Multi-modal Interface with Voice and Head Tracking for Multiple Home Appliances

Eiichi Ito

Kanagawa Rehabilitation Center
516 Nanasawa, Atsugi, Kanagawa, 243-0121, Japan

ito@kanagawa-rehab.or.jp

Abstract: In this paper, we describe a multi-modal interface that allows use of voice and gesture commands for controlling distributed home appliances used by people with disabilities. The main objective of this study is combined with nonverbal and verbal interface for intuitive and efficient control that uses hands-free operation. The pointing gesture by facing as nonverbal interface represents selecting one of the home appliances. The voice commands as verbal interface represent button operation of the remote controller such as the power on/off, the channel select and the volume up/down. The prototype system can provide a hands-free remote controller for people with quadriplegia who do not have to send verbal commands for selecting home appliances.

Keywords: Input Devices, People with Special needs, Gesture based I/O, Speech and voice, Home Appliances

1 Introduction

In the daily living of people with disabilities, they have to use lots of home appliances by them limited motor function. There are Environmental Control Systems with learning IR commander operated by voice for people with disabilities. In this system, user has to set a unique voice label for each physical object, like "TV" for the TV set. If the user would like to watch the TV, then they may speak out the target label and control command. They often have troubles when they have lots of home appliances for using. Because the user may forget the unique voice label of target home appliance, then they cannot select it. Thus, the control command of interface should not use the voice label that is unknown to the public or difficult to remember.

The study of multi-modal interface have shown intuitive and efficient input method for the human computer interaction(Bolt, 1980). The real-world interface can be used for controlling home appliance, such as FieldMouse(Siio et al, 1999). The application of multi-modal interface can assist to people with disabilities for desktop computing (Malkewitz, 1998). Our approach is combination of real-world interaction and multi-modal interface that may help to people with disabilities.

2 System description

The system we have designed allows that people with quadriplegia can control distributed home appliances with intuitive and hands-free operation.

We assumed that the intuitive operation of controlling home appliances is combined with the selecting target as nonverbal command and getting instruction as verbal command.

2.1 Selecting target

The selecting target module(STM) can distinguish which home appliance is selected by user's head direction. The head tracking system can detect the 3D positions and 3D angles of user’s head posture.

The STM can calculate the location of where the user looks toward by own head. The physical 3D locations of all home appliances have stored in the STM before. The STM can distinguish that which home appliance is selected by user.

2.2 Getting instruction

The getting instruction module(GIM) can detect the command for controlling home appliance by voice.

The user’s purpose is sensed by the Japanese voice recognition system. The voice recognition can detect the voice command from microphone. The

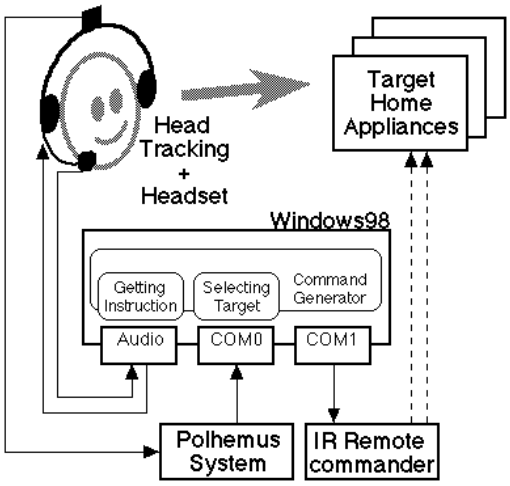

Figure 1: System configuration

GIM can interpret generalized control command from that recognized voice.

Many of remote controllers have same button, such as "power" "play" "stop" "up" "down" "channel". Thus, we have integrated from same function button to the generalized command.

2.3 Command generator

The command generator module(CGM) can send a proper command to each home appliance.

The STM has shown a selected home appliance. The GIM has shown a generalized command. The CGM can generate a proper command of each home appliance from the STM and the GIM. After the CGM send the proper command to the IR remote commander, the target home appliance detects own command then execute a command. This system configuration is shown in Figure 1.

3 Trials of practice

3.1 Participants

We have some trials of practice for controlling home appliances by a normal and a quadriplegia. Both of them have enough wide range of neck movement and enough volume of voice. While the practice, the user is sitting down on a stool or a wheel chair. But, user does not have to keep them location within the limits of head tracking system.

The user are wearing a headset mounted the Polhemus receiver. A TVset, a VCR, a DVDplayer, a Radio-cassette and a Lamp with IR relay are setting up to appropriate racks.

3.2 Results

The prototype system can provide a hands-free remote controller for people with quadriplegia who do not have to use verbal commands for selecting home appliance. The system can also reduce the verbal command by generalization.

Therefore, the system only can control to the home appliance with IR controller. The user often loss sight of the location, that made from the head tracking system. Because, the user cannot confirm except system have detected “lock on target”.

In this trial, we acknowledged that the prototype system needs 2 functions.

- Connect to the home network, such as IEEE1394, HAVi, Bluetooth.
- Set visual or sound marker for confirmation of sensor direction.

4 Conclusions

The trial has shown the prototype system can provide a hands-free remote controller using by people with disabilities. For hands-free operation, system has a head tracking for selecting target, voice recognition for getting instruction and a computer with IR commander for command generator.

The combination interface of nonverbal and verbal communication, such as head tracking and voice recognition, are useful for reducing verbal command. Application of the multi-modal interface that allows use of voice and gesture can assist for the daily living of people with disabilities.

Acknowledgements

We would like to thank Prof. Michiaki Yasumura (Keio Univ.) for his important suggestion. This research was supported from the Telecommunications Advancement Foundation.

References

Bolt,R.A.(1980), Put-That-There: Voice and Gesture at the Graphics Interface, ACM SIGGRAPH'80 Computer Graphics, 14(3), pp.262-270

Siio,I., Masui,T. & Fukuchi,K.(1999), Real-world Interaction sing the FieldMouse, ACM UIST'99, pp.113-119

Malkewitz,R.(1998), Head Pointing and Speech Control as a Hands-Free Interface to Desktop Computing, ACM ASSETS'98, pp.182-188