Human Interface Systems
Using Intentional and Unintentional Behaviors

Yoshinori Kuno, Tomoyuki Ishiyama, Yoshihisa Adachi, and Yoshiaki Shirai

Department of Computer-Controlled Mechanical Systems, Osaka University
2-1, Yamadaoka, Suita, Osaka 565-0871, Japan

Abstract

Human interfaces are usually designed to respond only to intentional human behaviors. However, humans show unintentional behaviors as well. They can convey useful information to realize user-friendly human interfaces. This paper presents two human interface systems using both types of behaviors: a gesture-based interface and an intelligent wheelchair. In the first system, only intentional hand movements are chosen using unintentional behaviors. In the second system, almost unintentional behaviors following intentional behaviors can be used to control the wheelchair movements. Experimental systems working in real time have been developed. Their operational experiments prove our approach promising.

1 Introduction

Human behaviors can be classified into two groups: intentional behaviors and unintentional behaviors. When we want to do a particular thing, we exhibit behaviors that are required to carry out the purpose. We do such behaviors intentionally and are aware of what we are doing. These behaviors belong to the first group. Conventional user interfaces including perceptual user interfaces use only behaviors in this category. However, we move our body and body parts unconsciously even when we do not have any intention to inform others of certain things. Still, such behaviors convey information about our attention, emotion and others. This information may be helpful to realize more user-friendly human interfaces. This paper proposes two ways to use unintentional behaviors in addition to intentional behaviors.

One is to use unintentional behaviors to choose intentional parts from nonverbal behaviors. We present a human interface system in which users can design an object by combining parts and changing the sizes of parts through hand movements which are expected to do so in the real world. When we use this kind of system, we may move our hands when we do not intend to manipulate objects in the virtual world. We use the face direction so that the system may not respond to such unintentional movements. The assumption here is that we watch the object when we manipulate it. Chino et al. [1] has proposed a similar method in their multimodal interface. It recognizes voice only when the user is looking at the display. Although their assumption is reasonable, we sometimes talk without looking at each other. However, it is rare to manipulate an object without watching it.

The other is to use unintentional behaviors to guess the user's intention. Humans sometimes show a particular pattern of movements when they want to do a certain task. If a computer can recognize this pattern, it can prepare useful information for the task and provide the information to the user. Or it can carry out the intended task autonomously. Pentland et al. have proposed this kind of intention understanding as the aid to car driving [2]. In this paper, we propose an intelligent wheelchair that can move as the user wishes with minimum intentional conscious operations.

2 Gesture-based Interface

We have presented a human interface system in which users can design an object by combining parts and changing the sizes of parts through hand gestures which are supposed to do so in the real world [3]. In this paper, we propose to add the capability to respond only to intentional gestures to the system. Conventional human interface systems assume that we move our hands only when we want to issue commands to the systems by hand gestures. However, in real situations, we often move our hands for other purposes than issuing commands, sometimes even unintentionally. These movements might be understood as some motion commands and objects in the virtual world might move accordingly. Thus, we need to be careful not to move our hands when we do not issue commands with conventional systems. To overcome this inconvenience, we introduce the capability to choose and to recognize only hand motions by which we intend to control objects in the virtual world. This can be done by observing the face direction. If we want to move an object in the virtual world, we will watch the object on the display screen. In other cases, we may not watch the screen. For example, if anyone enters the room and we want to say hello, we will turn toward the person. Based on this assumption, our system observes the face direction and chooses hand motions only when the face is directed to the screen. The chosen motions are recognized by the method using a state transition diagram.

2.1 Gesture Recognition

This section briefly describes our gesture recognition method. Details can be found in [3].

The purpose of our research is to develop a human interface system that enables a user to manipulate objects in a virtual 3D world. In the virtual space, he/she may point at an object, grasp it, move it, bring it into the work space, and may change its size or put it down on another object. These actions will not happen independently. We can narrow down the possible action candidates following each action to the small number. We represent these relationships by a state transition diagram. Since this diagram limits the number of possible recognition classes, the system can recognize the next gesture using simple image features.

Fig. 1 shows the state transition diagram used in the system. The transition condition from each state is represented by feature values and their changes. Except pointing gestures such as Direct and Move, the system uses feature vectors for several frames to recognize gestures. To cope with the variation of motion speed, the system examines in most cases whether particular feature values are increasing or decreasing rather than checks their exact values.

In Fig. 1, the transitions from and to the Rest state are shown by broken lines. If a new observation does not match any possible state transition conditions from the current state, the system changes it to the Rest state. While in the Rest state, if the system observes a feature vector that satisfies one of the possible transition conditions from the previous state, it changes the state according to the condition.

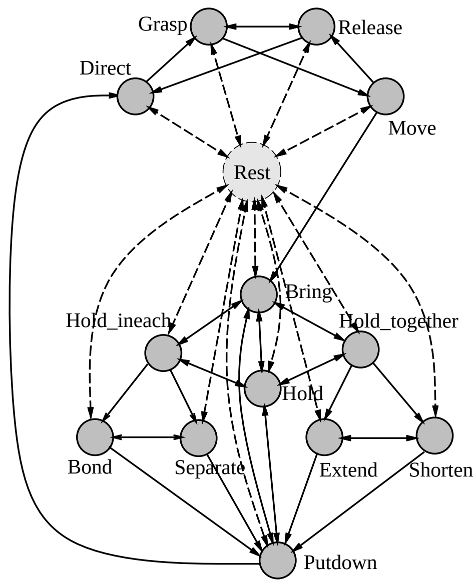

Figure 1: State transition diagram.

2.2 Intentional Gesture Selection by Face Direction

Although we have introduced the Rest state into our recognition method so that the system may respond only to intentional gestures [3], we have found this insufficient. Thus, we consider to use other nonverbal behaviors to choose intentional parts from hand motions. In this system, we consider gaze and eye contact information.

When we manipulate an object, we usually do this while watching it. Thus, it is a reasonable assumption that if we are not watching the object, the hand motions then are not intentional gestures for its manipulation. Although the reverse is not necessarily true, the use of this assumption with the introduction of the Rest state can help to solve the problem.

In the current implementation, the system uses the face direction instead of actual gaze direction. We need only rough gaze direction because we do not use gaze information to point at an object on the screen but to check whether or not the user is looking at the screen. In addition, the system should work in real time. Thus, it considers only the rotation of the head around the vertical axis and calculates the face direction using the following simple processes.

First, the system extracts a large skin color region as the face region. Second, it extracts face features such as the eyes and the eyebrows from the face region by thresholding. Third, it computes the projection onto the vertical axis to determine the vertical range of the eye-eyebrow region. If two close peaks are found in the projection result, they are considered for the eyes and the eyebrows. Fourth, the system computes the projection onto the horizontal axis to determine the horizontal range of the eye-eyebrow region. In this region, it examines the black parts, considering the lower two as the eyes. Then, it measures the distance, L_r, between the tail of the right eye and the right edge of the face on the line connecting the centroids of the eyes. It also calculates the counterpart to the left eye, L_l. Finally, it computes D by $D(L_l, L_r) = |L_l - L_r|/min\{L_l, L_r\}$. If D is smaller than 4, which means that the face direction is approximately within 20 degrees from the normal of the screen plane, the system considers that the user is facing toward the screen, thus recognizing the hand motions as intentional gestures.

2.3 Experimental System

We have developed an experimental human interface system using the proposed recognition method. It has three cameras: two for gesture recognition and the other for face direction computation. The current system uses two personal computers. One carries out face and gesture recognition with the aid of a real-time image processing board consisting of 256 processors developed by NEC [4]. The other displays two views for the user: a usual perspective view of the virtual world and a top view of the virtual world to help the user to understand the positional relationships of objects.

3 Intelligent Wheelchair

As the number of senior citizens has been increasing year by year, demand for human-friendly wheelchairs as mobility aids has been growing. Recently, robotic wheelchairs have been proposed to meet this need[5][6][7]. These wheelchairs help humans with the aid of ultrasonic, vision, and other sensors to avoid obstacles, to go to pre-designated places, and to pass through narrow or crowded areas.

A wheelchair with a human in it can be considered as a system consisting of the human and the machine. Conventional robotic wheelchairs see the outside of the systems, realizing the autonomous mobile functions as mentioned above. Although these functions are indispensable, we would like to point out the importance of seeing the inside of the system, that is, looking at the user to realize a human-friendly system. For example, if the machine can tell the user's intention of turning left in advance by observing his/her behaviors, it can give helpful information to him/her, such as a caution message if he/she is not paying necessary attention around him/her. Or if it is certain about the user's intention, it will turn left autonomously. This paper presents an intelligent wheelchair in which we aim to realize such intention understanding capability.

3.1 Motion Control by Face Direction

Our ultimate goal is to realize a wheelchair that can move as the user wishes even though the user does not explicitly show his/her intention. However, this is difficult to be achieved. Thus, as the first step toward this goal, we consider to use human behaviors which are done intentionally but so naturally that the user may not feel any burden. We use the face direction. The wheelchair moves in the direction of the user's face. For example, it turns left if the user turns the face to the left. The user should know this fact and needs to turn the face intentionally. However, it is a natural behavior to look in the direction where he/she intends to go. It is also a natural behavior turning back the face to the frontal position as the turn is getting completed. Although this behavior can be done almost unconsciously, it can control the wheelchair appropriately. When we use a steering lever or wheel to control a vehicle motion, we need to operate it intentionally all the time. Using the face direction can remove a considerable part of such intentional operations, realizing a user-friendly interface.

The problem, however, is that we move our heads in various other occasions than in controlling the wheelchair's direction. The intelligent wheelchair system needs to separate wheelchair-control behaviors from others. This is the intention understanding problem in the current system. We assume that one moves the head slowly and steadily when one is told that the wheelchair moves in the face direction. Thus, the current system ignores quick head movements, only responding to slow steady movements. There are, however, cases that we look in a certain direction steadily without any intention of controlling the wheelchair. For example, when we find a poster on a corridor wall, we may look at it by turning our face in the direction while moving straight. The environmental information around the wheelchair can be helpful in such cases. In the example case, even if we turn our faces to the left to look at the poster, the system can consider this not to show a left-turn intention if it knows that a wall exists on the left and that it cannot turn left. We have made a simple experiment on this.

3.2 Experimental System

We have developed an experimental wheelchair system that can move in the user's face direction computed from video camera images. Fig. 2 shows an overview of the system.

In the current system, we use the following simple methods for intention understanding. We apply a smoothing filter to the face direction calculation results by averaging them for a certain number of frames. If this number is large, the system will not be affected by quick unintentional movements of the head. However, the user may feel uneasy at the slow response of the system. We have made actual running experiments by changing the number of frames used for filtering from 1 to 30. We have decided to use 10 frames for averaging from the experimental result. In addition, we use thresholding to remove small changes of the face direction.

To understand the intentions in slow head movements, the system uses the environment information. We give a map data to the current system. It can tell the existence of walls and other obstacles using the dead reckoning data and the map data. Experimental results have shown the usefulness of the method. Even when the user turns the face to look at a poster on a wall, the system does not turn toward the wall.

We have made several other running experiments. These experimental results confirm that we can control the wheelchair by the face direction. Owing to smoothing and thresholding, the system does not respond to small and/or quick movements of the head. Thus, if we look for a while in the direction where we want to go, we can move the wheelchair to meet our intention. However, this slow response means that we cannot make quick precise control of the wheelchair. We may not be able to avoid an obstacle if it suddenly comes close. Thus, to make the system practical, we need to introduce an autonomous motion function for obstacle avoidance. The idea is that we show our intention by the face direction and after that the system can navigate autonomously while avoiding obstacles. We are now developing such a function using ultrasonic sensors. The sensor information around the wheelchair is also helpful for intention understanding. Instead of a map data, we are working on the use of ultrasonic and vision sensors.

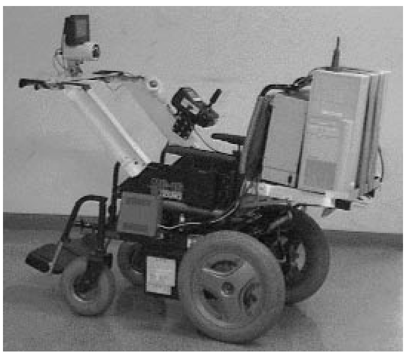

Figure 2: System overview.

4 Conclusion

This paper presents two human interface systems using both intentional and unintentional behaviors. Although the current systems are still in the early stage of research, we have obtained promising results.

More detailed information including video clips is available at http://www-cv.mech.eng.osaka-u.ac.jp /research/hi.html.

This work has been supported in part by the Ministry of Education, Science, Sports and Culture under the Grant-in-Aid for Scientific Research (09555080), the Kurata Foundation under the Kurata Research Grant, and the Kayamori Foundation of Informational Science Advancement.

References

[1] T. Chino, K. Fukui, O. Yamaguchi, K. Suzuki, and K. Tanaka, "GazeToTalk: A nonverbal interface system with speech recognition, gaze detection, and agent CG output," *Proc. Interaction'98* , pp.169-176, 1998 (in Japanese).

[2] A. Pentland and F. Liu, "Towards augmented control systems," *Proc. IEEE Intelligent Vehicles '95*, pp. 350-355, 1995.

[3] K.H. Jo, Y. Kuno, and Y. Shirai, "Manipulative hand gesture recognition using task knowledge for human computer interaction," *Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp.468-473, 1998.

[4] S. Okazaki, Y. Fujita, and N. Yamashita, "A compact real-time vision system using integrated memory array processor architecture," *IEEE Trans. on Circuits and Systems for Video Technology*, 5(5), 1995, 446-452.

[5] D. P. Miller and M. G. Slack, "Design and testing of a low-cost robotic wheelchair prototype," *Autonomous Robotics*, vol. 2, pp. 77-88, 1995.

[6] R. C. Simpson and S. P. Levine, "Adaptive shared control of a smart wheelchair operated by voice control," *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2 , pp. 622-626, 1997.

[7] H. A. Yanco and J. Gips, "Preliminary investigation of a semi-autonomous robotic wheelchair directed through electrodes," *Proc. Rehabilitation Engineering Society of North America 1997 Annual Conference*, pp. 414-416, 1997.