HAND GESTURE RECOGNITION USING BINARY LANGUAGE FOR AUDIO BASED COMMUNICATION

Subhashini.S[1], Dr.S.Revathi[2]
[1]Assistant Professor, [2]Associate Professor
B.S Abdur Rahman Crescent University, Chennai

Abstract

The gesture based interaction systems are becoming more and more popular both at workplace and home. Non-verbal way of exchanging information is gestures. It can be done through innumerable signs. The proposed work intends to develop a system that can recognize hand gestures which can be used as an input command to interact with the PC that can be applied to get audio output. One of the key areas that need to be looked while developing such systems is the image processing stage. It facilitates human control similar to robotic monitoring systems.

The proposed work is to design and implement a system that can perform general image processing of the user image of binary language captured in real time. The outcome of this system is that can detect the hand gesture captured in the user image and detect contours around the detected skin regions. Implementation of the code is developed with combination of java and the Open CV library.

Keywords: Binary Sign Language, Human-Computer Interaction (HCI), Hand gesture recognition, User interface.

I. INTRODUCTION

The proposed method is implementing a system to understand gestures of binary sign language by the computers and develop a user friendly human computer interfaces (HCI) to generate audio output for the hand gesture.

Information are exchanged non-verbally through gestures. Innumerable gestures can be performed by the person at a time. Computer vision researchers find this subject as their interest since human gestures are perceived through vision. The Proposed work aims to recognize binary sign of human gestures by creating an HCI. Coding of these gestures into binary language demands a better way of implementing an algorithm. To access any type of information. We use mouse, keyboard which results in waste of time and it is inconvenient to handle. So importance of hand gesture recognition increases. It creates the natural interaction between computer and human. Keyboard and mouse has been replaced by hand, face like gestures. Natural input devices like these attract more attention because it is powerful, more effective, and does not require extra connection than any other devices. Hand gesture using doppler effects [1] can be applied to control applications.

Human–computer interaction (HCI) makes the process to interact between users (people) and computers. As an alternative to the traditional human-computer interaction interface (HCI) such as the keyboard and the mouse, the use of human movements, involving face, the whole body and specially hand gestures has attracted more and more people in recent years. A computer that can recognize and respond to the users gestures could provide a natural interface. The diverse logical and physical capabilities of users (e.g. elderly, children or people with disabilities) also require human-computer interfaces that are easily learnable and usable, instead of traditional interaction techniques such as the mouse and keyboard, which require a certain kind of skills, and restrict the user in a certain kind of physical mode. One of the latest developments made in the field of gaming using such interaction technique is the Microsoft's Kinect for the Xbox 360 console [4]. It enables users to control and interact with the Xbox 360 without the need to touch a game controller like remote,

such gesture based have become extremely popular among today's community.

Hands are one of the most multipurpose tools in our human body to accomplish different tasks. They are one of the important features in our user interfaces and interaction applications. Interest in the field of computer vision based hand gesture recognition has increased in recent times due to its potential application in the field of Human Computer Interaction. The most important feature of this technique is that the system uses hand gesture recognition as an input, through which the users can control the system or devices without having to touch any external interaction devices such as mouse or keyboard. Also, it gives users a sense of freedom and ease to accomplish different tasks. This serves as a motivating factor to carry out this proposed method.

This proposed work aims to develop a system, which captures certain hand gestures as an input from the users using a web camera, and performs the task associated with the gesture recognized. A program based on an open source library, which looks into image processing, and hand gesture recognition developed to accomplish the aim of this proposed method.

II. HUMAN COMPUTER INTERACTION

Human–computer interaction (HCI) enables the user friendly communication [6][7][8] .The aim of HCI is to improve the interactions between computers and users by making computers more usable and more responsive to user's needs.

A. Interaction Method

There are different interaction methods with which we can interact with the computer the most common being the use of a mouse. The mouse was developed at Stanford Research Laboratory (now SRI) in 1965 to be a cheap replacement for light pens, which had been used at least since 1954 [10, p. 68]. The mouse was then famous as a practical input device by Xerox PARC in the 1970's.

Another interaction method that has been increasingly popular in the recent years is using gestures for interacting with the computer. Instead of learning completely different new ways to interact, the users may prefer to adopt the natural ways of communication that they are familiar with in everyday life. These demands have resulted research in which the user interfaces take advantages of the natural ways that people interact with each other and the physical world, e.g. eye gaze, speech, gesture and physical tools. Such systems accept gestures as an input form the user recognize the inputted gestures and perform the task associated with that gesture. This project will look into the gesture-based interaction in real time.

Based on their purpose,two main groups of classification of gesture based applications are multidirectional control and a symbolic language.

B. Sign Language

Sign languages is the most common natural form of all languages which makes to arise of the human civilization, . It has been practised as communication language when no spoken language were in existence. Even now the sign language adopted as an integral part of our day to day communication process. Sign languages are very useful for communication of deaf and dumb people and also in other fields where sign lanuage talks. Gestures are one of the first forms of communication when a child learns to express its needs. It enhances the emphasis of spoken language and helps in expressing thoughts and feelings effectively.

Sign language plays an vital role in human life. There are some common sign gestures used by the all people to communicate. A simple gesture with one hand has the same meaning either 'hi' or 'goodbye' all over the world. Many people travel to foreign countries without knowing the official language of the visited country and still manage to perform communication using gestures and sign language. These examples show that gestures can be considered international and used all over the world. In many jobs , there are predefined sign languages which has special meaning .

III. HAND GESTURE RECOGNITION

Gestures can be described as different types of human movements. These can be two-dimensional or three-dimensional and can be specific to the hand, arm or body movements as well as facial expressions. Gesture recognition enables humans to interface with the machine and interact naturally without any external devices such as the keyboard. It is a method of assigning commands to the computer (machine) to perform specific tasks.

A hand gesture us a sequence of hand postures connected by continuous hand or finger movements over a short period of time. Hand gestures provide a separate complementary modality to speech for expressing ones ideas. So, hand gestures recognizing system can be a natural way of communicating between the computer and humans.

There is basically two approaches to hand gesture recognition; Vision based and Non-vision based approaches.

The vision based approach uses camera as an input device, thus facilitating a natural interaction between users and computers without the use of any extra devices.

IV. VISION BASED METHOD

Bare-hand gestures are probably the most straightforward interpretations when people think about gestures. Here the gestures that are defined entirely by the movements of the user's hands and/or fingers. Typically the bare hand gestures are captured using computer vision techniques, i.e. cameras watching the user's movement it can be a single camera, stereo or multiple cameras depending on the application and settings[3].

This work requires a gesture recognition method that is easy to use and allows the user a certain level of freedom. This project uses a single camera (web camera) as an input device to capture gestures performed by the users.[2] The vision based gesture recognition seems to be a better option due to its advantages over non-vision based method. The devices used in non-vision methods are expensive and bring weighty experience to the users. Also, the devices are generally connected to the computer, which restricts free movement of the users to perform the activity they want. Whereas in vision method, the users are free to perform gestures without any restrictions.

Similar research of gesture recognition has been done throughout the past covering a wide variety approaches with successful outcome. For example, Segen describes a system with two cameras that can recognize three gestures and tracking hand in 3D. The system detects two fingers (thumb finger and pointing finger) by extracting the feature points on hand contour and output their poses. In Quek describes a system called finger mouse that can replace mouse with hand gestures for certain actions. The system defines only one gesture (pointing gesture), and the SHIFT key on the keyboard is adopted to register a mouse button press. In Triesch presents a robust gesture recognition algorithm using multiple cues including motion cue, color cue and stereo cue. This algorithm is used to build a gesture interface for a real robot grasping objects on a table.

A. Image Processing

"Images are stored as a collection of pixels." Color Images consists of a red, green and blue value, which is combined to allow colors to be represented. Grayscale images are different however; "as pixels are represented by a single number ranging from 0 to 255, where 0 is very black and 255 is very white."

Image processing in computing is used to extract useful information form images to perform some specific tasks. Image processing generally involves three basic steps. Image segmentation, which involves image conversion between different color spaces to minimize the complexity of image. Skin detection, which gets rid of any unwanted background objects and noises associated with the image. Contour detection, to locate an object in the image.

.

B. HSV (Hue Saturation Value)

Hue-saturation based color space is another popular color space that is based on human perception color. Hue defines the major color of an area. Saturation measures the colorfulness of

an area in proportion to its brightness. The value is related to the color luminance. It was introduced for users who need to define the color properties numerically. It is easier to implement and also can be converted to and from RGB anytime.

C. Image Segmentation

In computing, segmentation refers to the process of partitioning an image into multiple segments (sets of pixels). The main aim of segmentation is to simplify and/or change the representation of an image into something more meaningful and easier to use and analyze.

Image segmentation is one of the first step involved in the process of gesture recognition in our case hand gesture recognition. The image captured by the camera cannot be used to track hand or recognize gestures as the image consists of other background objects and exists generally in RGB color space which makes skin detection process complex due to the involvement of different color pixels in the image. So, in order to make the skin detection process simpler, the image needs to be converted to a simpler color space which is easier to analyze and which involves lesser color pixels. After the HSV skin color model is built, it can be used for skin detection.

D. Skin Detection

Skin color is one of the most important features in the humans. There are lots of color spaces that have been used in early work of skin detection, such as RGB, YCbCr, HSV (Yoo et al (1999)). Although RGB color space is one of the most used color spaces for processing images, it is not widely used in skin detection algorithms because the chrominance and luminance components are mixed. Some work has been done to compare different skin color space performance in skin detection problems. According to Zarit et al., HSV gives the best performance for skin pixel detection. When building a system that uses skin color as a feature for detection, several points must be kept in mind like what color space to choose and how to model the skin color distribution.

In this paper, "a skin color model based on HSV color space will be built because it has only two components (H, S) which help to speed up the calculations and also the transformations from RGB color space into HSV color space is done using simple and fast transformations"

E. Contour Detection

The term contour can be defined as an outline or a boundary of an object. Therefore, contour detection deals with detecting various objects in an image .Use of contour detection in image processing is to locate objects and their boundaries in Images. Also, output of contour detection shows only the prominent region boundaries leaving behind unwanted edges in the image. Hence, detection of specific objects in the image is only possible through contours.

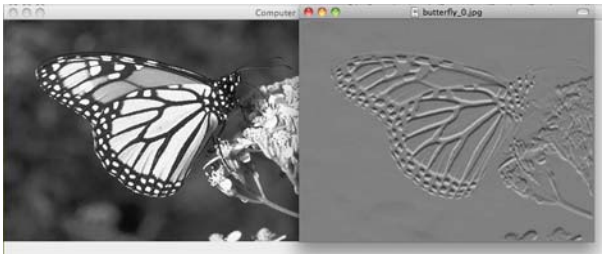

Fig: 1.0

Figure 1.0 shows the contour or boundary detection of an input image.

So in this project, it is very important to detect contours of the hand before we can extract the hand features from the image taken from the camera.

V. SYSTEM ARCHITECTURE

To implement this work , camera capture of hand gesture videos (frame slides) converted into grey scale image which in turn preprocessed to match with gesture database.The overall system architecture is shown in the figure 2.0

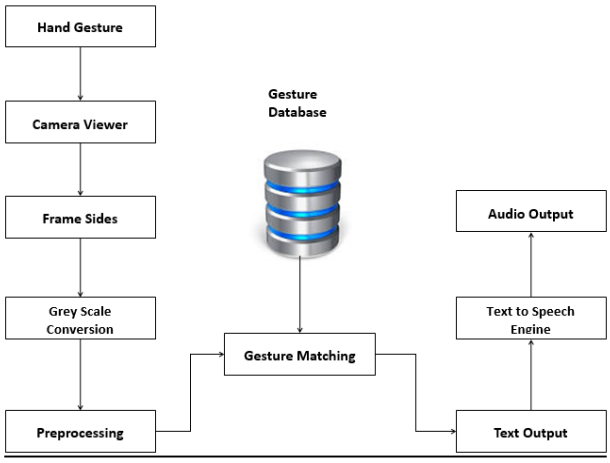

Fig: 2.0

VI. PROJECT OVERVIEW

A hand gesture recognition based audio generation system clearly identify the signs of each finger and hand orientation in real time. It should be robust on various changes in hand measurements, rotation, color and lighting. This is a very complex problem and requires advanced image processing and computer vision concepts. In the proposed work , a novel method is used to recognize hand gestures of binary sign language in real time which provide accurate results clearly.

A framework is used to detect and track palm region and recognize gestures. The proposed method uses a camera to capture both [5] the features of the ROI from the finger and hand gestures and it is implemented using OpenCV API in java language.

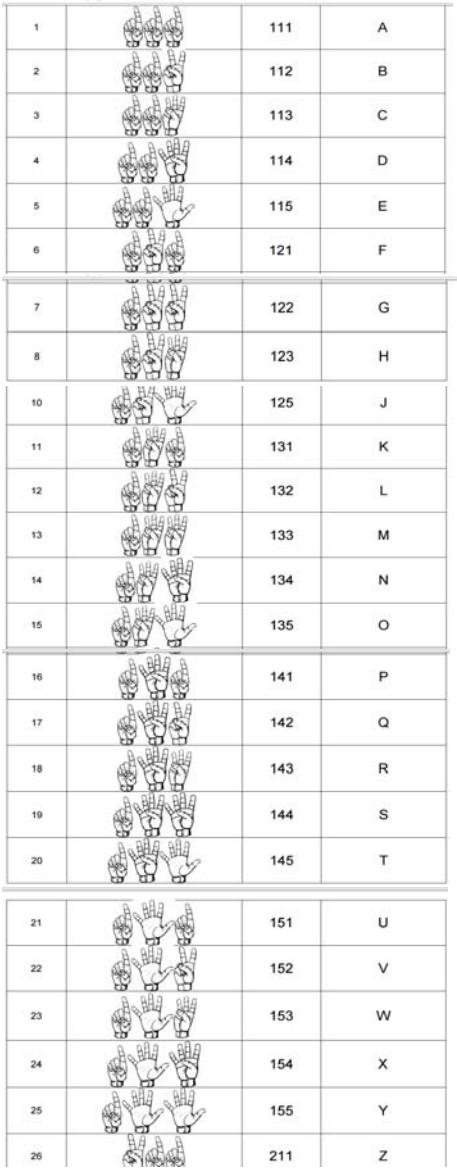

Fig 3.0

Figure 3.0 explains the hand gesture of binary language and coresponding data values (A-Z).

A. Hand Detection and Tracking

The predefined background is set up to caputre the video through camera. A black background is used to clear identication of the hand gesture. Caputerd video frame defines the depth of data. Depth based thresholding is performed to remove the background followed by the segmentation based on depth data for the object closest to the camera.

Here, P is the set of pixels (x, y) represent the hand region. D(x, y) represents the depth of pixel (x, y) from the camera. The RGB valued pixel images is converted into binary images. A morphological filtering is performed on these pixels to check if these actually represent the hand pixels based. If these are not recognized to belong to the hand region, then the algorithm waits for the next frame.

$P = \{(x,y): d < D(x,y) < d+c \}$

Where, P' is the set of color filtered pixels and I(x, y) represents the intensity of pixel (x, y.), S are the pixels recognized as hand pixels, then the back projection of the detected region is passed on to mean-shift algorithm to track the hand region. This provides us with the hand mask. Multipart figures

Figures compiled of more than one sub-figure presented side-by-side, or stacked. If a multipart figure is made up of multiple figure types (one part is linear, and another is grayscale or color) the figure should meet the stricter guidelines.

B. Hand Region Analysis

This section defines the segmention of arm from the palm region of the hand. To achieve this, first the fingers are distinguished using the triangular calculation method. In the palm detection, the color of skin, background subtraction, hand image extraction, edge detection and histogram analysis are used to achieve the goal.

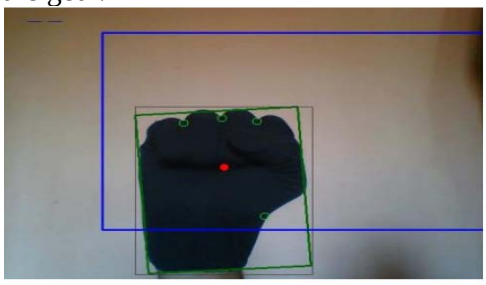

Fig 4.0

Figure 4.0 explains the hand detection of black identification of gestures after contour detection. In fingers distinguis h; it record the tips and valley of the fingers by means of calculating the histogram of the palm image firstly.

Next, it find out the original point which is the center of the gravity of the palm using the area that the palm image gets rid of the fingers part. Successively, it draw the original point and center of the cut line of the palm use as the base line, means zero angle line.

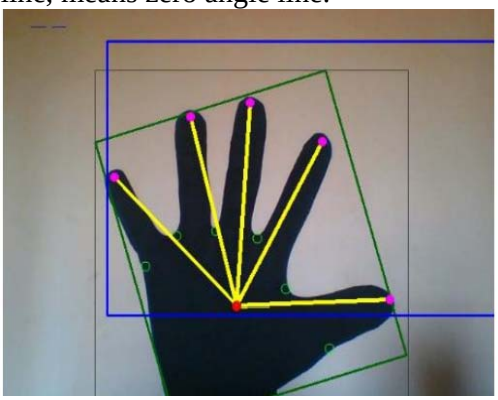

Fig 4.1

Figure 4.1 explains the fingers tips detection of feature selection.

Meanwhile, it draw another line from tip of finger to the original point called tip line. Finally, the system calculate the angle between base line and tip line use as the finger angular. Since the fingers have different angle, so the fingers are easily be distinguished..

C. Gesture matching

Gesture recognition module is the core part of the proposed method. This starts by first creating a closed contour C1 for the hand mask. This is done using simple chain approximation. A vector V1 representing the orientation of the hand in the image plane is created.

$$V1 = (x2-x1)i+(y2-y1)j$$

Where, (x1, y1) is the center of rectangle R1 and (x2, y2) is center of rectangle R2. A vector perpendicular to V1 is also created. A line L1 passing through the center of the points (x1,y1) and (x2, y2) with direction of V2 is created, and is used to create a contour C2 which consists of points that lie above L1.Resolution .

Gesture matching describes efforts to compare two sets of collected data. This can be done in many different ways, but the process is often based on algorithms or programmed loops, where processors perform sequential analyses of each individual piece of a data set, matching it against each individual piece of another data set, or comparing complex variables like strings for particular similarities.

VII. CONCLUSION

Human recognition has continued to be an active area of research and has thus rightfully attracted much attention from the researchers over the years. Important application domains, such as automatic video indexing and archiving, video surveillance, human-computer interaction, augmented reality, user interface design, and human factors would benefit immensely from a robust and efficient solution to this problem.

This project captures live video stream and pass it as input and process it to detect humans and gestures that occur in the video. This work proposes and evaluates alternative choices to extract patches densely. Beyond simple strategies derived from regular interest region detectors, we propose approaches based on super pixels, edges as detectors. The different approaches are evaluated on recent image retrieval and fine-grained classification benchmarks. In this project a unique method for detecting moving object based on edge detection, frame difference and Feature matching is proposed. The feature extraction method will be applied on edge difference image which is followed by feature matching technique.

The Text to speech synthesizer will take the output generated by the engine and convert it in to audio format which can be heard by the opponent.

REFERENCES

[1] Youngwook Kim and Brian Toomajian "Hand Gesture Recognition Using Micro-Doppler Signatures With Convolutional Neural Network", October 8 2016, IEEE.

[2] Rishabh Agrawal and Nikitha Gupta "Real Time Hand Gesture Recognition for Human Computer Interaction" - 6th

International Conference on Advanced Computing, 2016 IEEE.

[3] Zhiqi Guo,Chuvan Lv,Dong Zhou,Zili Wang, "Human-Computer Interaction in Immersive Virtual Maintenance" - 12th World Congress on Intelligent Control and Automation (WCICA), 2016.

[4] Liang Guo, Chenxi Liu, Xiaoyan Wen, Haohua Chen, Jianghui Zhang "A Control System of Human-Computer Interaction Based on Kinect Somatosensory Equipment" - 28th Chinese Control and Decision Conference (CCDC), IEEE 2016.

[5] Kassim Mwitondi et al, "Automatic Prediction and Detection of Affect State Based on Invariant Human Computer Interaction and Human Physiological Response" - Seventh International Conference on Computational Intelligence, Modelling and Simulation 2015.

[6] Hidetoshi Ishiyama et al, "Monochrome Glove: A Robust Real-Time Hand Gesture Recognition Method by using a Fabric Glove with Design of Structured Markers" - Virtual Reality Conference, IEEE 2016.

[7] Sriparna Saha et al, "Gesture Based Improved Human-Computer Interaction Using Microsoft's Kinect Sensor" – IEEE 2016.

[8] Abeer Alnuaim et al, "Enhancing Student Learning of Human-Computer Interaction Using a Contextual Mobile Application" - SAI Computing Conference 2016 IEEE.