GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

Gesture and Voice Based Real Time Control System

Priyanka Shende[1] and Prof. Archana Dehankar[*2]
[1]Student (ME-WCC) PCOE, [2]Prof. (ME-WCC) PCOE
Archana_dhenkar@rediffmail.com

ABSTRACT

Human Computer Interaction keeps moving toward interfaces which are more natural and intuitive to use, in comparison to traditional keyboard and mouse. Hand gestures are an important modality for human computer interaction (HCI). Compared to many existing interfaces, hand gestures have the advantages of being easy to use, natural, and intuitive. Speech and gesture are two types of multimodal inputs that can be used to facilitate more natural human machine interaction in applications for which the traditional keyboard and mouse input mechanisms are in appropriate, however the possibility of their concurrent use raises the issue of how best to fuse the two inputs.This new real time system combining the gesture and voice recognition in one system, which allows the users to interact with computers through hand postures, being the system adaptable to different light conditions and backgrounds and control the Operating System. Its efficiency makes it suitable for real-time application

Keywords: Hand Gesture, Voice Recognition, SAPI, Karhunen-Loeve (K-L) Transform

I. INTRODUCTION

Human–Computer Interaction (HCI) is the study, planning, and design of the interaction between user and computers. Human Computer Interaction keeps moving toward interfaces which are more natural and intuitive to use, in comparison to traditional keyboard and mouse. Hand gestures are an important modality for human computer interaction (HCI). Compared to many existing interfaces, hand gestures have the advantages of being easy to use, natural, and intuitive. Gesture recognition is interface with computers using gestures of the human body, typically hand movements. Body language is an important way of communication among humans, adding emphasis to voice messages or even being a complete message by itself. Thus, automatic posture recognition systems could be used for improving human machine interaction. This kind of human-machine interfaces would allow a human user to control remotely through hand postures a wide variety of devices.

Thus here we are combining the Voice and Gesture recognition in one system. Its efficiency makes it suitable for real-time application

Aim is the proposal of a real time control system for its application within visual interaction environments through gesture and voice recognition. The basic approach is able to deal with a large number of hand shapes against different backgrounds and lighting conditions, and a recognition process that identifies the hand posture from the temporal sequence of segmented hands & implementing a system that can interpret a user's gestures in real time with Speech Application Programming Interface (SAPI) to control windows media player. The Objective is to study the Human Computer Interaction. (HCI), implementing the combination of Gesture and voice recognition for HCI in one system.

This new real time system combining the gesture and voice recognition in one application that is windows media player which allows the users to control Windos Media Player that is Open, Pause, Play, Stop the song.

In this project the fusion of hand and speech recognition are used to control windows media player.

II. LITERATURE REVIEW

1. Used fast algorithm for vision-based hand gesture recognition

They proposed a fast and simple algorithm for a hand gesture recognition problem. Given observed images of the hand, the algorithm segments the hand region, and then makes an inference on the activity of the fingers involved in

the gesture. They have demonstrated the effectiveness of this computationally efficient algorithm on real images that acquired. Based on their motivating robot control application, they have only considered a limited number of gestures.

2. Hand gesture system based on evolutionary search.

They consider a vision-based system that can interpret a user's gestures in real time to manipulate windows and objects within a medical data visualization environment. A probabilistic neighborhood search algorithm is employed to automatically select a small number of visual features, and to tune a fuzzy c-means classification algorithm. Test results on four interface tasks showed that the use of these simple features with the supervised FCM yielded successful performance rates of 95 to 100%.

3. Advanced recognition techniques

A fast processing process and a robust matching carried out through an approach; a visual memory system and resolution of non-rigid distortions have been presented for hand posture detection and recognition problem. Different light conditions, backgrounds and human users have been tested in order to evaluate system's performance. The recognition rates show that the system is robust against similar postures.

4. Hand Gesture Interface for Browsing Medical Images

A vision-based system that can interpret user's gestures in real-time to manipulate windows and objects within a medical data visualization environment is presented. The system is user independent due to the fact that the gamut of colors of the users hand or glove is built at the start of each session. Hand segmentation and tracking uses a new adaptive color-motion fusion function. Dynamic navigation gestures along with zoom, rotate, and system sleep gestures are recognized.

III. SPEECH APPLICATION PROGRAMMING INTERFACE

SAPI is an API developed by Microsoft to allow the use of speech recognition and speech synthesis within Windows applications. To date, a number of versions of the API have been released, which have shipped either as part of a Speech SDK, or as part of the Windows OS itself. Applications that use SAPI include Microsoft Office, Microsoft Agent and Microsoft Speech Server.

In general all versions of the API have been designed such that a software developer can write an application to perform speech recognition and synthesis by using a standard set of interfaces, accessible from a variety of programming languages. In addition, it is possible for a 3rd-party company to produce their own Speech Recognition and Text-To-Speech engines or adapt existing engines to work with SAPI. In principle, as long as these engines conform to the defined interfaces they can be used instead of the Microsoft-supplied engines.

In this project we are using Speech SDK 5.1 version.

IV. CANNY EDGE DETECTION

This is one of the best edge detection techniques but little complex than other edge detection techniques. The major advantage of this technique is its performance. In case of other edge detection techniques only one threshold is used, in which all values below the threshold were set to 0. Thus, we must be very careful while selecting the threshold. Selecting the threshold too low may result in some false edges which are also known as false positives. Whereas if the threshold selected is too high, some valid edge points might be lost, this is also known as false negatives. But canny edge detection technique uses two thresholds: a lower threshold, TL and a higher threshold, TH thus eliminating problem of false positive and false negative. Steps involved in this type of detection are:

- The input image is smoothened with a Gaussian filter after which the Gradient magnitude and angle images are computed.

- Non-maxima suppression is applied to the gradient magnitude image.

•And finally detection and linking of the edges is done using double thresholding and connectivity analysis.

V. KARHUNEN-LOEVE (K-L) TRANSFORM

K-L Transform is used to translate and rotate the axes and new coordinate is established according to the variance of the data. The K-L transformation is also known as the principal component transformation, the eigenvector transformation or the Hoteling transformation. The advantages are that it eliminates the correlated data, reduces dimension keeping average square error minimum and gives good cluster characteristics. K-L Transform gives very good energy compression. It establishes a new co-ordinate system whose origin will be at the center of the object and the axis of the new co-ordinate system will be parallel to the directions of the Eigen vectors. It is often used to remove random noise. The steps involved in the process are:

• Firstly we consider an input data matrix say X and then we find out the mean vector say M

$$M = E\{X\}$$

• The next step is finding out the covariance matrix C of X. Mathematically, covariance is given by

$$C = E\{\_X - M\_\_X - M\_\$.$$

• The eigen values and eigen vectors are found out from such that the eigenvalues are arranged in the

Descending order and corresponding eigen vectors are obtained.

• Matrix A is obtained in such a manner that the first row represents eigen vector corresponding to maximum Eigen value and so on.

• K-L Transform is given by:

$KLT = A * \_X - M\_$ where A is the matrix consisting of Eigen vectors arranged in rows such that they are Arranged in decreasing order of Eigen value.

VI. CONCLUSION

In this project, the steps that we have used for recognizing different hand gestures are skin filtering, edge detection, K-L transform and finally a proper classifier, where we have used
Euclidean distance based classifier. In which we have performed command on windows media player like Open, Pause, Play and Exit song by using the fuse of gesture and speech.

VII. REFERENCES

[1] *Yuvraj V. Parkale,"Gesture Based Operating System Control",IEEE-2012*

[2] *T.N.Shanmugam and Priya Rajendran, "An Enhanced Content- Based Video Retrieval System Based On Query Clip",International Journal of Research and Reviews in Applied Sciences, Volume 1, Issue 3, December 2009.*

[3] *Juan Wachs, Helman Stern, Yael Edan, Michael Gillam "Real-time hand gesture system based on evolutionary search" , 2010*

[4] *Mr. Chetan A. Burande, Prof. Raju M. Tugnayat, Prof. Dr. Nitin K. Choudhary,Vivek Tyagi "Advanced recognition techniques for human computer interaction "in IEEE International Conf. on Automatic Face and Gesture Recognition-2010.*

[5] *Juan Wachs, Helman Stern, Yael Edan,Michael Gillam "Real- Time Hand Gesture Interface for Browsing Medical Images",IEEE 2007*

[6] *Antonis A. Argyros and Manolis I.A. Lourakis "Vision-Based Interpretation of Hand Gestures for Remote Control of a Computer Mouse",IEEE 2006.*

[7] *Asanterabi Malima, Erol Özgür, and Müjdat Cetin, "A fast algorithm for vision-based hand gesture recognition for robot control"*