Mrs Reena R Ambekar*, Ms. Vishakha Jadhav*, Ms. Dhwani Doshi*, Mr. Ashwinkumar Gothi*, Mr. Dishant Jani* (*-WIEE&CT, Worli , Mumbai) / International Journal of Engineering Research and Applications (IJERA)
ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 2, Mar-Apr 2012, pp.1443-1447

HAND RECOGNISED SET TOP BOX

Mrs. Reena R Ambekar
A.P., EXTC DEPARTMENT
(WIEE&CT, WORLI, MUMBAI)

Ms. Vishakha Jadhav
B.E. Student EXTC (WIEE&CT)

Ms. Dhwani Doshi,
B.E. Student EXTC (WIEE&CT)

Mr. Ashwinkumar Gothi
B.E. Student EXTC (WIEE&CT)

Mr. Dishant Jani
B.E. Student EXTC (WIEE&CT)

ABSTRACT

Television control By Hand Gestures is a concept by which we can control television via hand movement. In this, different gestures like thumbs up, a clenched fist, Waving, Jazz Hand and also alphabets will be assigned specific tasks and stored in the system. A Webcam can be used to recognize the gesture performed by user and that gesture is compared with gesture stored in the system and if both the gestures match, then the specific task assigned to it will be performed. So we can control the system with different hand gestures. Tasks like ON/OFF, Volume control, opening an application, controlling media players, television settings like brightness, color etc , and changing channel and many more can be easily performed using this application.

Also almost all consumer electronic equipment today uses remote controls for user interfaces. However, the variety of physical shapes and functional commands that each remote control features also raises numerous problems, the difficulties in locating the required remote control, the confusion with the button layout, the replacement issue and so on. The control of living room environments containing televisions and set-top boxes the existing research has failed to provide a flexible solution for controlling such devices by hand gestures.

INTRODUCTION

Human–computer interaction has become an increasingly important part of our lives because of massive technological infusion into our lifestyles. Whether it is our living room, bedroom or office room, there could be a range of electronic equipment that needs commands to perform some valuable tasks. It could be the television set, the VCR or the set-top box waiting for our command to provide us with music or perhaps news and the command may reach them with a push of a button of a remote controller or a keyboard. Humans communicate Mainly by vision and sound, a man-machine interface would be intuitive if it made greater use of vision and audio recognition. Another advantage is that the user not only can communicate from a distance, but need have no physical contact with the computer. The visual system chosen was the recognition of hand gestures. The amount of computation required to process hand gestures is much greater than that of the mechanical devices People have long tried to replace these items using voice recognition or glove-based devices with mixed results Glove-based devices are tethered to the main processor with cables which restricts the user's natural ability to communicate. Many of those approaches have been implemented to focus on a single aspect of gestures, such as hand tracking, hand posture estimation or hand pose classification using uniquely colored gloves or markers on hands/fingers.

ARCHITECTURE

Fig 1. Architcture

(a) 3D DEPTH CAMERA

A new high resolution 3D camera used in which real-time depth measurements based on variable gain and gating techniques. The camera uses a combination of Phase measurement and Time-of-Flight (TOF) measurement techniques. The new 3D camera brings the advantage of a considerable high resolution when it comes to the data it has to provide. This helps to make the result of gesture understanding stable enough for reliable control. The camera does not require costly and/or exotic devices such as image intensifiers and is controlled simultaneously along with an illuminator. The gesture control of TVs and other electronic devices is accomplished by reconstructing a 3D model of the user's hands based on the data received from the 3D camera. The camera will capture the hand gestures which are then registered, normalized and feature-extracted for eventual classification to control the remote controller .This camera with the illuminator form the 3D depth image which will process by the controller or PC.

(b)ILLUMINATER

The captured hand gestures from a real-time video stream need to be processed before they can be interpreted by a computer. Foreground and Background segmentation is the most important part for successful gesture recognition and the importance of it cannot be overemphasized. The output of this stage is a cropped image of the person. Normally the area of interest is less than 1/3rd of the whole image and it saves a huge amount of time for the rest of the computationally intensive stages. It is important that the output of this stage is accurate because it directs the focus of the whole system to the cropped area generated by this stage. There are well established algorithms for this purpose which has been developed for outdoor video surveillance.

(c) CONTROLLER

Controller is an embedded system made up of an electronic board for the image sensor, a controller board, an illuminator board, and an image processing board. The controller and the image processing are synchronized to produce the 8-bit depth map from the image sensor output. The human hand is a complex articulated object consisting of many connected parts and joints which, in motion, can be characterized by approximately 27 degrees of freedom. This makes the recognition problem complex enough to make place for errors. The problem is alleviated if three dimensions of the hand gesture are acquired and measured correctly. The information contained in a 3D image might have enough power of separation such that the computer can clearly distinguish the context in which the gesture means a command.

The principle of detecting hand gestures as follows; the hand and fingers are captured by the 3D camera and a 3D image is formed. Image processing algorithms which detect the convex hull of the hand and the irregularities present when fingers are apart are used to detect one, two, and more fingers as they are used.

(d) IR BLASTER

The IR BLASTER used to emit the necessary remote control signals from a PC, two Windows Media Center USB infrared blasters and a modified driver is used. They provide access to the sending and receiving capabilities of the blasters through an API ie application program interface. By pointing the existing TV and receiver remotes to the receiving end of the infrared blasters, the infrared signals emitted by the physical remotes were recorded for all of the necessary commands. These infrared signals could then be retransmitted by accessing the infrared blasters from the gesture-processing code and playing back the appropriate recording based on the gesture that the user performed so this system allows for the control of practically any television set and set-top box on the market today, as well as devices such as AV receivers and Blu-ray players.

BLOCK DIAGRAM

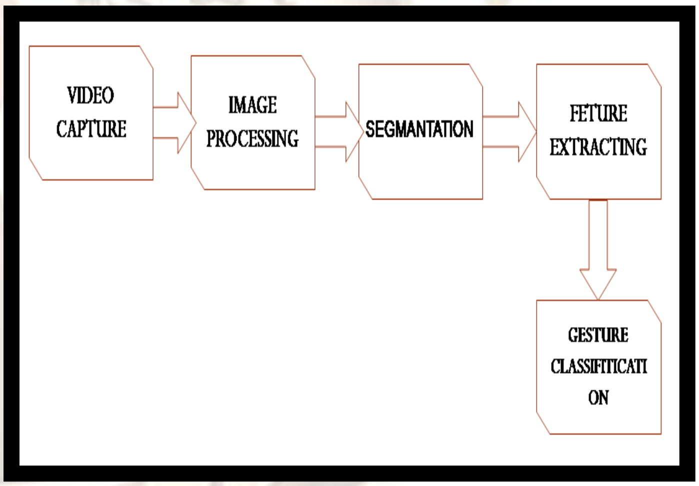

Fig 2.Block Diagram

(a)VIDEO CAPTURE

Vision based gesture method established the planar model of the hand gesture or restore the 3D model which is uses for following feature extraction and classification stages. such an approach would be ideal for real-time interaction between computer and human because it requires simpler input devices and less computational task due to using one single camera, which is vital for enhancing the speed of whole system. this will help to recognize three natural gesture and track two fingers in 3Ddimension.this system uniquely used a point light source working together with camera generating the shadow of hand on a table, which is uses to restore the 3D information.

By computing the 3D parameters acquired from projection of hand and it's shadow, this system identifies one dynamic and two static hand gesture fast and stably. However, the problem is that due to the lack of more details in the image of shadow, one of the gestures has a very low recognition rate. Moreover, the position of hand in front of camera is restricted at a certain position because the occlusion occurs when the hand is close to the table, which leads to the failure of acquisition of shadow. At last the calibration of angle between the source light and camera limits the system tremendously.

(b) IMAGE PROCESSING

When the original image of hand gesture are converted into image which can be identified by the computer, the generation and transformation procedure of image may be distributed and effected by various noise as a result the quality of image with distortion in various degrees is produced. the objective of image processing is to remove the noise and enhance the useful information in an image , and finally a completed and noise free segmented hand region will obtain. Image processing is a filtering process where the needed parts are retained and needless part is filtered. Because the imaging device will broaden the viewfinder range when reading image of hand gesture, a large number of useless pixels are incorporated together with hand gesture images. This adds the computational load since the useless information will be computed if they are not removed. Thus it is the first task to extract the useful information namely hand region from an input image. the difficulty is how to determine whether certain parts are to be retained and how to do it accurately .otherwise during the next stage the analysis and estimation will be seriously affected. The characteristic of human skin, noise processing and hand region segmentation are analyzed.

(c)SEGMENTATION

Skin segmentation is the cornerstone of many applications such as gesture recognition, face detection, and objectionable image filtering. The principal idea of most existing approaches is based on the assumption that skin color is quite different from colors of other objects and its distribution might form a cluster in some specific color-spaces. The captured hand gestures from a real-time video stream need to be processed before a computer can interpret them. It is extremely important that the captured image is registered as a hand gesture using skin segmentation after removing the background of the image. The skin segmentation techniques used in this research involves converting the image from RGB format to YCbCr format.

The threshold filter is applied to remove 'non-skin' components. The major advantage of this is that the influence of luminosity can be removed during the conversion process. Thus it makes the segmentation less dependent on the lightning condition, which has always been a critical obstacle for image recognition. The threshold values were obtained using our own data set. Then, the converted picture in YCbCr format is viewed in the 'imtool' of MATLAB so that every pixel and its associated values such as x, y coordinates and intensity can be determined accurately. A number of sample points that represent skin patches and non-skin patches are obtained .The property of the Y component implies that we are not filtering the luminance of the image, but the other two remaining components, Cb and Cr. For better visualization, the Cb and Cr values of both skin and non-skin patches are plotted in a graph to find the region that they are likely to fall in. . A similar approach is repeated on another image in fluorescent lighting conditions both the original image and the skin-segmented image are used to observe and verify the accuracy of the segmentation algorithm.

Images with low lightning condition are also tested there are always a number of noisy spots in the filtered images, regardless of the lightning condition. This distortion becomes more pronounced in low lighting conditions so, the skin-segmented image is noisy and distorted and is likely to result in incorrect recognition at the subsequent stages. These distortions, however, can be removed during the gesture normalization stage.

Fig 3.Segmentation & Normalization

Skin segmentation testing was mainly aimed at evaluating the performance of the skin segmentation and normalization modules of the control system. It concerned the shadow of the hand and the body did not have any effect on the filtering process. The remaining noise and unfilled pixels were removed by the normalization filter, which resulted in a smooth and clear region.

In the above figure in the first two images taken under incandescent light, the hand was segmented along with some parts of the guitar and the edges of the wardrobe, forming a fairly distorted image. However, after being passed through the normalization filter, the resultant images only consisted of the largest region found in filtered images, in this case effectively the hand region. This also implied that larger region of 'skin-like' objects might result in incorrect segmentation and should be carefully considered. The last two images taken under fluorescent light, on the other hand, showed significantly less noise than the first two images. This could be explained in terms of the difference in physical characteristics of the two light sources. In particular, incandescent light generated a yellowish glow which modified the look of objects that was captured. The resultant effect was that the object might have been recognized as a skin region as its color had been modified.

(d)FEATURE EXTRACTION

It is not too difficult to realize that effective real-time classification cannot be achieved using attempts such as template matching. Template matching itself is very much prone to error when a user cannot reproduce an exact hand gesture to a gesture that is already stored in the library. It also fails because of variance to scaling as the distance to the camera may produce a scaled version of the gesture. The gesture variations because of rotation, scaling and translation can be circumvented using a set of features that are invariant to these operations. Moment invariants offer a set of features that encapsulate these properties.

The moment invariants algorithm is most effective methods to extract descriptive features for object recognition applications. the algorithm derives a number of self-characteristic properties from a binary image of an object. These properties are invariant to rotation, scale and translation. the advantages of the moment invariants algorithm for gesture classification. For each specific gesture, moment invariants always give a specific set of values. These values can be used to classify the gesture from a sample set. The set of chosen gestures have a set of unique moments. Moment invariants are invariant to translation, scaling and rotation. Therefore the user can issue commands disregarding orientation of hand. The algorithm is susceptible to noise. Most of this noise, however, is filtered at the gesture normalization stage. The algorithm is moderately easy to implement and requires only an insignificant computational effort from the CPU.

(e)GESTURE CLASSIFICATION

The extracted data set from an image of a user hand gesture. However, this data set remains meaningless unless the program can interpret it into a preset command to control the electronic device. After the feature extraction stage, each group of the sample images that represent the same gesture produces a certain range of F1, F2, F3 and F4. These ranges are then used as preset values to classify a random input image. The nearest-neighbor classifier is more computationally intensive than the neural network. A neural network classifier proves itself more effective and more efficient. Neural networks have been applied to perform complex functions in numerous applications, including pattern recognition, classification, and identification and so on. Once implemented, they can compute the output significantly quicker than the classifier. Neural networks also encompass the ability to learn and predict over the time. This property enables the system to be viewed more as a human-like entity that can actually 'understand' the user, which is also one of the major objectives of our research.

The system is designed to capture one image frame (static image) every second and is then segmented for skin region detection and other pre-processing before the invariant moments are calculated. These invariant moments will be the input to the neural network for classification and the subsequent action using the remote control and the feedback system there are only three layers because of the limited number of hand gestures to be classified. More networks that are complex could be possibly designed and implemented, but it is neither practical nor necessary for our research.

For better visualization, the network can be use in where W represents the weighting function in which each input is weighted with an appropriate w, and b represents the bias coefficient and it is set to 1 in this design.

APPLICATIONS

(a) Virtual Reality
(b) Robotics and Tele presence
(c) Desktop and Tablet PC Applications
(d) Games
(e) Sign Language

RESULT AND CONCLUSION

The importance of gesture recognition lies in building efficient human–machine interaction. Its applications range from sign language recognition through medical rehabilitation to virtual reality. Given the amount of literature on the problem of gesture recognition and the promising recognition rates

reported, one would be led to believe that the problem is nearly solved. Sadly this is not so. A main problem hampering most approaches is that they rely on several underlying assumptions that may be suitable in a controlled lab setting but do not generalize to arbitrary settings.
Several common assumptions include assuming high contrast stationary backgrounds and ambient lighting conditions. In addition, recognition results presented in the literature are based on each author's own collection of data, making comparisons of approaches impossible and also raising suspicion on the general applicability. To ameliorate these problems there is a need for the establishment of a standard database for the evaluation and comparison of techniques.

Gesture-based technology has gained significant academic and commercial interest lately with the goal of allowing users to control computers and electronic devices with hand gestures. We are presenting an intelligent gesture interface for reliably commanding home televisions and set up boxes through a user-defined gesture language. A camera will be used to obtain accurate data from the environment, and this data can be processed for gestures.

In summary, a review of vision-based hand gesture recognition methods has been presented. Considering the relative infancy of research related to vision-based gesture recognition, remarkable progress has been made. To continue this momentum, it is clear that further research in the areas of feature extraction, classification methods and gesture representation are required, to realize the ultimate goal of humans interfacing with machines on their own natural terms.

FUTURE SCOPE

Future research will focus on adding support for gesture sequences, thereby further increasing the flexibility of the gesture language that can be defined. In addition to emitting infrared commands, plans exist to explore the control of any PC application by adding the ability to invoke keyboard, mouse and joystick actions on the PC. Finally, it is worth exploring the development of a custom user interface that is designed for gestures and can be superimposed on top of a live TV image by using a TV tuner connected to a PC. This way, the infrared blasters are eliminated and an even more optimal gesture control experience can be provided.

BIBLIOGRAPHY

- B. A. Myers, "A Brief History of Human Computer Interaction Technology,"in ACM Interactions, vol. 5, no. 2. New York, NY, USA:ACM, March 1998, pp. 44–54."Telecommunications Policy Review," Porter/Novelli Survey, July 1994.
- Q. Cai and J. Aggarwal, "Tracking Human Motion Using MultipleCameras," in Proc. of 13th Int. Conf. on Pattern Recognition, vol. 3.IEEE Computer Society, August 1996, pp. 68–72
- (2010) Xbox.com Kinect. Microsoft Corp. [Accessed: April2011].[Online]. Available: http://www.xbox.com/en-US/kinect
- P. Grimshaw. (2010, February) MCE Remote Replacement Driver FAQ. [Accessed: April 2011]. [Online].Available:http://www.byremote.com.au/Hip/mce remote faq.htm164
- An Intelligent Gesture Interface for Controlling TVSets and Set-Top BoxesDan Ionescu, Bogdan Ionescu, Cristian Gadea, Shahidul IslamUniversity of Ottawa, Mgestyk TechnologiesOttawa, Ontario, CanadaEmail: fdan, bogdan, cgadea, sislamg@ncct.uottawa.ca May 2011
- Konstantinos G. Derpanis. "A Review of Vision-Based HandGestures" (2004).
- Sushmita Mitra, and Tinku Acharya, "Gesture Recognition:A Survey", IEEE Transactions on Systems, Man andCybernetics–Part C: Applications and Reviews, 37(3)(2007).
- Davis & Shah, 1994] Visual gesture recognition. In proceedings IEEE Visual Image Signal Process, 1994, vol.141, No.2, pages 101-106.
- A Gesture Recognition Interface for A Set-Top Box Natesh Srinivasan Pace Micro Technologies India.
- M. Kölsch. Vision Based Hand Gesture Interface for Wearable Computing and Virtual Environments. Ph.D. Dissertation, August 2004.