

# Image Processing based Language Converter for deaf and dumb

S.N.Boraste

Department of Electronics and Telecommunication  
Engineering (PG Student)  
KCT's Late G. N. Sapkal College of Engineering  
Anjaneri, Nashik, India  
Swatiboraste@rediffmail.com

K.J.Mahajan

Department of Electronics and Telecommunication  
Engineering  
KCT's Late G. N. Sapkal College of Engineering  
Anjaneri, Nashik, India  
Kjmahajan264@gmail.com

**Abstract** – *This application helps the deaf and dumb person to communicate with the rest of the world using sign language. Suitable existing methods are integrated in this application. Computer recognition of sign language is an important research problem for enabling communication with hearing impaired people. The Computer based intelligent system will enable deaf & dumb people significantly to communicate with all other people using their natural hand gestures.*

**Index Terms**— *This application helps the deaf and dumb person to communicate with the rest of the world using sign language.*

## I. INTRODUCTION

Deaf and Dumb people are usually deprived of normal communication with other people in the society. It has been observed that they find it really difficult at times to interact with normal people with their gestures, as only a very few of those are recognized by most people. Since people with hearing impairment or deaf people cannot talk like normal people so they have to depend on some sort of visual communication in most of the time. Sign Language is the primary means of communication in the deaf and dumb community. As like any other language it has also got grammar and vocabulary but uses visual modality for exchanging information. The problem arises when dumb or deaf people try to express themselves to other people with the help of these sign language grammars. This is because normal people are usually unaware of these grammars. As a result it has been seen that communication of a dumb person are only limited within his/her family or the deaf community. The importance of sign language is emphasized by the growing public approval and funds for international project. At this age of Technology the demand for a computer based system is highly demanding for the dumb community. However, researchers have been attacking the problem for quite some time now and the results are showing some promise. Interesting technologies are being developed for speech recognition but no real commercial product for sign recognition is actually there in the current market. So, to take this field of research to another higher level this project was studied and carried out. The basic objective of this research was to develop a computer based intelligent system that will enable

dumb people significantly to communicate with all other people using their natural hand gestures. The idea consisted of designing and building up an intelligent system using image processing, data mining and artificial intelligence concepts to take visual inputs of sign language's hand gestures and generate easily recognizable form of outputs in the form of text & Voice.

## II. RELATED WORK

The literature survey is carried out as a part of the project work. It has provided review of the past research about image processing based language converter and other researchers. The past research effort will properly guide to justify the scope and direction of the present effort.

This chapter reviews the literature on sign language & its application in various national and international journals. Not many Researches have been carried out in this particular field, especially in Binary Sign Language Recognition. Few researches have been done on this issue though and some of them are still operational, but nobody was able to provide a full fledged solution to the problem. Christopher Lee and Yangsheng Xu developed a glove-based gesture recognition system that was able to recognize 14 of the letters from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode, with a rate of 10Hz. Over the years advanced glove devices have been designed such as the Sayre Glove, Dexterous Hand Master and Power Glove [1].

The most successful commercially available glove is by far the VPL Data Glove [2]. It was developed by Zimmerman during the 1970's. It is based upon patented optical fiber sensors along the back of the fingers. Star-ner and Pentland developed a glove-environment system capable of recognizing 40 signs from the American Sign Language (ASL) with a rate of 5Hz.

Another research is by Hyeon-Kyu Lee and Jin H. Kim presented work on real-time hand-gesture recognition using HMM (Hidden Markov Model). Kjeldsen and Kendersi devised a technique for doing skin-tone segmentation in HSV space,

based on the premise that skin tone in images occupies a connected volume in HSV space. They further developed a system which used a back-propagation neural network to recognize gestures from the segmented hand images [1].

Etsuko Ueda and Yoshio Matsumoto presented a novel technique a hand-pose estimation that can be used for vision-based human interfaces, in this method, the hand regions are extracted from multiple images obtained by a multi viewpoint camera system, and constructing the “voxel Model”[6]. Hand pose is estimated. Chan Wah Ng, Surendra Ranganath presented a hand gesture recognition system, they used image furrier descriptor as their prime feature and classified with the help of RBF network. Their system’s overall performance was 90.9%. Claudia Nolker and Helge Ritter presented a hand gesture recognition modal based on recognition of finger tips, in their approach they find full identification of all finger joint angles and based on that a 3D modal of hand is prepared and using neural network.

Their complex spatial grammars are remarkably different from the grammars of spoken languages [1], [2]. of sign languages, such as ASL (American Sign Language), BSL (British Sign Language), Auslan (Australian Sign Language) and LIBRAS (Brazilian Sign Language) [1], are in use around the world and are at the cores of local deaf cultures. Unfortunately, these languages are barely known outside of the deaf community.

Depth information makes the task of segmenting the hand from the background much easier. Depth information can be used to improve the segmentation process, as used in [4], [5], [6], [7]. Recently, depth cameras have raised a great interest in vision computer community due to their success in many applications, such as pose estimation [8], [9], tracking [10], object recognition [10], etc. Depth cameras were also used for hand gesture recognition [11], [12], [13]. Uebersax et al. [12] present a system for recognizing letter and finger spelled words. Pugeault & Bowden [11] use a Microsoft.

Kinect TM device to collect RGB and depth images. They extracted features using Gabor filters and then a Random Forest predicts the letters from the American Sign Language (ASL) finger spelling alphabet. Issacs & Foo [14] proposed an ASL finger spelling recognition system based on neural networks applied to wavelets features. Bergh & Van Gool [15] propose a method based on a concatenation of depth and colorsegmented images, using a combination of Haar wavelets and neural networks for 6 hand poses recognition of a single user.

Several techniques have been reported on gesture recognition which includes skin segmentation using color pixel classification [1], region growing by exemplar-based hand segmentation under complex background [2], Parametric Hidden Markov models for gesture recognition [3], statistical database comparison method [4], accelerometer-based gesture recognition

system [5], orientation histograms for gesture recognition [6], Finger Detection for Sign Language Recognition[7] etc. Most of the gesture recognition systems use special devices like hand glove [11]. The gloves get connected to the computers using a lot of cables. So these devices are cumbersome and expensive. In order to overcome these difficulties, alternatively vision-based approaches involving camera and image processing for recognizing gestures are being explored. In 2005, there were two significant proposals [1, 2] for treating color to grayscale conversion as an optimization problem involving either local contrasts between pixels [1] or global contrasts between colors [2]. However, unconstrained optimization over thousands of variables [1] or constrained optimization over hundreds of variables [2] is a slow process prone to local minima. It can prove impractical in applications that demand real-time, automated performance with mathematical guarantees on the results. Our algorithm satisfies these requirements without resorting to numerical optimization. Hence, our algorithm could easily be embedded in the driver software that renders color images on a grayscale printer. Also, its design is suitable for hardware implementation. Our running time is linear in the number of pixels, unlike the Gooch et al. algorithm [1], and is independent of the number of colors in the image, unlike the Rasche et al. algorithm [2]. Owing to their complexity, the optimization algorithms [1,2] offer few guarantees to aid in the interpretation of their results. For instance, they can fail to map pure white in the color image to pure white in the grayscale image, causing potential problems for pictures with white backgrounds. While our algorithm does not purport to model human visual adaptation, it ensures that converting color to grayscale leaves achromatic pixels unchanged. The closest precedent for our approach to global color to grayscale conversion is the classic work of Strickland et al. [3] on saturation feedback in unsharp masking, a local color image enhancement technique used to sharpen images. However, our method is designed to take advantage of both hue and saturation differences when augmenting the luminance channel.

### III. SYSTEM DEVELOPMENT

The goal of proposed method is to convert RGB to text message. Fig.1 shows the overall idea of proposed system The system consists of 3 modules. Image is captured through the webcam. The camera is mounted on top of system facing towards the wall with neutral background. Firstly, the captured Colored image is converted into the gray scale image which intern converted into the binary form. Coordinates of captured image is calculated with respect to X and Y coordinates. The calculated coordinates are then stored into the database in the form of template. 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 and textual 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.

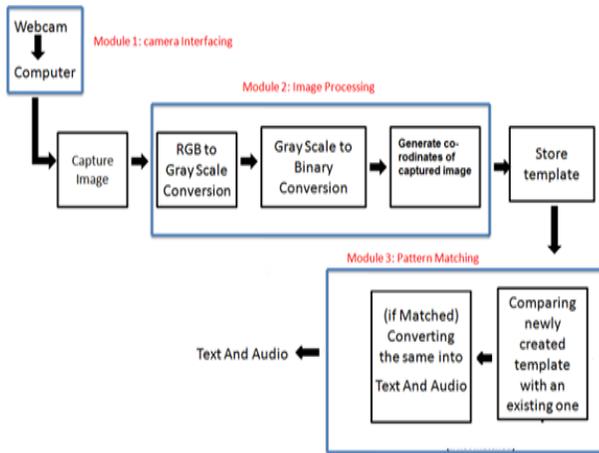


Figure 1: System block diagram

V. EXPERIMENTAL RESULTS

The experimental results of our project are as shown below:- Up to explain how system works theoretically. How RGB colors are detected & and how convert RGB to binary. As per experimental results explain how Matching template Output of system in text and audio.

The algorithm is implemented using C# & .net using various real time and standard images. The real time images are captured from web camera. The experimental results show the robustness of algorithm in both real time and standard images even in case of rotation, missing data. The main motive of this work is to Convert RGB to binary in form of audio & text.

A. RGB Color Recognition

The above Fig.2 indicates the red image captured using web camera & is the target image for binary conversion.



Figure 2: Captured Red Image using web camera

B. Conversion of RGB Color to Pixel data

After capturing all above images using webcam, different pixel values calculated as shown in below image.

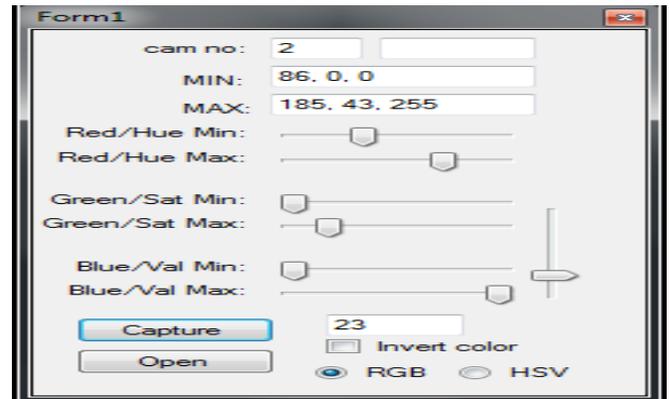


Figure 3: Captured Pixel Values for Red Image using web camera

C. Color image to Binary image conversion

For above Pixel values below is the binary image for red color.

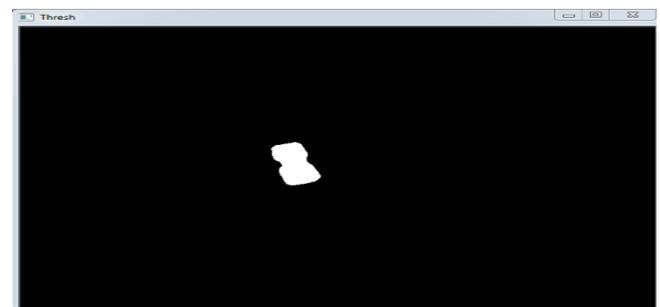


Figure 4: Captured Binary Image for Red Image using web camera

	Purple	Parrot gre	Blue	Dark Gree	Red
Power of Two	2 <sup>4</sup> .	2 <sup>3</sup>	2 <sup>2</sup>	2 <sup>1</sup>	2 <sup>0</sup>
Value	16	8	4	2	1

D. Coordinate Mapping

After getting the marker pixels that are now highlighted as a white color pixels, coordinates of that area for each color is generated. The newly generated coordinates are the compared with the stored coordinates in the database for the purpose of output generation using pattern matching technique explained in the next section.

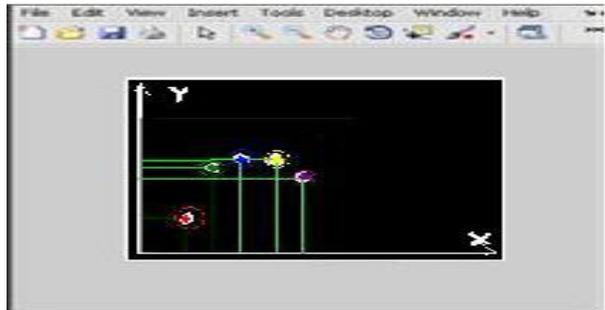


Figure 5: Co-ordinate Mapping

In this method, the input image after processing is set to the pixel values (3) of each color to be used such as Red\_new (Rx, Ry), Green\_new (Gx, Gy), Blue\_new (Bx, By), Purple\_new (Px, Py), Parrot Green \_new (PGx, PGy). Pixel values comprise of the minimum and maximum values of each color pixel or can be called as co-ordinates. The generated values of these co-ordinates will be then compared with the values stored in the templates stored in the database. To obtain these values the general idea is firstly to find the area of each color pixel and the coordinates (Yx,Yy) by using the equation: Area= count number of white pixels obtained by Thresholding.

Table 2. Alphabets code

SR NO	Purple	Parrot gre	Blue	Dark Gree	Red	Text
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
27	1	1	0	1	1	Space

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

E. Alphabet Recognition

Following table 1 shows the values assigned to each finger. Binary Alphabet calculation: It is possible to display total (2<sup>5</sup>-1) i.e.32 gestures using the fingers of a single hand.

Table.1 Values assigned to each Color

F. RGB to Text conversion

Putting pixel values to run the program for above different colors giving text messages as per above alphabate codes. Below images are the example shown for Red color. While showing Red input as shown in below input Fig 6 in front of webcam text form shows 'a' alphabate as shown in below output image 4.8. In the similar way deaf & dump can talk with other person using different colors. Below output image 4.9 is the example of one of the

statement.



Figure 6: Input Image

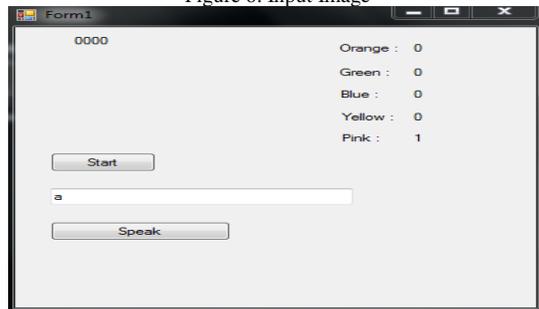


Figure 7: Output Image  
V. CONCLUSION

In this way see that

1. How system block diagram works theoretically.
2. How RGB detect .
3. How RGB to binary conversion work its result.
4. How Binary to text message & audio result.

#### VI. ACKNOWLEDGMENT

Acknowledgment not just being a hollow formality, I take this opportunity to express whole-heartedly, my sincere thoughts to all those who were responsible for the successful completion of this dissertation.

With a deep sense of gratitude, I wish to express my sincere thanks to my guide, Prof. K.J. Mahajan E&TC Engineering Department, without whom this dissertation would not have been possible. The confidence and dynamism with which she guided the work requires no elaboration. With her profound knowledge, the timely suggestions given and encouragement shown, made me more confident in successfully carrying out this project work. The opportunities given by her have stimulated many of my interests and enabled me to gain more exposure in the area of my work.

I am also thankful to Prof. Bagal S. B., Head of Electronics & Telecommunication Engineering Department, for giving me an opportunity to carry out the dissertation work within the college premises and allowing me to available the departmental facilities.

I express my thanks to Principal Dr. Balapgol B. S., for his continuous support. I am also thankful to all teaching & non-

teaching staff of Electronics & Telecommunication Department for their kind co-operation and guidance for preparing and presenting this project work.

This acknowledgement would not be complete without the mention of all my colleagues to whom I am grateful for all their help, support, interest and valuable hints. Finally, I would like to thank my parents for giving their most patient support and encouragement to my work.

#### VII. REFERENCES

- [1] Christopher Lee and Yangsheng Xu, "Online, interactive learning of gestures for human robot interfaces" Carnegie Mellon University, the Robotics Institute, Pittsburgh, Pennsylvania, USA, 1996
- [2] Richard Watson, "Gesture recognition techniques", Technical report, Trinity College, Department of Computer Science, Dublin, July, Technical Report No. TCD-CS-93-11, 1993
- [3] Ray Lockton, "Hand Gesture Recognition using computer Vision", 4th year project report, Ballilol College, Department of Engineering Science, Oxford University, 2000
- [4] Son Lam Phung, Member, IEEE, Abdesselam Bouzerdoum, Sr. Member, IEEE, and Douglas Chai, Sr. Member, IEEE, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison", IEEE transactions on pattern analysis and machine intelligence, vol. 27, no. 1, january 2005
- [5] Ma De-yi, Chen Yi-min, Li Qi-ming, Huang Chen, Xu Sheng, "Region Growing by Exemplar-Based Hand Segmentation under Complex Backgrounds"
- [6] Andrew D. Wilson, Student Member, IEEE Computer Society, and Aaron F. Bobick, Member, IEEE Computer Society, "Parametric Hidden Markov Models for Gesture Recognition", IEEE transactions on pattern analysis and machine intelligence, vol. 21, no. 9, september 1999
- [7] Chance M. Glenn, Divya Mandloi, Kanthi Sarella, and Muhammed Lonon, "An Image Processing Technique for the Translation of ASL Finger-Spelling to Digital Audio or Text"
- [8] Ahmad Akl, Student Member, IEEE, Chen Feng, Student member, IEEE, and Shahrokh Valaee, Senior Member, IEEE, "A Novel Accelerometer-Based Gesture Recognition System", IEEE transactions on signal processing, Vol.59, No.12, December 2011.
- [9] Meenakshi Panwar, Hand Gesture Recognition System based on Shape parameters, International Conference on Recent Advances in Computing and Software Systems, pp. 80-85, February 2012.
- [10] Meenakshi Panwar, P.S. Mehra, Hand Gesture Recognition for Human Computer Interaction, IEEE International Conference on Image Information Processing, pp. 1-7, November 2011.
- [11] Rohit Verma, Ankit Dev, Vision based Hand Gesture Recognition Using Finite State Machines and Fuzzy Logic, IEEE International Conference on Ultra Modern Technologies, pp 1-6, October 2009.
- [12] V. Frati and D. Prattichizzo, Using Kinect for hand tracking and rendering in wearable haptics," in Proceedings of the IEEE World Haptics Conference (WHC). IEEE, 2011, pp. 317-321.
- [13] Y. Li, "Hand gesture recognition using Kinect," in Proceedings of the 3rd IEEE International Conference on Software Engineering and Service Science (ICSESS). IEEE, 2012, pp. 196-199.
- [14] Z. Mo and U. Neumann, "Real-time hand pose recognition using lowresolutiondepth images," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2. IEEE, 2006, pp. 1499-1505.
- [15] N. Pugeault and R. Bowden, "Spelling it out: Real-time ASL fingerspelling recognition." in Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops). IEEE, 2011, pp. 1114-1119.