factors. Histogram based technique is one of the important technique in image processing. This technique is also used for image enhancement. Also, this technique is much less expensive as compared to other methods. This technique is based on equalizing the histogram of the image and increasing the dynamic range corresponding to the image [1]. Here, we use the concept of reference image. All the input images will be equalized to the reference image. So all the input images will have equal amount of brightness and other values.

C. Finding ROI using Region growing [1]

The main goal of image segmentation is domain independent partitioning of an image into a set of disjoint regions that are visually different, homogeneous and meaningful with respect to some characteristics or computed property such as grey level, texture or colour to enable easy image analysis [4]. Region based segmentation method is used for partitioning the image into similar areas of connected pixels through the application of similarity criteria among candidate set of pixels. Each of the pixel in the region is similar with respect to some characteristics. For this approach, we use Connected Component algorithm [3] in order to identify the gesture region and to make region growing computationally efficient, a block based updating process is used. This method works well even on input image with noise, shadow and occlusion. First, given an input image, the system selects the centre pixel to be the seed pixel for the hand region to be grown. The remaining region is defined as the source region φ (φ=1- Ω).). Next, we set the size of the template window Ψ as 3*3 pixels. Once these parameters are defined, the region-growing proceeds automatically. The four main steps of our hand segmentation algorithm are as follows:

Step 1: Computing patch priorities.
Given a patch p Ψ centre at the point p for some p, where denotes the boundary of . This paper defines its priority P(p) as the product of the three terms:

$$P(p)= D(p).C(p).M(p)$$

Where D(p), C(p) and M(p) respectively depict the data term, the confidence term and the absolute value of mean difference term.

International Research Journal of Emerging Trends in Multidisciplinary
ISSN 2395-4434
Volume 1, Issue 5 July 2015
www.irjetm.com

Step 2: Updating Region Growing
Once all priorities on the region front have been computed, the patch up with the least priority is found. Region growing data is extracted from the source region Φ. After updating region growing, the hand segmentation algorithm updates the front related to patch Ψ. Then update the data term D(p), which is near to the patch Ψ .

Step 3: Updating Confidence Values After the patch Qi has been updated with new pixel values, the confidence C(p) is updated as follows:

$$C(p)=C(q),$$

$$\forall p \in \Psi q \cap \Omega$$

The rule is simple, but it measures the relative confidence of patches on the region front without specific image parameters.

Step 4: Updating the Absolute Value of Mean Difference After updating the region growing mean, then the algorithm updates the absolute value of mean difference on the region front .

D. Mapping using Correlation based approach. [1]

The part of image that we get after finding region of interest by using region growing method is further used for feature extraction or selection. Correlation based approach is mainly used for matching this extracted features from inputted image to the gesture image that is stored in database. We find the correlation co-efficient between the database image and the input image so that the best match is obtained. The correlation coefficient is computed between A and B, where A and B are matrices or vectors of the same size.

Correlation Co-efficient

$$\text{Correlation}(r) = \left[\frac{N\Sigma XY - \Sigma X.\Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}}\right].[N\ \Sigma\ Y^2 - (\Sigma Y)^2]$$

Where

N = Number of values or elements
X= First Score
Y= Second Score
ΣXY = Sum of the product of first and Second Scores
ΣX = Sum of First Scores
ΣY = Sum of Second Scores
ΣX2 = Sum of square First Scores
ΣY2 = Sum of square Second Scores

E. Relevant audio generation using Google TTS API.

At this phase, we have mapped sign language gesture in text format. Generally, thinking it is pretty good processing if successful conversion of sign language gestures to relevant text data is performed. But, for better Human Computer Interaction, it is preferable to further process the text and convert it to relevant natural language speech. The Google TTS engine gives us all provisions for conversion of text to relevant audio natural speech.

IV. MODULE 2

A. Generation of text from real time audio speech using Google STT API

This module aims at making the normal person able to communicate successfully with the deaf and dumb person. The normal human being is capable of giving natural language as input from his side. So, in order to process this natural language, the first step is to generate text data out of audio speech data. It is easy to process to text data as compared to audio speech signals. The Google SST (Speech to Text) API provides us with all the functionalities and provisions to convert audio speech to corresponding text for further processing.

B. Mapping text to gesture images in database.

In this phase the text which we get after applying Google STT Engine is again extracted for its specific characters. All this extracted characters are then matched with the ISL gesture images that are stored in database. Then finally user will get sequence of gesture images for the corresponding text.

C. Sequencing the images to form video

This step deals with collecting together all the images obtained in the previous step and playing them together at a constant rate. The rate may vary depending on the application in which the video is to be used.

D. Display animated gesture video.

This phase deals with displaying the animated gesture video with proper display settings. Additional features like video display accuracy and other such factors are considered here.

V. CONCLUSION

The proposed system Two Way Communicator between Deaf and Dumb People and Normal People aims at bridging the communication gap between two strata of the society. The above specified strategies prove efficient in terms of time and accuracy. Further enhancements, can be made in terms of implementing the communicator with other sign languages like American Sign Language, accent recognition for various accents throughout the globe, emotion recognition in sign language and language translation.

VI. REFERENCES

[1] Ashish Sethi, Hemanth S,Kuldeep Kumar,Bhaskara Rao N,Krishnan R, SignPro-An Application Suite for Deaf and Dumb, IJCSET —May 2012— Vol 2, Issue 5,1203-1206 ISSN 2231-0711.

[2] Ms. Rashmi D. Kyatanavar, Prof. P. R. Futane ,Video Gesture Classification Using Fourier Descriptors and Genral Fuzzy Min Max Neural Network, International Journal of Emerging Technology and Advanced Engineering. ISSN 2250-2459, Volume 2, Issue 5, May 2012.

[3] Ma De-yi, Chen Yi-min, Li Qi-ming, Huang Chen, Xu Sheng, "Region Growing by Exemplar-Based Hand Segmentation under Complex Backgrounds".

[4] ShilpaKamdi, R.K.Krishna,Image Segmentation and Region Growing Algorithm, International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 2, Issue 1 ISSN 2249-6343.

[5] Zeno GeradtsJurrienBijhold Rob Hermsen, Use of Correlation algorithms in a database of spent cartridge cases of firearms.

[6] D.Sasirekha, E.Chandra Text to speech International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, March 2012.

[7] P.Rajesh Kumar P.V.V.Kishore Segment, Track, Extract, Recognize and Convert Sign Language Videos to Voice/Text, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 3, No.6, 2012.

[8] Kamalpreet Sharma1, Naresh Kumar Garg1, kamalpreet Sharma hand gestures recognition for deaf and dumb International Journal of Computer Application and Technology (s), May - 2014, pp. 1013 ISSN 2349-1841.

[9] Aditi Kalsh, N.S. Garewal Sign Language Recognition for Deaf and Dumb International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 9, September 2013 ISSN: 2277 128X.

[10] AnupNandy, SoumikMondal, Jay Shankar Prasad, Pavan Chakra borty and G.C.Nandi Recognizing and Interpreting Indian Sign Language Gesture for Human Robot Interaction Intl Conf. on Computer and Communication Technology —ICCCT10—.