Different Approaches Of Gesture Recognition Techniques

Reshma Kalepatil¹, Prof. V. D. Shinde²

¹P.G Student, Department of Electronics and Telecommunication, RMDSSOE Pune, India
²Assistant Professor, Department of Electronics and Telecommunication, RMDSSOE Pune, India

Abstract—Hand gesture recognition has become one of the key techniques of HCI (Human Computer Interaction). Gesture recognition pertains to recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. Hand gestures provide a separate complementary modality to speech for expressing one’s ideas. Hand gesture is a method of non-verbal communication. Hand gesture recognition has greater importance in designing an efficient human computer interaction system. The approaches for hand gesture recognition, such as vision-based, glove-based and depth-based, are contrasted briefly in this paper. Currently, the new finger identification and hand gesture recognition technique with Kinect depth data is the most popular research direction. This paper is intended to review different approaches along with a brief survey of various methods.

Index Terms- Glove based gesture recognition, Hand gesture recognition, Kinect, vision based gesture recognition.

I. INTRODUCTION

Hand gesture recognition has become one of the key techniques of human-computer interaction (HCI). It provides an intelligent, natural, and convenient way of human computer interaction. Many researchers are devoted in this field. There are several techniques introduced for the new Era of human-computer interaction such as face detection, speech recognition and motion gesture. As an example in mixed reality (MR), it opens a new direction for human-computer interaction which combined with computer vision techniques and it is possible to build an advanced input devices. Hand gesture provides a natural and intuitive communication mode for human dialog. Hand gestures recognition also has the least limitation for user in human computer interaction (HCI). Gesture-based control translates gestures performed by human subjects into controlling commands as the input of terminal devices, which complete the interaction approaches by providing acoustic, visual, or other feedback to human subjects. Many previous researchers investigated various systems which could be controlled by hand gestures, such as media players, remote controllers, robots and virtual objects.

II. LITERATURE SURVEY

Generally speaking, two methods for hand gesture recognition are accepted by us. One is static hand gesture recognition and another is dynamic hand gesture recognition. Static hand gesture recognition method could only recognize the preinstall hand gesture. But we could get a clear meaning by the gesture act in the method of dynamic hand gestures recognition. The second method is more practical than the first one, even though it has more difficulties. Researchers have developed a variety of ways to achieve the gesture recognition so far. Vision-based, glove-based and depth-based are widely used in hand gesture recognition [1].

Laura Dipietro and Angelo M. Sabatini have introduced sensorized gloves to the nonspecialist readers interested in selecting one of these devices for their particular application. Further they gives pertinent information on such devices, including measurement performance, is scattered across the engineering and scientific literature and, even when located, can be inaccessible to the nonspecialist. This makes it difficult for a novice to determine whether and how well a particular glove suits a particular application. A thorough study describe how gloves were applied for different uses, can then help this matching process, at the same time highlighting practical issues that may arise during it [2].

The computer vision devices can be implemented and upgrade to the new input devices in the future. It gives the input command to the computer rather than just a function of taking photo or record video. We can do more implementation to transform the computer vision devices to become an input command device to reach the function as keyboard or mouse. One of the ways to give signal to computer vision devices is by using hand gesture. More specifically hand gesture is used as the signal or input modality to the computer. Certain signal can be recognized by computer as an input of what computer should do [3].

Articulated hand-tracking systems have been widely used in virtual reality but are rarely deployed in consumer applications due to their price and complexity. Robert Y. Wang Jovan Popovic propose an easy-to-use and inexpensive system that facilitates 3-D articulated user-input using the
hands. This approach uses a single camera to track a hand wearing an ordinary cloth glove that is imprinted with a custom pattern. The pattern is designed to simplify the pose estimation problem, allowing us to employ a nearest-neighbor approach to track hands at interactive rates. They describe several proof-of-concept applications enabled by this system that hope will provide a foundation for new interactions in modeling, animation control and augmented reality [4].

III. HAND GESTURE RECOGNITION APPROACHES

Present hand gesture recognition approaches can be classified into various categories.

A. Data glove based approach:
The development of the most popular devices for hand movement acquisition, glove-based systems, started about 30 years ago and continues to engage a growing number of researchers. The original gesture recognition researchers were focused on doing a dedicated hardware device for input, such as the method of data glove. Data Glove, based approach uses a glove-type device which could detect hand position, movement and finger bending. The original gesture recognition researchers were focused on doing a dedicated hardware device for input. Such as the method we’ve mentioned before, data glove. The data glove is applied to collect the data of joints in order to extract the gesture, and we usually use the neural network to analyze these data. Then, the meaning of the gesture will show up. This method own the advantages of less input data, high speed, and it can get 3D information about hands or the information about the fingers movement directly. So it can also identify a variety of gestures, do a real time identification [2].

We can also define a glove-based system as a system composed of an array of sensors, electronics for data acquisition/processing and power supply, and a support for the sensors that can be worn on the user’s hand. Typically, it is a cloth glove made of Lycra where sensors are sewn. As worn by the user, it records data related to his/her hand configuration. These approaches can easily provide exact coordinates of palm and finger’s location and orientation, and hand configurations. The first glove prototypes included the 1) Sayre Glove 2) The Massachusetts Institute of Technology (MIT) - LED glove and 3) The Digital Entry Data Glove. The Data Glove-like systems also include the commercial Space Glove, Cyber Glove, Human glove, 5DT Data Glove, TCAS Glove, and the more recent Strin Glove and Didji glove as well as prototypes developed by research laboratories around the globe, such as the TUB-Sensor glove. Although these above gloves differed from the original Data Glove in terms of sensor technologies, locations, and mounting, they all shared three basic design concepts with it: they measured finger joint bending, used a cloth for supporting sensors, and were usually meant to be general-purpose devices.

The original version of the Data Glove used thin flexible plastic tubes sewn on a cloth and light sources and detectors to record joint angles. Fig. 1 shows the example of glove based gesture input. A new fiber optics version was developed and commercialized in 1987 by Visual Programming Language Research, Inc. It came equipped with 5 to 15 sensors.

The main advantage of this approach is high accuracy and fast reaction speed but this approach can be quite expensive. Despite their widespread use, they suffered from several drawbacks. Major limitations originated from the cloth support, which acted as a constraint on the user’s hand, and from the need for a tedious user-specific calibration procedure.

B. Vision based approaches:
The development of the computer hardware improved a lot in this recent decade make the hardware cheaper and offers a better performance in computing. The vision-based hand gesture slowly replaces the role of data glove with non wearable devices because it is more naturalness without using any devices in the hand and user friendly which is important in human-computer interaction. Compared with data glove, it look cumbersome and limitation in the movement of hand. Because vision is one of the six physical media that computer must be instantiated perceptibly when communicated to humans. Hence, vision-based approach is more than wearable devices in hand gesture recognition [3].

In this approach user not require to wear anything. Instead the system requires only camera(s), which are used to capture the images of hands for interaction between human and computers. Vision based approach is simple, natural and convenience [7]. Normally, most of the gesture recognition systems divided into three stages which are image pre-processing follow by tracking and finally recognition stage as shown in Fig.2.
Image pre-processing part belongs to the preprocessing stage, which locates the task-relevant regions of the images and eliminates the background. However, there are still several challenges to be addressed, for instance, illumination change, background clutter, partial or full occlusion etc. The challenge of hand tracking mainly comes from the fact that the hand is a highly articulated object and cannot be treated as a rigid object, and the speed of gesture can be rather fast. The final step recognition includes classification. It uses the information extracted and processed by previous steps and outputs the recognition results.

However, there are still several challenges to be addressed, for instance, illumination change, background clutter, partial or full occlusion etc.

C. Color glove based approaches:
Color glove based approaches represent a compromise between data glove based approaches and vision based approaches. Marked gloves or colored markers are gloves that worn by the human hand with some colors to direct the process of tracking the hand and locating the palm and fingers, which provide the ability to extract geometric features necessary to form hand shape [6].

The custom pattern on the glove facilitates faster and more robust pose estimation. While our approach is not as accurate as traditional optical mocap, it requires only a single camera. The result is an unrestricted and inexpensive hand-tracking device suitable for desktop virtual reality. Instrumented gloves systems such as the P5 Data Glove and the Immersion Cyber glove have demonstrated precise capture of 3-D input for real-time control. However, these systems are expensive and unwieldy. They rely on exoskeletons or embed more than a dozen sensors into a glove, which can be restrictive to movement. This use of a color glove is inspired by advances in cloth mocap, where dense patterns of colored markers have enabled precise deformation capture. A variety of color patterns may be appropriate for particular design is motivated by the tracking [4].

This particular design is motivated by the limitations of consumer-grade cameras, the significant amount of self-occlusion in natural hand motion, and the speed requirements of the inference algorithm. The disadvantages are similar to data glove based approaches: They are unnatural and not suitable for applications with multiple users due to hygiene issues.

D. Depth-based hand gesture recognition:
Depth-based recognition method has a high robust. It is a promising research direction because it has the characteristics of real-time identification and high precision. But depth camera based on the technologies such as time of flight (TOF) structure light, 3d laser scanning is so expensive that its utility has been limited. So when Microsoft launched the Kinect device, then researcher’s interest was greatly inspired.

This system uses the Kinect’s depth sensor to capture the movements of user. In this system, the user stands in the front of a Kinect sensor and performs a gesture. The system takes a gesture as a sequence of frames, sampled at small regular intervals. This sequence of frames is given to the system as an input and the system extracts features from each frame, and then after performing computations the system outputs the most appropriate meaning (translation) of that gesture [5].

The major specialty of the system lies in the approach of extracting some novel and highly robust features for the purpose of learning and classification. The internal architecture of the proposed system consists of two stages of filters for the translation of a gesture. The internal architecture of Kinect is shown in Fig. 4. It has the infrared camera and Prime Sense sensor while it is used when RGB camera charms the image in order to calculate the depth of the object. System first records a test gesture and puts it in the first stage filter along with information of all pre-recorded gestures (main dictionary) and the first stage filter reduces this main-dictionary to a reduced dictionary.
Which contains information of very few appropriate gestures. Then this reduced dictionary along with the recorded frames of the gesture are fed as an input to the second stage filter, which finds the most appropriate translation for the test gesture.

Until it finds the user performing any gesture, the system remains in idle mode. The system assumes that when any of the hand of the user comes above the torso, it is the moment when user starts performing any gesture and the system starts recording. The system keeps recording the sequence of frames, until both the hands go below the torso of the user. Then this whole recording sequence of frames (which is a single gesture) is used by the system for the purpose of translation.

IV. CONCLUSION

For improving the interaction between human and computer the hand gesture interaction plays important role. The evolutions of computer technology have reached the level in which the human can interact with computer by using non-verbal language. Therefore, there is need of a robust and efficient approach for the different types of hand gesture recognition that can work well in a real-time environment. After the survey on the approaches used for hand gesture recognition systems, we can give an opinion about the technologies and systems.

There are many approaches for hand gesture recognition and every approach has its strength and weaknesses. Here, we reviewed several existing methods for supporting vision-based human computer interaction based on the recognition of hand gestures like data glove based method. Glove devices has increased and grown over the past three decades. However vision based method and colored glove method are also better in some applications.

Last method discussed here is gesture recognition based on Kinect camera. It is capable of working in the dark, invariant to signer's skin color, clothing, and background lighting conditions. The system is robust against clutter in the background.

REFERENCES


