

Review of Hand Gesture Recognition Study and Application

Yanan Xu

Department of Game and Multimedia Engineering
PaiChai University, 155-40 Baejae-ro, Seo-gu, Daejeon, 35345, South Korea

Yunhai Dai*

Department of General education and International Trade
Namseoul University, 91 Daehak-ro Seonghwan-eup Sebuk-gu Cheonan-si
Chungcheongnam-do, 31020, South Korea

*Corresponding Author

Copyright © 2017 Yanan Xu and Yunhai Dai. This article is distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Human-computer interaction is an essential part of most people's daily life. The traditional human-computer interaction mode from the original keyboard to the current mouse, joystick, and wireless input devices, greatly facilitates the interaction between people and computers and makes it easier for people to operate the computer and improve work efficiency. However, this kind of interaction mode cannot completely meet the demands of human-computer interaction due to the dependence on additional input hardware devices. Hand gesture can be defined as a variety of gestures or movements produced by hands or arms combined, it is always capable of expressing a signer's intention, so it can act as a means of natural communication between human and machine. Studies on hand gesture recognition is very important for the development of new human-centered human-computer interaction. This paper reviewed the current study status and application of gesture recognition aiming to summarize the commonly used hand gesture recognition methods, analysis their strength and weak points, and list the challenging problems in current research of hand gesture recognition.

Keywords: Hand gesture, Human-computer interaction, Study status, Application

1 Introduction

In recent years, with the rapid development of machine vision technology, especially the progress of image processing and recognition technology, people's attention is no longer limited to the improvement of traditional input methods of human-computer interaction, how to make use of the biological characteristics of human beings to study more natural interaction technologies, so that human and computer can interact directly has becoming the current research focus of human-computer interaction.

According to the current development trend of human-computer interaction technology, human centered human-computer interaction technology is bound to replace the computer centered ones, this kind of studies include gesture recognition, facial expression recognition, face recognition, lip reading recognition, limb movement tracking, eye gaze tracking and pose recognition [1], etc. Among the characteristics can be used as human-computer interactive medium, Hand gesture is vivid, intuitive, and contains rich information, with same expression ability as natural languages such as spoken language and written language, it can act as a means of natural communication between human and machine, and plays an important role in the field of human-computer interaction. However, due to the complexity, diversity, ambiguity and uncertainty of hand gesture, hand gesture recognition has been becoming a challenging interdisciplinary research topic.

2 Hand gesture recognition overview

Hand gesture recognition is generally divided into static gesture recognition and dynamic gesture recognition, static gesture recognition is the recognition of hand shape, read out the meaning of hand expression, and dynamic hand gesture recognition is the recognition of hand motion trajectory in space, and then perform the corresponding operation based on obtained trajectory parameters, such as for the playing courseware on the projection, hand gestured can be used to flip up and down, pause, start, etc.

The traditional gesture recognition was through the use of wearable technology, allowing users to do some hand gestures with special data gloves on [2], the data gloves would transfer user's gestures and location information to the computer and help it comprehend the gestures and behaviors of uses. Figure 1 shows a multi-function virtual reality device composed of many sensors on the glove called Immersion CyberGrasp. Through the software mapping, the virtual objects can be shifted, clutched and rotated by the glove with the ability of "reach into the computer".

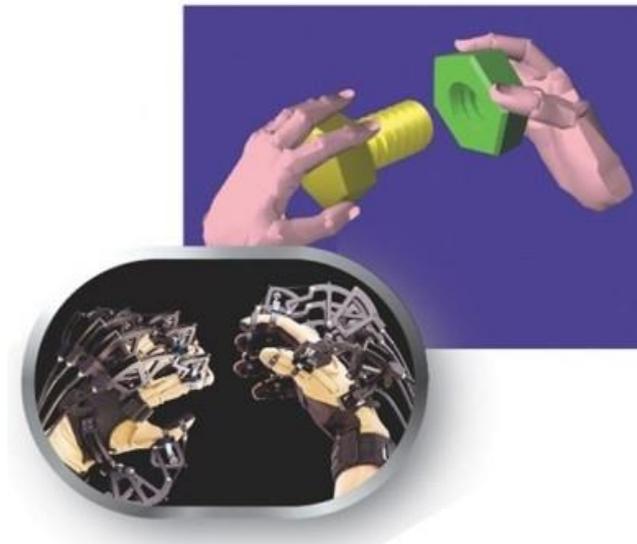


Figure 1. Immersion CyberGrasp data glove

The glove can transmit hand gesture to the computer in real time accurately, and then receives feedback from the virtual environment to the operator. It provides users with a direct and universal human-computer interaction mode [3] with advantages of high accuracy, simple data and fast processing speed, etc., but because of the shortcomings of expensive equipment, inconvenient operation, and not suitable for long-distance control, this kind of interaction model is hard to get promotion.

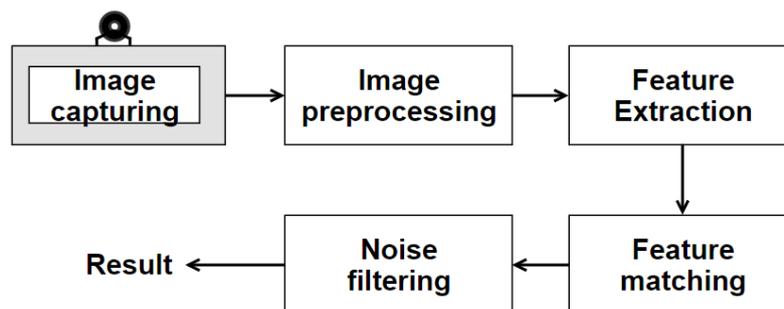


Figure 2. Vision based hand gesture recognition system

Vision based gesture recognition as shown in figure 2 takes use of the camera to capture hand gestures to system, and after image preprocessing such as detection and segmentation to extract features of extracted image sequence to understand and describe its behavior. When one or more cameras obtain the video stream of user gestures, the system will monitor whether there are hand gestures contained in the data stream according to the interactive mode of gesture, if there are, separates them. Then choose appropriate method to detect and extract features,

and choose appropriate classifier to recognize the gesture in current image. The greatest advantage of vision based hand gesture recognition is that the input is simple with lower dependence on equipment, and it is in line with the people's daily interaction, therefore, vision based hand gesture recognition is bound to be the new pursuit of human-computer interaction.

3 Research status

The most commonly used hand gesture recognition algorithms include HMM (Hidden Markov Model) model based on statistics [4], algorithm based on genetic algorithm [5] and artificial neural network [6].

Statistics based HMM takes advantage of causal relationship between visual features obtained from prior knowledge to deal with the inherent problem of uncertainty in video processing, not only can build the probabilistic modeling of dependencies among different features corresponding to multiple random variables in every moment, but also consider the transition probability between every moment, which can well reflect the timing relationships between features. However, it needs to maintain a certain size of the sample library and ask for large computational quantity, even though the larger the size of the sample library, the closer to the actual situation and the higher the accuracy of hand gesture recognition will be, moreover, this method also needs data smoothing technology to enlarge the value of small probability. The genetic algorithm is used to discretize images, control the discrete points and then convert image recognition problems into combinatorial optimization problems of a series of discrete points. But it is not able to take use of the feedback information from network in time, in addition, this method is troubled with slow search speed, large training sample and long training time. Artificial neural network has a large number of simple processing units (neurons) which are widely connected to form a complex information processing network, it mimics the function of information processing, storage and retrieval of human brain neural system in a certain extent and level, its processing efficiency is high with small samples can be satisfied, but the training process needs the participation of human, and the accuracy of recognition is influenced by the subjective factors.

Many techniques on HOG (Histogram of Gradient) like [7] have been proposed in the past which employ edge and gradient based descriptors for hand gesture recognition. But they are only able to detect hand gestures in a simple background and are liable to fail when the background is cluttered. Paper [8] extracted the edge pixels of hand gesture, took use of the idea of model-based matching using Hausdorff distance to realize the recognition of Chinese alphabet, the method proposed had advantages of small computation and strong adaptability but disadvantages of ignoring the situation of rotation, scaling and skin color interfere. The authors in [9] putted up with a method based on hand characteristic curves, the result of combination of color, motion and edge information, this method can reduce the dependence on hand segmentation, but the computation is too complex, and the real-time performance is poor.

In addition, there are also many other research works that successfully harvest some certain achievement of hand gesture recognition. Paper [10] used the maximum difference feature to classify the gestures after the segmentation of MDF (The Most Discriminating Feature) space, the algorithm can adapt to the occasion with complex background. In [11] an inductive learning system was introduced, this system could extract rules from DNF (Disjunctive Normal Form), and it obtained the recognition rate of 94%. The elastic curve matching method introduced by paper [12] was less independent on segmentation, and could reach the recognition rate of 85% in a complex background. Bjorn et al. [13] used color and motion features to detect and track the hand, combined the method of template matching and nearest neighbor classifier to recognize hand gesture. In [14] the authors divided the hand into 21 different regions and train a SVM classifier to model the joint distribution of these regions for various hand gestures so as to classify the gestures.

4 Application and foreground



Figure 3. Canesta virtual keyboard

In December 2003, the Cybernet company located in Michigan, USA, introduced the world's first fully functional weather map management system called "GestureStorm" that utilizes both body tracking and gesture recognition technology, with which the meteorologist can control the pace of the visuals and even incorporate spontaneous close-ups with simple hand movements that seem natural to the viewing public. Canesta Company launched a new device in late 2004 as shown in figure 3, it can ensure the personal digital assistant (PDA) with the function of gesture recognition. This device can project the image of keyboard on the desktop through internal lens of PDA, in the meanwhile, it will also shoot a beam of red light over the "keyboard", by detecting the interval time consumed by the infrared pulse after leaving transmitter, being rebound by user's finger and return back to PDA sensor eventually to accurately sense the position of a user's finger at any moment. A research team in Georgia Institute of Technology developed a device named as "Gesture Panel" to replace the instrument panel device usually used on automobile, with which, the driver only needs to make gestu-

res in the designed area to adjust the temperature or sound volume of the car, without transferring the attention of the road. This device based on hand gesture recognition can make good contribution to reduce traffic accidents.

In the aspect of vision based hand gesture recognition, the representative research results include: J. Davis and M. Shah [15] used gestures captured from visual gloves with bright marks on the part of fingertips as the input of system, which can recognize 7 kind of hand gestures. Starner et al. achieved a recognition rate of 99.2% on the recognition of short sentences composed of 40 words with part of speech in American Sign Language. Vogler [16] and Metaxas applied these two methods to the recognition of American Sign Language, by using a position tracker and three mutually perpendicular cameras as the gesture input device they completed the recognition of 53 isolated words with recognition rate of 89.9%. On the side of method based on sensor dependent devices, Christopher Lee and Xu in CMU developed a gesture control system for manipulator operation [17]. M. W. Kadous [18] took use of Power Gloves as the sign language input device to identify the vocabulary set composed of 95 isolated words, and the recognition rate was up to 80%. The main application fields of gesture recognition are as following:

- 1 Education and life improvement for children, elderly and deaf-mutes. Through some human-computer interfaces, can complete the natural communication between children, the elderly or deaf-mutes and the computers, and thus improve their ability of education. In addition, by using hand gesture recognition based human-computer interfaces, a communication channel between the normal peoples and deaf-mutes can be established, so that normal peoples can “listen” and understand the “words” of deaf-mutes.

- 2.Environment and virtual objects control in virtual reality environment. Use hands to complete the browsing, selection or manipulation of the virtual environment on the operation interface. Use different definitions of gestures to control the movement and turning of virtual objects, or through the movement of real hands to control the movement of the mirror images of hands in augmented real environment.

- 3.Application on smart home appliances and control field. Among the computer controlled means, hand is regarded as a flexible and efficient controlling way, application on hand gesture in the field of control has obtained some results, such as video cameras controlled by gesture command such as “zoom”, “panoramic” and “tilt”. Jaguar company [19] introduced a vehicle control system based on hand gesture recognition, figure 4 is a conceptual image of this system, it shows how to response to the incoming phone call with natural hand gestures.

- 4.Demonstration and study of robot. By studying the mechanism of human visual language from the perspective of cognitive science to improve the human language understanding ability of computer.

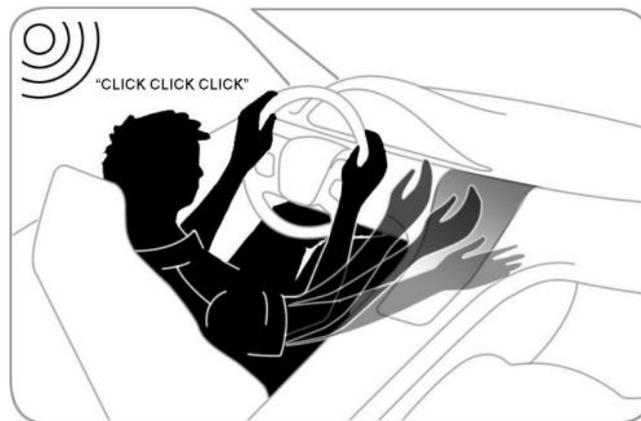


Figure 4. Conceptual map in response to incoming phone call by driver gestures

5 Challenges of hand gesture recognition

At present, although the study of hand gesture recognition has made great progress and achieved high recognition rate in different areas, it is still facing many challenges, such as: extraction of invariant features, transition model between gestures, minimal sign language recognition units, automatic segmentation of recognition units, recognition approach with scalability about vocabularies, auxiliary information, signer independent and mixed gestures recognition, etc.

Thereinto, the static gesture recognition based on vision is the current trend of hand gesture recognition and mainly has the following two aspects of technical difficulties:

1. Difficulties in target detection

Target detection is to capture the target from the image stream in a complex background, which is to extract the object of interest. In vision based hand gesture recognition methods, it is always a difficult problem to divide the human hand area and other background areas in the image, this is mainly due to the variety of background and unforeseen environment factors.

2. Difficulties in target recognition

Hand gesture recognition is to explain the high-level implications contained according to the posture and changing process of hand. In view of the following characteristics of hand gesture, the key technology of hand gesture recognition is to extract the geometric invariant features.

1) Hand is an elastic object, there may big differences exist between same gestures and high similarity between different gestures. Human hand has more than 20 degrees of freedom, its movement is very flexible and complex. Therefore, same gestures made by different people may vary, and gestures made by the same person at different time or place may also differ.

2) Hand has a lot of redundant information, since the key part of hand gesture recognition is to identify finger features, so palm feature is one of the redundant information.

3) The position of hand refers to the projection of hand from three-dimensional space to two-dimensional, so the projection direction is really related.

4) It is easy to produce shadows due to the non-smooth surface of hand.

Because of these characteristics of hand, the above two problems have not been solved well yet, so it is necessary to add some restrictions when doing hand gesture recognition.

6 Conclusion

In this paper commonly used hand gesture recognition methods are reviewed include HMM, genetic algorithm, neural network, HOG, etc., both of their advantages and disadvantages are discussed.

Application and foreground of hand gesture recognition show that the study on hand gesture recognition is not only useful to improve and enhance the studying and working conditions of deaf peoples, to provide them with a better quality of service, but also can apply to computer aided sign language teaching, TV program bilingual broadcast, virtual human research, special effects production in film, animation production, medical research, games, entertainment, etc.

Because of the complexity and diversity of hand gestures, the variety of background and unforeseen environment factors, the study of hand gesture recognition is still facing many challenges.

Acknowledgements. Funding for this paper was provided by Namseoul University.

References

- [1] N. D. Binh, T. Ejima, Real-Time Hand Gesture Recognition Using Pseudo 3-D Hidden Markov Model, *5th IEEE International Conference on Cognitive Informatics*, **2** (2006), 820-824. <https://doi.org/10.1109/coginf.2006.365596>
- [2] Y. Han, A Low-Cost Visual Motion Data Glove as an Input Device to Interpret Human Hand Gestures, *IEEE Transactions on Consumer Electronics*, **56** (2010), 501-509. <https://doi.org/10.1109/tce.2010.5505962>
- [3] L. Dipietro, A. M. Sabatini, P. Dario, Survey of Glove Based Systems and Their Applications, *IEEE Transaction on Systems, Man and Cybernetics, Part C: Applications and Reviews*, **38** (2008), 461-482. <https://doi.org/10.1109/tsmcc.2008.923862>
- [4] X. Y. Wang, G. Z. Dai, X. W. Zhang, Recognition of Complex Dynamic Gesture Based on HMM-FNN Model, *Journal of Software*, **19** (2008), 2302-2312. <https://doi.org/10.3724/sp.j.1001.2008.02302>

- [5] X. H. Wang, H. J. Bao, Gesture Recognition Based on Adaptive Genetic Algorithm, *Journal of Computer-aided Design & Computer Graphics*, **19** (2007), 1056-1062.
- [6] F. F. Zeng, Y. Wang, G. J. Huang, Fuzzy Neural Network Based Hand Gesture Recognition, *Mini-Micro Computer Systems*, **21** (2000), 706-709.
- [7] H. J. Lee, J. H. Chung, Hand Gesture Recognition Using Orientation Histogram, *Proceedings of the IEEE Region 10 Conference*, **2** (1999), 1355-1358. <https://doi.org/10.1109/tencon.1999.818681>
- [8] L. G. Zhang, J. Q. Wu, Hand Gesture Recognition Based on Hausdorff Distance, *Journal of Image and Graphics: A*, **7** (2002), 1056-1062.
- [9] H. B. Ren, G. Y. Xu, Hand Gesture Recognition Based on Characteristic Curves, *Journal of Software*, **13** (2002), 987-993.
- [10] Y. Cui, D.L. Swets and J.J. Weng, Learning-based Hand Sign Recognition Using SHOSLIF-M, *Proceeding of 5th International Conference on Computer Vision*, (1995), 631-636. <https://doi.org/10.1109/iccv.1995.466879>
- [11] F. Quek, T. Mysliwiec, M. Zhao, Finger Mouse: a Free-hand Pointing Interface, *IEEE International Conference on Automatic Face and Gesture Recognition-FGR*, (1995), 372-377.
- [12] J. Triesch, C. V. D. Malsburg, A System for Person-independent Hand Posture Recognition against Complex Backgrounds, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **23** (2001), 1449-1453. <https://doi.org/10.1109/34.977568>
- [13] B. Stenger, Template Based Hand Pose Recognition Using Multiple Cues, *Lecture Notes in Computer Science*, Vol. 3852, Springer-Verlag Berlin Heidelberg, 2006, 551-560. https://doi.org/10.1007/11612704_55
- [14] C. Keskin, F. Kirac and Y. E. Kara, Lale Akarun, Real Time Hand Pose Estimation Using Depth Sensors, *In Proceedings of the IEEE International Conference on Computer Vision Workshops*, (2011), 1228–1234. <https://doi.org/10.1109/iccvw.2011.6130391>
- [15] J. Davis, M. Shah, Visual Gesture Recognition, *IEEE Proceedings on Vision, Image and Signal Processing*, **141** (1994), 101-106. <https://doi.org/10.1049/ip-vis:19941058>
- [16] C. Vogler, D. Meraxas, Adapting Hidden Markov Models for ASL Recognition by Using Three-dimensional Computer Vision Methods, *IEEE*

- International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, **1** (1997), 156-161.
<https://doi.org/10.1109/icsmc.1997.625741>
- [17] C. Lee, Y. Xu, Online Interactive Learning of Gesture for Human/Robot Interfaces, *Proceedings of IEEE International Conference on Robotics and Automation*, **4** (1996), 2981-2987. <https://doi.org/10.1109/robot.1996.509165>
- [18] M. W. Kadous, Machine Recognition of Auslan Signs Using PowerGloves: Towards Large-lexicon Recognition of Sign Language, *Proceeding of the Workshop on the Integration of Gesture in Language and Speech*, (1996), 165-174. <https://www.researchgate.net/publication/2828963>
- [19] C. A. Pickering, K. J. Burnham, A Research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction, *3th Institution of Engineering and Technology Conference on Automotive Electronics*, (2007), 1-15. <http://ieeexplore.ieee.org/document/4383638>

Received: March 26, 2017; Published: April 19, 2017