Review of Gesture Recognition for Virtual Automotive Training Applications

Sarah N. Zamalik, Juliana A. Abubakar, and Abdul Razak Yaakub

Abstract— In a virtual system, gesture is considered as a natural way for communication between human and computers. Gesture recognition is the process of recognizing and interpreting a stream of continuous sequential gesture from a given set of input data. The main goal of gesture recognition is to create a system which can recognize specific human gestures and use them to convey information or to control a device and/or application. The objective of this paper is to review gesture-based recognition methods and its applications as an approach to improve user interaction particularly for virtual automotive training. Review of applications and techniques of gesture recognition system will be provided with emphasis on hand gesture expressions. The major techniques include HMMs, ANN, and fuzzy clustering will be reviewed and analyzed.

Index Terms— Gesture Recognition, Artificial Neural Networks, Fuzzy Clustering Algorithm, Hidden Markov Model, Histogram Based Feature, Virtual Automotive Training

1 Introduction

USING the gesture for natural human-computer interaction became one of the important issues in the recent few decades, to enclose human life appliance. Sign languages usually used among people to explain specific meaning or deliver a meaningful message, for this reason gestures motivate to simulate the natural interaction between humans, but in this time between human and computers by modelling, analyzing the gesture, and finally recognize it. Gesture recognition relates to recognizing meaningful expressions of motion by a human, involving the head, face, hands, arms, and/or body. It is of utmost importance in designing an intelligent and efficient human–computer interface.

For many years, gestures have been used in human-computer interaction (HCI). Earlier, hardware based gesture recognition was more common. The user had to wear gloves, helmet and other heavy equipment. Sensor actuator and accelerometer were used for gesture recognition. But the whole process was difficult in real time environment. According to Mitra & Acharya [16], gesture can be subdivided into two types; static (the user assumes a certain pose or configuration) and dynamic (with prestroke, stroke, and poststroke phases). However, some gestures also can have both static and dynamic elements, as in sign languages.

The automotive industry is one of the leading industry in this environmentally conscious manufacturing and product recovery [25]. In Malaysia, the establishment of *Proton* in 1985 and consequently *Perodua* in 1993 acted as a catalyst to the development of the automotive sector in Malaysia including the development of local automotive component manufacturers [1].

There is concern among industry leaders that many of the instructors and trainers are not staying current, as required. There is also a concern that the resources are outdated. Since technology in the automotive industry is changing rapidly, training facilities must keep up with the pace of the technology. Modern science and technology requires high quality of automotive engineering education. Application using gesture-based interfaces for virtual reality automotive training is one of the possible ways for improving the educational process. This study attempts to develop gesture recognition for user interaction with virtual automotive training applications. Therefore, the objective of this paper is to review gesture recognition methods and its applications as an approach to user interaction particularly in virtual automotive training.

2 Literature Review

The literature review will discuss the analysis and existing or similar research that relates to this study. It will focus on the Gesture Recognition Based Applications, Gesture Recognition Methods and Virtual Automotive Training.

2.1 Gesture Recognition Based Applications

Gesture based applications are broadly classified into two groups on the basis of their purpose: multidirectional control and a symbolic language. Multidirectional control was applied in 3D design, tele presence and virtual reality while symbolic language applied in sign language.

3D Design: CAD (computer aided design) is an HCI which provides a platform for interpretation and manipulation of 3-Dimensional inputs which can be the gestures. For instance, Huang & Rai [7] developed a novel real-time gesture based conceptual computer aided design tool

- *Juliana A. Abubakar is with the School of Multimedia Technology & Communication, Universiti Utara Malaysia, Malaysia, Kedah 06010. E-mail: liana@uum.edu.my*
- *Sarah N. Zamalik is with the UUM M3DIA Lab, Universiti Utara Malaysia, Malaysia, Kedah 06010. E-mail: sarahnabila232@gmail.com*
- *Abdul Razak Yaakub is with the Centre of Foundation Studies in Management, Universiti Utara Malaysia, Malaysia, Kedah 06010. E-mail: ary321@uum.edu.my*

which enables intuitive hand gesture-based interaction with a given design interface. In the presented system, gestures are identified based solely on the depth information obtained via inexpensive depth-sensing cameras. Another research is done by Jaiswal et al. [9] on gesture-based conceptual computer-aided design (C-CAD) exploration tool for scaled 3D product family models that were obtained by using depth sensing (RGB-D) camera, and also available in large online repository such as Google Warehouse and TurboSquid.

Telepresence: Telepresence robots allow the remote user to control a remote robot avatar, exploring the environment and interacting with people in a more natural way, for instance a telepresence robot developed by Tee et al. [24] that automatically directs attention with speech source localization, head tracking, as well as gesture recognition. Another research is a Telepresence Mobile Robot using a Kinect sensor as the main interface device done by Berri et al. [3]. The proposed gesture recognition method tracks the hands positions and movements, when moving it forward towards the robot, and then recognizing a set of predefined gestures/commands.

Virtual reality: Virtual reality applies to computer-simulated environments that can simulate the physical presence in places in the real world, as well as in imaginary worlds. Most current virtual reality environments are primarily visual experiences, displayed either on a computer screen or through special stereoscopic displays [10]. Some related work was done by Sucar et al. [23] on Virtual Reality-Based Motor Rehabilitation Platform that provide support for rehabilitation principles: promote repetition, task-oriented training, appropriate feedback, and a motivating environment by gesture therapy. In other hand, Liu et al. [14] presents a hand gesture recognition method in interaction of virtual reality. Proved by experiments, the hand gesture segmentation in YC_bC_r color space is very accurate and the HU moments of the hand gesture can well characterize the hand geometric features.

Sign Language: Sign language recognition is a growing research area in the field of gesture recognition. Research on sign language recognition has been performed around the world using many sign languages [26]. Martínez-Camarena et al. [15] presented a method focused on representing each sign by the combination of responses derived from hand postures and hand gestures. The experiments proved that modeling hand gestures by considering spatio-temporal relations between different parts of the body brings improvements over only considering the global trajectories of the hands.

2.2 Gesture Recognition Methods

There are different tools for gesture recognition, based on the approaches ranging from statistical modeling, computer vision and pattern recognition, image processing, connectionist systems, etc. Most of the problems have been addressed based on statistical modeling, such as Artificial Neural Networks (ANN), Histogram Based Feature, Fuzzy Clustering Algorithm and Hidden Markov Model (HMM) [8].

A. Histogram-Based Feature

Table 1 describes the list of projects that employs Histogram Based Feature. These projects aim to improve hand gesture recognition methods using hand segmentation and feature extraction.

TABLE 1. GESTURE RECOGNITION BASED ON HISTOGRAM BASED FEATURE

Author(s)	Objective	Notable Findings
Fang et al. [4]	To develop an accurate and robust hand gesture recognition method.	By combining fast hand tracking, hand segmentation and multi-scale feature extraction. If there is obstruction passes the recognition area, the histogram will deviate and segmentation results will be rapidly degraded.
Zhou et al. [27]	To present hand gesture recognition system based on local orientation histogram feature distribution model.	Proposed feature distribution based recognition algorithm captures the distinct hand shape without requiring a clean segmentation.

Based on [4], it takes advantage of color and motion cues acquired during tracking to implement adaptive hand segmentation. On the basis of segmentation, multi-scale feature extraction is executed and gestures are recognized with palm-finger decomposition.

Meanwhile, Zhou et al. [27] presented hand gesture recognition system based on local orientation histogram feature distribution model. Skin color based segmentation algorithm was used to find a mask for the hand region.

B. Artificial Neural Networks (ANN)

Table 2 describes the list of projects that employs Artificial Neural Networks (ANN). Apart from hand gesture, facial expression may also adopt ANN to recognize human emotion.

TABLE 2. GESTURE RECOGNITION BASED ON ARTIFICIAL NEURAL NETWORKS (ANN)

Author(s)	Objective	Notable Findings
Razuri et al. [19]	To focus on a system of recognizing human emotion through facial expression analysis in merged images based on an Artificial Neural Network.	Face recognition for human emotions using pixel technique.
Hasan [6]	To explore the utility of two feature extraction methods, namely, hand contour and complex moments to solve the hand gesture recognition problem by identifying the primary advantages and disadvantages of each method.	A recognition algorithm to recognize a set of six specific static hand gestures, namely: Open, Close, Cut, Paste, Maximize, and Minimize. The hand gesture image passes through three stages, namely, preprocessing, feature extraction, and classification.

Razuri et al. [19] used a pixeling technique where the original large-size face image is represented by another reduced-size image, keeping the basic features of the face so that the person could even be recognized by this image to recognize human emotions based on face images. In [6], the hand contour is used as a feature which treats scaling and translation problems (in some cases). The complex algorithm is, however, used to describe the hand gesture and treat the rotation problem in addition to the scaling and translation. The results show that the hand contour method has a performance of 71.30% recognition, while complex moments have a better performance of 86.90% recognition rate.

C. Fuzzy Clustering Algorithm

Table 3 describes the list of projects that employs Fuzzy Clustering Algorithm. Hand gesture recognition is used in a healthcare system and for interacting with mobile robot.

TABLE 3. GESTURE RECOGNITION BASED ON FUZZY CLUSTERING ALGORITHM

Author(s)	Objective	Notable Findings
Saha [20]	To recognize 12 health care linked gestures from young individuals of 20-40 years of age group	Recognize 12 gestures arising from pain and stiffness in joint and muscles.
Li, X. [12]	Presented fuzzy c-means clustering algorithm to recognize hand gestures in a mobile remote.	Focus on the image processing stage and how to use the FCM algorithm to do pattern recognition. Found that a one-meter distance between the user and the camera is a good distance for the robot to recognize the hand gesture.

Saha [20] elaborates a method to recognize the early stage gestures related to healthcare in young person using FCM clustering with accuracy of 96.0201% in 0.0439 second obtained using a dataset whose dimension is reduced from 171 to 50 using PCA. Li [12] describes the algorithm, implementation and testing of a static hand gesture recognition system. Experimental results showed that the system satisfies the requirement for a robot implementation and a user-friendly input device. The Fuzzy C-Means algorithm provided enough speed and sufficient reliability to perform the desired task.

D. Hidden Markov Model (HMM)

User interaction techniques are normally associated with Hidden Markov Model (HMM) approach. Table 4 describes the list of projects that employs HMM.

TABLE 4. GESTURE RECOGNITION BASED ON HIDDEN MARKOV MODEL (HMM)

Author(s)	Objective	Notable Findings
Pansare et al. [18]	To introduce an application, "Gestuelle", that recognizes dynamic hand gestures using hidden Markov model to control windows applications.	8 observation symbols have been used in the Markov model. Through experimentation, it realized that there was no need for so many symbols and that 8 symbols were sufficient for performing most rectangular gestures and also, gestures

		involving curves.
Nguyen-Duc-Thanh et al. [17]	To introduce an approach for HRI in gesture recognition for human-robot interaction by using the 2-stages Hidden Markov Model method.	The 2-stage HMM is implemented for the robot to recognize prime gestures and then classify what kind of task the robot should execute.

Results from [18] observed that it is relatively easier to capture the linear gestures, rather than inclined or circular ones, but still the algorithms implemented perform effectively and the overall recognition rate achieved by the system is high because of the use of the Markov model and adequate training. Nguyen-Duc-Thanh et al. [17] performed three kinds of experiments. The first experiment aims to obtain the optimal number of hidden states and mixture components. The second one is to compare the recognition rate of the first stage with that of the other method that uses the *diagonal covariance Gaussian mixture hidden Markov model (DCGM-HMM)*. Lastly, an experiment was conducted for the 2-stage HMM on iRobot Create to classify two kinds of tasks: cleaning and transporting.

2.3 Virtual Automotive Training

Stork et al. [22] presented a technical approach, how to enable large-scale virtual assembly training in and beyond the automotive industry by introducing the VISTRA project. The VISTRA Training Simulator (VTS) represents an interactive virtual assembly simulation, which is used to train and test manual assembly processes. The VTS provides the actual training functionality to the trainees. It communicates internally to the VKP in order to request and restore training content. To enhance the acceptance and hence the learning effect of the intended virtual training system for procedural assembly knowledge, they propose a game-based approach allowing natural user interaction as Fig. 1 below.

Fig. 1. Natural interaction for grabbing a tool. Image courtesy from [22].

The basic interaction prototype employs a *Microsoft Kinect* camera and a *WiiMote* controller to allow for easy-to-use and easy-to-learn user interaction. The features and information provided by the *Microsoft Kinect* and the corresponding SDK enable the system to capture and track the body posture and actions of the user and directly reflect these within the virtual scene. In combination with a handheld *WiiMote* controller the user can directly manipulate virtual objects using the corresponding real-world actions performed during the assembly process, e.g. rotating and twisting an object. In addition the device provides a means for triggering "virtual" actions, e.g. button presses, as well as acoustic and haptic feedback through the included actuators.

Research and development using gesture recognition are mainly conducted in the application of virtual training in areas such as aircraft maintenance [2], machine operations [13], surgical operations [11, 21], and the military [5]. There have been virtual training systems which were developed to aid the acquisition of procedural skills related to assembly tasks. Most systems are aimed at supporting the training of maintenance tasks in which knowledge of both assembly and disassembly is part of what is acquired during the training of the tasks.

One key challenge to address when developing virtual automotive training is the extensive authoring effort for setting up virtual environments. Although knowledge of the product and manufacturing design is available and could be used for virtual training, a concept for integration of this data is still missing. It is also unlikely to change the automotive materials. Since technology in the automotive industry is changing rapidly, training facilities must keep up with the pace of the technology. The rate of introduction of any automotive gesture recognition system is more likely to be dictated by the rate of user acceptance and not the timing of the technical issue resolution. Also, it is proposed to have interactive applications for tackling classroom applications in education. Most of the model design systems are more focused in the areas of training, medicine, architecture and computer games as it is used for commercial purposes.

3 Conclusion & Future Work

Building an efficient human-machine interaction is an important goal of the gesture recognition system. Many applications of gesture recognition system ranging from virtual reality to sign language recognition and robot control. In this paper, review of applications and techniques of gesture recognition system have been provided with emphasis on hand gesture expressions. The major techniques include HMMs, ANN, and fuzzy clustering have been reviewed and analyzed. Future works will focus on using symbols or gestures used for virtual automotive training system and also the use of such system in classroom for education purposes. Although the presented solution is closely oriented towards the needs of the automotive sector, the virtual training system is anticipated to be effectively used in other related industries with similar product complexity and production structure, such as the aerospace and monorail industry.

This study is expected to help educators, instructional designers and software developers that interested in providing virtual technology to students or users for the future. In summary, results from this study can contribute an up-to-date references to the used of gesture recognition in learning and training purposes. In addition, the results of this study can create awareness about the gesture recognition potential in current technology for educational purposes.

ACKNOWLEDGMENT

The authors would like to express their sincere appreciation to Ministry of Education Malaysia for partly funding this project under Fundamental Research Grant Scheme, FRGS (S/O Code : 12897).

REFERENCES

[1] Association, M. A., "Summary of Sales and Production Data," http://www.maa.org.my/info_summary.htm. 2011.

[2] Barnett, B., Helbing, K., Hancock, G., Heininger, R., & Perrin, B., "An evaluation of the training effectiveness of virtual environments," Proc. Interservice/Industry Training, Simulation & Education Conference (I/ITSEC), 2000.

[3] Berri, R., Wolf, D., & Osorio, F., "Telepresence Robot with Image-Based Face Tracking and 3D Perception with Human Gesture Interface Using Kinect Sensor," Proc. Robotics: SBR-LARS Robotics Symposium and Robocontrol (SBR LARS Robocontrol), 2014.

[4] Fang, Y., Wang, K., Cheng, J., & Lu, H., "A real-time hand gesture recognition method," Proc. Multimedia and Expo, IEEE International Conference, 2007.

[5] Gerbaud, S., Molle, N., Ganier, F., Arnaldi, B., & Tisseau, J., "GVT: a platform to create virtual environments for procedural training," Proc. Virtual Reality Conference VR'08, IEEE, 2008.

[6] Hasan, H. S., "Static hand gesture recognition using artificial neural network," University of Malaya, 2014.

[7] Huang, J., & Rai, R., "Hand Gesture Based Intuitive CAD Interface" Proc. ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering, 2014.

[8] Ibraheem, N. A., & Khan, R., "Survey on various gesture recognition technologies and techniques," International Journal of Computer Applications, vol. 50, no. 7, pp. 38-44, 2012.

[9] Jaiswal, P., Bajad, A. B., Nanjundaswamy, V. G., Verma, A., & Rai, R., "Creative Exploration of Scaled Product Family 3D Models Using Gesture Based Conceptual Computer Aided Design (C-CAD) Tool" Proc. ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2013.

[10] Jakkula, V., "Tutorial on support vector machine (svm)," School of EECS, Washington State University, 2006.

[11] Larsen, C. R., Soerensen, J. L., Grantcharov, T. P., Dalsgaard, T., Schouenborg, L., Ottosen, C., Shroeder, T. V., Ottesen, B. S., "Effect of virtual reality training on laparoscopic surgery: randomized controlled trial," vol. 338, pp. b1802, 2009.

[12] Li, X., "Gesture recognition based on fuzzy C-Means clustering algorithm," Department Of Computer Science, The University Of Tennessee Knoxville, 2003.

[13] Lin, F., Ye, L., Duffy, V. G., & Su, C.-J., "Developing virtual environments for industrial training," Information Sciences, vol. 140, no. 1, pp. 153-170, 2002.

[14] Liu, Y., Yin, Y., & Zhang, S., "Hand gesture recognition based on HU moments in the interaction of virtual reality," Proc. Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2012 4th International Conference, 2012.

[15] Martínez-Camarena, M., Mogrovejo, O., Antonio, J., & Tuytelaars, T., "Towards Sign Language Recognition based on Body Parts Relations," Proc. ICIP, 2015.

[16] Mitra, S., & Acharya, T., "Gesture recognition: A survey," IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 37, no. 3, pp. 311-324, 2007.

[17] Nguyen-Duc-Thanh, N., Lee, S., & Kim, D., "Two-stage hidden Markov model in gesture recognition for human robot interaction," International Journal of Advanced Robotic Systems, vol. 9, no. 39, 2012.

[18] Pansare, J., Bansal, M., Saxena, S., & Desale, D., "Gestuelle: A system to recognize dynamic hand gestures using hidden Markov model to control windows applications," International journal of computer applications, vol. 62, no. 17, pp. 19-24, 2013.

[19] Razuri, J. G., Sundgren, D., Rahmani, R., & Moran Cardenas, A., "Automatic emotion recognition through facial expression analysis in merged images based on an Artificial Neural Network," Proc. Artificial Intelligence (MICAI), 2013 12th Mexican International Conference, 2013.

[20] Saha, S., "A fuzzy C means clustering approach for gesture recognition in healthcare," 2014.

[21] Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., & Satava, R. M., "Virtual reality training improves operating room performance: results of a randomized, double-blinded study," Annals of surgery, vol. 236, no. 4, pp. 458, 2002.

[22] Stork, A., Sevilmis, N., Weber, D., Gorecky, D., Stahl, C., Loskyll, M., & Michel, F., "Enabling virtual assembly training in and beyond the automotive industry," Proc. Virtual Systems and Multimedia (VSMM), 2012 18th International Conference, 2012.

[23] Sucar, L. E., Orihuela-Espina, F., Velazquez, R. L., Reinkensmeyer, D. J., Leder, R., & Hernandez-Franco, J., "Gesture therapy: An upper limb virtual reality-based motor rehabilitation platform," Neural Systems and Rehabilitation Engineering, vol. 22, no. 3, pp. 634-643, 2014.

[24] Tee, K. P., Yan, R., Chua, Y., Huang, Z., & Liemhetcharat, S., "Gesture-based attention direction for a telepresence robot: Design and experimental study," Proc. Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference, 2014.

[25] Wahab, D., Amelia, L., Hooi, N., Haron, C. C., & Azhari, C., "The application of artificial intelligence in optimization of automotive components for reuse," Journal of Achievements in Materials and Manufacturing Engineering, vol. 31, no. 2, pp. 595-601, 2008.

[26] Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., & Presti, P., "American sign language recognition with the Kinect," Proc. of the 13th international conference on multimodal interfaces, 2011.

[27] Zhou, H., Lin, D. J., & Huang, T. S., "Static hand gesture

recognition based on local orientation histogram feature distribution model," Proc. Computer Vision and Pattern Recognition Workshop, CVPRW'04, 2004.

Juliana A. Abubakar received her B.Eng. in Electronic Engineering from University of Leeds and MSc. in Information Technology from University Utara Malaysia. She received her Ph.D from International Islamic University Malaysia. She is currently a Senior Lecturer at the School of Multimedia Technology & Communication, Universiti Utara Malaysia. She has won gold medals for innovative products in numerous exhibitions at national and international level and successfully secured handsome amount of national grants for the past five years. Her research interests include virtual reality design, development and evaluation; augmented reality; and virtual heritage.

Sarah N. Zamalik received her Bsc in Multimedia Computing from Universiti Teknologi Mara (UiTM) Shah Alam, Malaysia. Currently she is a Masters student at the School of Multimedia Technology & Communication, Universiti Utara Malaysia. She is a member of M3DIA Research Lab at School of Multimedia Technology & Communication, Universiti Utara Malaysia. Her research interests include virtual reality, gesture-based recognition; and virtual automotive training.

Abdul Razak Yaakub obtained his bachelor degree in Mathematics with Education from Universiti Teknologi Malaysia in 1985 and by master degree of Numerical Analysis, Department of Mathematics, Faculty of Science, University of Technology Malaysia (UTM), Kuala Lumpur, Malaysia in 1989. Followed by PhD degree of Scientific Computing from Department of Computer Science, Loughborough University of Technology, United Kingdom in 1996. Now, he is a Professor of School of Computing, College of Arts and Sciences, Universiti Utara Malaysia. He is a member of Scientific Computing , Virtual Reality and Virtual Environment research group. His research interests include mobile application technology, decision support system, numerical analysis and mathematical sciences.