

A Survey on Hand Gesture Recognition

Lingchen Chen, Feng Wang*, Hui Deng, Kaifan Ji

Computer Technology Application Key Lab of Yunnan Province

Kunming University of Science and Technology

Kunming, China

email : { chenlingchen, wangfeng, denghui, [jikaifan](mailto:jikaifan@cnlab.net) }@cnlab.net

Abstract—Hand gesture recognition has become one of the key techniques of human-computer interaction (HCI). Many researchers are devoted in this field. In this paper, firstly the history of hand gesture recognition is discussed and the technical difficulties are also enumerated. Then, we analyze the definition of hand gesture and introduce the basic principle of it. The approaches for hand gesture recognition, such as vision-based, glove-based and depth-based, are contrasted briefly in this paper. But the former two methods are too simple and not natural enough. Currently, the new finger identification and hand gesture recognition technique with Kinect depth data is the most popular research direction. Finally, we discuss the application prospective of hand gesture recognition based on Kinect.

Keywords: *Human-Computer Interaction (HCI), Hand Gesture Recognition, Kinect*

I. INTRODUCTION

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). When our hand is used as a direct input device of computer, the communication between human and computer will no longer need the intermediate media. Users can control the machine by the hand gestures which were defined before [1]. 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 [2]. 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 [3]. But the recognition method based on vision is hard to work well in bad conditions. And the recognition method based on glove also has an embarrassing situation. Although this method owns the advantages of less input data, high speed, and it can get 3D information about hands or fingers movement directly. It could also recognize a lot of hand gestures on time [1]. However, the recognition method based on glove is not able to leave the support of equipment. It's impossible for users to wear bloated gloves all the time in nature condition. This

obvious disadvantage destines its useless, so we need to develop a new technology to solve it. Depth-based recognition method has a high robust. And because it has the characteristics of real-time identification and high precision, it is a promising research direction. 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 researchers' interest was greatly inspired when Microsoft launched the Kinect device [4].

II. ABOUT HAND GESTURE

Hand gesture is a high practical value body language that its specific meaning is established through our language center by the palm and finger position and shape. Generally speaking, the gesture consists of static and dynamic hand gestures. As its name suggests, Static hand gestures refer to the individual shape of the hand, and the dynamic hand gestures are composed of a series of hand movements. Because the gesture has diversity and ambiguity, and it is affected by the time and the cultural backgrounds, different people have different definitions of gestures. In this paper, hand gesture is defined as the combination of all kinds of gestures and movements which produced by hand or hand and arm. It includes dynamic hand gestures, whose meanings are based on the track of the motion of hands, and static hand gestures in which the shape of hand gesture is used to express the meaning. The main difference between gesture and posture is that posture has more emphasis on the shape and state of the hand and the body, but gesture has more emphasis on the hand movement [5] [6].

III. RELATED WORK

At present, the technology difficulty about hand gesture recognition can be listed as follows:

1). *Difficulty of detecting the human gestures*

It's hard for us to extract the gesture from the target image because of the unpredictable environment and the complex background.

2). *Difficulty of hand gesture recognition*

We need to explain the profound meanings through the posture of hand and its changing process. Then, the key lies in how to extract a feature which has the geometric invariance. But as is known to all, our hand is an elastic body, and it has more than 20 degrees of freedom. Therefore, different people make the same gestures will exist a difference. It will be also different when one person made the gestures in different time or different place.

* Corresponding Author, E-mail: wangfeng@acm.org

These characteristics of gesture make the problem hard to be resolved, and the implementation still must take the limiting condition into consideration till now [7]. Even in these cases, the researchers have obtained a series of achievements. 1. Kenn [8] presents an interface for wearable computing applications using glove-based finger recognition. 2. Chu et al [9] presents self-portrait interface using vision-based hand motion gesture. 3. Yang et al, [10] proposed a hand motion gesture recognition system using Kinect depth data.

A. Glove-based hand gesture recognition

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.



Figure 1. 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 [11]. So it can also identify a variety of gestures, do a real-time identification [12]. But the glove is both expensive and heavy. This condition is not suitable for natural interaction [13]. In this respect, Kenn presents an interface for wearable computing applications using glove-based finger recognition. In his research, he conducted a large number of experiments to demonstrate the reliability of this method. And the finger identification for hand gesture was implemented in his paper. But because of its innate defects, we have to wear the data glove all the time. This method could only be used in the laboratory until we achieves major breakthrough in the field of materials science to make the data glove become light and handy.

B. Vision-based hand gesture recognition

Researchers used to focus their energy on vision-based gesture recognition (VGR), because there are some notable advantages exists [14]. And we take hand as the direct input equipment, communication between human and computer will need no more other intermediate media. Users can control the machine by the hand gestures which were defined before. The method of vision-based gesture recognition use camera to collect gesture image sequence [15], and identify the gesture

by processing and analyze the image. Chu et al presents self-portrait interface using vision-based hand motion gesture. Their research makes it possible for user to manipulate digital camera when taking self-portrait pictures [9] [16]. It is convenient to control camera functions through gestures. Actually, this method got a good precision. In order to assist the hearing and/or speech impaired people to communicate with hearing people, Madhuri, Y et al [17] presents a vision-based sign language translation device. According to their description, we could know it that this device is used as a translator for people that do not understand sign language, then avoiding by this way the intervention of an intermediate person and allow communication using their natural way of speaking. Through their hard work, the device has good real-time performance. This allows for almost instantaneous recognition from finger and hand movements to translation. And it is able to recognize one handed sign representations of alphabets (A-Z) and numbers (0-9).

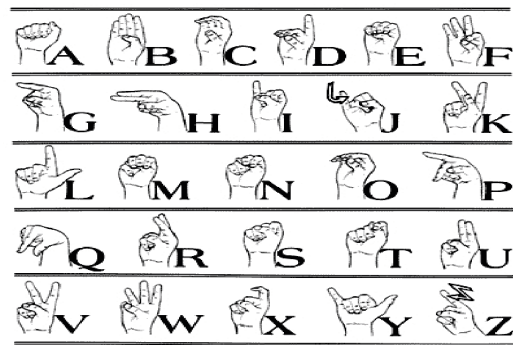


Figure 2. Database of One-Handed Alphabet Signs



Figure 3. Database of One-Handed Number Sign

But, these methods are still constrained by the following conditions:

- Lighting sensitivity is a main weakness for vision-based hand gesture recognition.
- The distance from the user to the camera is another issue.
- The optical-flow measurements need a cluster background.

In fact, this is also the common defect of the vision-based gesture recognition system.

C. Depth-based hand gesture recognition

With the development of the advanced techniques of human-computer interaction (HCI), gesture recognition is becoming one of the key techniques of HCI. But years of research has proved the limitation of conventional method. Being a newly developed distance measuring hardware, the depth camera technology opens a new epoch for 3D geometric information acquisition [18]. There are two main approaches employed currently in depth camera technology before 2010. The first one is based on the time of flight (TOF) principle. And the second approach is based on the light coding. Compared with color images, the depth image could reflect the 3d feature directly. So it would not be affected by the factors such as illumination, shadow and color. Even if there is a covered part between two objects, by using different distance information which we've got from the depth image, or take advantage of the stratification of grey value, we could separate the covered object and the different parts of the same object. It could not be done by the visible light image [19]. But at this time, depth camera is too expensive to apply. Thanks to the cost of the depth camera drops rapidly, especially when Microsoft launched the Kinect device, researcher's interest was greatly inspired. The internal architecture of Kinect is shown in Figure 1. 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.

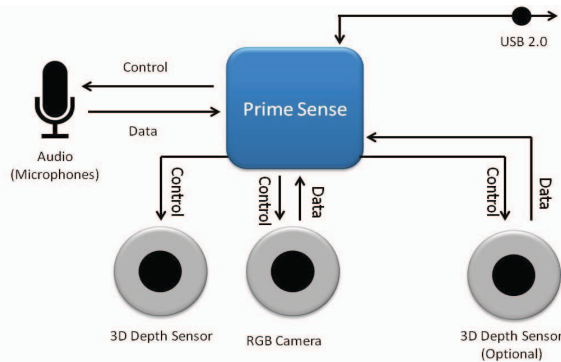


Figure 4. MS Kinect Architecture [20]

Compared to the traditional device, Kinect have advantages in resolution and cost. Zhang Yi et al [21] proposed a gesture track recognition method based on Kinect depth image information. This method adopted Hidden Markov model (HMM) trained the sample. And they completed the basic control of the intelligent wheelchair by tracking five different gestures. However, the implemented system was not able to recognize fingertip. That is to say, this method could not to recognize the delicate changes of our finger. Raheja et al, [22] proposed a method to recognize and track fingertips and center of palm using Kinect. Through their experimental data, we can find it that they were able to identify fingertips and center of palm very accurately, even when the fingers were bent by large angle. But this method also have a shortage, they don't implement any interactions, gesture and finger identification. Therefore, finger identification and gesture recognition methods provide natural

interactions and interface by using fingers is a promising research direction.

IV. APPLICATION PROSPECT

Generally, the application of hand gesture recognition can be listed as follows [23]:

- Virtual environment interaction techniques
- Sign language recognition techniques
- Multi-channel, multimedia user interface
- Manipulator-snatching

At the same time, the Kinect sensor has been widely used in medicine, business, computer science, and many other fields. Then we'll show the prospects of gesture recognition based on the Kinect.

A. The application in clinical surgery

In clinical surgery, it is necessary to keep the surroundings aseptic. However, in the surgery, the surgeon is also need to view the patient's clinical imaging data from the computer, and the computer must sterile. But the current way of human-computer interaction is not convenient for staff to operate during the surgery. Actually, it increases the workload and the operation personnel. So it is difficult to ensure a timely, accurate, safe surgery [24]. This disadvantage could be solved by applying the method of hand gesture recognition based on the Kinect. A team at Microsoft Research Cambridge consisting of social scientists, computer scientists, and designers has used the Kinect for Windows hardware and SDK to simplify this process. It enables doctors to use simple hand gestures to change, move, or zoom in on CT scans, MRIs, and other medical images [25].

B. The application in robot

In recent years, robot technology has developed rapidly. But there still exists a problem we have to face. That is how to improve the robot's ability to understand the surrounding environment. Sensor is the critical component to improve the robot's ability to understand surroundings. Kinect have advantages in resolution and cost, and it can capture 3d images. In order to control the robot's behavior timely and effective, we want to recommend the method of hand gesture recognition based on Kinect [26]. It will be a new research hotspot in robot's visual and control. Japan's well-known robot manufacturer Yakawa, install the kinect on a service robot whose name is Smartpal. Then the users can control the robot by their gesture, and let it simulate the users to do the same gesture [27]. We believe it that with the development of the time, by means of the hand gestures to control the robot to accomplish various functions is no longer a dream for us.

V. CONCLUSION

Though the gesture recognition is not a new subject, with the emergence of the strong vitality and diversity application technology, Kinect, hand gesture recognition has achieved an important breakthrough. And as the further mature and fully

dig of the Kinect, its complete implement in the natural conditions is no longer far away from us.

ACKNOWLEDGMENT

We would like to thank all the authors of the literatures which we surveyed. We appreciate the support from National Natural Science Foundation of China (61063027). We thank the members of Yunnan Computer Technology Application Key Lab in Kunming University of Science and Technology for their support and assistance.

REFERENCES

- [1] HU You-shu. "Gesture recognition technology: A Review", Published by China Science and Technology Information, 2005.
- [2] GAN Chuang. "Dynamic Gesture Recognition Based on 3D Kinect", Published by ELECTRO-OPTIC TECHNOLOGY APPLICATION, pp.59-62+67, 2012.
- [3] D. Wigdor and D. Wixon. "BRAVE NUI WORLD", Published by Morgan Kaufmann, pp.9-15, 2011.
- [4] LI Hong-bo, DING Lin-jian, RAN Jian-yong. "Analysis of human body recognition based on the Kinect depth image", [J]. Digital Communication, pp.23-28, 2012.
- [5] FANG Zhi-gang. "Computer Gesture Input And Its Application in Human Computer Interaction", Published by the small micro computer system, 1999.
- [6] S. Mitra, T. Acharya. "Gesture Recognition: A Survey", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, pp. 311-324, 2007.
- [7] SUN LI-juan, ZHANG Li-cai, GUO Cai-long. "Technologies of Hand Gesture Recognition Based on Vision", [J]. Computer Technology and Development, pp.220-222+227, 2008.
- [8] H. Kenn, F. Megan and R. Sugar. "A glove-based gesture interface for wearable computing applications", In Proceedings of the IFAWC 4th international forum on applied wearable computing 2007, pp. 169-177, 2007.
- [9] S. Chu and J. Tanaka. "Hand Gesture for Taking Self Portrait", Human-Computer Interaction, Part II, HCII 2011, LNCS 6762, pp. 238-247, 2011.
- [10] Y. Cheoljong, J. Yujeong, B. Jounghoon, H. David and K. Hanseok. "Gesture recognition using depth-based hand tracking for contactless controller application", In Consumer Electronics (ICCE), 2012 IEEE International Conference on, pp. 297-298, 2012.
- [11] WENG Han-liang, ZHAN Yin-wei. "Vision-Based Hand Gesture Recognition with Multiple Cues", [J]. Computer engineering and science, pp.127-131, 2012.
- [12] Yang Ji-guang. "Sign language recognition method based on data glove", [A]. In Intelligent Computer Interface and Application Progress [C]. Beijing: Publishing House of Electronics Industry, 1997.
- [13] WANG Zhao-qi. "Virtual human synthesis: A Review", [J]. Journal of Graduate School, Academia Sinica, pp.90-99, 2000.
- [14] S. Lenman, L. Bretzner and B. Thuresson. "Computer Vision Based Hand Gesture Inter-faces for Human-Computer Interaction", Technical Report CID-172, Center for User Oriented IT Design, pp.3-4, 2002.
- [15] A. Bobick, J. Davis. "Real time recognition of activity using temporal templates", Proc. of Third IEEE Workshop on applications of computer vision, Florida, 1996, 39-42.
- [16] Samsung Self-Portrait Camera: TL225 (2009). <http://www.samsung.com/us/photography/digital-cameras/>.
- [17] Yellapu Madhuri, Anitha.G, Anburajan.M. "VISION-BASED SIGN LANGUAGE TRANSLATION DEVICE", Information Communication and Embedded Systems (ICICES), 2013.
- [18] XIANG Xue-qin, PAN Zhi-geng, TONG Jing. "Depth Camera in Computer Vision and Computer Graphics: An Overview", [J]. Journal of Frontiers of Computer Science & Technology, pp.5-16, 2011.
- [19] ZHU Yan-ming, YANG Zhi-bo, YUAN Bo. "Vision Based Hand Gesture Recognition", Published by International Conference on Service Science, 2013.
- [20] The teardown, Engineering & Technology, Vol.6, no.3, pp. 94-95, 2011.
- [21] ZHANG Yi, ZHANG Shuo, LUO Yuan, XU Xiao-dong, [J]. Computer application research, pp.353-356, 2012.
- [22] J. Raheja, C. Ankit and S. Singal. "Tracking of fingertips and centers of palm using KINECT", 2011 Third International Conference on Computational Intelligence Modelling Simulation, pp. 248-252, 2011.
- [23] LI Qing-shui, FANG Zhi-gang. "Gesture recognition technology and its application in the human-computer interaction", Published by Chinese Journal of Ergonomics, 2002.
- [24] Ni Chen, Qiu Peng et al. "The Research And Application of Kinect Technology in the Field of Human-computer Interaction", Information Technology and Informatization, 2013.
- [25] Douglas Gantenbein. Kinect Launches a Surgical Revolution [R/OL]. <http://research.microsoft.com/en-us/news/features/touchlessurgery-060712.aspx>.
- [26] ZHU Tao, JIN Guo-dong, LU Li-wu. "Overview the application of Kinect and its development prospects", Published by Modern Computer, 2013.
- [27] Yakawa. "Terrence O'Brien. Yakawa Electric's SmartPal VII Lets you Clean up Grandma's House Using Kinect". <http://www.engadget.com/2011/11/21/yakawa-electrics-smartpal-vii-lets-you-clean-up-grandmas-house/>.