iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4414

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

Real Time Sign Language Processing System

Arfat Shaikh[1], Aafreen Sayyed[2], Shreeraj Shettigar[3], Ajay Sirsat[4]
[1, 2, 3, 4,] *Computer Science Department, Mumbai University* St . john college of engineering and management

Abstract*: The only way the speech and hearing impaired (i.e dumb and deaf) people can communicate is by sign language. The main purpose of this project is to ease communication for mute people in situations where they have to interact with people who are unable to understand sign language. The main idea of this project is to build a system using which mute people can communicate with all other people using their normal gestures. The system does not require the background to be perfectly black. It works on any background. The project uses artificial intelligence, machine learning and image processing system to identify, sign language used by the deaf people to communicate and converts them into text so that normal people can understand.

Keywords: Sign language; artificial intelligence; machine learning; image processing.

I. INTRODUCTION

Communication is the foundation of human relationship. It plays a vital role in human life. Communication and community are significant parts of human life. According to the World Health Organization (WHO), there are approximately 70 million that are mute. Mute people are isolated from the most common forms of communication in today's society such as warnings, or any other form of oral communication between people in regular daily activities such as visiting the doctor or communicating in the street. In other words, mute people can often feel disassociated and thus find it hard to get information or help in daily activities or even encountered in emergency situations. As a consequence, mute people are twice as likely as hearing people to be affected by depression, anxiety and similar problems. Sign language is primary means of communication in deaf and dumb community. It is a medium of communicating your thoughts and ideas through the facial expressions, hand shapes and finger spellings. The project builds an application for translating sign language to natural language. The project will be delivering output corresponding to the gestures. The project uses custom created sign language for better success rate.

II. METHODOLOGY

In this paper a robust and efficient method of sign language detection is presented. The method has five main phases consisting of image capture and processing, co-ordinate mapping, thresholding [3] and pattern matching [2]. The methodology does not make use of data gloves instead detection is done using real time image processing. By using a threshold value while converting the image from Grayscale to Binary form, this system can be used in any background and is not restricted to be used with Black or White Background.

III.SYSTEM ARCHITECTURE

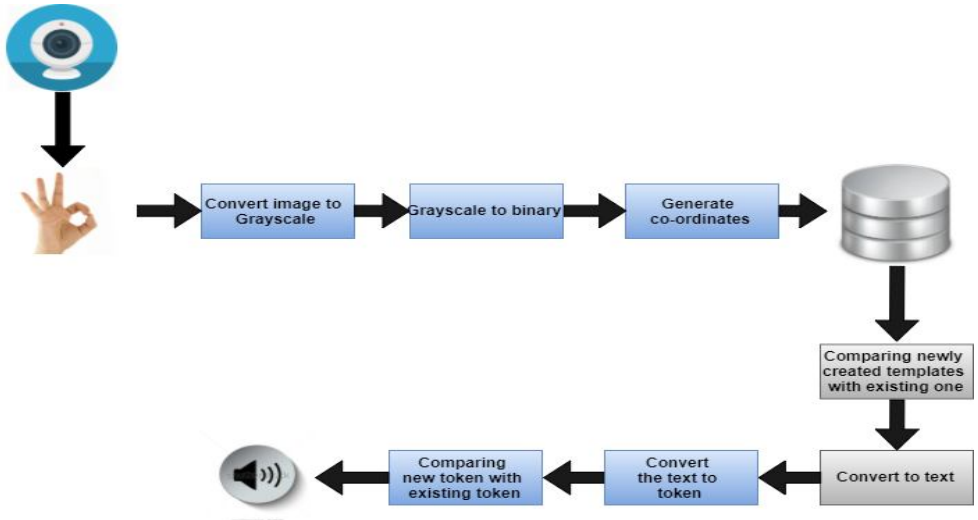


Fig. 1 System Architecture

Fig. 1 shows the overall idea of proposed system. The system consists of 4 modules. Images are captured through the webcam. The camera is placed on the front facing the ground with neutral background. Firstly, the captured Colored image is converted into the gray scale image which intern is converted into the binary form. With respect to X and Y coordinates the coordinates of captured image is calculated and then stored in the form of template into the database. The templates of newly created coordinates are compared with the existing one. If comparison leads to success then the same will be converted into audio form. The system works in two different mode i.e. training mode and operational mode. Training mode is part of machine learning where we are training our system to accomplish the task for which it is implemented i.e. Alphabet Recognition.

A. Interface and Image Acquisition

The first block i.e the camera interface block is hardware block that acts as an interface and provides a standard output that can be used as input for subsequent image processing.

1) Camera Orientation: This is a very important decision to make so as to minimize the background noise and have maximum success rate. There are two realistic options available, either point the camera towards the wall or towards the floor. The camera is pointed towards the floor since in that case we will have less background noise to deal with.

2) Camera specification: Logitech C270 HD Webcam is used, a high utility device. The Logitech C270 webcam is 3 MP camera. The webcam gives clear video imaging and can work even in the darkness. The upper portion of the webcam is Movable depend on the need. The resolution of captured image is 640x480 having frame rate up to 30fps and the format of image is RGB 24, 1420.

B. Image Processing

1) RGB Recognition

Fig. 2 Input

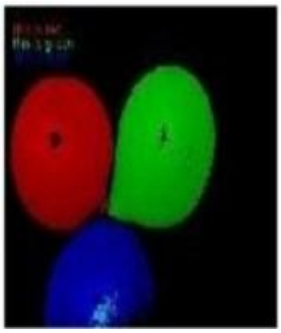


Fig. 3 Output

RGB recognition is the process of segmenting the colors and identifying the segmented colors. It is the one of the human to computer interaction. Here the colors are acting as interface between human and computer. At initial, primary color modal is used for a recognition process, it only recognizes primary colors. In every color images, primary colors are segmented and segmented colors are recognized to identify its name like red, green, blue. After segmenting every three colors their respective pixels are counted. The given colors are recognized once it gets above 300 pixels of these three colors(RGB). The color image is taken and the required portion of image is made as white by using Thresholding technique(as explained below) and garbage part that is background as black. Then we get black and white image and it is compared with the stored template [1].

2) Thresholding

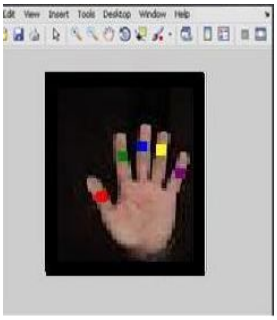


Fig. 4 Original Image

Fig. 5 Binary Image

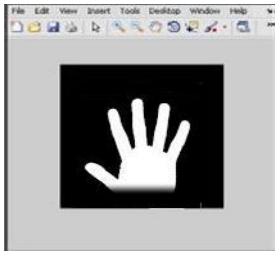


Fig. 6 Binary image after masking

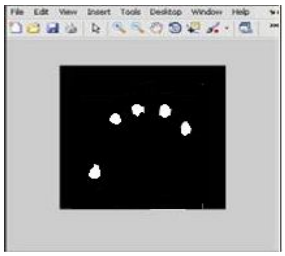


Fig. 7 Only the area of interest is preserved

Thresholding is the simplest method of image segmentation. In this method the RGB image is converted to Binary image. Fig.4, Fig.5, Fig.6, Fig.7 shows the details of image processing. Binary image is digital image and has only two values, 0 or 1. A threshold is applied to the grayscale image and the gray levels below the minimum value of the threshold are converted into black while the ones above the threshold value is selected such that it represents skin colour in RGB form of image. Thus, the system is not restricted with only black or white background and can work in any background.

3) Pattern Matching: The output image after thresholding is now compared with the images present in the database. The comparison is done between the binary image and the binary images present in the database. Comparing algorithm is used(Explained in section 4) to compare captures image with all the images from database. A single gesture is captured from more than 2 angles to improve success rate. Pixels of captured image are with pixels of images in database, if 90 percent of the pixel values are matched then the audio corresponding to the gesture(present in the database) is played. If no match found then the pixels are compared with the next image in the database. This process keeps going on till match found, if no match is found then that image is discarded and next image is considered for pattern matching.

4) Pattern matching algorithm: The processed image is set to the pixel values [2] of each color to be used such as Red_new (Rx, Ry), Green_new (Gx, Gy), Blue_new (Bx,By), Purple_new (Px,Py), Yellow_new (Yx,Yy). Pixel values consists values of each color pixel also can be called as co-ordinates. The generated values of these co-ordinates will be then compared with the values present in database. These values are obtained by firstly finding the area
of each color pixel and coordinates (Yx,Yy) by using the equation:

Area = count number of white pixels obtained

$$X = \frac{\text{Moment in X direction}}{\text{Area}}$$

$$Y = \frac{\text{Moment in Y direction}}{\text{Area}}$$

Each newly generated pixel value then gets compared with the previously stored template value in the database. Algorithm proceeds until the comparison leads to success or failure. If algorithm returns positive results then the sign will be converted into corresponding audio, if comparison fails then the algorithm will move to the next image.

5) Text To Speech Conversion: After the successful completion of the template matching operation the matched image is then converted into audio format. This is done by linking each audio files with corresponding gestures(pixel values) present in the database [3].

IV. ALPHABET RECOGNITION

Each finger tip is assigned a value, thus giving rise to various hand gestures possibilities, for single hand use there are 31 possibilities. If both the hands are used then possibilities goes up to 1023 gestures.

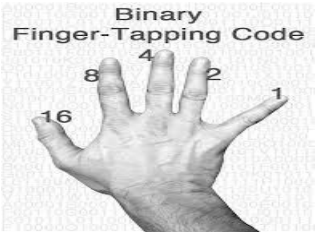


Fig. 8 Finger Tapping

SR NO	RED	GREEN	BLUE	YELLOW	PURPLE	ALPHABETS
1	0	0	0	0	1	A
2	0	0	0	1	0	B
3	0	0	0	1	1	C
4	0	0	1	0	0	D
5	0	0	1	0	1	E
6	0	0	1	1	0	F
7	0	0	1	1	1	G
8	0	1	0	0	0	H
9	0	1	0	0	1	I
10	0	1	0	1	0	J
11	0	1	0	1	1	K
12	0	1	1	0	0	L
13	0	1	1	0	1	M
14	0	1	1	1	0	N
15	0	1	1	1	1	O
16	1	0	0	0	0	P
17	1	0	0	0	1	Q
18	1	0	0	1	0	R
19	1	0	0	1	1	S
20	1	0	1	0	0	T
21	1	0	1	0	1	U
22	1	0	1	1	0	V
23	1	0	1	1	1	W
24	1	1	0	0	0	X
25	1	1	0	0	1	Y
26	1	1	0	1	0	Z

Fig. 9 Alphabet Code

Fig. 10 Gesture of alphabet G

V. FUTURE SCOPE

In future work, the proposed system can be developed more to recognize signs that involve motion. The system can be implemented on Rasberry pie computer, to make the entire system portable. The sequences of gestures are converted into text first and then into speech.

VI. CONCLUSION

The project aims to make communication simpler between mute and normal people by introducing Computer in communication path so that sign language can be automatically captured, recognized, translated to speech(audio).

VII. ACKNOWLEDGMENT

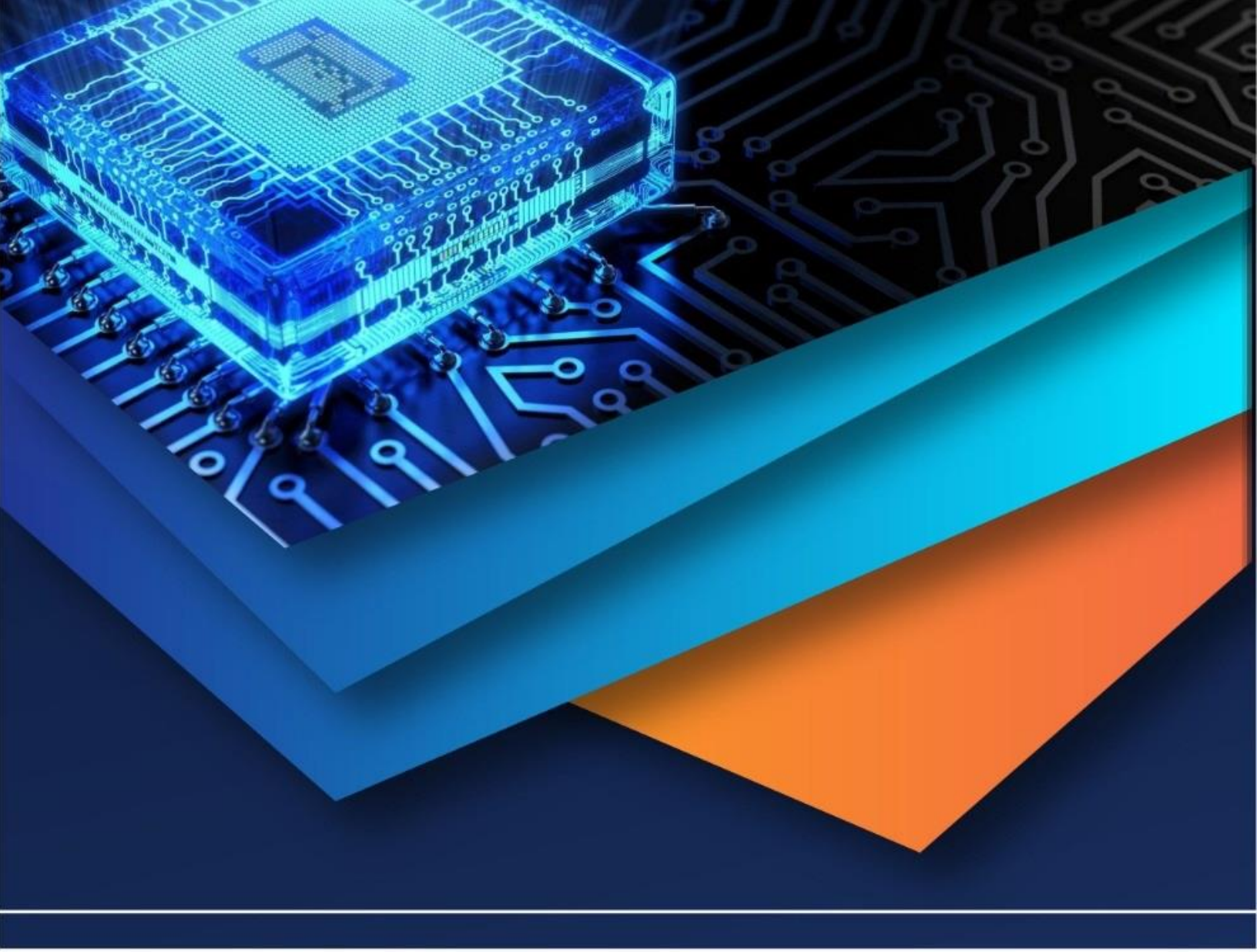
The authors would like to thank Mr. Ajay Sirsat, Assistant professor at computer science department, St John College of Engineering and Management for his continuous support and guidance. we thank all the faculty members and entire computer science department who contributed directly or indirectly to this work.

REFERENCES

[1] D.Senthamaraikannan, S.Shriram, Dr.J.William "REAL TIME COLOR RECOGNITION" International Journal Of Innovative Research In Electrical, Electronics, Instrumentation and control Engineering Vol. 2, Issue 3, March 2014

[2] Aleksandra Mojsilovic, Jianying Hu "A METHOD FOR COLOR CONTENT MATCHING OF IMAGES" Technical Report RT-0030, IBM Research, Tokyo Research Laboratory, Nov. 1989.

[3] P.Viola and M.Jones, "Rapid object detection using a boosted cascade of simple features" in Computer Vision and Pattern Recognition, Cambridge, MA(USA),2001.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY