

Sixth Sense Technology-A New Innovation

Prof. Mr. D.S. Patil¹, Mr. Shahak Patil²

Digital Electronics & Communication System, Electronics & Communication Dept., Visvesvaraya Technical University
Belgaum, Karnataka, India

¹ dsp_ece@mtec.edu.in, ² shahak.p@gmail.com
Maratha Mandal College of Engineering & Technology
Belgaum, Karnataka, India

Abstract – This paper deals with a Sixth Sense device which works of the principles of gesture recognition and image processing to capture, zoom (in and out), toggle pictures and control mouse pointer with ease just by the help of colored caps worn on the fingertips of the user. Sixth Sense can also obey hand gestures. The camera recognizes objects around a person instantly and performs various gesture related applications depending upon the scenario. Also can access or manipulate the information using fingers.

I. INTRODUCTION

Sixth sense, or subtle perception ability, is our ability to perceive the subtle-dimension or the unseen world. It also includes our ability to understand the subtle cause and effect relationship behind many events, which is beyond the understanding of the intellect. Extrasensory perception (ESP), clairvoyance, premonition, intuition are synonymous with sixth sense or subtle perception ability. We perceive the gross or seen world through the five physical senses (i.e. smell, taste, sight, touch and sound), our mind (our feelings), and our intellect (decision making capacity). When it comes to the unseen world or the subtle-world, we perceive it through the five subtle-senses, the subtle-mind and the subtle-intellect – more popularly known as our sixth sense. When the sixth sense is developed or activated, it helps us to experience the subtle-world or subtle-dimension.

We use our five natural senses to perceive any information; that information helps us make decisions and choose the right actions to take. But arguably the most useful information that can help us make the right decision is not naturally perceivable with our five senses, namely the data, information and knowledge that mankind has accumulated about everything and which is increasingly all available online. Sixth Sense Technology bridges this gap, bringing intangible, digital information out into the tangible world, and allowing us to interact with this information via natural hand gestures. 'Sixth Sense' frees information from its confines by seamlessly integrating it with reality, and thus making the entire world your computer.

From last few years the demands of image processing applications have increased considerably. Especially the one concept given by Pranav Mistry about his application 'Sixth Sense'. Yes, in this paper you will get the brief idea about that and also you can build your own sixth sense device. He interfaced a camera and projector with one of the best processors.

Image processing is just another phrase to make any computer or a device not just to see objects but also to understand what it represents. For example to find a colour proportion in an image,

for finding any cracks or flaws on a wall or even finding a particular colour from the image, etc. There are mainly two types divided on the basis of their use. The first is just the simple image from a camera or any other source. It is also called static image. The other is a video or an array of images (the one which we are going to use). We all know this type of format and a video is the array of images called frame. This kind of array is used by dividing them into many images and processing one frame at a time. So here also we are going to do the same firstly we will process one image than doing it continuously for upcoming frames from a camera and develop various applications such as controlling mouse pointer using finger, capturing of images and interfacing image (zoom in & out/rotating). For doing that we need some components as follows: PC camera / webcam, colour markers and laptop or PC.

II. COMPONENTS

The devices which are used in Sixth Sense Technology are:

- A. Web Camera.
- B. Coloured Markers
- C. Laptop or Personal Computer.

A. Camera:

It grabs the motion of the coloured markers worn on the fingertips and tracks the users hand gesture. The camera recognizes individuals, images, pictures, gestures that user makes with his hand. The camera then sends this data to the laptop or PC for processing. The camera performs a basic interface device between laptop and the physical world.



Fig1: Web cam

B. Coloured Marker:

These colour markers placed at the tip of users fingers. Marking the user's fingers with red, yellow, green and blue coloured tape helps the webcam to recognize the

hand gestures. The motion of the colour markers placed at the tip of the hand is recognized and several gestures are fed depending upon the combination of colours present in the laptop as the form of preloaded data.



Fig2: colour markers

C. Laptop or Personal Computer:

The Laptop can be used for running or implementing the code written on the MATLAB tools for executing the concept of image processing. The various gestures done by the colour marker are tracked by the camera and hence given to the laptop for further processing of the various applications involved.



Fig3: Laptop

III. TECHNOLOGIES RELATED TO SIXTH SENSE TECHNOLOGY

Sixth Sense technology takes a different approach to computing and tries to make the digital aspect of our lives more intuitive, interactive and, above all, more natural. We shouldn't have to think about it separately. It's a lot of complex technology squeezed into a simple portable device. When we bring in connectivity, we can get instant, relevant visual information projected on any object we pick up or interact with the technology is mainly based on hand augmented reality, gesture recognition, computer vision based algorithm etc.

A. Augmented reality:

(AR) is a term for a live direct or indirect view of a physical real world environment whose elements are augmented by virtual computer generated imagery. It is related to a more general concept called mediated reality in which a view of reality is modified (possibly even diminished rather than augmented) by a computer. The augmentation is conventionally in real-time and in semantic context with environmental elements.

B. Gesture recognition:

Gesture recognition can be seen as a way for computers to begin to understand human body language, thus building a

richer bridge between machines and humans than primitive text user interfaces or even GUIs (graphical user interfaces), which still limit the majority of input to keyboard and mouse. Gesture recognition enables humans to interface with the machine (HMI) and interact naturally without any mechanical devices.

C. Computer vision:

On the other hand, studies and describes the processes implemented in software and hardware behind artificial vision systems. The software tracks the user's gestures using computer-vision based algorithms. Computer vision is, in some ways, the inverse of computer graphics. While computer graphics produces image data from 3D models, computer vision often produces 3D models from image data. There is also a trend towards a combination of the two disciplines, e.g., as explored in augmented reality.

IV. CONSTRUCTION AND WORKING

The methodology shown below in the algorithm used is based on the Sixth Sense Technology where user has to make several gestures using the finger worn colour markers and perform real time actions whose images are preloaded in the program. Our aim is to move mouse cursor as the user moves his/her fingers, zooming of images and capturing of photos using fingers. For this purpose, three components of Sixth Sense are used i.e. Camera, Coloured Caps and MATLAB installed in Laptop.

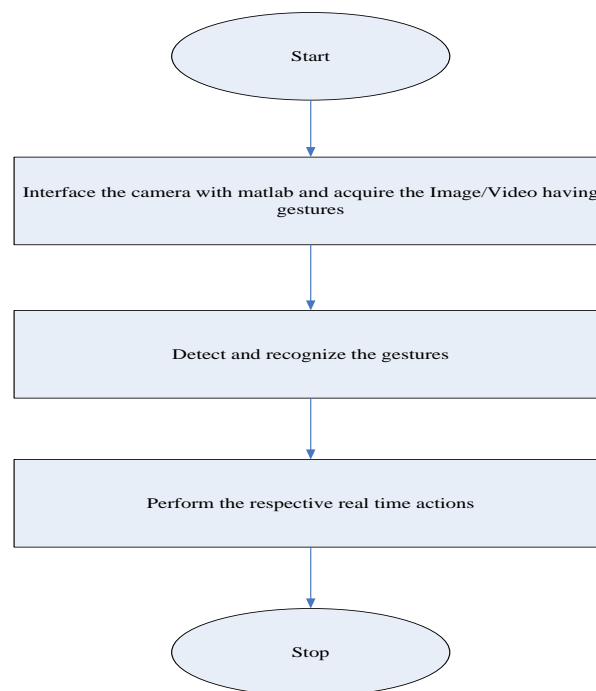


Fig4: Algorithm of Working

The approach works in a continuous manner where camera takes the live video, sending to the laptop, and MATLAB installed in laptop processes the input and recognizes the colours at the finger tips of the user. Following figure shows the algorithm we used in our approach to move mouse cursor on screen, capturing images using gestures and interfacing the image (zoom in & out/rotating).

Here, in our proposed methodology, first interaction with the physical world is done by camera. Camera takes the video and starts recording the live video and in continuation of recording it sends the live video to MATLAB which is already installed in laptop which is connected with the camera. In MATLAB, code is prepared which convert the incoming live video from camera into frames of images or slicing of video is done in the form of images. These images that are obtained from the slicing of video are then processed for colour recognition process. The output of the colour recognition process are the images that contains those colours of which colour caps are present at the finger tips of the user, the background of the image and shadow if present. The fingers of user are not shown in the output images. For this purpose, RGB values of the colour caps are set prior in the code so that no other colour will be detected in the image after colour recognition except the caps colours and the background. The output images are displayed in continuation and at the same speed as the speed at which slicing of video is done, so that it looks like a continuous movie in which the input is physical world and the output is only those colours which are present at the fingertips of the user.

The colour is then associated with the mouse cursor in code so that whenever the colour moves in the output image from one position to another, the mouse cursor gets attached at the same position where the colour is now displayed. In the same manner the combination of yellow, green and blue, red is detected and hence by the action performed, we can click the images. Similarly various gestures of the finger marker are processed and allow the user to interface the image (zoom in & out/rotating).

V. APPLICATION

- A. Taking Pictures: The implementation of gestural camera is done by forming gesture of yellow, green and blue, red when these colours are positioned on the output image the action of gestural camera is done.



Fig 5: capture of object

- B. Moving mouse pointer using fingers: With the help of colour markers the image processed on the camera implements the mouse. The yellow colour marker helps in moving the mouse pointer while blue marker acts as right click and red marker acts as left click.

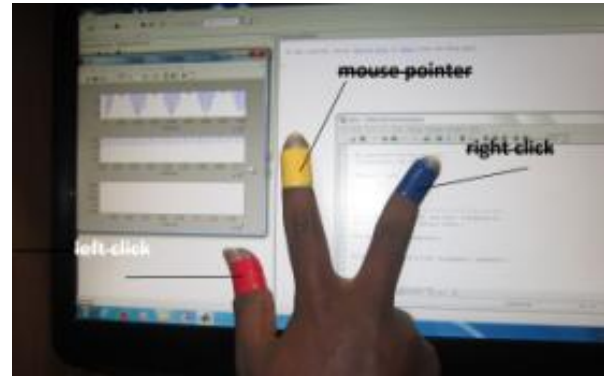


Fig6: mouse pointer description

- C. Resizing of images: The gesture helps us to zoom in and zoom out, rotate the given image using the concept of hand gesture and image processing. The distance and angle between the markers is calculated. Depending on the distance between markers zoom in and zoom out is done. The max. distance performs zoom in and min. performs zoom out The change of angle between colour markers performs the function of rotating the image.

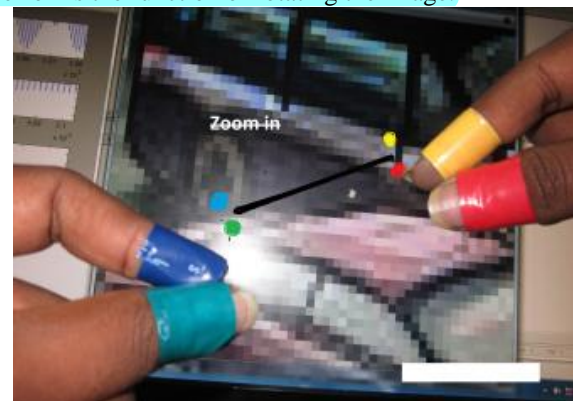


Fig7: zooming of image

IMPLEMENTATION & RESULTS

The image processing and colour recognition was simulated in the MATLAB to process the image from the live video from camera to obtain the final image showing required colours only. Video from the web camera at 30 frames per second is applied to the video input block. Resize block enlarge or shrinks image size. Captured video will be in RGB format. It is converted into chroma and luma components. Luma represents the brightness in an image and it represents the achromatic image without any colour while the chroma component represents the colour information. Image split block splits the image into number of blocks. Each splitted block is resized using bilinear interpolation technique. The Design is carried out using video and image processing block set in MATLAB simulink version R2007b here how images are splitted and various parameters required achieving the desired task.

The individual pixel values in most images do actually correspond to some physical response in real 2-D space (e.g. the optical intensity received at the image plane of a camera or the ultrasound intensity at a transceiver). However, we are also free to consider images in abstract spaces where the coordinates correspond to something other than physical space and we may also extend the notion of an image to three or more dimensions. For example, medical imaging applications sometimes consider full three-dimensional (3-D) reconstruction

of internal organs and a time sequence of such images (such as a beating heart) can be treated (if we wish) as a single four-dimensional (4-D) image in which three coordinates are spatial and the other corresponds to time.

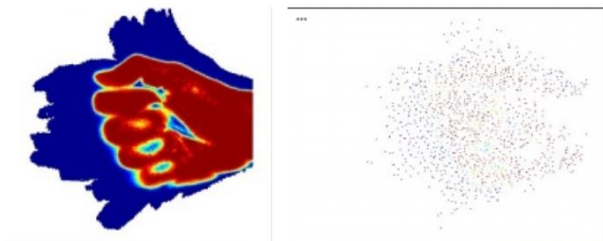


Fig8: conversion of image into pixels

The raw data provided by the video card was in the RGB (red, green, blue) format. However, since the detection system relies on changes in colour (or hue), it could be an advantage to use HSL (hue, saturation, luminosity- see Glossary) to permit the separation of the hue from luminosity (light level). To test this maximum and minimum HSL pixel colour values of a small test area of skin were manually calculated. These HSL ranges were then used to detect skin pixels in a subsequent frame (detection was indicated by a change of pixel colour to white). The test was carried out three times using either hue, saturation or luminosity colour ranges to detect the skin pixels. Next, histograms were drawn of the number of skin pixels of each value of hue, saturation and luminosity within the test area. Histograms were also drawn for an equal sized area of non-skin pixels.

Since the above colour recognition till shows some other colours of the background of the image, hence working on the live video during colour recognition process, it will show background colours also which can create ambiguity for the cursor to detect the exact colour or colours to get associated with to change its position. For this purpose, some more optimization and RGB pixel threshold values need to be modified in such a way that it will take only and only those colours which user want to show. The figure shown below demonstrate the optimized colour recognition output image showing only three colours and background is taken to be black so that visualization doesn't create any kind of ambiguity to the user.

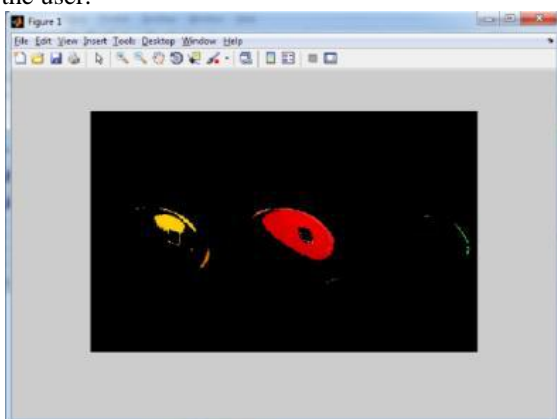


Fig9: Optimized output image obtained after colour recognition process done for three colours i.e. yellow, red and green

VI. CONCLUSION

The implantation of image processing and gesture recognition using MATLAB is practically proved and hence the resizing of image, capturing of photo and movement of mouse pointer is

achieved with good precision. Clearly, this has the potential of becoming the ultimate "transparent" user interface for accessing information about everything around us. If they can get rid of the coloured finger caps and it ever goes beyond the initial development phase, that is. But as it is now, it may change the way we interact with the real world and truly give everyone complete awareness of the environment around us.

ACKNOWLEDGEMENTS

I would like to express heartfelt gratitude towards my faculty guide Mr. D.S.Patil, Faculty, Electronics and communication Engineering Department, Maratha Mandal Engineering College for his guidance and help in completing this work successfully with good results. I would also like to express my gratitude to my Head of the Department Prof. S.S. Gurav for his timely support and help.

REFERENCES:

- [1] Relevant feature for video based continuous sign language recognition. Department of Technical Computer Science, Aachen University of Technology Aachen, Germany, 2000.
- [2] Building temporal models for gesture recognition. In proceedings British Machine Vision Conference, 2000, pages 32-41.
- [3] Use your hand as a 3-D mouse or relative orientation from extended sequences of sparse point and line correspondences using the affine trifocal tensor. In proceedings 5th European Conference on Computer Vision, 1998, pages 141-157.
- [4] O. Castillo, O. and P. Melin, "A New Approach for Plant Monitoring using Type-2 Fuzzy Logic and Fractal Theory", International Journal of General Systems, Taylor and Francis, Vol. 33, 2004, pp. 305-319.
- [5] S. Furui, "Cepstral analysis technique for automatic speaker verification", IEEE Transactions on Acoustics, Speech and Signal Processing, 29(2), 1981, pp. 254-272.
- [6] Intelligent Image Processing, John Wiley and Sons, pp.384, 02001NOV02, ISBN 0-471-40637-6 sixth sense. Pranav Mistry.