

# Augmenting Digital and Physical World through Gestures and Image Processing To Develop New Software to Express the Feeling of Those People Who Can't Speak Naturally

Goutam Prasad Sahay, Suchismita Samanta

**Abstract**— we come across lot of computing devices in our day-to-day life with the incredible development in computing and mobile platform. Since the computer technology continues to grow up, the importance of human computer interaction is enormously increasing and the Sixth Sense Technology is making its way into our lives which is a revolutionary way to connect the physical world with the digital world. This technology is implemented in the project where color detection and object tracking method is used to express the feelings of those people who can't speak.

**Index Terms**— sixth sense technology; computer vision; object detection and corresponding voice play

## I. INTRODUCTION

In this Project the focus is on bridging the crevice between the two different worlds i.e. Physical and Digital world using sixth sense technology. Sixth Sense is a set of wearable devices that acts as a gestural interface and aggrandize the physical world around us with digital information and lets the users to use natural hand gestures to interact with the digital information through it. The goal is to make a user friendly system for those people who can't speak, so that through the system they can express their feelings easily. The newly born technology named Sixth Sense Technology is used for this purpose. The color detection and object tracking methods are used in this system where, the color of the fingertip is detected using a real-time camera and the number of fingers are counted and using that count numbers, various voice notes can be played through which the people with speaking disability can express their feelings.

The project has been developed using MATLAB. Recognition and pose estimation in this system are user independent and robust as color tapes or custom made LED gloves are used on finger to perform actions.

## II. SYSTEM OVERVIEW

The implementation consists of three main components, that collectively acts as a system in itself and each device has its important role in the system. The devices include a

webcam or digital camera, colored caps or LED gloves, a laptop. The camera is used to capture the object in sight range and detect the finger of the user using color tracking as colored caps or LED gloves are attached at the fingertips of the user. Then the data is sent to the laptop connected with it, from this data specific features are retrieved and the number of fingers are counted and using that information different voice notes are played which are already there in the system through which those people having speaking disability can express what they want to say or give instructions easily. Camera acts as a digital eye connecting the user to the digital world. In this study, color pointers have been used for the object recognition and tracking.

## III. SYSTEM DESCRIPTION

The work has been completed through different steps. The first step of the project is to capture the image with a webcam or acquire the image, i.e. Image Acquisition.

### A. Image Acquisition

Image acquisition is the digitization and storage of an image. In order to process the image, the images should be acquired with the help of image acquisition hardware, a camera which act as a digital eye to the image processing system. In this model, the incumbent resource is used in a laptop. The first stage of any vision system is the image acquisition stage. After the image has been obtained, various methods of processing can be applied to the image to perform the many different vision tasks required. Here MATLAB software is used as a software package for processing the data from the digital camera.

1) *Creating Video Object*: For a successful integration, a video object is created which helps to proceed with further processing of the data from the digital eye. Creation of video object is done in MATLAB with the help of integration of windows "winvideo" function with MATLAB. The following command helps to create a video object using

Video=videoinput (winvideo, 1) (1)

Here in the above command, Video is a variable name of the video object and the index number 1 represents the identity of camera. It can supports up to n numbers of input.

2) *Setting Frames*: The frames per second of the camera and other properties of the video object are also set. Frame grab interval is also set for better transition.

3) *Color Space Conversion*: Color space conversion is much needed for processing digital images. For applications regarding color recognition it's much better to use RGB color

**Manuscript received September 20, 2014.**

Goutam Prasad Sahay, Computer Science and Engineering, Dr. Sudhir Chandra Sur Degree Engineering College, Kolkata, India

Suchismita Samanta, Computer Science and Engineering, Dr. Sudhir Chandra Sur Degree Engineering College, Kolkata, India

space format. RGB is a basic type of color image which comprises of three components namely R, G, and B (grayscale type). Once the image has been converted into RGB image, then store it in a variable for further processing and feature extraction concepts to be applied.

4) *Preview The Video Output*: To view the video output in the user preferred size (window size) following command is used.

$$\text{Preview (Video)} \quad (2)$$

5) *Snap the Video Data*: After a successful integration and preview of the video, the images from the real time video data are snapped. To get a snap of the present video data, "get snapshot" command is used along with the respective video object. A loop is set to get the acquisition frames continuously.

### B. Color Recognition

Recognizing a color is a complex task. While recognizing a certain color in a scene, its background information is always trying to affect our region of interest. Another one issue is that, it always depends on the background illumination and luminance, chrominance of the scene. Upon overcoming these kinds of issue a perfect color recognized module is used with background subtraction method.

To detect the color of the pointer in the fingertip, MATLAB's built in "imsubtract" function has been used. "imsubtract" function can be used as,

$$Z = \text{imsubtract}(X, Y) \quad (3)$$

Where, it subtracts each element in array Y from the corresponding element in array X and returns the difference in the corresponding element of the output array Z. X and Y are real, non-sparse numeric arrays of the same size and class, or Y is a double scalar. The array returned, Z, has the same size and class as X unless X is logical, in which case

Z is double. So with the help of "imsubtract" function the background is subtracted from the image and the specific colored region remains as the color code is specified in the image from which the background is subtracted. The image is converted to gray scale here. In this way color in the fingertips are detected. The example is shown in Figure1.

Once extracted the object roughly, to give more intelligence to the system to redefine the size of the object which is extracted, morphological techniques like median filtration, binary conversion, removal of small unexpected pixels and like are used. These will redefine and enhance the size and shape of the object.

### C. Filtering the Noise

After detecting the blue color in the input image, a median filter has been used to filter out the noise. Median filtering is a nonlinear operation often used in image processing to reduce "salt and pepper" noise. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges. The example is shown in Figure2.

### D. Converting Grey Scale Image into Binary Scale Image

To convert the gray scale image to binary scale image MATLAB's built in "im2bw" function has been used. Function can be used as;

$$BW = \text{im2bw}(I, \text{level}) \quad (4)$$

Where; it converts the grayscale image I to a binary image. The output image BW replaces all pixels in the input image with luminance greater than level with the value 1 (white) and replaces all other pixels with the value 0 (black). Specify level in the range [0, 1]. This range is relative to the signal levels possible for the image's class. Therefore, a level value of 0.5 is midway between black and white, regardless of class. In this study, the threshold 0.15-0.18 gave the best result for the large range of illumination change. The example is shown in Figure3.

### E. Removing All the Small Areas

To get the best accurate number of the object detected in the image, all the areas other than the pointer need to be removed. To do this, MATLAB's "bwareaopen" function is used.

$$BW2 = \text{bwareaopen}(BW, P) \quad (5)$$

Where; it removes from a binary image all connected components (objects) that have fewer than P pixels, producing another binary image, BW2. A threshold of 500 pixels for this study is set. The example is shown in Figure4.

### F. Labeling

After removing all the connected components (objects) other than the pointer, using MATLAB's "bwlabel" function the pointer can be labeled. In other words the region can be detected.

### G. Feature Extraction

After recognizing the color of the object i.e.; the color of the tapes or the LEDs in the fingertips, some information about the recognized one needs to be extracted. To reduce the dimensionality of the object to give shape to the virtual object which originates in the physical world, feature extraction principles are used.

To get the features of the detected region such as center point or bounding box etc., MATLAB's built in "regionprops" function can be used as;

$$\text{STATS} = \text{regionprops}(BW, \text{properties}) \quad (6)$$

Where; it measures a set of properties for each connected component (object) in the binary image, BW. The image BW is a logical array; it can have any dimension. In this application the center of the object is used only.

### H. Counting the Number of the Object

After the feature extraction the numbers of the objects are counted with the help of labeling. Labeling is necessary to count, same number of color objects extracted. In labeling, a connected region of white pixels by pixel operation can be found. Once the connected pixels and its boundary are identified, it is being stored by a variable at its front which shows the number of objects available in the input image. Here, the object represents the recognized color of physical objects.

### I. Reading Voice Notes

After counting the number of the detected fingertips these numbers are used to play the voice notes with the help of which people can express their feelings. There are some voice notes which are already in the system. Now when one

object i.e. the fingertip is detected and the number of the count is one then a particular voice note will be played .When two objects will be detected and the count is two then another voice note will be played and so on.



Figure 1. (a) Input image, (b) after using "imsubtract" and detect the blue color



Figure 2. After filtering with the median filter the little noise which was in Figure 1 (b), is gone



Figure 3. After converting the gray scale image to binary scale image.



Figure 4. After removing the small pixels from the image.

## II. RESULTS

The output of this work is shown below. Here the red colored object is the one which is recognized and used. The color of the object i.e.; the fingertips are detected and counted. When the number of object detected is one then a particular voice note is played and when the number of detected objects will be two then another voice note will be played and so on. These images are snapped version of the recorded output.

- A LED glove shown in Figure5 is used which the user can wear in his hand. The Glove has colored LEDs attached in the fingertips which will help to detect the fingertips of the user.
- When the number of the detected object is one then a particular voice note is played. For example here a voice note saying "Hello" will be played when the count is one. So if the user wants to say "Hello" he just have to show one of the fingers to the camera and the camera will detect the fingertip with the help of the color and then the voice note will be played. It is shown in Figure6.
- In this way when the number of detected objects is counted two then another voice note saying "Good Bye" will be played. In this way those who can't speak can express their feelings through voices. It is shown in Figure7

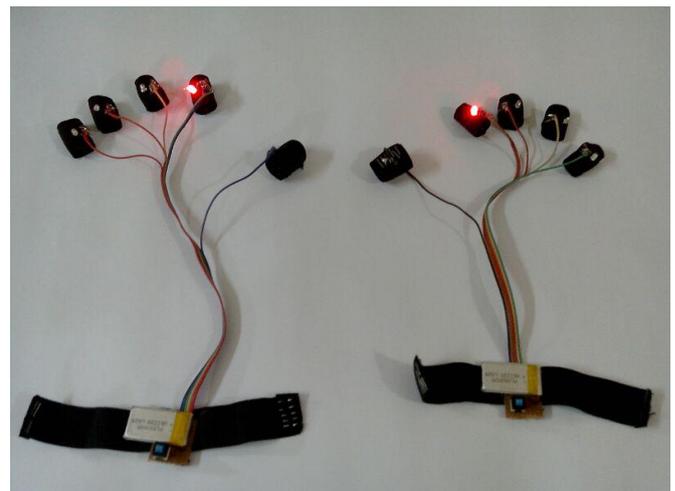


Figure 5. The LED glove with colored LEDs attached to the fingertips.

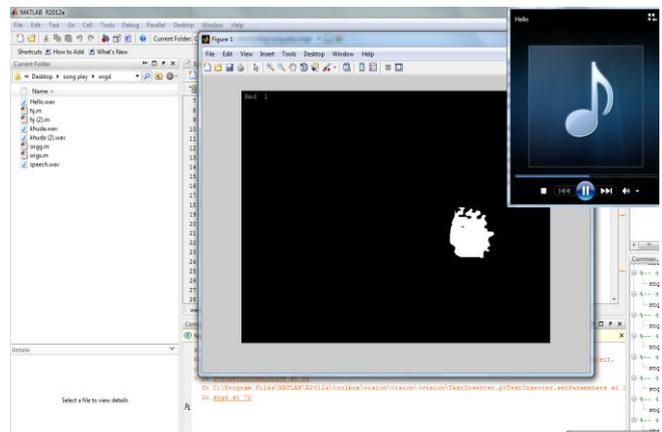


Figure 6. When the number of the object counted is one voice note saying "Hello" is played.

# Augmenting Digital and Physical World through Gestures and Image Processing To Develop New Software to Express the Feeling of Those People Who Can't Speak Naturally

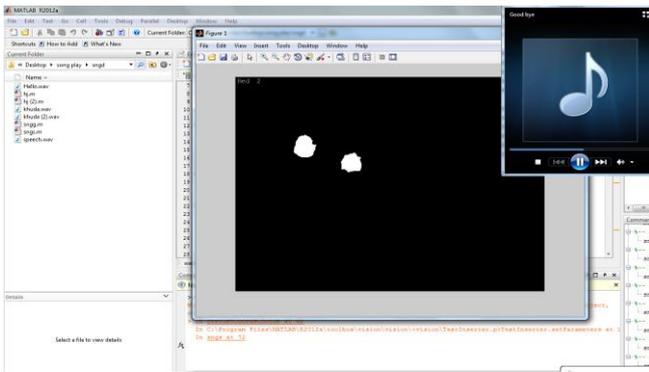


Figure 7. When the number of the detected object is two a voice note saying "Good Bye" is played.

## IV. CONCLUSION

The use of image processing and color recognition in MATLAB for the implementation of the proposed approach proved to be practically successful. The approach has huge potential once it gets further optimized, as its time complexity is higher, with the help of hardware having better specifications. This approach has much high potential for future advanced applications which can have the ability to change the mobile world also. Also this would lead to a new era of Human Computer Interaction (HCI) where with the help of computer technology the people having speaking disabilities can find their voices.

## ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression, "One of us (R.B.G.) thanks . . ." Instead, try "R.B.G. thanks". Put applicable sponsor acknowledgments here; DO NOT place them on the first page of your paper or as a footnote.

## REFERENCES

- [1] E. Ul Haq, S.J.H.pirzada, M.W Biag and Hyunchal Shin "A new hand gesture recognition method for mouse operation", The 5th International Midwest symposium in circuits and systems IEEE, (MWSCAS), Pp. 1 – 4, 2011.
- [2] A.Pandit, D.Dand, S.Mehta, S.Sabesan and A.Daferly,"A simple wearable hand gesture recognition device using iMEMS", International Conference on soft computing and recognition (SOCPAR'09), Pp., 2009.
- [3] K Manchanda and B Bing, "Advanced mouse using trajectory-based gesture recognition", proceeding of the IEEE southeastcon (southeastcon), Pp. 412-415, 2010.K. Elissa, "Title of paper if known," unpublished.
- [4] S Kumar, A Agarwal, A Rai and N Bachani,"Vision based human interaction system for disable",2nd international conference on Image processing theory tools and application (IPTA), Pp. 441-446, 2010.
- [5] Young wook kim, "Development of headset-type computer mouse using gyro sensors for the handicapped", Electronics letters, volume 38, issue 22, Pp. 1313 – 1314, 2002.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [6] A W Mohamed and R Koggalage,"Control of mouse movements using human facial expression",3rd international conference on Information and Automation for sustainability(ICIAFS),Pp. 13-18, 2007.
- [7] D. Kornack and P. Rakic, "Cell Proliferation without Neurogenesis in Adult Primate Neocortex," Science, vol. 294, Dec. 2001, pp. 2127-2130, doi:10.1126/science.1065467.

- [8] T , Palleja, E Rubion, M Teixido, M Tresanchez, A F Delviso, C Rebate and J Palacin, "Simple and robust implementation of a relative virtual mouse controlled by head movements", Conference on Human System interactions, Pp. 221 – 224, 2008
- [9] D Ionescu, "Gesture control and the new and intelligent Man-Machine interface", 6th international symposium on applied computational intelligence and informatics (SACI), Pp. 291, 2011.
- [10] Jilin Tu ,T Huan and Hai Tao, "face as mouse through visual face tracking ", 2nd Canadian conference on computer and robot vision, Pp. 339 – 346, 2005.
- [11] M Betke , J Gips and P Fleming, "the camera mouse : visual tracking of body features to provide computer access for people with severe disabilities", IEEE transactions on Neural systems and Rehabilitation Engineering , Volume 10, Issue 1, Pp. 1 – 10 ,
- [12] Tsang W M and Kong – Pang Pun,"A finger – tracking virtual mouse realized in an embedded system", International symposium on Intelligent signal processing and communication system (ISPACS), Pp. 781 – 784 , 2005.
- [13] M Atyabi, M S K Hosseini and M Mokhtari, "A webcam mouse: visual 3D tracking of body features to provide computer access for people with severe disabilities", Annual IEEE conference India, Pp. 1 – 6 ,2006.
- [14] Hojoon Park, "A Method For Controlling The Mouse Movement using a Real Time Camera", 2008, Brown University, Providence, RI ,USA, Department of computer science
- [15] A. Erdem, E. Yardimci, Y. Atalay, V. Cetin, A. E. "Computer vision based mouse", Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASS). IEEE International Conference
- [16] A New Spatial-Color Mean-Shift Object Tracking Algorithm with Scale and Orientation Estimation Chung-Wei Juan and Jwu-Sheng Hu, Member, IEEE.