Hand gesture recognition Technology in Human-computer interaction

Xueyuan Li
School of Computer Science
University of Birmingham
Email: XXL418@student.bham.ac.uk

Abstract—Hand gesture recognition has gradually become a hot topic in Human-Computer Interaction (HCI) in recent years. It is a technology that can mimic how our human beings communicate with each other to interact between human and computers in natural and intuitive ways. Computer recognition of hand gestures is widely used in many applications, such as entertainment system, medical systems, crisis management and Human-robot interaction, since hand gesture recognition provide a natural and intuitive computer interface to people. Firstly, this paper intends to illustrate some technical issues related hand gesture recognition, such as hand gesture types and input devices. Subsequently, this paper will focus on various applications in personal and industry domains respectively. For individual areas, I only focus on the entertainment system. In industry domain, medical systems and Human-robot interaction will be discussed. At the end of this article, some shortcomings of each area are revealed and the future trends of the hand gesture recognition are concluded as three guidelines.

Index Terms—Human-computer interaction, Hand gesture recognition, Data glove, Depth-aware cameras, Stereo cameras, Computer games, Medical system, Human-robot interaction.

I. INTRODUCTION

Since its appearance, the computer has become an indispensible part of our human society. The technology will increasingly influence our everyday life because of the popularity of personal computers and other devices, such as iPhone, Xbox 360 and intelligent home. The efficient use of these intelligent devices means that more interaction between human and computer are required. Therefore, HCI has become an extremely active branch in the field of computer science research[1]. It is obvious that future Human-computer interaction modes are more similar to the natural, intuitive interaction. In other words, it is closer to the way of communication between people. Gesture recognition is considered as a critical technique in HCI. Compared to other parts of the body, hand gesture seems more flexible and natural[2]. Therefore, it is a suitable tool for control the machine.

Hand gesture recognition has been applied to various fields. I divide these applications into two main classes personal domain and industry domain. In the field of personal, entertainment is definitely the significant role[3]. In addition, hand gesture recognition is widely developed about the medical system and human-robot interaction in the industry domain[6]. For each of these applications, I describe human factors that need to be considered first. Then usability consideration is anther critical factor which motivates the users to interact with the devices.

This paper is organized in the following way: Section 2 introduces some technical issues in hand gesture recognition; Section 3 presents some applications where hand gesture recognition has been developed into the entertainment system. Section 4 illustrates some researches in the medical system and human-robot interaction. Section 5 draws the conclusion and exposes the future trends about hand gesture recognition.

II. TECHNICAL ISSUES IN GESTURE RECOGNITION

A. Hand gesture type

Hand gesture type, from simple to complicated, can be roughly divided into two categories: static and dynamic hand gesture[5]. Based on that, hand gesture recognition technology also contains three levels: two-dimensional static hand gesture recognition, two-dimensional dynamic hand gesture recognition and 3D hand gesture recognition[6].

The first two kinds of hand gesture recognition technology, is based on the two-dimensional level, they do not contain the depth data as the input sources. The third technology, however, is based on three-dimensional. The most fundamental difference between 3D hand gesture recognition and two-dimensional hand gesture recognition is that 3D hand gesture recognition requires depth data. Consequently, this makes the 3D Hand-Gesture recognition in two aspects of hardware and software is more complex than two-dimensional Hand-Gesture recognition[7].

As noted above, special hardware is needed to get the depth data. At present, there are three main hardware ways, Structure Light, Time of Flight and Multi-camera, to realize this goal in the world[8]. With the addition of the new advanced computer vision software algorithm, 3D gesture recognition can be realized. The next subsection talks about three main input devices which can track the hands movements and recognize the hand gesture.

B. Input device

Although there are numbers of vices based hand gesture recognition, the most useful and commercial input devices are as follow.
1) Data gloves: In some gesture interaction scenarios such as film shooting, accurate physical data, like bending of fingers is needed to capture. Data glove is an essential input device to the hand gesture recognition.

Normally, lots of small sensors are attached to a glove that can be worn on the hand to capture the position data of the hand[9] (as is show in Figure 1). Moreover, some kinds of data gloves can detect the movement of the joints of the fingers, or even give the user some feedback, such as vibration[10]. Then, the movement information of the hand is collected and transmitted to the computer in a wired or wireless way. And then the computer through the collation and analysis of data to identify the user’s gestures, so as to achieve human-computer interaction. One of the main application scenarios it is virtual reality.

2) Depth award cameras: Instead of the normal single cameras, depth award cameras establish a depth map through specialized cameras[11]. Kinect, as an accessory to Xbox 360, has a depth sensor called infrared emitter to detect the depth data of the picture (Figure 2). The depth sensor works with normal camera. They are complementary and provide us a better entertainment experience. Involved in the hardware principle, structure light and time of flight are two novel techniques, getting the depth data[12]. The structure light has been applied to the Kinect generation one. Nevertheless, the Kinect new generation adopts the time of flight, since the structure light can just detect 1 to 4 meters.

3) Stereo cameras: The basic principle is to use two or more cameras at the same time to take pictures[13], it seems to be human eyes, insects compound eyes to observe the world (Figure 3). By comparing the differences of images obtained from these different cameras at the same time, using an algorithm to calculate the depth information, the 3D image is obtained. This technique has the minimum requirement for hardware. However, it is also hard to implement the computer vision algorithm.

As a conclusion, the advantage of the data glove is to recognize the gesture more accurately. On the contrary, data glove is only suitable for specific areas such as animation and surgical operation, because of its high cost. Depth award camera relies on advanced hardware technique to capture the depth data. Instead of the depth award camera that consumes much more power, the stereo camera completely relies on computer vision algorithms. Generally speaking, no single method can be suitable for every application. Different input devices depend on the complex circumstances. In my opinion, the two latter should be the future development trend.

III. HAND GESTURE RECOGNITION IN PERSONAL DOMAIN

With the development of human-computer interaction and virtual reality, personal computer games gradually became a promising and commercially rewarding area. Because of the entertaining nature of the interaction, Users have a great interest in trying new game interface(figure 4), and they will be immersed in the virtual environment of the game[13].

As for computer vision-based, the system of hand-gesture-controlled games should respond to users gestures as quickly as possible. This is the so called fast-response requirement[14]. In games, the recognition performance should have the highest priority. In addition, applications sometimes do not have a real-time requirement, so the computer-vision algorithms must be efficient and robust enough. Another technical challenge is to distinguish between effective gestures and invalid gestures[14]. One possible solution is to choose a specific gesture to mark the beginning of the recognition, as in the use of iPhone where
users make the slide gesture to mark the start of operating the mobile phone. To finish the interaction, the user can set an end mark, for example, to put hands on both sides of the body. Zhang et al[15] have presented a novel hand gesture recognition system to test user-friendly interaction between human and computers. The system is based on two specific sensors, using Hidden Markov Models to recognize hand gestures. After that, a novel game called virtual Rubik's Cube is controlled to estimate the performance of the system. There are five subjects to test the system. Each subject performs about 60 gesture to the hand gesture recognition system. As for the result(Figure 5), It can accurately identify the gestures of an average of 57. In addition, The overall accuracy was 91.7%. That is to say, the system can effectively recognize human gestures, but it would take a lot of time to communicate well between the two sides.

Intuitiveness is another important requirement in hand-gesture recognition and entertainment systems[16]. In the commercial field, Most of the Wii games are able to mimic the real trajectory of human motion in the sports games, such as table tennis, tennis, and golf. Since users need to hold the handle of the game machine rather than directly interacting with the computer by the hands(Figure 6), Wii games can easily meet the requirement of intuitiveness. The Kinect sensor which is an attached sensor for Microsofts Xbox360 breaks through this limitation and implement the interaction through natural gestures(Figure 6). These interfaces use hand- body gesture recognition to increase the users gaming experience.

User acceptance is also a problem that needs to be considered in these applications[16]. They need to remember a few words that would be used in the process of interaction with the game. Before the start of a game, a new tutorial is necessary to teach users how to use gestures. For a beginner, it takes some time to get familiar with the process. Experienced users can quickly learn how to manipulate the game through gestures.
As the above analysis, in the field of the game industry, hand gesture recognition is facing some special challenges. These issues need to be resolved before the entertainment system is accepted by the majority of users eventually. Among the challenges mentioned above, intuitiveness should be given a lot of attention. In the future research, it should also focus on solving this problem.

IV. HAND GESTURE RECOGNITION IN INDUSTRY DOMAIN

Compared with the game industry, medical system and Human-robot interaction are more meaningful and have a promising prospect. So let us get a deeper understanding of the development of hand gesture recognition in these two areas.

A. Medical systems and assistive technologies

As the technology has become mature, hand gesture recognition has been developed to help improve medical systems. Hand gestures can be used to control the medical instruments, interact with the visualization displays, and help the disable people to recover quickly[17]. Some of these situations have become a reality.

Starner et al[18] have presented a wearable device that can control the medical monitoring and home automation systems by hand. More specifically, the wearable device is a gesture pendant(Figure 8), hanging around the neck. As users rise the hand or walk around, the system can also analyze their motion data. Once get the information, the system can help with the medical diagnosis and emergency situations. Similarly, another research investigated a controller-free system, exploring medical image data via Kinect[19]. In this system, Microsoft Xbox Kinect is the only input device. The surgeons can interact with the device at a distance through hand gestures to navigate the patients body (Figure 9). As we can see, the main challenge in this situation is that surgeons need to interact without any physical touch. Therefore, the controller-free system is a significant and necessary system to be applied in the medical industry. Moreover, a doctor-computer system has been developed by Graetzel et al[20]. In the system, standard input devices, like mouse and keyboard, are replaced by computer vision with using hand gestures. It is more convenient for the surgeons to use the computers during a surgery.

These systems mentioned above illustrate that hand gesture interaction techniques how to play a critical role in the traditional medical industries. Nevertheless, there are still many problems in these systems. Health care is an industry that requires a high degree of accuracy. In the process of gesture recognition, any small errors can lead to serious consequences. Still, it is necessary to have addition research to improve the accuracy and robustness of the system.

B. Human-robot interaction

Whether for fixed or mobile robot, the technique of gesture recognition technology is one of the important aspects[21]. Greater attention also needs to be given to combining the gestures and voice commands, improving robustness and accuracy of the system. Rogalla et al[22] have come up with a robot-assistant interaction system (Figure 10). In this system, gesture recognition and voice are combined. Firstly, the system learns gestures through gesture recognition and voice, then it will do the command. This system just only trains six gestures in the implementation. As a result, the system reported 95.9%

Secondly, hand gestures include many valuable geometric features in the robot task that requires navigation[24]; for instance, the pointing gesture can command a mobile robot walk to a place. For a fixed robot, The users may use a specific hand gesture to command it pick up an item and place it somewhere (Figure 11).Hand movements can be used to control mechanical actions, as the human hand can imitate the form of the robot joint.

Third, for people with disabilities, when other traditional interactive devices, such as keyboard and mouse, can not be a good interaction, they can control the robot by hand gestures
or voice[25]. For example; Moon et al[23] have developed an intelligent robotic wheelchair, combining the gestures and voice commands. The user can control the wheelchair by gesture and voice. When the information is inputted in the system, the wheelchair will go to the destination correctly. In additionally, wheelchair will automatically avoid obstacles since sonar sensors are used to improve robustness of the system.

Most of the methods mentioned above use the stereo camera to get hand gestures. Moreover, some systems employ voice detection to improving recognition accuracy. The majority of the system can only recognize the static hand gestures and just remember no more than about 10 gestures. In my view, the dynamic interaction of two hands is a promising field of future research.

V. Conclusion and future trends

The implementation of hand gesture recognition technology involves all kinds of the challenges. It requires fast response time in the entertainment industry; it requires a high precision in the medical field; it requires the user to quickly learn gesture operations in the field of robotics. At the same time, gesture recognition technology helps explain why some vision-based hand gesture systems can be more stable and mature for human-computer devices.

Compared with the entertainment industry, although people have little interest in medical system and Human-robot interaction, the significance of the two latter to the human beings is enormous. It will change every aspect of human life, make our life better and better. On the other hand, the market reaction showed that the gesture recognition technology will continue to expand its influence in various fields. This technology can make the human communicate with computers in extremely natural and intuitive ways.

As for the future hand gesture interfaces, there are three guidelines helping us evaluate the possibility of the acceptance:

The first one is validation. As we said above, different areas have different definitions for the same gesture. For example, we slide our finger to the right to unlock the phone, while may use the OK gesture to indicate the beginning of the XBOX games. To think deeply, statistical methods can be used to assess the validation of a system.

Another guideline is user independence. Data glove can be a vivid description of the concept. When you use a data glove, you need to take it with you. Furthermore, it sometimes drags a heavy cable. In addition, different people have different hand size. User independence is a critical guideline, since it can determine users acceptability.

The third one is usability criteria. There is a lot of evaluation criteria for the usability for one system. In entertainment system, taking a beginner how long to start the game is a good example. In medical system, does a surgeon can check the patient successfully using a hand gesture recognition system.

ACKNOWLEDGMENT

The author would like to thank Robert Hendley, for providing useful advice in the research. Additionally, the author would also like to thank Tang yajing who offers encouragement and support.

REFERENCES
