International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Volume 3, Issue 6 (November-December, 2015), PP. 216-221

GLOVE BASED AND ACCELEROMETER BASED GESTURE CONTROL: A LITERATURE REVIEW

Priya Matnani
OTH (Department of Information Technology)
Maharashtra, India.
matnanipriya@gmail.com

Abstract— **This paper reviews the technology for using hand, body and facial gestures as a means for interacting with computers and other physical devices. It discusses the rationale for gesture based control technology, methods for acquiring and processing such signals from human operators, applications of these control technologies, and anticipated future developments. Today, there is a growing interest in research and development of new human-machine interaction systems that are more natural and ergonomic for the users. The gesture recognition plays an important part of human-machine interaction systems. The focus is done in systems that are based on accelerometers, and on glove based equipments. Based on several papers, the process for different types of approach of gesture control is described. Some applications of these technologies are also presented.**

Index terms- **Gesture, control, recognition, accelerometer, glove-based.**

I. INTRODUCTION

There have been growing attempts in finding natural ways for human–machine interaction (HMI) for multimedia entertainment [18]. That is because traditional ways, e.g., by using a mouse or a keyboard, are inherently limited by speed and space, and lack of an immersed sense within the simulated environments [19]. In recent years, gesture recognition for HMI has become popular because it helps overcome the limitations of traditional ways and enriches personal experience during the interaction between human and machine for entertainment [3],[20]. Gesture recognition is also important in automated surveillance [14] and human monitoring applications [15], where they can yield valuable clues into human activities and intentions [1].

Gestures are defined as human motion sequences with trajectories [11] performed in a short interval of time. Human gestures are a natural means of interaction and communication among people. Gestures employ hand, limb and body motion to express ideas or exchange information non- verbally [3].

Gestures can be divided into two groups: static and dynamic gestures [12]. According to input device, the gesture recognition technique can be divided into three categories: glove-data based, vision- based and accelerometer-based [3].

Glove-data based gesture recognition systems require users to wear gloves and cumbersome devices to record the movement status [3]. The performance of vision-based systems is relied heavily on the operation environment, e.g., the background and the lighting condition. In addition, these systems face the occlusion problem, let alone the low sampling rate issue.

With the rapid development of micro– electro-mechanical system technology (MEMS), the accelerometer-based gesture recognition becomes increasingly popular [13], [3] and already shows potential in practical applications.

This paper investigates the research works of gesture controlled technology for user interactions. Gesture type, use of different parts of the body, gesture commands, chronological evolution, gesture application, interface, technology, user type, issues addressed, tasks and final result have been listed and described to give the background of gesture based technology development.

II. GESTURE CONTROLLED SYSTEM

Humans naturally use gesture to

Paper	Primary method of recognition	Number of gestures recognised	Background to gesture images	Additional markers required (such as wrist band)	Number of training images	Accuracy	Frame rate
[Bauer & Hienz, 2000]	Hidden Markov Models	97	General	Multi-coloured gloves	7-hours signing	91.7%	-
[Starner, Weaver & Pentland, 1998]	Hidden Markov Models	40	General	No	400 training sentences	97.6%	10
[Bowden & Sarhadi, 2000]	Linear approximation to non-linear point distribution models	26	Blue screen	No	7441 images	-	-
[Davis & Shah, 1994]	Finite state machine / model matching	7	Static	Markers on glove	10 sequences of 200 frames each	≈98%	10

Table 1. Existing gesture recognition system

communicate. It has been demonstrated that young children can readily learn to communicate with gesture before they learn to talk. A gesture is non-verbal communication made with a part of the body [7]. We use gesture instead of or in combination with verbal communication. Using this process, humans can interface with the machine without any mechanical devices. Human movements are typically analyzed by segmenting them into shorter and understandable format [7]. The movements vary person to person. It can be used as a command to control different devices of daily activities, mobility etc. So our natural or intuitive body movements or gestures can be used as command or interface to operate machines, communicate with intelligent environments to control home appliances, smart home, telecare systems etc [7]. In this paper we also review the different types of technologies of gesture controlled system.

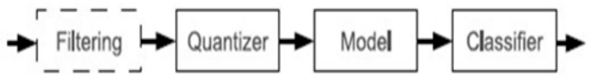

Figure 1: The process of gesture recognition with Accelerometer based device.

A. Existing Systems

A simplification used in one of the projects, which was not found in any recognition methods researched, is the use of a wrist band to remove several degrees of freedom.

This enabled three new recognition methods to be devised. The recognition frame rate achieved is comparable to most of the systems in existence (after allowance for processor speed) but the number of different gestures recognised and the recognition accuracy are amongst the best found. Table 1 shows several of the existing gesture recognition systems along with recognition statistics and method.

With the rapid development of micro– electro-mechanical system technology (MEMS), the accelerometer-based gesture Recognition becomes increasingly popular and already shows potential in practical applications.

III. ACCELEROMETER BASED

In this section we focus on devices that use accelerometers to measure movements. In [4] and [5], a SoapBox is used as accelerometer based device. In [6], a Wii Controller (Wiimote) is used as accelerometer based device.

A. Overview

In [6], the famous Wiimote from Nintendo is used for input of the user movements. This device is the main controller for the Wii console. An accelerometer in the controller is responsible to measure acceleration along three axes. An extension that contains a gyroscope can be added to the controller to improve rotation motions. The controller also contains an optical sensor allowing to determine where it is pointing. For that, a sensor bar highlighting IR LEDs is used. The connectivity of the controller is done via Bluetooth.

In the two others papers [4] and [5], a SoapBox is used. SoapBox is defined as Sensing, Operating and Activating Peripheral Box. This box is small and has low power consumption. It is equipped with 3-axis accelerometer, an illumination sensor, a electronic compass and a optical proximity sensor. For communication purpose, it is wireless with RF technologies. All the development can be written in C, utilizing the API offering.

In both papers, it is the 3-axis accelerometer component that was responsible to measure the movements.

The types of gestures to be recognized are for example: the trace of a square, a circle, a Z.

B. Process:

There is a system allowing the training and recognition of arbitrary gestures with the use of 3-axis accelerometers. With this type of device, we must deal with spatial and temporal data. We need a mathematical process to exploit these signals. Gestures are represented with data vectors representing the current acceleration of the controller in 3-axis. Theses vectors are analyzed to train and to recognize patterns for distinct gestures [9].

Figure 2: The process of gesture recognition with accelerometer based device. The process consists of 3 phases as seen in figure 2. First we have the "Quantizer" that is responsible to cluster the data using a k-mean algorithm. Then, the "model" is a discrete HiddenMarkovModel (HMM) that is used to train/recognize characteristic patterns for distinct gestures[9]. Finally, a "Bayesclassifier" is used to select the appropriate gesture.

Before these 3 phases, a filtering is applied on data for simplification purpose.

Filtering

In [6], two filters are applied to the vector data to get a minimum representation of a gesture. The first filter is a simple threshold eliminating all vectors which do not have a significant contribution to the characteristic of a gesture.

The second filter is responsible to eliminate all vectors which are roughly equivalent to their previous.

Quantizer

Acceleration sensor produces too much data. Before putting them in a HMM, we need to cluster and abstract them. A k-means clustering method is used to partition the n observations into k clusters. Each observation belongs to the cluster with the nearest mean. The k is the number of clusters, their collection forms the codebook.

In [6], true 3D gestures are evaluated with a k of size 14 which was found empirically.

In [4] and [5], the size of the codebook k is 8 which were also found empirically. But, in contrast to [6], 3D data are converted to 1D vector.

Model and Classifier

The gesture recognition system work in two phases: training and recognition. The training consists of several repetitions of each gesture that must be recognized later.

In [4], [5] and [6] a discrete Hidden Markov Model is used for the training and the recognition of gestures. A HMM is stochastic signal modeling method in which the system being modeled is assumed to be a Markov process with unobserved states. The global structure of the recognition system is composed with trained HMM for each gesture. The classification is done with a naive Bayes classifier that is a simple probilistic classifier [9].

Another famous method used is also taken up in this paper.

IV. GLOVE BASED

In [21] and [22] a gesture recognition based on gloves is described

A. Overview

A data glove is a glove-like input device often used for virtual reality environments. It is equipped with various technologies such as a system for detection of bending of fingers. Often a motion tracker is attached to capture the global position/rotation data of the glove.

In [21], a P5 Glove from Essential reality was used. It is an inexpensive (∽ 50 Euro) glove with integrated 6 DOF tracking designed as a game controller. 6 DOF means six degrees of freedom, in fact the ability to move forward/backward, up/down, left/right (translation in three perpendicular axes) combined with rotation about three perpendicular axes (pitch, yaw, roll). The glove consists of five bend sensors to track the flexion of the wearer’s fingers. An infrared-based optical tracking system is used to compute the glove position and orientation without the need for additional hardware. The glove is connected with a cable to the base station.

In [22], their gesture recognition system is based on two different components. First, two ”Cyber Glove” from Virtual Technologies are used, for each hand. This glove has flexible sensors that measure the position and movement of the fingers and wrist. Then, five ”Flock of Birds” from Ascension Technology Corporation are used for six degrees-of-freedom tracking. This 6DOF tracking system sensor is based on magnetic technology. Note that the electro-magnetic field is distorted by metallic objects. Two are attached to the mounting point of each glove. Another one is mounted on a light-weight helmet worn by the user. The two others are attached to the subjects upper arms to register the position and orientation of the humerus.

B. Process

In [21], an important aspect is that a gesture is seen as a sequence of successive postures. Postures in the recognition engine are composed of the flexion values of the fingers, the orienation data of the hand and an additional value to indicate the relevance of the orientation for the posture. These postures are taught to the system by simply performing them, then associating an identifier with the posture.

The recognition engine is divided into two components: the data acquisition and the gesture manager [9].

Data acquisition

The data acquisition component is responsible for processing the received data and then transmit them to the gesture manager. First, a set of filter is used to optimize the data [9].

For example, the position/orientation information is very noisy due to dependance of lighting conditions. Thus, orientation data that exceed a given limit are discarded as improbable and replaced with their previous values. This type of filters are applied: deadband filter, dynamically adjusting average filter. Note that to be recognize as a posture, the user has to hold a position between 300 and 600 milliseconds in order to allow the system to detect a posture.

Gesture manager

The gesture manager is the principal part of the recognition system. This library maintains a list of known postures. The system try to match incoming data with existing posture. This is done by first looking for the best matching fingers constellation. Five dimensional vectors represent the bend values of the fingers and for each posture definition the distance to the current data is calculated. Then, the position/orientation data is compared in a likewise manner.

Finally, in this gesture recognition system, a gesture is just a sequence of successive postures. For example, let’s consider the detection of a ”click” gesture. This gesture is defined as a pointing posture with outstretched index finger and thumb and the other fingers flexed, then a tapping posture with half-bent index finger.

V. APPLICATION EXAMPLES

A. TELEOPERATION AND ROBOT CONTROL

Remote manipulation of objects because of weight or exposure risks, e.g., radioactivity, has been performed for many years using direct mechanical linkages or electric motors that permit force amplification. Even though these systems do not actually include a computing system, they involve the transmission of gestural information. In that respect they are forerunners of a number of object manipulation applications.

Hale [23] used a DataGlove to control a robot arm in a task that required retraction, slewing and insertion of a block in a test panel. He compared his results to another study that used a conventional six degree- of-freedom hand controller as the input device. He concluded that performance with the DataGlove compared favourably with the “standard” device and that it provided a natural and intuitive user interface.

Brooks [24] on the other hand, was less optimistic about the DataGlove for robot control; his evaluation involved more complex gestures and a neural network for gesture recognition. The reader should recall, however, that the DataGlove is very limited for precise manipulation tasks.

B. IRTUAL AND AUGMENTED REALITY

Several examples are provided by 2-D and 3-D displays in which the user can touch, grab and move objects by pantomiming these activities with glove-based sensors.

In these applications the user actually sees a computer rendering of their hand performing the object manipulations. Researchers at NASA/Ames have used this approach in a virtual wind tunnel to explore simulations of computational fluid dynamics.

Aeronautical engineers can put their hands and head into a simulated fluid flow and manipulate the patterns in real time [25].

The GROPE project at the University of North Carolina [26], [27] is among the first applications to use force feedback for interacting with a computer simulation. The application domain is the simulation and graphical representation of interactions between complex molecules. A specitically developed force feedback manipulating rod, allowing six degree-of- freedom movement is employed. When the user modifies the simulated position of one molecule by moving the rod, the simulation computes intermolecular forces and reflects them back through the feedback system. As a result of the computational time needed for the simulation, the system produces relatively low fidelity sensations. Nevertheless, this system allows one to begin to explore possible chemical bonds between molecules.

If an application requires the user to manipulate virtual objects in some way, accuracy of depth perception becomes an issue, particularly for computer-generated displays that lack the rich textural cues available in real life.

C. SIGN-LANGUAGE INTERPRETATION

Sign language interpretation continues to be a significant area for gesture research and development. This type of application is not the focus of this lecture and we will touch on it only briefly. Fels [28] and Hinton developed a hand gesture to speech system using a neural network. Their system mapped hand postures to complete root words, followed by a directional hand movement that modified the word ending (singular, plural, etc.) and controlled speech rate and emphasis.

Performance of a single "speaker" with a vocabulary of 203 words was evaluated following a network training phase. With near real-time speech output, the wrong word was produced less than 1 percent of the time and no word was generated approximately 5 percent of the time

D. COCKPIT APPLICATIONS

Ineson, Parker and Evans[29] compared a video-based finger tracker with several other designation mechanisms to select buttons on a virtual, head-down panel during simulated low level flight. Feedback for contact with the button was a colour change. Activation of the button required depressing a switch on the Hands- On Throttle and Stick (HOTAS) for confirmation.

The finger-tracker was poorly rated by the subjects since it removed the hand from the flight controls for a substantial period of time. Some subjects found the device awkward to use since it was necessary to keep the finger in clear view of the tracking cameras. Although the normal means of operating a button is to reach out and press it, the task is essentially two dimensional. Methods such as head pointing and stick-top cursor controllers are suitable mechanisms also, and both were preferred to the finger tracker. Finger pointing direction would have been more suitable than finger position, since it could have been operated with the hand on, or near, the controls. Voice control was the overwhelmingly preferred selection technique for this task.

A series of experiments carried out at Wright-Patterson Air Force Base, Ohio, USA required true 3-D selection of targets from a head-down, 3-D tactical map. In this case an electromagnetic tracker was strapped to the back of the hand, resulting in more robust and responsive tracking than the video based technique used by Ineson et al. The tracking volume was remote from the actual map so that hand movements were actually made in a space close to the aircraft controls rather than within the volume of the map. This hand volume was reduced in scale so that hand movements were small compared to the size of the map. The volume was divided into four depth planes, so accurate depth control was not required. The hand tracker worked well and was, in general, faster and more accurate than a three-dimensional joystick. If the tracking volume was made too small, selection accuracy was impaired.

VI. CONCLUSIONS AND FUTURE SCOPE

Technologies developed based on gesture are now really affordable and converged with familiar and popular technologies like TV, large screen. It's ubiquitous and non-intrusive as we can install a camera or remote with the TV. From this paper we can see the trends of gesture controlled communication systems. Easing of the technology use, affordability and familiarity indicate that gesture based user interface can open new opportunity for elderly and disable people. The older population (65+) numbered 36.3 million in 2004, an increase of 3.1 million or 9.3% since 1994 and it's growing over time [3]. There will be more elderly people and fewer younger ones to care for them. So we need to invest much more heavily in Assistive Living solutions. The research 'A gesture controlled communication aid for elderly and disabled people' can be a significant task for future**.** The two important aims of the research are to identify the different gestures of elderly and disabled people for communication and to design a rich augmented-reality interface for communication via ubiquitous device such as a television set.

REFERENCES

[1] Cuong Tran, Student Member, IEEE, and Mohan Manubhai Trivedi, Fellow, IEEE." 3-D Posture and Gesture Recognition for

Interactivity in Smart Spaces" . IEEE transactions on industrial informatics, VOL. 8, No. 1, February 2012.

[2] S. Sidney Fels and Geoffrey E. Hinton. "Glove-TalkII". IEEE transactions on neural networks, VOL. 9, No. 1, January 1998 205.

[3] Jun Cheng a,b,c, Wei Bian d, Dacheng Tao d,⇑ ." Locally regularized sliced inverse regression based 3D hand gesture Recognition on a dance robot". Information Sciences 221 (2013) 274–283.

[4] J. Kela, P. Korpipaa, J. Maentyjaervi, S. Kallio, G. Savino, L. Jozzo, and S. Marca. Accelerometer-based gesture control for a design environment. Springer. Personal And Ubiquitous Computing, 10(5):285–299, 2006.

[5] J. Maentyjaervi, J. Kela, P. Korpipaa, and S. Kallio. Enabling fast and effortless customisation in accelerometer based gesture interaction. In Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia, pages 25–31. ACM, 2006.

[6] T. Schloemer, B. Poppinga, N. Henze, and S. Boll. Gesture recognition with a wii controller. In Proceedings of the Second International Conference on Tangible and Embedded Interaction (TEI'08). ACM, 2008.

[7] Moniruzzaman Bhuiyan and Rich Picking ." Gesture-controlled user interfaces". Centre for Applied Internet Research (CAIR), Glyndŵr University, Wrexham, UK

[8] Barinder Pal Singh Ahluwalia MCA Scholar Institute of Information Technology and Management. "Gestural Interface Interaction: A Methodical Review". International Journal of Computer Applications (0975 – 8887) Volume 60– No.1, December 2012

[9] Damien Zufferey Department of Informatics (DIUF) University of Fribourg ." Device based gesture recognition"

[10] Grant R. McMillan Air Force Research Laboratory (AFRL/HECP). "The technology and applications of Geture-based control"

[11] Z. Zeng, J. Tu, M. Liu, T.-S. Huang, B. Pianfetti, D. Roth, S. Levinson, Audio– visual affect recognition, IEEE Trans. Multimedia 9 (2) (2007) 424–

[12] S. Mitra, T. Acharya, Gesture recognition: a survey, IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev. 37 (3) (2007) 311–324

[13] Z.F. Walter, S. Jones, D. Tjondronegoro, Detecting gesture force peaks for intuitive interaction, in: Proc. the 5th Australasian Conf. Interactive Entertainment, vol. 391, no. 2, 2008.

[14] M. M. Trivedi, T. L. Gandhi, and K. S. Huang, "Distributed interactive video arrays for event capture and enhanced situational awareness," IEEE Intell. Syst., vol. 20, no. 5, pp. 58–66, Sep./Oct. 2005.

[17] T. Schloemer, B. Poppinga, N. Henze, and S. Boll. Gesture recognition with a wii controller. In Proceedings of the Second International Conference on Tangible and Embedded Interaction (TEI'08). ACM, 2008.

[18] V. Pavlovic, R. Sharma, T.-S. Huang, Visual interpretation of hand gestures for human–computer interaction: a review, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 677–695.

[19] L.L. Scarlatos, TICLE: using multimedia multimodal guidance to enhance learning, Inform. Sci. 140 (1–2) (2002) 85–103.

[20] M.-G. Cho," A new gesture recognition algorithm and segmentation method of Korean scripts for gesture" – allowed ink editor, Inform. Sci. 176 (9) (2006) 1290–1303.

[21] M. Deller, A. Ebert, M. Bender, and H. Hagen. Flexible gesture recognition for immersive virtual environments. In Tenth International Conference on Information Visualization (IV 2006), pages 563–568. IEEE, July 2006.

[22] M. Froehlich and I. Wachsmuth. Gesture recognition of the upper limbs - from signal to symbol. I. Wachsmuth and M. Froehlich (eds.): Gesture and Sign Language in Human-Computer Interaction. Berlin: Springer-Verlag (LNAI 1371), pages 173–184, 1998.

[23] Hale, J.P., "Anthropometric Teleoperation: Controlling Remote Manipulators with the DataGlove", NASA-TM-103588 (NTIS: N92- 28521/2/XAB), 1992.

[24] Brooks, M., "The DataGlove as a Man-Machine Interface for Robotics", in "2nd IARP Workshop on Medical and Healthcare Robotics", Newcastle upon Tyne, United Kingdom, 1989, pp. 213-225.

[25] Bryson, S., and Levit, C., "The Virtual Wind Tunnel", IEEE Computer Graphics and Applications, 12, 4, 1992, pp 25-34.

[26] Brooks, F.P., "Grasping Reality Through Illusion: Interactive Graphics Serving Science", in "Human Factors in Computing Systems - Proceedings of CHI '88", 1988, pp l-1 1.

[27]Brooks, F.P., Ouh-Young, M., Batter, J.J., and Kilpatrick, J., "Project GROPE: Haptic Displays for Scientific Visualization", Computer Graphics, 24, 4, August 1990.

[28] Fels, S.S., and Hinton, G.E., “Glove- Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer”, IEEE Transactions on Neural Networks, 4, 1, 1993, ~~2-8.

[29] Ineson, J., Parker, C.C., and Evans, A., “A Comparison of Head-Out and Head-In Selection Mechanisms During Simulated Flight”, DERA Customer Report DERAIASISIDIS 1 O/CR971 53,1997.