ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

Simulation of Real Time Hand Gesture recognition for Physically Impaired

Monuri Hemantha[1], M.V.Srikanth[2]

M.Tech, E.C.E-Dept., Gudlavalleru Engineering College, Krishna-Dt, India [1]

Assistant professor, E.C.E-Dept., Gudlavalleru Engineering College, Krishna-Dt, India [2]

Abstract: We propose a vision based Hand gesture Recognition system that recognizes hand gesture in midair, especially for physically challenged people, and provides recognized character or number as text and corresponding sound. This system also detects face of the gesture requesting person, so as to satisfy the needs of the person. This system provides a convenient and easy to use human computer interaction (HCI) system for especially visually or physically challenged people. The system does not demand on sensors or markers attached to the users and allow unrestricted character and symbol input from any position. It is the result of combining hand gesture recognition based on Centroids with real time 3D hand tracking and face recognition using PCA. We evaluated the system for both quantitative and qualitative aspects. The system achieves recognition rates of 86:15% for character and 97:54% for symbol recognition. Limitations are due to slow and lower resolution cameras or physical strain of training dataset. Overall, the proposed hand gesture recognition system provides an easy to use and accurate text input modality without placing restrictions on the users.

Category: User Interfaces: Image Processing HCI,
Keywords: PCA, Hand gesture, Real time motion tracking, face detection.

I. INTRODUCTION

With current advances in technology, we see a rapidly increasing availability, but also demand, for intuitive HCI. Devices are not only controlled by mouse and keyboard anymore, but we are now using gesture controlled devices in public areas and at our homes. Distant hand gestures in particular removed the restriction to operate a device directly; we can now interact freely with machines while moving around. In this work, we are interested in HCI that does not force users to touch a specific device or to wear special sensors, but that allows for unrestricted use.
While there is a great variety of HCI techniques to interact with distant virtual objects, e.g. to select menu entries, there is still a lack for intuitive and unrestricted text input. Although there are ways to input text by using a virtual keyboard on a display or by speech recognition, there are situations where both are not suitable: the rest requires interaction with a display. Users must speak which is not always possible. demanding on the surroundings, especially for physically challenged people.

With this work, we extend the available text input modalities by introducing an intuitive hand gesture recognition system. In the remainder of this work, we present a system that combines vision based 3D hand gesture, motion tracking and face recognition with extended audiovisual facilities. It allows users to give characters/symbols or any assigned operations to the computer or any machine /robot just by waving their hands in air towards it. This extends the use of 3D hand gestures to allow for convenient and unrestricted text input. In Section 2, we discuss related work before introducing the proposed system in Section 3. In Section 4, we present the experimental evaluation before concluding in Section 5.

Fig.1.1.Hand gestures and their equivalent text- character outputs

The platform used for the programming is MATLAB. Initially, we used R2007a version but it raised problem in executing the command "imtool"[1] so we moved to R2007b but that raised problem in executing the commands of video capturing and taking snapshots from running video. Then finally we moved to R2010a version which supports both of these commands and it runs in 64-bit computers.

II. RELATED WORK

Hand gesture recognition is not limited to paper or digital surfaces, but has also been extended to the third dimension. Hand gesture recognition is a problem that has elicited significant attention and research as computational capabilities, camera performance, and computer-vision-style learning algorithms have rapidly improved over the past few years. Such research is driven by the tremendous growth and variety in development of applications that require some form of gesture comprehension: these

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

include remote hardware control, game controls, affective computing, and other endeavors in enhanced HCI [1, 2]. There are, however, fundamental limitations to most current systems for gesture detection based off training on a set of predefined gestures. Non-uniform lighting conditions and less- than-ideal camera resolution and depth of color limit the number and accuracy of possible gesture classifications in practice [3]. Moreover, the modeling and analysis of hand gestures is complicated by the variegated treatment required for adequate detection of static gestures, which represent a combination of different finger states, orientations, and angles of finger joint that are often hidden by self-occlusion [4, 5].

In computer vision (CV) based solutions such as ours, hand gestures are captured by web cameras which offer resolutions that allow only a general sense of the figure state to be detected [4]. On the other hand, modeling gestures as temporal objects (emphasizing an understanding the movement of the hand as a pointer to the nature of the gesture) allows for greater accuracy and better differentiation of gestures [5]. Additionally, the problem of hand-gesture recognition usually occurs in contexts where gestures involving finger conformation are accompanied by movement of the hands relative to the body and background.

In our work, we will combine a vision based hand gesture tracking system with face recognition to provide a text input modality that does not require additional devices. To the best of our knowledge, we are unaware of vision based hand gesture recognition of whole symbols based on concatenated individual character models.

III. VISION BASED HAND GESTURE RECOGNITION

The proposed hand gesture recognition system consists of three modules: vision based data acquisition, feature extraction, and hand gesture recognition that will now be introduced in more detail.

3.1. Proposed system

This proposed system presents a low-cost hand gesture HCI system for the users of visually and Hearing Impaired. The proposed system is practicable to integrate into Remote PC, ROBOT, TV, Virtual Games and other consumer electronics, benefiting from its three advantages:

a) The proposed system adds only a little to the hardware cost as just a USB webcam and audio player are used.

b) The proposed system can run secured and authenticable on mainstream and even highly populous area due to its face recognition feature.

c) The control commands issued by proposed system are converted to standard text display and Audio files simultaneously. Only drawbacks of this project are that it works only when it is trained in different environments.

3.2. Image acquisition

The input images are captured by a webcam placed on a table or laptop. The system is demonstrated on a conventional PC/ Laptop computer running on Intel Pentium Dual Processor with 4GB of RAM. Each image has a spatial resolution of 620 x 480 pixels and a grayscale resolution of 32 bit. The system developed can process hand gestures at an acceptable speed. Given a variety of available image processing techniques and recognition algorithms, we have designed our preliminary process on detecting the image as part of our image processing. Hand detection preprocessing workflow is showed in Fig. 3.2.

The system starts by capturing a hand image from signer with a webcam setup towards certain angle with black background. The next process will convert the RGB image into grey scale with either black (0) or white (1). The edge of each object is then computed against the black background. The object can then be segmented and differs greatly in contrast to the background images.

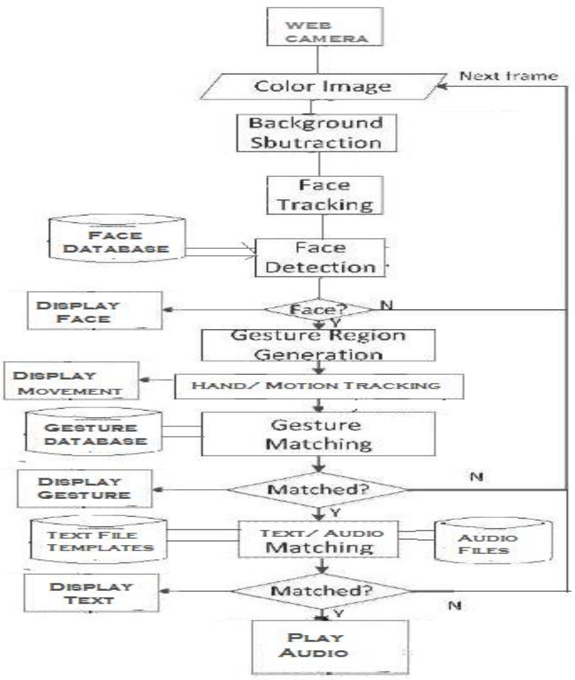


Fig.3.1. Block Diagram of proposed System

3.3. Preprocessing

Changes in contrast can be detected by operators that calculate the gradient of an image. One way to calculate the gradient of an image is the Sobel operator [7], [8], [9], which creates a binary mask using a user-specified threshold value.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

The binary gradient mask shows lines of high contrast in the image. These lines do not quite delineate the outline of the object of interest. Compared to the original image, the gaps in the lines surrounding the object in the gradient mask can be seen.

The proposed system recognizes characters and symbols that are given in midair with the hand acting as the hand main object, from a Video camera or webcam. Therefore, the gesture and motion of the hand is of interest. To compute this motion, we used a Dynamic time wrapping approach to reconstruct gesture [10]. Based on the user's pose, we extract the palm motion and gesture of the hand. (a) (b)
There are also other ways to acquire this data, e.g. with sensors based on structured light like the Microsoft Kinect. But a vision based system has the advantage that there is much broader variety of sensors to chose from with varying resolutions and frame rates which can be important for the overall performance (Section 4.3). To deal with sensor noise and occasional misclassifications, we applied a median filter to smooth the gesture. The system set gesture is described in Section 4.

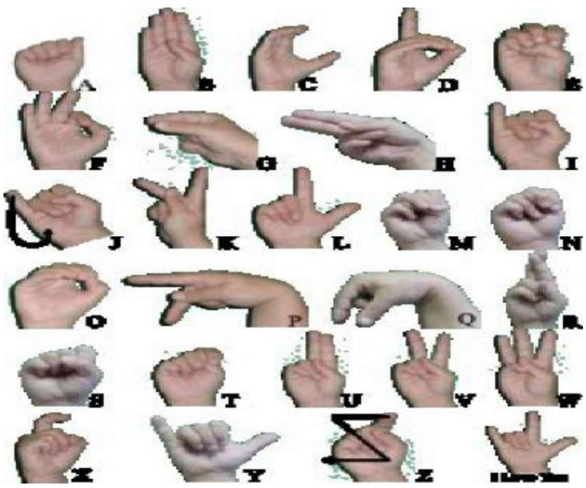


Fig.3.2.Alphabets for Hand Gestures

3.4. Image Segmentation
Image Pre-processing is necessary for getting good results. In this algorithm, we take the RGB image as input image. Image segmentation is typically performed to locate the hand object and boundaries in image. It assigns label to every pixel in image such that the pixels which share certain visual characteristics will have the same label.

3.5. Feature extraction
When gesture is given in air, users should tend to give on a video camera in front of them. We exploit this video by projecting in to different frames of the hand gesture on a 2D plane, in front of the users, which has the advantage that it reduces the feature space. The beginning and ending of hand gesture can be easily detected because system provides different time slot and notify the same to users, user move their hands, correspondingly. We found that there were hand and hand movements while giving as is the case in traditional hand gesture. Character and symbol recognition from hand gesture are both based on centroid, which is an established technique in hand gesture recognition [8].

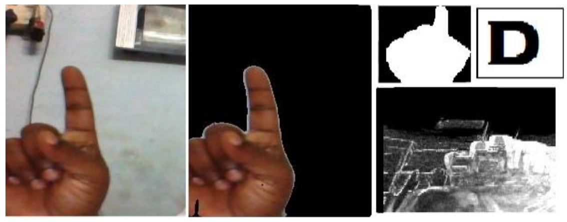
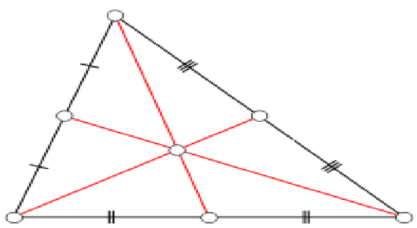


Fig.3.3. Hand gesture segmentation and detected BW image and text display.

3.4.1. Centroid:

In mathematics and physics, the centroid or geometric center of a two- dimensional region is, informally, the point at which a cardboard cut-out of the region could be perfectly balanced on the tip of a pencil (assuming uniform density and a uniform gravitational field). Formally, the centroid of a plane figure or two-dimensional shape is the arithmetic mean ("average") position of all the points in the shape. The definition extends to any object in*n*-dimensional space: its centroid is the mean position of all the points in all of the coordinate directions. we calculate the centroid for differentiating various hand pattern based on number of fingers. By assuming number of fingers and finger position as distinct from location of the centroid in the hand image we can differentiate many hand gestures.



The centroid of a non-self-intersecting closed polygon defined by n vertices (x_0,y_0), (x_1,y_1), ..., (x_{n-1},y_{n-1}) is the point (C_x, C_y), where

$$C_x = \frac{1}{6A}\sum_{i=0}^{n-1}(x_i + x_{i+1})(x_i\ y_{i+1} - x_{i+1}\ y_i)$$

$$C_y = \frac{1}{6A}\sum_{i=0}^{n-1}(y_i + y_{i+1})(x_i\ y_{i+1} - x_{i+1}\ y_i)$$

and where A is the polygon's signed area,

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

$$A = \frac{1}{2}\sum_{i=0}^{n-1}(x_i\ y_{i+1} - x_{i+1}\ y_i)$$

In these formulas, the vertices are assumed to be numbered in order of their occurrence along the polygon's perimeter, and the vertex (x_n, y_n) is assumed to be the same as (x_0, y_0). Note that if the points are numbered in clockwise order the area A, computed as above, will have a negative sign; but the centroid coordinates will be correct even in this case.

we calculate the centroid for differentiating various hand pattern based on number of fingers. By assuming number of fingers and finger position as distinct from location of the centroid in the hand image we can differentiate many hand gestures. The partition of the hand is very important. Hand gesture patterns which contain the index finger will fall in the left hand side of the centroid. And other hand patterns which contain little finger will find in the right side of the centroid location in image. Centroid always computed at the geometric center of the image and it is also called as center of mass if the image is uniformly distributed. Centroid is calculated using the image moment, which is the weighted average of pixel's intensities of the image.

3.5. Hand gesture recognition

Specially, the system is based on left to right Centroids. First, for each character, a separate c is trained with training data of multiple people with the extracted features described above as observations. Then, symbols can be modeled as a concatenation of character Centroids, i.e. for any given arbitrary symbol, and Centroids can be constructed. A database denotes the set of symbols that can be actually recognized. We examine two recognition tasks, the recognition of face and the recognition of character from hand gesture. The general block diagram for matching of Hand gesture is given below

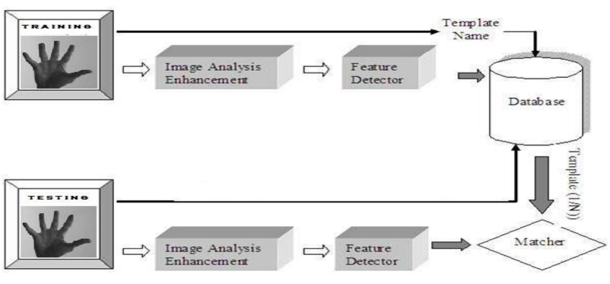


Fig 3.2.: Gesture Hand gesture matching Block diagram

IV. EXPERIMENTAL RESULTS

We have applied the above discussed algorithm and tested 360 images with 36 different patterns. By using effective shape based features and orientation, we can recognize and classify 36 different hand gesture patterns. On the basis of text and motion, we can communicate with computers meaningfully or assign different tasks to support human computer interaction or direct conversion of sign language to voice may be even better.

These hand gestures would always be unique based on their orientation, presence of thumb and pattern of hand gesture. However, the algorithm does not rotates the image even the orientation of hand gesture is horizontal, rather it requires training, or in other words the programme simply depends on training data set.

4.1. Hand gesture training system

The backbone of this algorithm is training system. This includes training of face, Hand gesture and even motion of any number of persons unlimited. From this training the system takes only main gesture and face and corresponding motion, then it automatically saves in the corresponding database location with a single mouse click.

Hand orientation, shape and centroids are only saved and remaining things are discarded, as a process of segmentation and feature extraction.

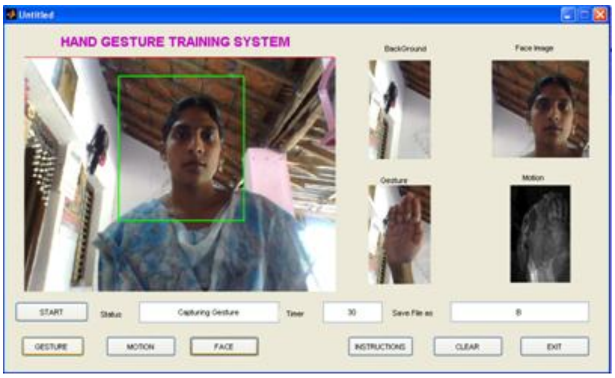


Fig.4.1. Output of hand gesture recognition system.

All further processing steps for generating of outputs and recognition of hand gesture will be depend on same orientation and shape of hand gesture. Figure 4.1.shows the training system output. Table II shows the result of 360 images. Out of 360 images tested through the algorithm, it has correctly recognized 339 images and falsely identified the remaining 21 cases. At an average it gives the success rate of 94% approximately with average computation time of 0.60 second. The algorithm is based on simple shape based feature calculation which provides us with the comfort of implementation. The algorithm discussed above is implemented on MATLAB

4.2. Hand gesture recognition system

This is the system where the input can be provided for the user needs of interaction with computer. Here if the face is not recognized then the system does not allow user to interact, then the system advises for the training.

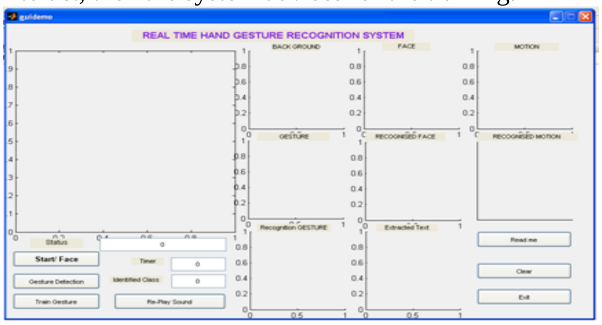


Fig.4.2. Real time hand gesture recognition system GUI.

This system interfaces directly with live video and detects face and gesture spontaneously.

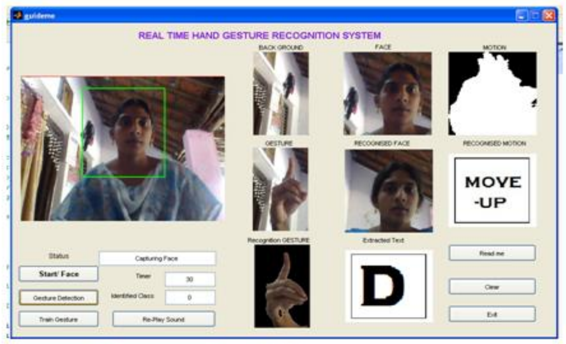


Fig.4.3.Complete recognition, including face, hand gesture, motion and their corresponding text characters and sounds.

4.3. Inputs and Outputs

The input to the system was only a **video** containing face, gesture and motion. The program splits the video into separate file frames and saved automatically.

A).Text Outputs: The distinct feature of this system is presenting text character display in conveniently big size to the users. The characters are pre designed in bit map formats to embed with code. The system is now ready to recognize 26 alphabets and 10 numbers, and convert them to text display as output.

B). Motion tracking: Another feature of this system is to identify the meaning of motion and conveniently convert them into Text format display as shown in Fig.4.3. The motion files are also pre designed in bit map formats to embed with code.

C).Audio output: Most important feature of this system is to convert a gesture in audio – audible format. The audio files are pre loaded in wav formats and as gesture is recognized audio file is copied to buffer and same audio file is played at end using Matlab Wavplay function, and GUI Provide another feature to repeat the audio file, using replay push button. This is useful for the hearing impaired person at minimum communicable standards.

4.4. Drawbacks

Each gesture system has some shortcomings since there is no perfect simulation of the nature, since all human researches fall into the attempt to simulate the nature, and, certainly, the simulation should has some sort of errors.

TABLE I
Detection rate and false alarm rate of hand motion gestures

Gesture	False Alarm Rate	Detection Rate
Right	9%	88%
Left	9%	91%
Down	8%	87%
Up	6%	85%
Zoom-out	0.6%	94%
Zoom-in	0.4%	96%

The hand must be in a static direction in [6], horizontal or vertical, and the hand object should be closest object to the camera since the segmentation process depends on this fact in [7], the input image should be the only visible part of the human skin in [4] with a care selection for adaboost filters, using only the right hand for processing in [6], a lot of storage space and processing time are required for matching stage in [10] since a total of 2280 templates have been employed, 56 sub-image, 4 features per sub-image, then, the database features are 510720 need to be matched at recognition time. The time delay is a factor of drawback in case of recognition methodologies, where in practical systems this could be effectively solved by the use of DSP processors for recognition over large number of data sets of stored databases.

V. RESULTS AND CONCLUSION

We proposed a simple yet powerful shape based approach for hand gesture recognition. Visually impaired people can make use of hand gestures for writing text on electronic documents like MS Office, notepad etc. Moreover, almost all deaf and dumb people communicate with each other by forming several hand shapes. Similarly, a visually impaired person would be able to work on computer through computer vision.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

The strength of this approach lies in the ease of implementation, as it does not require any significant amount of training or post processing and it provides us with the higher recognition rate with minimum computation time. The weakness of this method is, we define certain parameters and threshold values experimentally.

Table – II
Comparison table for gesture and their corresponding output text recognized with time elapsed.

Gesture	Input Images	Successful Cases	Recognition Rate	Elapsed Time
A	10	10	100	0.72s
B	10	10	100	0.56s
C	10	10	100	0.74s
D	10	09	90	0.90s
E	10	10	100	0.57s
F	10	10	100	0.41s
G	10	10	100	0.45s
H	10	10	100	0.43s
I	10	10	100	0.65s
J	10	09	90	0.76s
K	10	10	100	0.56s
L	10	10	100	0.41s
M	10	10	100	0.46s
N	10	09	90	0.50s
O	10	09	90	0.87s
P	10	10	100	0.72s
Q	10	09	90	0.45s
R	10	10	100	0.74s
S	10	10	100	0.67s
T	10	09	90	0.53s
U	10	09	90	0.64s
V	10	10	100	0.82s
W	10	10	100	0.54s
X	10	10	100	0.70s
Y	10	10	100	0.43s
Z	10	10	100	0.63s
0	10	09	90	0.66s
1	10	10	100	0.79s
2	10	09	90	0.51s
3	10	10	100	0.46s
4	10	10	100	0.52s
5	10	10	100	0.64s
6	10	09	90	0.52s
7	10	10	100	0.54s
8	10	09	90	0.61s
9	10	09	90	0.61s
.	10	09	90	0.61s
+	10	09	90	0.61s
All	380	366	93.87%	0.792s

Since it does not follow any systematic approach for gesture recognition, and maximum parameters taken in this approach are based on the assumption made after testing a number of images. If we compare our approach with our previous approach described in paper [2], the success rate has improved from 92.3% to 94%, the computation time decreased up to fraction of seconds. And we have removed some of the constraints needed to be followed in our previous approach which makes it simpler. In future the focus would be on improving the system by including some different backgrounds while enlarging the data set.

VI. REFERENCES

[1] Prince Nagar, Ashwani Sengar and Mayank Sharma: **"**Hand shape based gesture recognition in hardware", Archives of Applied Science Research, 2013, 5 (3):261-269: (http://scholarsresearchlibrary .com/archive.html)

[2] Mokhtar M. Hasan, and Pramod K. Mishra: **"**Hand gesture Modeling and recognition using Geometric Features: A Review": Canadian journal on image processing and computer vision vol.3 No.1, March 2012.

[3] Meenakshi Panwar, "Hand Gesture based Interface for Aiding Visually Impaired", Centre for Development of Advanced Computing Noida, Uttar Pradesh, India, 2012 International Conference on Recent Advances in Computing and Software Systems.

[4] Rayi Yanu Tara, Paulus Insap Santosa, Teguh Bharata Adj: "Hand Segmentation from Depth Image using Anthropometric Approach in Natural Interface Development": International Journal of Scientific & Engineering Research, Volume 3, Issue 5, May-2012

[5] Hervé Lahamy * and Derek D. Lichti, "Towards Real-Time and Rotation - Invariant America Sign Language Alphabet Recognition Using a Range Camera", 29 October 2012.

[6] M.S. Dias1,2, R. Bastos1,, J. Fernandes1, J. Tavares1, P. Santos2 , "Using Hand Gesture and Speech in a Multimodal Augmented Reality Environment".2011

[7] Title: "Real-time DTW-based gesture recognition external object for max/msp and pure data": proceedings of the SMC 2009 6th sound and music computing conference, July 2009, porta-portugal.

[8] Oi Mean Foong, Tan Jung Low, and Satrio Wibowo, "Hand Gesture Recognition: Sign to Voice System (S2V)", World Academy of Science, Engineering and Technology 42- 2008.

[9] Kiran Kunnur,Pallavi Guijal : "Real time hand gesture recognition using sift", international Journal of Electronics and Electrical Engineering, ISSN : 2277-7040 Volume 2 Issue 3 (March 2012).

[10] Shaun K. Kane, Jacob O. Wobbrock "Usable Gestures for Blind People: Understanding Preference and Performance", The Information School, DUB Group University of Washington Seattle, WA 98195 USA {skane, wobbrock}@uw.edu, May 7-12, 2011.

BIOGRAPHY

Monuri Hemantha is a post graduate student of M.Tech at Gudlavalleru College of Engineering, Krishna –Dist, under the guidance of M.V. Srikanth, Associate professor in E.C.E.Dept., and are strongly willing to work for the support of physically handicapped persons. And this paper is also a contribution in one of such kind. Author has started her research in B.Tech level and would like to continue up to Ph.D. level.