ISSN 2395-1621

Interactive Interface Between Human And Computer Through Pattern Recognition

[#1]Apurva A. Saudagar, [#2]Ashwini G. Kamble, [#3]Geetashri S. Date,
[#4]Prof. Sandhya Potadar

[1]apurva822@gmail.com
[2]ashwini.kamble40@yahoo.in
[3]geetashri.date@cummincollege.in

[#123]Student, Department of Electronics and Telecommunication
[#4]Prof. Department of Electronics and Telecommunication

MKSSS's Cummins College Of Engineering for Women Pune-52.

ABSTRACT

A system that provides a computer operating method that merges the advantages of pattern control. In our system, user can interact with computer in a natural way without remembering too much pre-defined gestures and the input of text can be much faster than simply using pattern control. The system not only should be available in various environments with high fault-tolerance but also has to ensure there is no conflict in pattern control.The pattern control computer using hand gesture to control the computer mouse and/or keyboard functions. The user will interact with the system by gesturing in the view of the camera. A more flexible and healthy operating method should be developed instead of the traditional one. Pattern recognition based man-machine interface is being developed vigorously in recent years. Due to the effect of lighting and complex background most visual pattern recognition systems work only under restricted environment. Gestures are expressive, meaningful body motions –i.e., physical movements of the fingers, hands, arms, head, face, or body with the intent to convey information or interact with the environment.

ARTICLE INFO

Article History

Received: 13th May 2016

Received in revised form :

13th May 2016

Accepted: 16th May 2016

Published online :

17th May 2016

Keywords: **Frame Extraction, Gaussian Blur, Rbg To Hsv, Hsv Thresholding.**

I. INTRODUCTION

With the rise of the technical inventions that let people communicate without barrier in the global scope, computer has become part of our lives, however, using the typical input method with keyboard and mouse has a great possibility bringing ‘obesity’, ‘dizziness’ and other syndromes that influences health. Long duration of using computer is beyond the capacity of the old and the complexity of traditional operating method cannot let them grab information quickly and efficiently. A more flexible and healthy operating method should be developed instead of the traditional one.

We conceived the idea of this project while in discussion about the advancements taking in human computer interfaces as Hand gesture recognition based man-machine interface is being developed vigorously in recent years. The project designs a simple, natural and real time system for pattern interaction between the user and computer for providing dynamic user interface. This technique may be also called as pattern recognition, human computer interaction.

The basis of our approach is real time tracking process to obtain the pattern from the whole image .A system that provides a computer operating method that merges the advantages of pattern control. In our system, user can interact with computer in a natural way without remembering too much pre-defined gestures and the input of text can be much faster than simply using pattern control.
Besides the convenience this system can bring user, it still faces challenges. The system not only should be available in various environments with high fault-tolerance but also has to ensure there is no conflict in pattern control.

The pattern control computer using hand gesture to control the computer mouse and/or keyboard functions .The user will interact with the system by gesturing in the view of the camera. However the operating environment is varied, the accuracy of recognition .Gestures are expressive, meaningful body motions – i.e. physical movements of the fingers, hands, arms, head, face, or body with the intent to convey information or interact with the environment.

II. LITERATURE SURVEY

There have been various computer-driven revolutions in the past: the widespread introduction of the personal computer (PC) was one, the invention of the graphical browser was another, and the Internet yet another. There have also been computer eras where one type of computer has dominated, having straightforward implications for whether the computers were shared or personal, and for whether they

Were specialized commodities or not (see diagram below). But the ways computers have altered our lives, all aspects of our lives, is more comprehensive than, at first blush, recollections of these technological revolutions or eras might suggest. Computers affect how we undertake the most prosaic of activities – from buying food to paying our bills – and they do so in ways we might not have imagined when the first personal computers arrived on our desks. They have also created wholly new experiences, for example, allowing us to inhabit virtual worlds with people from many different parts of the globe. In between these extremes, from the prosaic to the wholly new, computers have taken over from older technologies in ways that looked merely like substitution at first but which have ended up creating radical change. Photography, for example, has retained its familiarity despite moving from being chemically-based to being digital. At the point of creation, people still ‘point and shoot’ in much the same way as they used to. However, what one can do with images when they are digital is quite different. Whereas, before, we may have only printed one or two rolls of film, displaying the photos on the mantelpiece or in an album, digital images are now reproduced many times over, and are often broadcast around the world on websites. The activities we undertake and the goals we have in mind when we take photos and share them, then, are not at all the same now as they were even five years ago. It is not just in terms of user experiences, such as shopping, games, and picture-taking that the world has changed.

Computers have altered our sense of the world at large, letting us see images of far-away places, instantaneously and ubiquitously. The world, now, seems so much smaller than it was even a decade ago. In this section we begin to look at many different aspects of how computing technologies have changed and their impact on our lives.

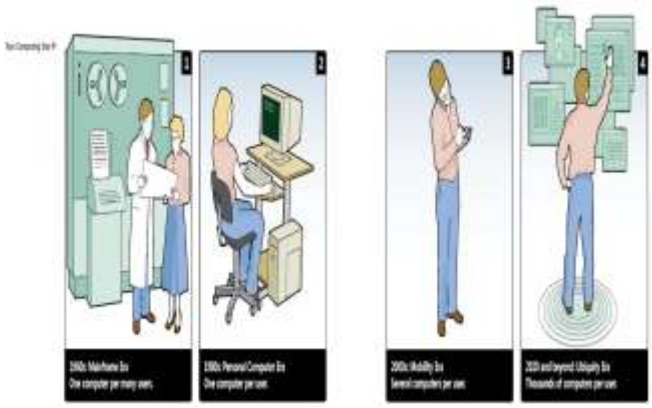


Fig 1. Human Computer Interaction

Human-computer interaction (HCI) is a multidisciplinary research area focused on interaction modalities between humans and computers; sometimes, the more general term human-machine interface (HMI) is used to refer to the user interface in a manufacturing or process-control system. In other words, the HCI discipline investigates and tackles all issues related to the design and implementation of the interface between humans and computers.

Although studies about HCI date back to 1975, recent technological advances in consumer electronics have opened exciting new scenarios: gestures, hand and body poses, speech, and gaze are just a few natural interaction modes that can be used to design affordable natural user interfaces (NUIs)[3].

For many years, humans have sent commands to "machines" primarily via the keyboard-mouse paradigm — also known as WIMP (windows, icons, menus, point-and-click devices). Here, the term machine is used in a very broad sense: in addition to the point-and-click devices that are usually associated with computers, we use a keyboard of sorts to dial numbers on a telephone, to interact with a TV, to select a wide range of functions on a car dashboard, and many other activities that employ key-based interaction modalities. In most cases, the machine's output to the user is then based on a display device such as a monitor.

As foreseen by Andy van Dam, in his vision paper published in IEEE Computer Graphics & Applications' first issue of the new century: "Post-WIMP interfaces will not only take advantage of more of our senses but also be increasingly based on the way we naturally interact with our environment and with other humans."

Sensors such as the Microsoft Kinect are a further step toward the implementation of fully natural interfaces in which the human body becomes the controller. The device lets users provide commands to the machine via gestures and body poses as embedded hardware performs real-time processing of raw data from a depth camera, thus obtaining a schematic of a human skeleton comprising a set of bones and joints. Recognizing the position and orientation of bones lets the hardware identify poses and gestures, which can be mapped to commands for the machine. Researchers have also proposed sensors that can track a user's hands. For instance, the Leap Motion can interactively track both hands of a user by identifying the positions of finger tips and the palm centre, and later computing finger joints using an inverse kinematics solver [3].

Human hand gestures provide the most important means for non-verbal interaction among people. They range from simple manipulative gestures that are used to point at and move objects around to more complex communicative ones that express our feelings and allow us to communicate with others. Hand gesture recognition based man-machine interface is being developed vigorously in recent years. Due to the effect of lighting and complex background, most visual hand gesture recognition systems work only under restricted environment. Many methods for hand gesture recognition using visual analysis have been proposed for hand gesture recognition.

Another method is proposed by E. Stergiopoulou and N. Papamarkos which says that detection of the hand region can be achieved through color segmentation [4].

Another very important method is suggested by Meide Zhao, Francis K.H. Quek and Xindong Wu. They have used AQ Family Algorithms and R-MINI Algorithms for the detection of Hand Gestures [5].

Chris Joslinet. al. have suggested the method for enabling dynamic gesture recognition for hand gestures . Rotation

Invariant method is widely used for texture classification and recognition [6].

TimiOjalaet. al. have suggested the method for texture classification using Local Binary Patterns[7].

III. PROPOSED SYSTEM

The aim of the project is to control computer's mouse by viewing the colour pointer in front of the camera. Image is captured using a webcam. The captured image is converted into frames. Image pre-processing using Gaussian blur, RGB to HSV, Thresholding, COG detection, Segmentation. Gesture recognition logic is applied by Detection of the fingertips. Tracking of the fingertips in consecutive frames to determine the motion. The following steps are involved in the process.

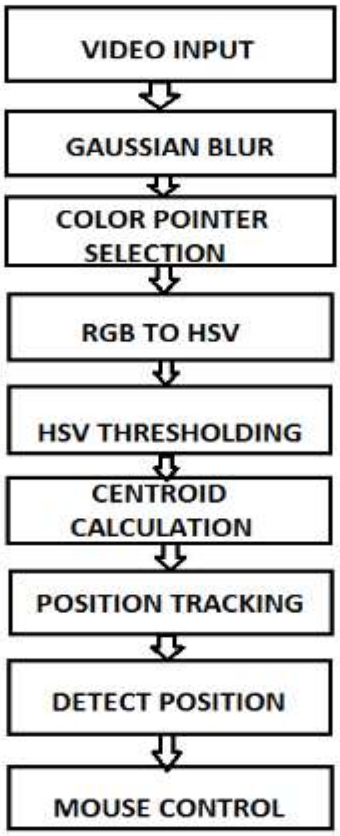


Fig 2. Block diagram

FRAME EXTRACTION:

Video segmentation and key frame extraction are the bases of video analysis and content-based video retrieval. Key frame extraction , is an essential part in video analysis and management, providing a suitable video summarization for video indexing, browsing and retrieval. The use of key frames reduces the amount of data required in video indexing and provides the framework for dealing with the video content . For video, a common first step is to segment the videos into temporal "shots," each representing an event or continuous sequence of actions. A shot represents a sequence of frames captured from a unique and continuous record from a camera. Then key frames are to be extracted. Video segmentation is the premise of key frame extraction, and key frames are the salient content of the video (key factors to describe the video contents).Figure illustrates the basic framework of our algorithm.[2]

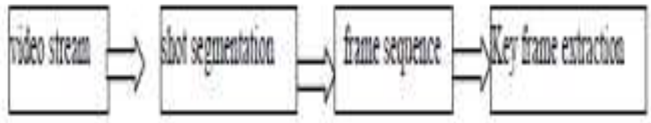


Fig 3. Framework of algorithm

GAUSSIAN BLUR:

A Gaussian blur (also known as Gaussian smoothing) is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise and reduce detail. Gaussian smoothing is also used as a pre-processing stage in computer vision algorithms in order to enhance image structures at different scales see scale space representation and scale space implementation .Gaussian blur is a low-pass filter, attenuating high frequency signals. Mathematically, applying a Gaussian blur to an image is the same as convolving the image with a Gaussian function. Since the Fourier transform of a Gaussian is another Gaussian, applying a Gaussian blur has the effect of reducing the image's high-frequency components; a Gaussian blur is thus a low pass filter. Gaussian blur can be used in order to obtain a smooth grayscale digital image. The Gaussian blur is a type of image-blurring filter that uses a Gaussian function (which also expresses the normal distribution in statistics) for calculating the transformation to apply to each pixel in the image. The equation of a Gaussian function in one dimension is

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \qquad \text{………………..} \quad \text{(I)}$$

in two dimensions, it is the product of two such Gaussians, one in each dimension:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad \text{……………… (II)}$$

Where x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis, and σ is the standard deviation of the Gaussian distribution Values from this distribution are used to build a convolution matrix which is applied to the original image. Each pixel's new value is set to a weighted average of that pixel's neighbourhood.

The original pixel's value receives the heaviest weight (having the highest Gaussian value) and neighbouring pixels receive smaller weights as their distance to the original pixel increases

This results in a blur that preserves boundaries and edges better than other, more uniform blurring filters.

RBG TO HSV:

RGB (Red, Green, Blue) colour model

To see how a RGB model works, we first have to look at the technical aspect of colour. The human eye perceives colour reaction through a mix of red, green and blue signals. It was a thought that each colour can be mixed with only these three primary colours.

HSV (Hue, Saturation, Value) colour model

The HSV colour model, also called HSB (Hue, Saturation, Brightness), defines a colour space in terms of three constituent components:

- Hue is the colour type (such as red, magenta, blue, cyan, green or yellow). Hue ranges from 0-360 deg.
- Saturation refers to the intensity of specific hue. Saturation ranges are from 0 to 100%. In this work saturation is presenting in range 0-255.
- Value refers to the brightness of the colour. Saturation ranges are from 0 to 100%. Value ranges are from 0-100%. In this work saturation and value are presenting in range 0-255.

Hue and saturation taken together are called Chromaticity, and therefore, a color may be characterized by its Brightness and Chromaticity.

The simple answer is that unlike RGB, HSV separates luma, or the image intensity, from chroma or the colour information. This is very useful in many applications.

For example, if you want to do histogram equalization of a colour image, you probably want to do that only on the intensity component, and leave the colour components alone. Otherwise you will get very strange colours.

In computer vision you often want to separate colour components from intensity for various reasons, such as robustness to lighting changes, or removing shadows.Note, however, that HSV is one of many colour spaces that separate colour from intensity (See YCbCr, Lab, etc.). HSV is often used simply because the code for converting between RGB and HSV is widely available and can also be easily implemented. For example, the Image Processing Toolbox for MATLAB includes functions rgb2hsv and hsv2rgb.

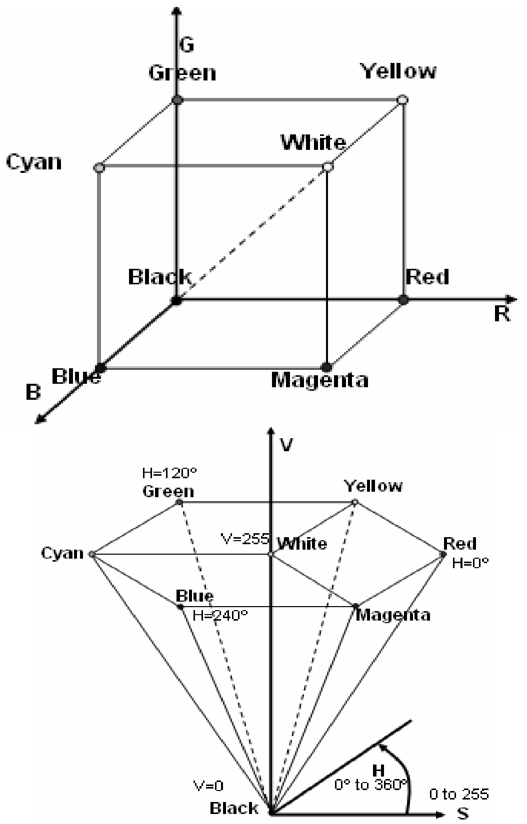


Fig 4 RGB TO HSV Block

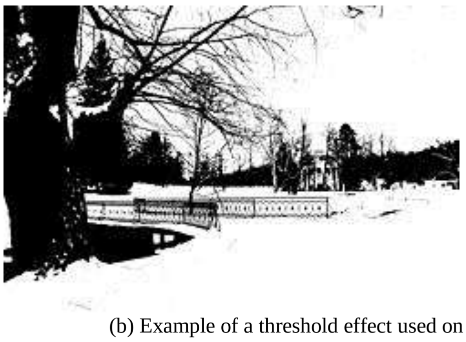
HSV THRESHOLDING:

The simplest thresholding methods replace each pixel in an image with a black pixel if the image intensity $I_{i,j}$ is less than some fixed constant T (that is, $I_{i,j} < T$), or a white pixel if the image intensity is greater than that constant. In the example image on the right, this results in the dark tree becoming completely black, and the white snow becoming complete white.

The class HSV Threshold is a tool for clustering problems in HSV (Hue, Saturation, and Value) colour space for colour-based pixel separation. The main purpose of this tool is to support semi-automated colour-based segmentation. Particularly, this tool provides users the decision support about valid pixel value ranges for specific types of object detection.

The class operates in three ranges of HSV values. Points within the user-specified value ranges are appeared in the result image, both as RGB and HSV images. In each scrollbar, a user can adjust the upper and lower limits of the pixel values, and can fix the dynamic value range by checking the “Fix Range” option. For example, the lower/upper limit will move automatically while keeping the same dynamic value range when changing the upper/lower limit.

(a)Original image



(b) Example of a threshold effect used on an Image

Fig 5. Thresholding Of Given Image

CENTROID DETECTION:

Object detection plays an important role in many image processing problems. It is important to determine the position of each detected object with the highest possible accuracy and precision. When the positions of the objects need to be known with sub pixel precision, accurate and robust estimates can be obtained by computing their centroid. In the field of computer vision, blob detection refers to mathematical methods that are aimed at detecting regions in a digital image that differ in properties, such as brightness or colour, compared to areas surrounding those regions. Informally, a blob is a region of a digital image in which some properties are constant or vary within a prescribed range of values; all the points in a blob can be considered in some sense to be similar to each other.Given some property of interest expressed as a function of position on the digital image, there are two main classes of blob detectors:

Differential methods, which are based on derivatives of the function with respect to position.

Methods based on local extrema, which are based on finding the local maxima and minima of the function. With the more recent terminology used in the field, these detectors can also be referred to as interest point operators, or alternatively interest region operators.

There are several motivations for studying and developing blob detectors. One main reason is to provide complementary information about regions, which is not obtained from edge detectors or corner detectors.

In early work in the area, blob detection was used to obtain regions of interest for further processing. These regions could signal the presence of objects or parts of objects in the image domain with application to object recognition and/or object tracking. In other domains, such as histogram analysis, blob descriptors can also be used for peak detection with application to segmentation. Another common use of blob descriptors is as main primitives for texture analysis and texture recognition. In more recent work, blob descriptors have found increasingly popular use as interest points for wide baseline stereo matching and to signal the presence of informative image features for appearance-based object recognition based on local image statistics. There is also the related notion of ridge detection to signal the presence of elongated objects.

IMAGE SEGMENTATION

Decomposition of scene into its components is called as image segmentation. Segmentation sub divides an image into its constituent regions or objects. Level of division depends on the problem being solved. For higher level of extraction segmentation is fundamental. In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).

GESTURE RECOGNITION

Gestures are expressive, meaningful body motions – i.e., physical movements of the fingers, hands, arms, head, face, or body with the intent to convey information or interact with the environment. There are several aspects of a gesture that may be relevant and therefore may need to be represented explicitly. Hummels and Stappers describe four aspects of a gesture which may be important to its meaning:

- Spatial information – where it occurs, locations a gesture refers to.
- Pathic information – the path that a gesture takes.
- Symbolic information – the sign that a gesture makes.
- Affective information – the emotional quality of a gesture.

In order to infer these aspects of gesture, human position, configuration, and movement must be sensed. Gesture recognition is the process by which gestures made by the user are made known to the system. Gesture recognition is also important for developing alternative human-computer interaction modalities .It enables human to interface with machine in a more natural way. Gesture recognition is a technique which used to make computers 'see' and interpret intelligently is becoming increasingly popular. Dynamic gesture recognition isn't something entirely new.

IV. ALGORITHM

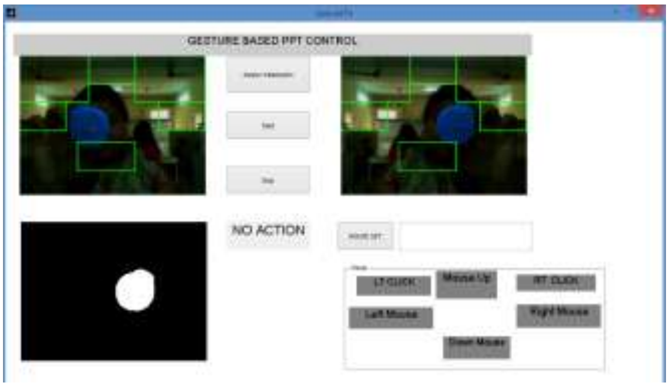
1. Declare variables
2. Extract video frame
3. Declare co-ordinates for frame
4. Take mirror image of video feed
5. Blur extracted video feed using Gaussian filter
6. Convert RGB to HSV
7. Also convert RGB to Binary Image
8. Apply HSV Thresholding
9. Perform Image Segmentation based on screen partitioning by counting number of pixel
10. Apply gesture recognition logic

V. RESULTS

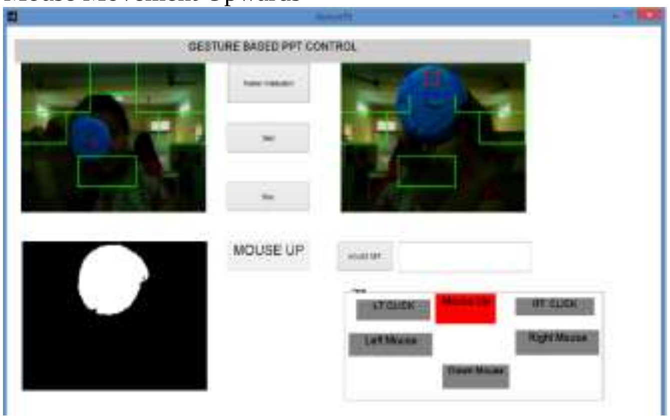
Master initialization (Frame extraction)

Colour Pointer Selection

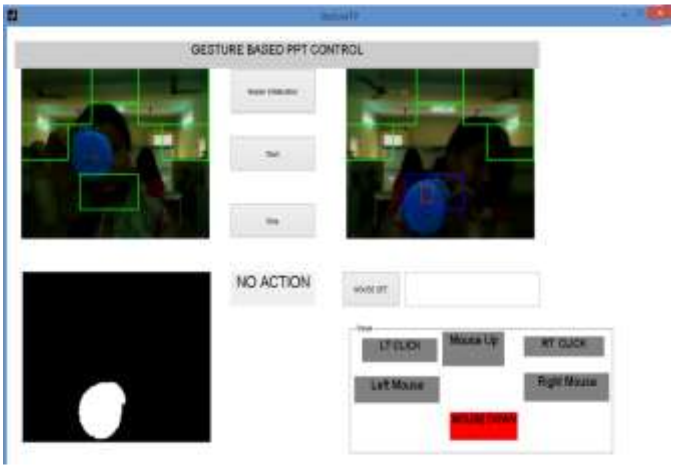
No Action Performed



Mouse Movement Upwards



Mouse Movement Downwards



VI. CONCLUSION AND FUTURE WORK

This new interaction paradigm's success will rely on future technological advances, which aim to transform interface devices into wearable and embeddable objects. Interfaces based on augmented reality (AR) technologies are clear examples of this transformation. Many applications for tourism, entertainment, maintenance, shopping, and social networks are already available for personal devices, but new wearable sensors might soon change our habits. Google Glass will be (massively) marketed in the near future, and new application fields are proposed daily. Human-machine interaction and human-machine "integration" are doomed to become very similar concepts, and indeed, Google Glass-like solutions could soon be replaced by contact lenses that implement natural eyewear-based interfaces.

The ability to sense our interaction without direct physical engagement with computer systems or input devices is also a growing trend. Eye movements have been used for many years as a way of supporting the disabled in interacting with computers, but now we are also seeing the advent of ‘brain-computer interfaces’. Such systems allow, for example, people with severe physical disabilities to use their brain waves to interact with their environment. Real-time brainwave activity is beginning to be used to control digital movies, turn on music, and switch the lights on and off. These interfaces can even control robot arms, allowing paralysed individuals to manipulate objects.

REFERENCE

1.Sang-woong Lee, Sone-hyang Moon, Seong-whan Lee, “Face recognition under arbitrary illumination using illuminated exemplars”, Pattern Recognition - PR , vol. 40, no. 5, pp. 1605-1620, 2007.

2.Guozhu Liu, and Junming Zhao College of Information Science & Technology, Qingdao Univ. of Science & Technology,266061,P. R. China Email: lgz_0228@163.com.

3.Guest Editors' Introduction Paolo Montuschi, Andrea Sanna, Fabrizio Lamberti, and Gianluca Paravati September 2014.

4.E.Stergiopoulou and N.Papamarkos: “A New Technique on Hand Gesture Recognition”, Proc of the IEEE International Conference on Image Processing, 2657-2660, 2006.

5. Meide Zhao, Francis K.H. Quek, Member, IEEE, and Xindong Wu, Senior Member, IEEE : “RIEVL: Recursive Induction Learning in Hand Gesture Recognition”, IEEE Transactions on Pattern Analysis and machine intelligence, vol. 20, no. 11, November 1998.

6. Chris Joslin, Ayman El-Sawah, Qing chen, Nicolas Georganas, “ Dynamic Gesture Recognition”, Proc. of the Instrumental and Measurement Technology Conference, pp 1706-1710, 2005.

7. Timi Ojala, Matti Pietikainen and Topi Maenpaa, “Multi-resolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, pp. 971-987, 2002.

