Charu Jain et al. International Journal of Recent Research Aspects ISSN: 2349-7688, Vol. 3, Issue 2, June 2016, pp. 21-24

A Novel Technique for Gesture Recognition System

Charu Jain

Amity School of Engineering and Technology, Amity University Haryana

***Abstract*— Gesture is a way of Human Computer Interaction. It is a technique that is mainly carried out to make impaired community communicate with normal humans. It's main aim is to convey human gestures to computing device. They are efficient for natural and intuitive human-computer interaction. To achieve this goal, computers should be able to visually recognize hand gestures from video input. We propose a new architecture to solve the problem of real-time gesture recognition. The fundamental idea is to use a combination of hand gestures to control various functionalities. This will include two parts, first is to track the movement of the hand and secondly is to track the orientation of the hand. This paper presents a robust and efficient technique for gesture recognition. The OpenCV library provides us a greatly interesting demonstration for object detection. Furthermore, it provides programs (or functions) that they used to train classifiers for their gesture detection system, called HaarTraining, so that we can create our own object classifiers using these functions. Our working environment is Visual Studio on Windows 7. The objective of this project is to develop an application for recognition of hand gestures with reasonable accuracy and thus creating windows media player compatible remote control.**

***Index Terms*— Gesture detection, User interface, Haar-Like features.**

I. Introduction

Computer is used by many people either at their work or in their spare-time. Special input and output devices have been designed over the years with the purpose of easing the communication between computers and humans, the two most known are the keyboard and mouse [1]. Every new device can be seen as an attempt to make the computer more intelligent and making humans able to perform more complicated communication with the computer. This has been achieved by creating successful human computer interface. The computer programmers have been incredibly successful in easing the communication between computers and human. The idea is to make computers understand human language and develop a user friendly human computer interfaces (HCI). Making a computer understand human gestures is one of the first some step towards it. Gestures are the non-verbally exchanged information. [2]A person can perform innumerable gestures at a time. Since human gestures are perceived through vision, it is a subject of great interest for computer vision researchers. The paper aims to determine human gestures by creating an HCI. Coding of these gestures into machine language demands a complex programming algorithm. This complex process is called gesture recognition system. Gesture recognition is the process of recognizing and interpreting a stream continuous sequential gesture from the given set of input data.

The aim of the research presented in this paper is to determine an application of a gesture recognition system based on HAAR-like feature. This paper is structured as follows. Section 2 provides the overview of Gestures and GRS. Section 3 describes the research methodology. Section 4 explains framework of our proposed gesture recognition system. Section 5 describes the discussion of these results in terms of benefits and limitations. Finally, Section 6 concludes the paper.

II. Literature Survey

A. Gesture recognition system

As per the context of the paper, gesture is defined as an expressive movement of body parts which has a particular message, to be communicated precisely between a sender and a receiver. It is hard to settle on a specific useful definition of gestures due to its wide variety of applications and a statement can only specify a particular domain of gestures. Many researchers had tried to define gestures but their actual meaning is still arbitrary. Bobick and Wilson [3] have defined gestures as the motion of the body that is intended to communicate with other agents. For a successful communication, a sender and a receiver must have the same set of information for a particular gesture. A gesture is scientifically categorized into two distinctive categories: dynamic and static [1]. A dynamic gesture is intended to change over a period of time whereas a static gesture is observed at the spurt of time.

B. Feature Invariant Approach

When choosing the object to train, it's often encounter at a different angle, or a different lighting condition. In this method, an invariant characteristic can be found, even when the angle and lighting conditions change. This paper focuses on feature invariants approaches. When the environment changes, the image character will change along with the environment, it is harder to detect the image. Especially, when posture or shooting angle changes, it will affect image detection seriously. Therefore, the algorithms need to search for particular characters which are the fixed structure feature especially when posture, visual angle or shooting angle varies, and use those features to locate the human hand. This method

is known as feature invariant approaches. To implement this approach, the image feature is observed, which does not change along with the environment. Different lighting condition may interfere with the skin color, reducing the accuracy of the detection process. However, if the studied image sample is big enough, it can locate the skin color inside a smaller color space and will be able to make the image detection simple. This report is based on Haar-like feature, which are invariant to scaling.

III. METHODOLOGY

A. Haar-like Feature

The process of image object detection deals with determining whether an object of interest is present in an image/ video frame or not. An image object detection system receives an input image/video frame, which will subsequently search to find possible objects of interest. Haar-Like feature helps in object detection, thus minimizes the overall computation time while achieving high detection accuracy. Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector. Historically, working with only image intensities (i.e., the RGB pixel values at each and every pixel of image) made the task of feature calculation computationally expensive. A publication by Papageorgiou et al. [4] discussed working with an alternate feature set based on Haar wavelets instead of the usual image intensities. [5] Viola and Jones adapted the idea of using Haar wavelets and developed the so-called Haar-like features. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image.

In our approach, the concept of "Integral Image" is used to compute a rich set of Haar-like features, which can significantly reduce the image processing time. The training algorithm of the Viola-Jones approach takes a set of "positive" samples, which contain the objects of interest (in our case: hand postures) and a set of "negative" samples, i.e. images do not contain the objects of interest. During the training process, distinctive Haar-like features are selected at each stage to identify the images containing the object of interest. When the trained classifier misses an object or detects a false object, adjustments can be made easily by adding more Haar-like features so that the mistakenly classified samples can be corrected. [5]The Viola and Jones algorithm is approximately 15 times faster than any previous approaches while achieving equivalent accuracy as the best published results. The major reason to use Haar-like features is that it can encode ad-hoc domain knowledge which is difficult to learn using a finite quantity of training data. Haar-like features are effective to catch the characters represented by the difference between the dark and bright areas within an image. Another b reason is that a Haar-like feature-based system can operate much faster than a pixel based system.

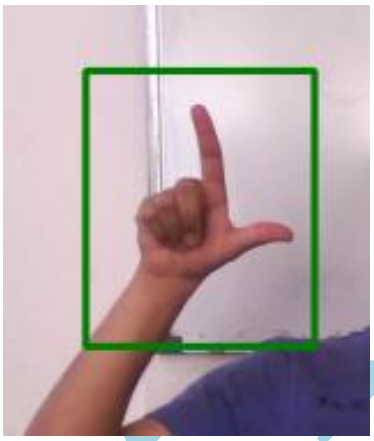

Fig.1 Gesture recognized for lock command

B. AForge.NET framework

There are many approaches for motion detection in a continuous video stream. [12]One of the most common approaches is to compare the current frame with the previous one. The Motion Detection application which is followed in the research is based on the AForge.NET framework, which stores all the filters and image processing routines used in this application. AForge.NET Framework is a C# framework designed for developers and researchers in the fields of Computer Vision and Artificial Intelligence. This framework has come in use for various applications such as digital image processing, genetic algorithms, neural networks, machine learning, robotics, etc.

Fig.2 Area under motion marked by red color

IV. SYSTEM OVERVIEW

Analysis through vision have been based on the way human beings understand or perceive information about their surroundings, still after years of research it is probably the most difficult task to implement in a satisfactory way. Many

different approaches have been tested .Traditional method is to create a 3-dimensional model of the human hand. This human-like model is matched to images of the hand by one or more cameras, and thus other parameters corresponding to palm orientation, joint angles and brightness are estimated. These parameters are observed and the results are then used to perform gesture classification.

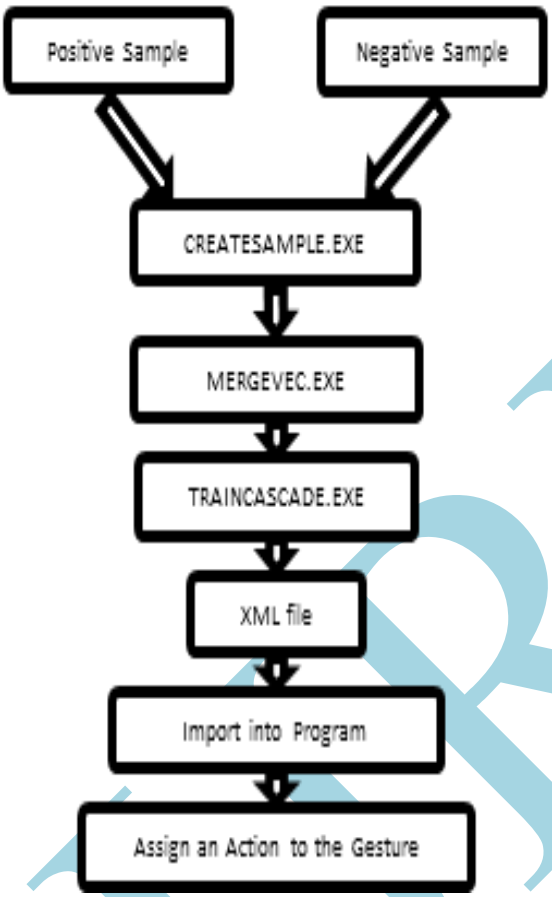

Fig. 3 Overview flow chart for the whole process

In our modern approach, we combine the OpenCV provided functions, application's basic operation describes as follows: data Preparation, image segmentation, create sample, training sample and finally use strong cascade. In different stage, the requirement is different. In the 'create sample' and 'training sample' stages, this project will apply the functions OpenCV provided to achieve the goal. In the early data preparation stage, it's required to implement the auxiliary program to detect using C# as the programming language. Previously, it was inconvenient to interact with the system during presentation and seminars. But this approach can effectively catch the appearance properties of the hand postures. To make the system robust against over cluttered backgrounds, noise removal and background subtraction are applied in this application

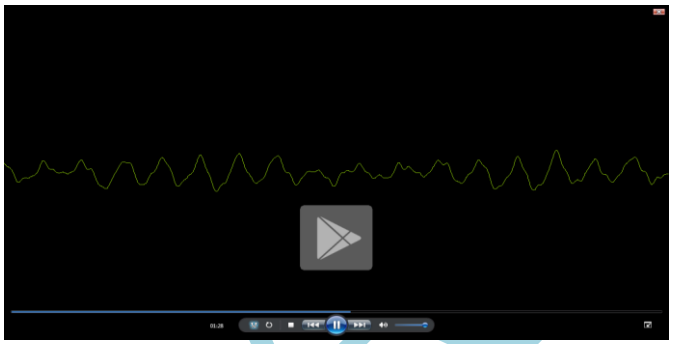

Fig. 4(a)

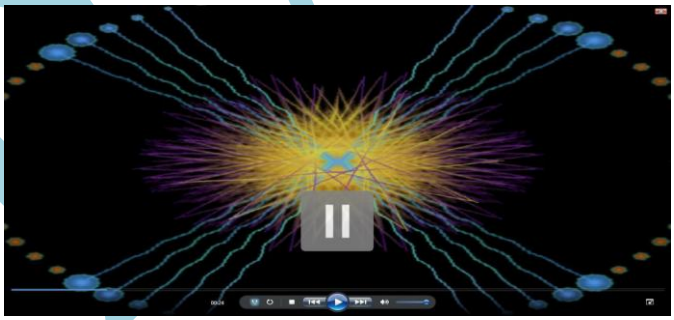

Fig. 4(a) & (b) Play/Pause button in wmp recognized

V. Discussion

A. *Benefits*

There are several advantages of gesture recognition system using OpenCV Haar object detection. The main advantage of using this technique is its speed. It detects gesture at 15 frames per second for 384*288 pixel images and the results obtained are 90% accurate with 10^{-6} false positive rate (FAR). Major issue in gesture detection in any system is background subtraction: refers to the complex background where there is other objects in the scene with the hand objects and these objects might contain skin like colour which would produce misclassification problem and Variation of illumination conditions: where any change in the lighting condition affects badly on the extracted hand skin region. Both issues are overcome in this gesture recognition system.

B. *Limitations*

The proposed method is susceptible to errors .System limitations restrict the applications such as; gestures are made with the right hand only, the arm must be vertical, the palm is facing the camera. In order to obtain accurate results the number of positive samples to be collected is very high.

Another issue involves the training time for each cascade in normal systems can range from 2 to 4 weeks.

VI. Conclusion

In this paper, we aimed to implement gesture recognition system. This was achieved through the design and development of an application Zauber. The windows based application is able to recognize gestures and act like a remote to windows media player at a high rate. Using Haar-like feature for gesture detection is a faster method to achieve results, when compared to traditional methods and provided correct results all of the time, a better result than humans performing the task alone.

For the future work, there are many possible improvements that can extend this work. First of all; more diversified hand samples from different people can be used in the training process so that the classifiers will be more user independent. The second improvement could be context-awareness for the gesture recognition system. By integrating this rich command set with other communication modalities such as speech recognition and haptic feedback, the human-computer interaction experience can be enriched greatly and be much more interesting. The system developed in this work can be extended into many other research topics in the field of computer vision. This project could trigger more investigations to make computers see and think better.

References

[1] Lalit Gupta and Suwei Ma "Gesture-Based Interaction and Communication: Automated Classification of Hand Gesture Contours", IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 31, no. 1, February 2001.

[2] Pravin R Futane, Dr. R. V. Dharaskar and Dr. V. M. Thakare. Cursor Controlling of Computer System through Gestures. International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 10, December 2013.

[3] Shigeo Abe, "Support Vector Machines for Pattern Classification, second edition", Kobe University, Graduate School of Engineering, 2nd edition, Springer-Verlag London Limited 2005, 2010.

[4] Papageorgiou, Oren and Poggio, "A general framework for object detection", International Conference on Computer Vision, 1998.

[5] Viola and Jones, "Rapid object detection using a boosted cascade of simple features", Computer Vision and Pattern Recognition, 2001.

[6] M. K¨olsch and M. Turk, "Robust hand detection," in Proc. 6th IEEE Conference on Automatic Face and Gesture Recognition, 2004.

[7] Alexander Kuranov, Rainer Lienhart, and Vadim Pisarevsky. An Empirical Analysis of Boosting Algorithms for Rapid Objects with an Extended Set of Haar-like Features.

[8] Rainer Lienhart and Jochen Maydt. An Extended Set of Haar-like Features for Rapid Object Detection. IEEE ICIP 2002, Sep. 2002.

[9] Z.NAN Face Detection Based on AdaBoost Beijing University 2005.

[10] Muharremmercimek, Kayhangulez and Tarikvelimumcu, Real object recognition using moment invariants, Yildiz Technical University, Electrical Electronics Faculty, ElectricalEngineering Department,34349 Besiktas-Istanbul,Turkey MS received 3 April2005,revised-2005.

[11] "Tutorial for haarclassifier",http://coding-robin.de/2013/07/22/train-your-ownopencv-haar-classifier.html.

[12] Aforge.NETcode,http://www.codeproject.com/Articles/16859/AForge-NET-opensource-framework.

