Gesture as a Means of Human-Friendly Communication between Man and Machine

Zeungnam Bien

Dept. of Electrical Engineering and Computer Science, KAIST
373-1, Kusong-dong, Yusong-gu, Taejon 305-701, Korea
Tel: +82-42-869-3419, Fax: +82-42-869-8410
E-mail: zbien@ee.kaist.ac.kr

Abstract: In this paper, the 'gesture' is discussed as a means of human-friendly communication between man and machine. We classify various gestures into two categories: 'contact based' and 'non-contact based'. Each method is reviewed and some real applications are introduced. Also, key design issues of the methods are addressed and some contributions of soft-computing techniques, such as fuzzy logic, artificial neural networks (ANN), rough set theory and evolutionary computation, are discussed.

1. Introduction

Webster's Dictionary defines gesture as "a movement, or movements collectively, of body, or of part of body, to express or emphasize ideas, emotions, etc [1]." "Gesture" is getting a lot of attention due to its communicational aspect via body movements. Psychological and social studies tend to narrow this broad definition and relate it even more to human's expression and social interaction.

Even though gesture plays a very important role with communicational aspect, the notion of gesture is also used in a somewhat different way in the domain of Human Computer Interaction (HCI). In a computer controlled environment, one wants to use human hands to perform tasks that mimic both the natural use of the hand as a manipulator, as well as its use in human-computer communication (control of computer/machine functions through gestures). Classical definitions of gestures, on the other hand, are rarely concerned with manipulative feature of gesture.

Gesture is surely a means of communication, similar to spoken language. A psychologist [2] reported that, for ordinary human-to-human communication, only 7% of a message is transmitted by linguistic language, while 35% is transmitted by paralanguage such as gesture and voice tone, and 55% by facial expression, as shown in Figure 1.

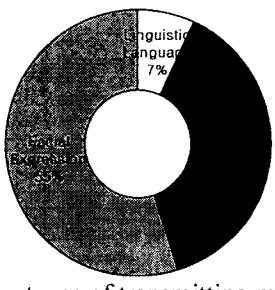
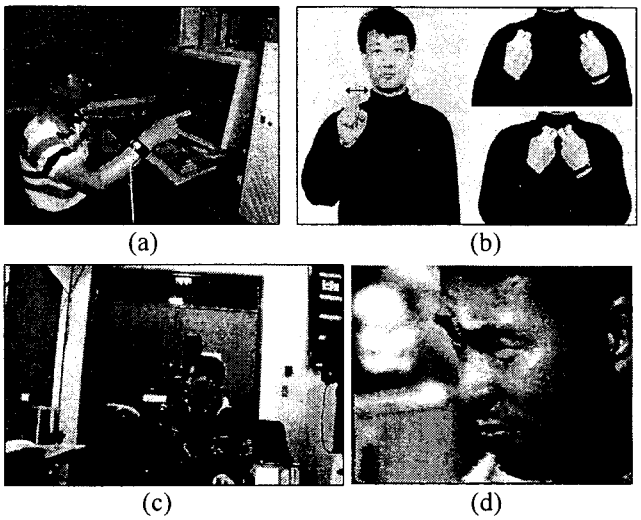


Figure 1. Importance of transmitting message by the nonverbal language

There are many kinds of gestures. They include virtual reality interaction gesture, sign language, pointing gesture, and facial emotional expression as shown in Figure 2.



(a) (b) (c) (d)

Figure 2. Kinds of Gestures: (a) Virtual Reality, (b) Sign Language, (c) Pointing Gesture, (d) Facial Emotional Expression

Several alternative taxonomies have been suggested in the literatures that deal with psychological aspects of gestures. The taxonomy that seems most appropriate within the context of HCI is the one which is recently developed by Quek [3]. A slightly modified version of the taxonomy is given in Figure 3.

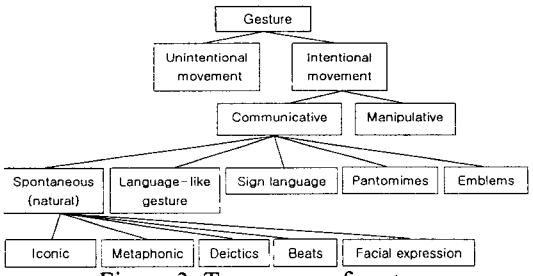


Figure 3. Taxonomy of gesture

According to Quek [3], all gestures are first classified into two major classes; intentional and unintentional movements. Unintentional movements are those hand/arm movements that do not convey any meaningful information. Intentional movements themselves can have two modalities; communicative and manipulative. For example, in Figure 2, (b) and (d) correspond to communicative gestures, whereas (a) and

(c) correspond to manipulative gestures.

Manipulative gestures are the ones used to act on objects in an environment (object movements, rotation, etc.). Communicative gestures, on the other hand, have an inherent communicational purpose. In this paper, we examine gestures with communicative and/or manipulative aspects with emphasis on communicative aspect. And we can divide gestures into one class with contact based interface and another with non-contact based interface. It depends on the form of interface when it is implemented. Our presentation will follow this categorization.

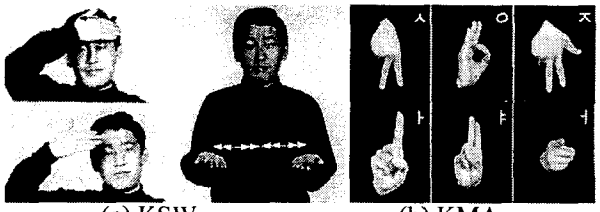
This talk is organized as follows. In Section 2, the contact based interfaces are considered. In Section 3, the non-contact based interfaces are examined. Finally, in Section 4, we present possible applications of gestures , and make a summary.

2. Contact based interface

There are many kinds of contact based interfaces for implementing gestures. In this paper, we consider two types of them. One is a wearable device type and the other is a patch type. We may say that CyberGlove-based system is a representative interface of wearable device type.

The Korean Sign Language (KSL) interpretation system for hearing-impaired [4] is an important application that uses CyberGlove as a main interface. And it is a communicative gesture-based application, if classified according to the taxonomy given in Figure 3.

Standard Korean Sign Language (KSL) is composed of both the Korean Sign Words (KSW) and the Korean Manual Alphabet (KMA) as shown in Figure 4.



(a) KSW (b) KMA

Figure 4. KSL

We have developed a KSL Recognizer that can recognize about 500 KSW and 31 KMA in real time with 92% accuracy. It uses CyberGlove and a camera vision for input device and Indigo2 workstation for the main computer program as shown in Figure 5.

Figure 5. System configuration of KSL Recognizer

The gesture of KSL is decomposed into four basic types of elements named as phoneme. These are hand motion-direction, hand posture, hand location, and hand orientation. The hand motion-direction is concerned with hand trajectory that a KSW shows from the starting point to the ending point in 3 dimensional space. Hand posture refers to geometric hand shape. Hand location indicates the part of the signer's body that he/she points to with his/her hands. Hand orientation is the rotational direction in 3 dimensional space.

KSL phoneme was studied by Kim and Suk [5]. We have established the four basic elements by analyzing KSL dictionaries [6]. We can classify KSL using these basic elements. And various soft computing techniques are used for this system such as fuzzy logic, ANN, etc.

Based-on this analysis, we can make a communication tool between man and machine more effectively. Users can express their intention using gesture, and machines can express or respond user's intention by generating gesture with 3 dimensional animated character as shown Figure 6.

Figure 6. KSL generator using 3 dimensional animated character and some generated facial expressions.

The EMG signal-based robotic arm control system [7] is another contact based interface for manipulative purpose. And as stated earlier, this system uses a patch type interface. EMG signal pattern recognition is considered a very difficult problem because there are many artifacts in the bio-signal measurement. For this system, soft computing techniques that can deal with uncertainties, ambiguous quantities have been used for effective human natural gestures.

8 primitive motions are defined for recognizing human's basic 4 degree-of-freedom movements as shown in Figure 7. Combining these motions, the user can manipulate a robotic arm for picking an object on the floor or on the table as shown in Figure 8.

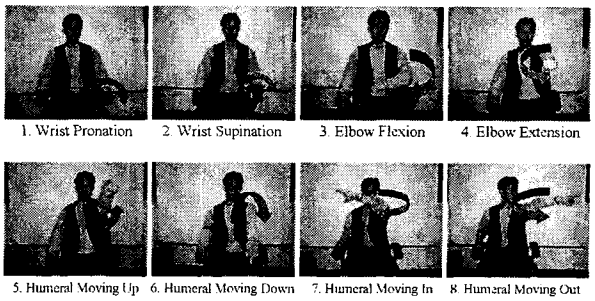


Figure 7. Pre-defined 8 primitive motions

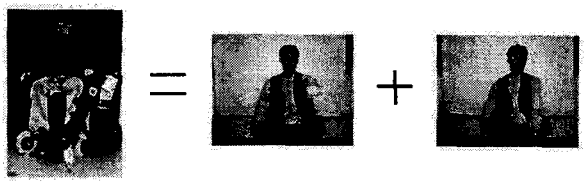


Figure 8. Performing a task for picking an object on the floor using combination of pre-defined primitive motions

Contact based interfaces provide a precise and accurate measurement, which may make a machine to recognize the user's gesture quite well. However, this type requires very expensive measurement devices and cannot provide comfort to its users. Nevertheless, many interesting results using this framework are available. Also, other contact based applications are developed for man-machine interface such as eye-mouse system, haptic suit system, the ShapeTape (the fiber optic curvature sensors) [8] and so on.

3. Non-contact based interface

One of the most popular non-contact based interfaces is the vision-based approach.

In the previous section, we have introduced Korean Sign Language (KSL) interpretation system for hearing-impaired as a contact based interface. However, additional vision information can improve the performance of this system effectively.

The facial emotional expression recognition system is one for that purpose [9]. Facial gesture is very important factor in human-to-human communication as shown in Figure 1. And as many researches have pointed out, it is also a useful communication tool for man and machine [9,10]. In fact, it can be one of the most fundamental and natural ways for the machine to communicate with humans.

In general, the problem of recognizing emotion from human face is known to be very complicated and difficult because individuality issue may arise in expressing and observing emotions [11,12]. To deal with such difficulties effectively, soft computing techniques are used for recognizing facial emotional expressions.

The system has a three layered architecture: at the high level, a fuzzy system is designed based on human linguistic expressions, and at the mid level, a fuzzy observer is designed to indirectly estimate the linguistic variables using available image features, while, at the low level, image features are extracted to characterize the facial features. A multi-layered ANN is employed to develop parameter adjustment of the fuzzy observer based on available crisp input-fuzzy output sample sets. Experiments are performed on a real image sequence to demonstrate the effectiveness and efficiency of the system. Figure 9 shows the overall structure of this system except feature extraction stage.

Users can express many facial expressions for communication with machines. And their intentions are best expressed with facial expressions. For facial communications, it is an essential task to recognize 6 universal facial expressions such as surprise, fear, anger, disgust, sadness and happiness as shown in Figure 10.

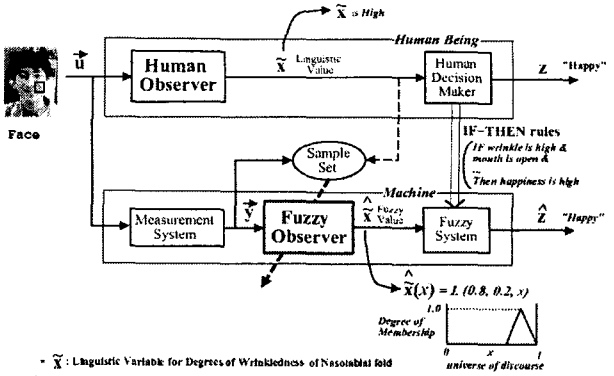


Figure 9. Overall structure with a fuzzy observer and a fuzzy subsystem

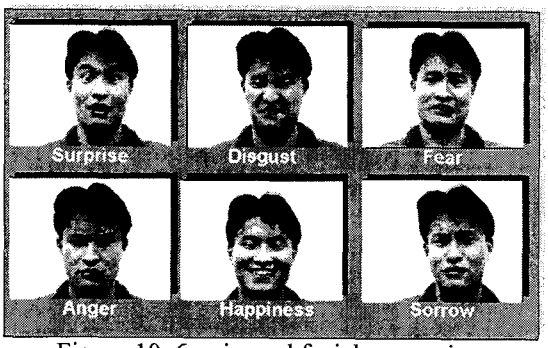


Figure 10. 6 universal facial expressions

Soft Remocon (remote controller) [13] is also a very interesting concept for man-machine interface. It is to operate various devices using gesture. User can select an electrical home appliance and operate it conveniently with pre-defined gestures. It will be very useful to operate an electrical home appliance that needs to know complex operation sequences.

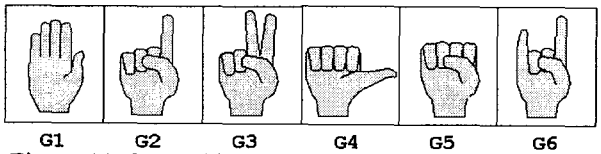


Figure 11. Several basic hand gestures for soft remocon

In the same vein, pointing gestures are used for pointing objects to move or for the similar actions. In this case, a laser pointer may be needed to indicate an object that is pointed by user for interactive response. And in this framework, assistive machine is needed to perform user's intention such as a robotic arm or a mobile platform. As shown in Figure 12, multiple cameras are needed to recognize pointing gestures.

Unlike the contact based interface, non-contact based interface can provide comfort for users. And in some cases, the user may not know its existence. So, the user can express their gesture without any constraints. In addition to this, the cost needed to implement is also relatively cheap. However, the most difficult fact about it is that it cannot provide precise and accurate information. Since it is the ultimate form for natural and comfortable communication between man and machine, there remain many things to resolve.

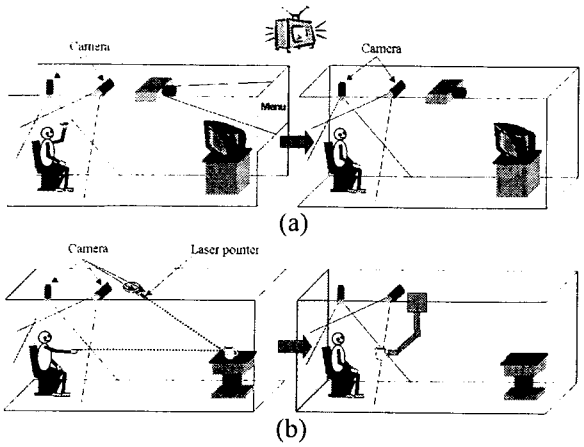


Figure 12. Manipulative gesture with non-contact based interface (a) turn on TV using soft remote controller (b) pointing an object and robotic arm picking it for user

4. Concluding Remarks

Gesture is a very effective means for communication between man and machine in natural and comfortable working environment. In this paper, we have examined some important gesture-based methods and their applications.

Contact and non-contact based interfaces are two main approaches for gesture recognition. Several successful examples are considered. However, there are still many difficulties to solve for recognition of gestures as a communication tool as humans do.

In order to recognize natural gesture effectively, the following three issues must be resolved [14]. First is the naturalness issue. We generate discrete gesture in a laboratory setting to recognize, but, basically, natural gesture is a continuous gesture. And it is not narrative gesture but conversational gesture. The second is an anthropological issue. As gesture has direct relation with culture, even the same gesture can be interpreted differently from countries to persons. The third is a system building issue. We must determine what kind of gesture we recognize and what type of interfaces.

Recent interest in gestural interface for HCI has been increasing due to the potential of gesture. Unlike the gestures in a natural environment, both manipulative and communicative gestures in HCI can be employed to direct manipulations of objects or to convey messages.

Hand gestures as a mode of HCI can simply enhance the interaction in classical desktop computer applications by replacing the computer mouse or similar hand-held devices. They can also replace joysticks and buttons in the control of computerized machinery [15] or they can be used to help the physically impaired to communicate more easily with others [16].

Some of the potential application examples are given in the following:

- Entertainment and musical conduction
- Rehabilitation Engineering
- Remote-control operation
- Teleoperation and Telemetry
- Recognition of sign language
- Virtual reality
- Simulation and Training
- CAD/CAM, Motion capture
- Military application.

Acknowledgement

The author likes to thank Mr. D.-J. Kim and J.-B. Kim for helping me in organizing this paper.

References

[1] D. B. Guralnik, Webster's New World Dictionary of the American Language, Koreaone press, 1989.

[2] A. Mehrabian, Communication without words, In Joseph A. DeVito, editor, Communication: Concepts and Process, pp.106-114, Prentice-Hall, Inc., 1971.

[3] K. H. Quek, "Toward a vision-based hand gesture interface," in Int. Proc. of the VRST'94, pp.17-31, 1994.

[4] J.-B. Kim, "Extension of the real-time Korean sign language(KSL) recognition system using colored hand tracker," M.S. Thesis, KAIST, 1998.

[5] D.-I. Suk, "A Linguistic Analysis of the Korean Sign Language," Ph.D, Thesis, Taegu University, Korea, 1989.

[6] S.-G. Kim, The Korean Sign Language Dictionary, Oh-Sung Press, 1993.

[7] J.-S. Han, J.-S. Kim, W.-K. Song, H. Lee, and Z. Bien, "EMG Pattern Classification using Soft Computing Techniques and its Application to the Control of a Rehabilitation Robotic Arm," 1st International Workshop on Human-friendly Welfare Robotic Systems, Taejon, KOREA, pp. 46-53, Jan. 20-21, 2000.

[8] http://www.measurand.com/, Measurand Inc.

[9] G.-T. Park and Z. Bien, "Neural network-based fuzzy observer with application to facial analysis," Pattern Recognition Letters, vol. 21, pp. 93-105, Feb., 2000.

[10] F. Hara and H. Kobayashi, "Use of face robot for human-computer communication," In IEEE Conf. on SMC'95, pp. 1515-1520, 1995.

[11] P. Ekman and W. V. Friesen, Unmasking the Face, Prentice-Hall, Inc., 1975.

[12] V. Bruce, Recognising Faces, Lawrence Erlbaum Associates, Publishers, 1988.

[13] Z. Bien, J.-B. Kim, J.-W. Jung, K.-H. Park, and W.-C. Bang, "Issues of Human-friendly Man-Machine Interface for Intelligent Residential System," 1st International Workshop on Human-friendly Welfare Robotic Systems, Taejon, Korea, pp. 10-14, Jan. 20-21, 2000.

[14] A. Wexelblat, "Research Challenges in Gesture: Open Issues and Unsolved Problems," In Bielefeld Gesture Workshop '97, pp. 1-11, 1997.

[15] J.-S. Kim, K.-H. Park, J.-H. Do, J.-B. Kim, K.-J. Song, and Z. Bien, "Study on Intelligent Autonomous Navigation of Avatar using Hand Gesture Recognition," accepted to SMC2000.

[16] J.-B. Kim, Z. Bien and W. Jang, "Intelligent Recognizer of Compound Words in the Continuous Korean Sign Language," World Automation Congress 2000, Maui, USA, June, 2000.