HIT-Wear: A Menu System Superimposing on a Human Hand for Wearable Computers

Hiroshi SASAKI, Tomohiro KURODA,
Yoshitsugu MANABE, and Kunihiro CHIHARA

Graduate School of Information Science,
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, JAPAN
hirosi-s@is.aist-nara.ac.jp

Abstract

The recent developments of computer hardware have stimulated the study of Wearable Computers. Many research works focus on various applications and hardware configurations of wearable computers. However, there are few research works on the interface system, which determines the usability of wearable computers.

Therefore, this paper presents "HIT-Wear" (***H***and ***I***nterface ***T***echnology for ***Wear***able Computers), an innovative input interface for wearable computers, utilizing visual feedback and physical senses with Augmented Reality technology. This system gives the user a truly intuitive non-contact input interface. As "HIT-Wear" requires only one camera mounted on a see-through head mounted display without any devices, this interface enables the users to easily operate wearable computers with empty hands. This paper presents "Hand-Menu System" as an example of "HIT-Wear". When the wide opened hand in the users sight calls up this hand-menu system, the menu appears on the fingertips of a hand through the head mounted display. The user selects a certain item from the menu by touching a certain fingertip with the index finger of the other hand. As the user touches on a menu item with his/her own finger physically, the user can be certain that he/she selects the menu item. Therefore, this hand-menu interface enables the user to select a menu intuitively without any additional devices.

Key words: Wearable Computer, Human Interface, Augmented Reality, Physical Sense, HIT-Wear

1. Introduction

The recent developments of computer hardware have stimulated the study of the Wearable Computers [1]. Many research works focus on various applications and hardware configurations of wearable computers. However, there are few research works on the interface system, which determines the usability of wearable computers.

This paper presents "HIT-Wear" (***H***and ***I***nterface ***T***echnology for ***Wear***able Computers), an innovative input interface for wearable computers, utilizing visual feedback and physical senses with Augmented Reality technology. This system gives the user a truly intuitive non-contact input interface. As this interface requires only one camera mounted on a see-through head mounted display without any devices, this interface enables the user to operate wearable computers with empty hands.

2. Input Interfaces

The input interfaces for wearable computers proposed till now can be divided into two types, contact devices and non-contact devices, such as:

1. Contact devices
   - Keyboard, Mouse [2]

     Handykey Twiddler™ (Figure 1)[3]
   - Pen interface with touch panel

     Sharp ZAURUS™ (Figure 2)[4]
   - Wearable sensor [5]

     NTT HI-Lab. FingeRing (Figure 3)[5]
2. Non-contact devices
   - Speech recognition [6]

     MIT Media Lab. Nomadic Radio(Figure 4)[6], IBM ViaVoice™ [7]
   - Image recognition [8]

The input interface system should enable users to use wearable computers easily, anytime and anywhere. However, users may feel troublesome and restricted to wear these contact devices when using and carrying these devices with themselves.

On the other hand, foregoing non-contact devices require user's efforts; the speech recognition system makes users to teach their voice to the system by reading many sentences beforehand, and the motion recognition system makes users to learn enormous gestures beforehand.

The above discussion gives the insight about the conditions of the favorable input interface for wearable computers. The interface needs to be...

- Non-contact

  To require users to wear no special device for input interface

Figure 1: Handykey Corporation's Twiddler™ [3]

Figure 2: Sharp Corporation's ZAURUS™ [4]

- Instant

  To require users no special preparation

- Intuitive

  To require users no preceding knowledge

This paper introduces an innovative concept of input interface, which satisfies above conditions.

Figure 3: NTT HI-Lab. FingeRing [5]

Figure 4: MIT Media Lab. Nomadic Radio [6]

3. Conceptual Design

In order to develop intuitive interface, the interface system must utilize the natural interaction mechanism of human beings. The interaction is the repetition of exchanging information, that is, to send and to receive information.

Human beings send the information through two media; voice and motion. The most flexible and useful media is voice. We mainly use this communication method to interact with another person. However, the information given by voice is too flexible to use as a media to communicate with computer; we can express same things in various ways. Therefore, usually we use motion communication, to put it more concrete, touching the buttons to interact with computers. Actually, hands are the most useful media for human beings to interact with machines, which allow limited numbers of inputs.

On the other hand, human beings utilize the five senses, sight, hearing, smell, taste, and touch to receive information. Because of the difficulty to supply chemical interaction, computers use physical three senses, sight, hearing, and touch, to give information for users. Among these three senses, sight carries on the most important roll to receive information from surroundings. Therefore, most of wearable computers utilize head mounted display (HMD) such as SONY's PC Glassron™ (Figure 5)[9], to give visual information.

The conventional way to interact with machines including computers is along this discussion; a machine gives the information through visual display, and a user inputs by touching on some buttons. In some cases, a user may need to input pictorial information. In that case, a user just needs to write down the picture with his/her index finger or a pen-type input device. Thus, one of the best available interfaces is "touch-screen". A touch-screen allows a user to indicate a favorable item directly and to write pictorial information down directly.

Figure 5: Sony Corporation's PC Glasstron™ [9]

This paper introduces touch-screen like interface without any sensory pad. A human hand has five fingers and palm. Five fingers are comparable with selection items of menu interface, and a palm is comparable with a touch pad to get pictorial information. If visual information from a wearable computer is projected on fingertips or around a palm, it is just the same system as touch-screen interface.

This total interface system including "Hand-Menu System" and "Hand-Pad System" named "HIT-Wear" such as shown in Figure 6, gives an intuitive and instant interface. Augmented Reality (AR) technology using see-through HMD and image recognition technique make this interface non-contact. User needs to wear just a see-through HMD with small video camera as in the case of usual wearable computers, such as XYBERNAUT's MOBILE ASSISTANT® IV [10], IBM's wearable PC (Figure 7) [11], and the one of MIT Media Lab.[12].

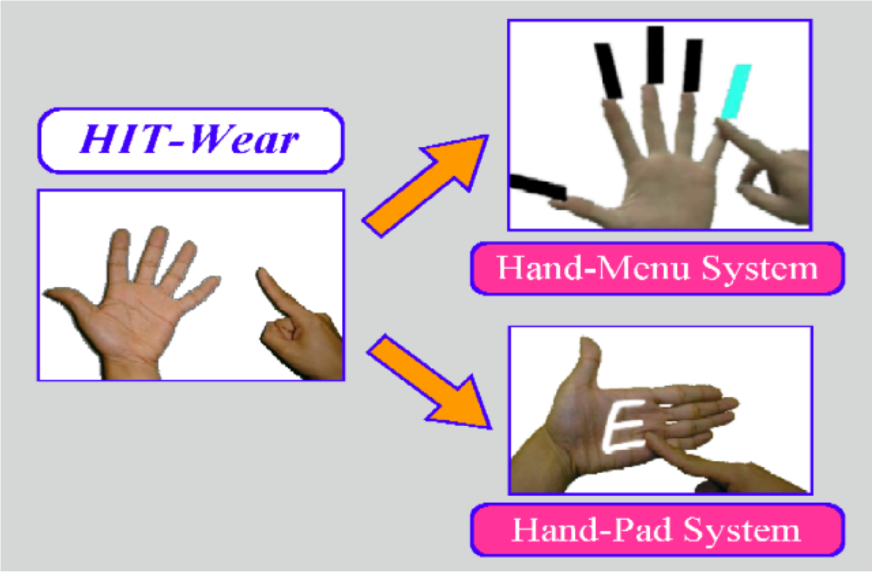

Figure 6: *H*and *I*nterface *T*echnology for *Wear*able Computers

Figure 7: IBM Corporation's wearable PC [11]

A user just needs to show his/her weak hand to recall interface system and he/she have visual and touch feedback when he/she inputs with the index finger of his/her strong hand. Comparing with systems in which user inputs information by touching virtual pad or menu, the passive touch feedback gives reliability for the HIT-Wear system. Additionally, user can recall the interface system by just looking on his/her weak hand only when he/she wants to input.

This paper introduces the implementation about "Hand-Menu System" in the following sections as an example of HIT-Wear.

4. Hand-Menu System

The conceptual design of "Hand-Menu System" is shown in Figure 8.

In the following part, we call the weak hand as "Menu-Hand" because menu is projected on the weak hand in this system. In the same manner, we call the strong hand as "Selection-Hand".

When the wide opened menu-hand in the user's sight calls up this hand-menu system, the menu appears on the fingertips of the menu-hand through the see-through HMD. The user selects a certain item from the menu by touching a certain fingertip of the menu-hand with the index finger of the selection-hand. As the user touches on a menu item with his/her own finger physically, the user can be certain that he/she selects the menu item.

All of above processes is performed through the image recognition technique. Therefore, the hand-menu interface enables the user to select a menu item intuitively without any additional devices.

5. System Procedure

In order to realize the hand-menu system as a non-contact interface system, the system needs to acquire all selecting motion through image recognition techniques. Additionally, the image processing must not exhaust whole computational resources of the wearable computer as the system is just a part of the wearable computer system. Therefore, this paper utilizes some simple image recognition techniques for the hand-menu system.

The hand-menu system recognizes the user's input in the following manner:

1. Hand-area Extraction
2. Fingertips Extraction
3. Distinguishing the Menu-hand and the Selection-hand
4. Displaying Menu
5. Menu Selection

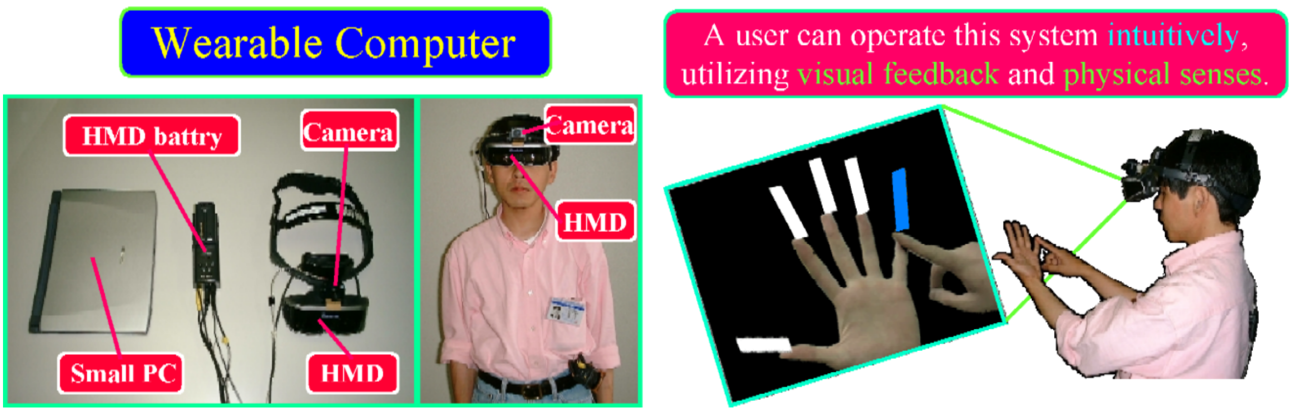

Figure 8: Conceptual design of hand-menu system

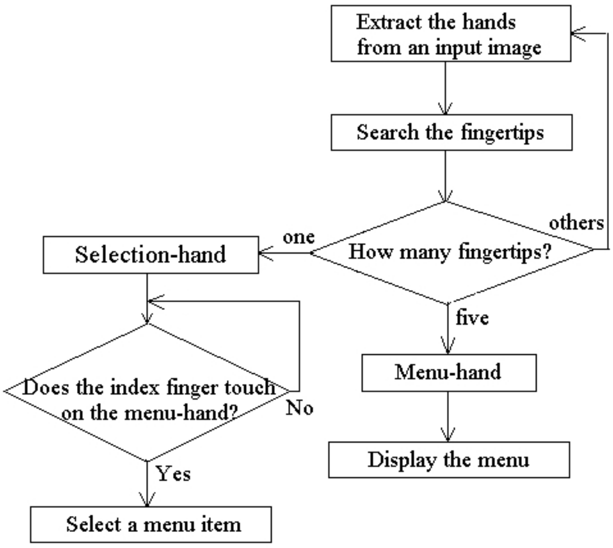

Figure 9: Flowchart of the hand-menu system

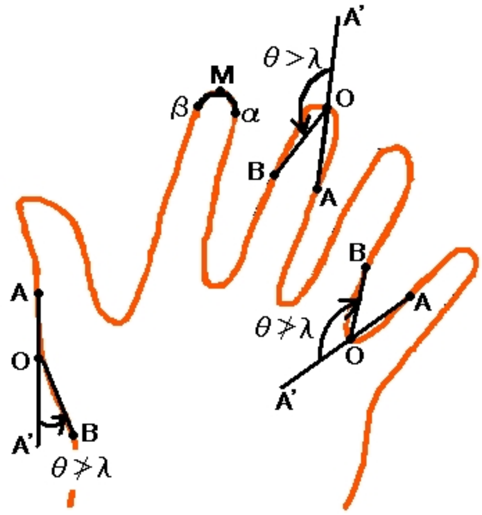

Figure 10: Tracking algorithm of fingertips

The flowchart of the hand-menu system is shown in Figure 9. In order to make the hand-menu system practical, all of these processes should be done in real-time. The following sections explain detailed techniques.

5.1. Hand-area Extraction

At first, the system requires to detect the menu-hand and the selection-hand. To make the system simple, these areas are detected through color matching technique. The obtained RGB color image is projected HSV color space. The look-up table developed in advance speeds up this process. Using H and S value, skin color regions are extracted from the image. The biggest connected region and the second biggest connected region are selected as the menu-hand or the selection-hand.

5.2. Fingertips Extraction

Consequently, the system needs to extract the fingertips of the hands in order to know the place to show menu items and to know the pointer. Fingertips are extracted through the outline tracking.

Figure 10 shows the tracking algorithm. The point O is the tracking point. The points A and B are the points on the outline in a certain distance from O. The triangle AOB becomes the fingertip tracker. When the exterior angle θ becomes bigger than a certain threshold λ, the point O is marked as a candidate point of a certain fingertip. Through this process, the candidate section $\alpha\beta$ is extracted.

In order to avoid selecting the bottom of a ravine as a fingertip, the direction of the angle θ is decided as the anti-clockwise direction. Thus, the angle θ becomes negative and the ravine cannot be a candidate.

At last, the system regards the middle point M of the candidate section as a fingertip.

5.3. Distinguishing the Menu-hand and the Selection-hand

The method to distinguish the menu-hand and the selection-hand is quite simple. The system recognizes the region with five fingertips as the menu-hand, and the region with one fingertip as the selection-hand.

5.4. Displaying Menu

As soon as the system recognizes that the menu-hand is exposed, the system superimposes the menu on the menu-hand in the user's sight through the see-through HMD.

Menu-boxes are displayed as shown in Figure 11. The point α, β, and M in Figure 11 are the same points as in Figure 10. The point N is the center point between the point α and β.

A menu-box is a rectangle of 25x125pixels. The bottom line of a menu-box is the line, which is parallel to the vector $\alpha\beta$ and is 15pixels apart from the point M. The height direction of the menu-box equals the direction of the vector NM.

5.5. Menu Selection

The system recognizes the menu selection based on the area of extracted regions. As shown in Figure 12, the system recognizes that a certain menu item is selected, when following formula is satisfied:

$$|\alpha(t+1)-\{\alpha(t)+\beta(t)\}|\leq\gamma$$

$\alpha(t)$ is the size of biggest hand-color area and $\beta(t)$ is the size of the second biggest hand-color area at a certain frame t.

In fact, the system recognizes that the menu is selected when the size of α is fixed in 3 continuous frames, that is:

$$\alpha(t+3)\approx\alpha(t+2)\approx\alpha(t+1)\approx\alpha(t)+\beta(t)$$

The system executes the nearest menu item to the pointer, that is the fingertip of the selection hand, at the frame t.

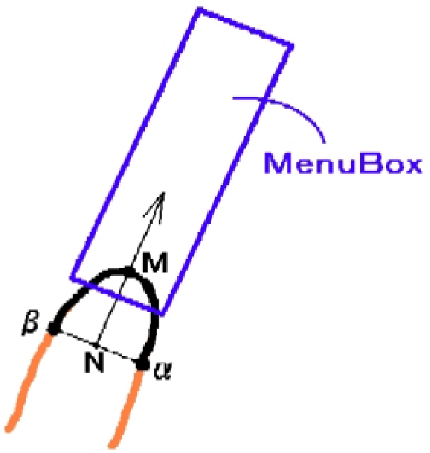

Figure 11: Displaying a menu-box

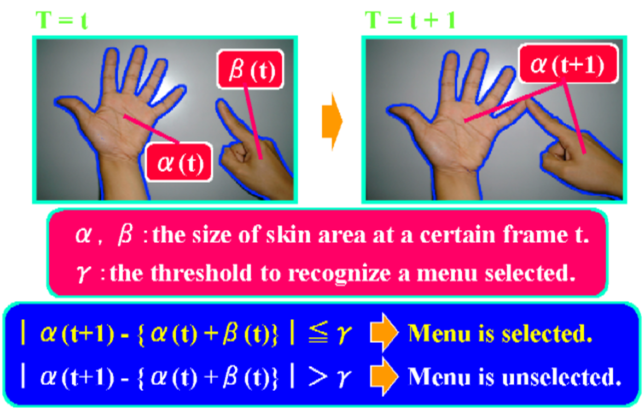

Figure 12: Selecting algorithm of a menu

6. Prototyping

The prototype interface system is developed on a Toshiba DynabookSS 3300, SGI O2, and SGI Onyx2 with infinite reality2. Figure 13 shows the output of the prototype. The features and the performances of each prototype systems are as follows:

- Toshiba DynabookSS 3300

  CPU: Intel Mobile PentiumII 266PE MHz
  Video Capture Card: RATOC REX9590
  320x240pixels, 8bit-colors
  Video Output: NeoMagic MagicMedia256AV
  Performance: 3.57fps

- SGI O2

  CPU: MIPS R5000 180MHz
  Video Input: 640x480pixels, 32bit-colors
  Performance: 2.17fps

- SGI Onyx2 InfiniteReality2

  CPU: MIPS R10000 195MHz 16processers
  Video Input: 640x480pixels, 32bit-colors
  Performance: 4.76fps

Figure 13: Displaying result

7. Discussions

The performance of the prototype system on notebook PC(Toshiba DynabookSS 3300) was 3.57fps. For smooth communication, the system should perform more than 10fps. This prototype utilizes only still image based image processing techniques. Therefore, the introduction of the motion image processing techniques such as motion tracking to eliminate the scanning area may result in the better performance. Also, the skin-color area extraction is quite time consuming. Therefore, the active image sensing technique projecting infrared light to eliminate the view volume of the camera may help much for the better performance.

Additionally, this active sensing technique makes the system robust to the environmental factors such as lighting conditions. As the proposed method is based on the size of selected area, this method cannot recognize the menu selection when the selection hand intersects over the menu-hand as shown in Figure 14. This system is available for only the case of Figure 15. The edge detection technique to extract whole figure of the index finger of the selection-hand can overcome this problem. The range sensors such as stereo camera or active range finder may be also helpful.

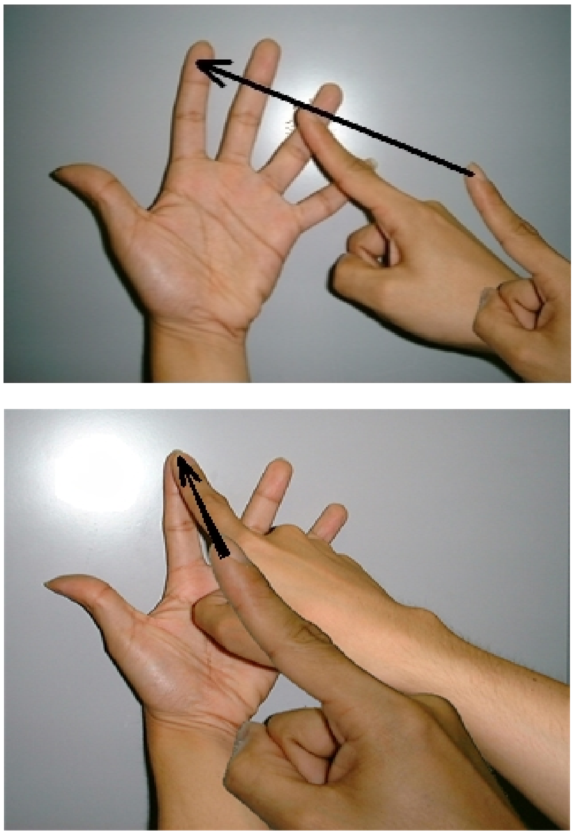

Figure 14: Menu selection crossing on the menu-hand which cannot obtained proposed method

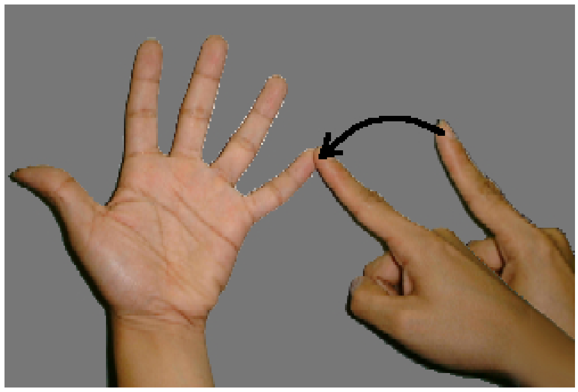

Figure 15: Menu selection which can obtained by proposed method

8. Future Work

The main factor to make the hand-menu system unpractical is the robustness for the lighting conditions. Therefore, the authors will introduce a certain active image sensing technique for the prototype. On the other hand, the authors are introducing some motion image processing techniques to get better performance.

9. Conclusions

This paper presented the HIT-wear, which is truly intuitive, non-contact input interface for wearable computers.

This paper explained the design of the hand-menu system, a part of HIT-wear, and experimented the prototype.

For better performance, the active image sensor and the motion image processing techniques should be introduced to the system. Many researches propose so many applications of wearable computers to make our lives convenient. HIT-wear can increase the usability of wearable computers dramatically. The authors believe that the realization of HIT-wear accelerates the realization of wearable computers.

Acknowledgement

This research is partly funded by International Communications Foundation.

References

1. Thad Starner, Steve Mann, Bradley Rhodes, Jeffrey Levine, Jennifer Healey, Dana Kirsch, Rosalind W.Picard, and Alex Pentland: Augmented Reality Through Wearable Computing. M.I.T. Media Laboratory Perceptual Computing Section Technical Report No.397. Presence, Special Issue on Augmented Reality, 1997

2. Bruce Thomas, Susan Tyerman, Karen Grimmer. Evaluation of Three Input Mechanisms for Wearable Computers. Presented at the IEEE International Symposium on Wearable Computer, 1997.

3. Handykey Corporation. Twiddler™: http://www.handykey.com/

4. Sharp Corporation. ZAURUS™: http://www.sharp.co.jp/

5. Masaaki Fukumoto, Yoshinobu Tonomura. "Body Coupled FingeRing": Wireless Wearable Keyboard. http://www.acm.org/sigchi/chi97/proceedings/paper/fkm.htm

6. Nitin Sawhney, Chris Schmandt. Speaking and Listening on the Run: Design for Wearable Audio Computing. In Proceedings of the International Symposium on Wearable Computing, 1998.

7. IBM Corporation. ViaVioce™: http://www.ibm.co.jp/voiceland/

8. James Davis, Mubarak Shah. Recognizing Hand Gestures. Computer vision – ECCV '94, 1994

9. Sony Corporation. PC Glasttron™: http://www.sony.co.jp/sd/ProductsPark/Consumer/Peripheral/PLM-S700/index.html

10. Xybernaut Corporation. MOBILE ASSISTANT® IV. http://www.xybernaut.com/

11. IBM Corporation's wearable PC: http://www.ibm.co.jp/News/leads/980912/

12. Lizzy: MIT's Wearable Computer Design 2.0.5. http://lcs.www.media.mit.edu/projects/wearables/lizzy/index.html