Design and Implementation of a Natural User Interface using Hand Gesture Recognition Method

Jishmi Jos Choondal, C. Sharavanabhavan

Abstract*—This paper presents a system which is capable of communicating PC using natural gestures. This system integrates the physical surroundings of a person with real time computer generated information. It creates an enhanced perception of the surrounding environment. To achieve this, a Natural User Interface is designed and implemented by using vision Based hand gesture recognition method. It should be capable of tracking the hand gestures and provides a feedback according to the recognized gesture. The user will be able to use his/her hand movements in order to control the operations which are usually carried out with a mouse. According to different hand movements, the system will respond and carry out the respective operations that are available. Here the natural gestures are recognized through analyzing the image frames from the web camera which is focused to the computer monitor. And it relies on a user being able to carry out relatively natural motions, movements or gestures that they quickly discover and control the computer application or manipulate the on-screen content. This system provides an efficient way of communication with PC.

***Index Terms*— Hue Saturation Value color space, Human computer interface, Natural user interface, Vision based hand gesture recognition**

I. INTRODUCTION

Augmented vision and reality is a relatively new field that has seen a lot of activity in recent days. Current user interaction approaches with keyboard, mouse and pen are not natural enough with the development of ubiquitous computing. On PC platforms, there are applications like interactive entertainments with augmented reality which requires more natural and intuitive interfaces. For mobile or hand held devices, their size limitation leads to limited input space and encumbered experience with tiny keyboard or touch screen. Herein the significant importance of hand gestures has come. It is an important component of body language in linguistics. If hand gestures can be used for communication between human and computing devices, then a natural interaction can be achieved.

Vision based hand gesture interface has been attracting more attentions in this field. It requires no extra hardware requirement except camera, which is very suitable for ubiquitous computing and emerging applications. This idea is aimed at reducing amount of use of the physical interfaces of the PC. It is simple and easy to use for a layman because of making use of hand gestures. This system leads to be a part of paving a new path for the future in computers. Although this paper is not exactly the refined system, it will be a stepping stone towards the new path. Although the specified hand gestures for certain actions must be learned in order to use the system as it is easy to learn. The system is designed in such a way that the users can easily interact with the system and manipulate application programs with simple actions.

In this system, spatial hand gestures are used for interaction instead of a Mouse or a Keyboard. That is, the system is capable of tracking the hand gestures and then provides a feedback according to the recognized gesture. The recognized gestures will invoke the required operations based on its functionality.

To achieve the system, literature survey is carried out throughout the entire phases of this system. It includes many technical papers and relevant web documents. Based on the background survey, design analysis and the requirement specification are identified for the proposed system. The objectives of the proposed system are arrived based on the requirements. It includes the identification of the hardware and software requirement specifications. A detailed study of the software description has to be carried out. Then the system is implemented using modular approach.

II. LITERATURE REVIEW

The following session gives an idea of different systems existing in the relevant area. Physical interfaces are commonly used with PCs but have weight, size and related problems in mobility. So the best options are the use of hand gestures. There are various ways of developing the natural interfaces using gestures. Some of them have discussed below.

This paper, Vision based Hand Gesture Recognition, presents a complete interaction of 3D application without any device. The system uses a vision based hand gesture recognition method. In this method hands has to be isolated from the other parts of the body and from the other background objects. Then the isolated hands are used for recognizing the proper gestures. Recognizer has to be trained for the proper grammar. Then it is used for driving the user application. [13]

This paper, Enhancing Hand Gesture Recognition using Fuzzy clustering based Mixture of Experts Model, presents a method to improve the hand gesture recognition method. This system uses a fuzzy c means clustering algorithm and uses opinion of multiple local experts derived from them.

Revised Manuscript Received on March, 2013.

Jishmi Jos Choondal, Department of Computer Science and Engineering, AMS Engineering College, Anna University Chennai, Bangalore, Karnataka.

Prof. C Sharavanabhavan, Department of Computer Science and Engineering, AMS Engineering College, Anna University Chennai, Namakkal, Thamil Nadu.

The main idea is that it divides the whole hand gestures in to several groups. The similar ones from each group is trained by local specialized experts. [9]

This paper, Finger Tips Detection and Gesture Recognition, presents a method for hand tracking and finger tips detection in order to create a natural interface between human and PC. This system identifies the hands and finger tips with the help of video frames. [3]

This paper, Dynamic Hand Gesture Recognition Using the Skeleton of the Hand, presents the system which uses vision based method for interaction with the help of gestures. The motivation of the development of system is to control Robot from a distance. This paper presents the idea of two types of gestures; static and dynamic. Here dynamic hand gestures are used by extracting 2D skeleton of the hand. Dynamic hand gestures are represented by a sequence of images, each one corresponding to a hand posture. The recognition of a gesture is performed using the fusion of a static and a dynamic recognition technique. The static recognition is used to detect the start and end positions of the hand during the gesture. The dynamic recognition is only performed between the start and stop images. The gesture recognition process is identified with four main procedures. First the Images are received from a camera. The main procedures are i) Image Pre-processing stage; which includes binarization, enhancement and hand region extraction to focus on gesture, ii) Feature Extraction; it consists of two techniques called static gesture recognition and dynamic gesture recognition, which includes Orientation histogram and hand region Skelton, iii) Training; which is to use trained set of data to recognize both static and dynamic gestures, iv) Gesture Classification; which is used to classify the gestures received based on the feature training set [5].

This paper, Hand Gesture Recognition in natural State Based on Rotation Invariance and OpenCV Realization, presents a hand gesture recognition method using rotation invariance. This system can realize multi-angle and posture changing hand gesture recognition in natural state. This system used OpenCV, which is image processing software, in order to design and realize hand gesture recognition. This system uses a method called gesture segmentation based on the colour space of hands. So the HSV colour space segmentation was done for the skin colours. [6]

This paper, Real Time Hand Gesture Recognition System for Dynamic Applications, presents a system which is capable of controlling applications like virtual games, browsing images etc. without the use of mouse. This system has three common stages for gesture recognition. They are i) Image Pre-processing ii) Tracking iii) Recognition. This system application architecture design is done as follows; it starts with the acquisition phase. In this phase natural hand gestures in the real world are used instead of standard input peripheral devices like keyboard, pointing device etc. Here real time constraints are used for the design of the system. So the incoming unnecessary information is removed. In particular, a background suppression procedure has been performed in the HSV colour space. Here human skin colour is used for identification. [14]

This paper, Hand Gesture Interaction Using Colour Based Method for Tabletop Interfaces, presents a hand gesture recognition method by detecting colour markers mounted on hands. Here four types of colour markers; red, blue, yellow and Green are used. With this posture, different hand gestures are performed on a virtual keyboard. Then it provides best interaction between human and PC. This system is followed different steps to achieve this. First, the video received from the camera is cut in to different images. From these images the hand is to be detected and tracked. Then it can go for gesture recognition. Gesture recognition is divided in to two methods like Hand Detection Tracking and Hand Gesture Recognition. In Hand Detection tracking, hand positions are detected in order to find the colour markers. In order to detect the coloured markers, a colour segmentation method is implemented using the Hue Saturation Value (HSV). Once the markers have been found, these have to be followed in order to detect the movement it has done. Hand Gesture Recognition is divided in to two parts; Hand posture Recognition (HPR) and Hand Gesture Recognition (HGR). HPR is static and it is used for a single image. HGR is dynamic and it is followed over a sequence of images. This system is designed and implemented using a web camera and a video projector. These are mounted on support with an overlapping view of tabletop. [1]

III. EXISTING SYSTEMS

Different types of hand gesture recognition systems are available but have their own limitations. In this session, some of the existing systems are discussed which are using hand gestures for communication but in different ways.

A . Existing Systems

Some of the existing hand gesture recognition systems estimate hands 3 model based on image features such as contour or skin texture. But it is difficult to estimate the wrist rotation because the contour and the texture data do not have enough information to distinguish hands sides. Figure 1 shows the hand gesture recognition in rotation variance.[13]

Figure 1 Hand Gesture Recognition in Rotation State [13]

To solve this problem a new 3D hand posture estimation system that uses data of nail positions is used. Nail positions are an important factor to recognize hand sides using nail positions. It becomes possible to detect whether the camera is facing palm or dorsum. In addition, nail areas can be robustly extracted from a skin area by a simple image processing technique. [13] The system uses a database consisting of data sets of the hands contour, the nail positions and the finger joint angles. To estimate the hand posture, the system first extracts the hands contour and the nail positions from the captured image. Then search for a similar data set from the database. System then outputs the finger

joint angles of the searched data set. The following Figure 2 shows the fingertips recognition for hand gesture recognition.[3]

Figure 2 Fingertips Recognition for Hand Gestures

A common technique is to instrument the hand with a glove which is equipped with a number of sensors which provide information about hand position, orientation, and flex of the fingers. The first commercially available hand tracker, the Dataglove uses thin fiber optic cables running down the back of each hand, each with a small crack in it. Light is shone down the cable so when the fingers are bent light leaks out through the cracks. The Dataglove could measure each joint bend to an accuracy of 5 to 10 degrees, but not the sideways movement of the fingers (finger abduction). Figure 3 shows glove of glove based hand gesture recognition. [9]

Figure 3 Glove Based Hand Gesture Recognition [9]

Some of the systems are using the method of real time hand gesture recognition for dynamic applications like virtual games. First, the image taken is converted in to gray scale. Then this system finds and extracts the biggest contour area. From this gesture is identified and executed the corresponding command. Figure 4 shows the gesture contour in virtual environment. [14]

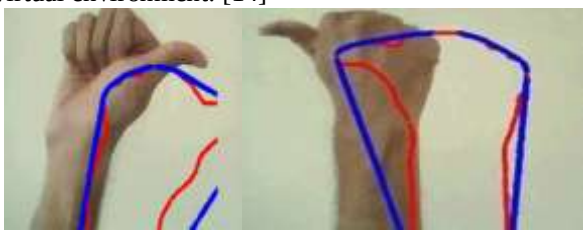

Figure 4 Gestures Contour in Virtual Environment [14]

Another method of using hand gestures is with the help of colour markers for table top interfaces. Colour markers are identified from the image taken from camera. Then the positions of these colour markers are identified. By using these positions gesture posture is identified. Figure 5 shows gesture recognition with colour markers. [1]

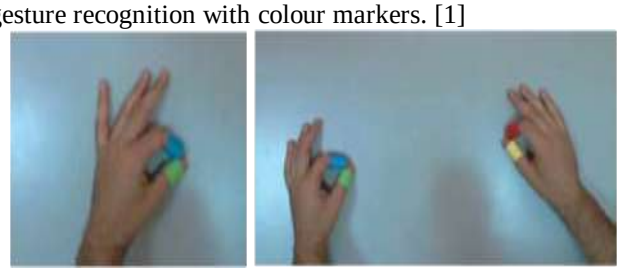

Figure 5 Gesture recognition with Color Markers [1]

B. Limitations of Existing Systems

The main difficulties involved in these existing systems are mainly comes from the gesture identification in complex background. The colour that matches the skin cannot be identified easily in case of vision based hand gesture recognition.

Fingers and nail positions are difficult to track as they occlude each other and are occluded by the hand itself. It is very difficult to distinguish each finger from others.

Many researchers have studied and used glove based devices to measure hand location. In general, glove-based devices can measure hand position with high accuracy and speed, but they aren't suitable for some applications because the cables connected to them restrict hand motion. Although instrumented gloves provide very accurate results they are expensive and encumbering. Some glove based systems without wires could not able to execute commands in a real time manner.

Computer vision techniques can also be used for gesture recognition overcoming some of these limitations. The resolution of video cameras is too low to both resolve the fingers easily and cover the field of view encompassed by broad hand motions. The 3 or 60 frames-per-second conventional video technology is insufficient to capture rapid hand motion.

IV. PROPOSED SYSTEM

In this paper, Design and Implementation of a Natural User Interface using hand gestures, presents a system which is capable doing applications with the help of natural hand gestures.

A. Proposed System

Natural user interface (NUI) is an emerging paradigm shift in man machine interaction of computer interfaces. It refers to a user interface that is effectively invisible, or becomes invisible with successive learned interactions, to its users. Here the natural interaction is used. That is, most computer interfaces use artificial control devices whose operation has to be learned. A Natural User Interface relies on a user being able to carry out relatively natural motions, movements or gestures that they quickly discover control the computer application or manipulate the on-screen content. This paper discusses a system which is created capable of tracking the hand gestures and then providing a feedback according to the recognized gesture.

Thus this paper brings the real world and virtual world together. The user will be able to use hand movements in the place of a regular mouse in order to control the operations usually carried out with a mouse. According to the different hand movements, the system will respond and carry out many operations that are available. Integration of reality and the computer system is apparent in this paper where a computer can be operated in a simplified manner using our hand.

Some common hand gestures have to be identified in order to develop the system. The hand gesture identified should be simple and easy to use. It should be the most frequent way of doing that operation. The most common and frequently used gestures have to be identified.

Vision based hand gesture recognition technique is used in this system. To achieve this, colour markers are mounted on hands. Here four colour markers like blue, green, yellow and red are used. These markers are used for hand gesture recognition. First, the image captured from the camera is converted to HSV colour space. One method to identify the gestures is done by identifying the colour markers and its positions In order to detect the positions of coloured markers; an algorithm is designed and implemented. By identifying the positions of the colour markers, the appropriate gestures are identified.

This system uses evolutionary process model for the system development. A process model is that views development as a series of hills and each representing as a separate loop of the spiral model. This model has been chosen, because, if any failures occur, they are accepted and the action is redone from where the failure has occurred and not from the start. This model incorporates efficiency of including the user feedback and also additional requirements during the development of the system. The Evolutionary models take the concept of evolution into engineering paradigm. Therefore evolutionary models are iterative. They are built in a manner that enables the software engineers to develop increasingly more complex versions of the software. This Incremental Model combines the elements of the Linear Sequential Model with the iterative philosophy of prototyping. When the incremental model is used the first increment is the core product. The subsequent iterations are the supporting functionalities or add on features that a consumer would like to see.

B. Advantages of the Proposed System

This system helps to interact in a real time manner. The significance of the input device is reduced. All most all the basic functionalities can be implemented by using hand gesture recognition method. The people who are dump and deaf can easily learn the gesture movement and also can do the proper functionalities of the system. It also aids portability to the system. The layman can also do the proper interaction with this type of system. So it is user-friendly. It has many advantages over the other systems which uses data glove method.

Data glove method restricts the hand movement in some directions but it is not. The vision based approach is used for the proposed system because it requires only camera rather than some sensor devices in data glove method. The major limitations of existing system based on skin texture recognition can be avoided here because it is using a colour based approach for the hand gesture recognition. This way allows the user to be more flexible when he/she performs a set of tasks. Also, this proposed used approach reduces errors due to bad detection and the brightness variation and also offers an economic and practical way of interaction.

Advantages in using symbolic gestures for interaction, including Natural Interaction, in which gestures are a natural form of interaction and easy to use, Terse and Powerful in which a single gesture can be used to specify both a command and its parameters and direct interaction in which hand act as input device eliminates the need for intermediate transducersHelpful Hints

V. SYSTEM DESIGN

This section gives the detailed design of the system. In order to design the system, the requirements for the system should be clearly identified.

A. System Requirements

Some specified hand gestures for certain actions must be learned in order to use the system which should be easy to learn. The main requirements for the system are two kinds of gestures - Static gestures and Interactive gestures.

- ***Static Gestures***

The following are the detailed activities to be performed as a part of the static gesture. The activities are like creating a Photo Frame with the hands takes a photo using the webcam, creating an Eye shape with the left hand invokes the Image Viewer to browse the images taken by the program and Optical Drive can be ejected by creating a Drive Shape using both hands. Based on theses gestures, the images can be browsed easily.

- ***Interactive Gestures***

The following are the detailed activities to be performed as a part of the interactive gesture. They are like moving the Index finger moves the mouse on screen. When the thumb is shown a double click is triggered. There is also a Drag and Drop feature to move files, text, windows around. There are Drawing an X closes the currently active application, drawing a C launches the user defined Web Browser, drawing a V launches the user defined Media Player, Zoom & Pan photos using Pinch gesture, minimizing the active window, launching the user specified location in the Explorer, switching Windows, navigating through Files or Folders using Mouse & Keyboard gestures, launch presentation and navigate through the slides and play the game like Angry Birds.

B. System Design

After identifying the requirements of the system, its design is done. This system is capable of tracking the hand gestures and then providing a feedback according to the recognized gesture. The overall design of the system is done as follows.

The system mainly consists of a camera and a processor. Identified natural gestures are recognized through analyzing the image frames from the web camera which is focused to the computer monitor. For that, video stream is received through the camera, which contains the colour markers. Then the received input video is converted to frame sequences and each frame is processed separately to get the position of the markers. The position of the markers is used to identify the appropriate gestures. After identifying the gesture, the corresponding operation or command execution is done. The system is designed in such a way that the users can easily interact with the system and manipulate user application programs with simple actions. This process is illustrated in the following Figure 6.

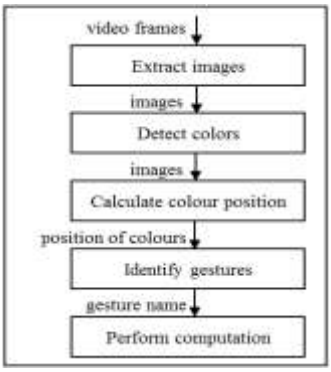

Figure 6 Natural User Interface System

The system architecture starts with the capturing of video frames from a webcam and processing it frame by frame Each frame is checked for the existence and locations of the four colour markers, ie, red, yellow, blue and green, making use of the built in functions provided by the OpenCV library. The position of the colour markers are tracked continuously and based on their movements and relative positions, various gestures are identified and corresponding actions are triggered.

VI. SYSTEM DESIGN

This section gives the detailed design of the system. In order to design the system, the requirements for the system should be clearly identified.

A. System Implementation

The entire work is done in different modules. One module starts with capturing frames, then go through extraction of colours like blue, green, yellow and red. This module ends with finding the position of these four colours. The other module is for calculating the distance between these four colours which is required for hand gestures. The next module is for the execution of commands related to the identified gestures. Major features to be provided in each module are decided in advance.

The following section gives a detailed study of the module breakup. The main modules of the system are the initial settings module, distance calculation of colours for gestures and command execution related to gestures.

B. Initial Settings Module

This is one main module which reads the user settings stored in a CSV file, captures the frames, locates the color markers, displays the captured frame and also invokes all the other sub modules.

This consists of placing colour markers directly over the hands fingertips. So, in order to detect the hand's positions, it's enough to detect the placed markers positions. This module includes the image capturing and the conversion to HSV colour space. This module reads the user settings stored in a CSV file, captures the frames, convert in to HSV colour space. It also includes the identification of position of colours locates the colour markers and displays the captured frame in a CSV file. In order to detect the colour markers, a colour segmentation method is implemented using the Hue Saturation Value (HSV).A big advantage of this colour space is that it is often more natural to think in hue, saturation and value. This colour space is less sensitive to shadow and uneven lighting. Picking colours and matching it is easier compared with RGB colour space. To segment the image in different colour images there are three different thresholds used, one in each component of HSV (hue, saturation and value).this module gives the input for the next module

C. Distance calculation of colours for gesture

This module calculates the distance between these four colours. One algorithm is designed and implemented in order to find the distance between the colours. This algorithm is based on calculating the coordinate values of the colours identified. These distances are used for identifying the gestures for the user. This module is interacting with the next module for its input.

D. Command execution related to gestures

The functionalities like taking snap shot, invoking the image viewer, interfacing mouse or keyboard, launching explorer, optical drive eject and checking interactive gestures are done here. Most of these sub modules contain sub-sub modules for extended functionalities and improved intuitiveness. Here some condition that is defined for identifying a particular gesture is checked and then corresponding event is executed if the matching is right. It invokes the modules for taking snap shot, invoking the image viewer, interfacing mouse or keyboard , launching explorer, optical drive eject and checking interactive gestures etc. Any number of functionality can be added as an extension to the system. The following figures show interfaces and gestures used for different applications

Figure 7 Gesture for Zoom Operation

Figure 7 shows one gesture, which is used for zoom operation. This gesture with the four mounted markers is capable of doing zoom operation which is doing by mouse device.

Figure 8 Gesture for Mouse Activation

Figure 8 shows the gesture for mouse activation. This gesture can be used to activate mouse for its operations. Figure 9 shows the gesture for toggling windows in a system. Since the system is a multitasking system, number of windows or tasks can be handled at the same time. This can easily possible with the help of this gesture.

Figure 9 Gesture for Toggling Windows

Figure 10 shows the gesture for close operation. This close operation is commonly used in all applications. This is one of the frequently used gesture.

Figure 10 Gesture for Close Operation

Figure 11 shows the gesture used for playing an interactive game called Angry Birds. This is one of the commonly used games in the recent days. This can be achieved with the help of a simple gesture. Most of the application based operations can be performed by using simple gestures in this system.

Figure 11 Gesture to Play "Angry Birds" Application

Figure 12 shows a gesture used to open an application like LibreOffice 3. This is also one commonly used application nowadays.

These are some of the screen shots that has included as part of the implementation. Any number of gestures can be developed in order to direct the system for more functionality.

VII. CONCLUSION

Natural user interface is the interface used to communicate with the system. It is designed and implemented based on the vision hand gesture recognition method. The requirements for the proposed system are arrived through the literature survey. The design of the system is done according to the arrived requirements. The appropriate hardware and software requirements are also done. The appropriate modules for its functionality are identified and implemented. The system is verified for its functionality based on the appropriate test cases. And its performance is analysed.

FUTURE ENHANCEMENT

This system can be enhanced with the support of a video projector. The projector and the camera are mounted on support with an overlapping view of the tabletop. In this setup, the camera can see the user hands during the interaction. The camera is overlooking the projected content on the tabletop and the hands as well. Users need no other devices for interaction; hence the system provides direct, natural and intuitive interactions between the user and the computer.

Another direction for future work is to improve the system performance under different light conditions and backgrounds. Future research will concentrate on investigating efficient matching and studying other robust and efficient methods about face and hand location in order to integrate the components of the system into a gesture interface for an anthropomorphic autonomous robot with an active vision system and into virtual environment applications.

REFERENCES

1. A. BellarbiI, S. Benbelkacem, N. Zenati-Henda, M. Belhocine, "Hand Gesture Interaction Using Color-Based Method For Tabletop Interfaces", IEEE 2011.
2. A. Downey, J. Elkner, and C. Meyers, "How to think like a computer scientist: Learning with python," Green Tea Press, Wellesley, Massachusetts,2008.
3. A. Gupta and K.A. Pati, "Finger Tips Detection and Gesture Recognition", Indian Institute of Technology Kanpur, 2009.
4. A. Sepehri, Y. Yacoob, and L. S. Davis, "Employing the hand as an interface device," Journal of Multimedia, 2006.
5. B. Ionescu and D. Coquin, "Dynamic hand gesture recognition using the skeleton of the hand," EURASIP Journal on Applied Signal Processing, 2005.
6. B. Zhang, R. Yun, and H. Qiu, "Hand gesture recognition in natural state based on rotation invariance and opencv realization", Digital Entertainment Research Center, Nanjing Normal University, 2010.
7. E. Tamaki, "A robust and accurate 3d hand posture estimation method for interactive systems". Graduate School of Interdisciplinary Information Studies, University of Tokyo, 2010.
8. G. Bradski and A. Kaehler, "Learning opencv: Computer vision with the opencv library,"O'Reilly Media, 2008.
9. J. V. Yoon and J. K. Min, "Enhancing hand gesture recognition using fuzzy clustering based mixture of experts model.", Yonsei University,134 Shinchon-dong.
10. M. Billinghurst, "Gesture-based interaction", Buxton, 2011.
11. M. Porta, "Vision-based user interfaces: methods and applications," International Journal of Human-Computer Studies, 2002.
12. "Open CV Reference Manual", Release 2.4.3, November 2012.
13. P. Garg and N. Aggarwal, "Vision based hand gesture recognition.", World Academy of Science, Engineering and Technology, 2009.
14. Siddharth S. Rautaray1, Anupam Agrawal, "Real Time Hand Gesture Recognition System for Dynamic Applications", International Journal of Ubicomp, Vol.3, No.1, January 2012.
15. Utarakash, "Tracking colored objects in opencv", 2010.