Wearable Computers: Field-Test Results and System Design Guidelines

Chris Esposito

Virtual and Augmented Reality Group
Boeing ISS Research and Technology
PO Box 3707, MS 7L-48
Seattle, WA 98124 USA
chrise@atc.boeing.com

ABSTRACT Since the summer of 1994, a group of partners, led by Boeing and including Carnegie Mellon University, Virtual Vision, and Honeywell have developed or provided significant design input on several generations of wearable computer systems, head-mounted displays (HMDs), and field-tested several application prototypes using these systems. Some of these applications include KC-135 skin inspections at McClellan Air Force Base in Sacramento, California, autopilot troubleshooting for the Boeing 777, and fuel system problem diagnosis and repair for the Boeing 757. This paper will address three questions: Why would someone wear a computer? What is a wearable computer? How does application and user interface design for these systems differ from that for more conventional office desktop applications?

Keywords Wearable Computers, Head Mounted Displays, Input Devices, Speech Interfaces

1 INTRODUCTION

Since the summer of 1994, a group of partners, led by Boeing and including Carnegie Mellon University, Virtual Vision, and Honeywell have developed or provided significant design input on several generations of wearable computer systems, head-mounted displays (HMDs), and field-tested several application prototypes using these systems. Some of these applications include KC-135 skin inspections at McClellan Air Force Base in Sacramento, California, autopilot troubleshooting for

the Boeing 777, and fuel system problem diagnosis and repair for the Boeing 757. A separate branch of this project has been investigating the use of 'Augmented Reality' in several manufacturing applications, although this work will not be reported on here - see (Caudell, 1992) or (Janin, 1994) for more information on this area.

This paper will address three questions: Why would someone wear a computer and its peripheral devices? What is a wearable computer? How does application and user interface design for these systems differ from that for more conventional office

desktop applications? Since this is a relatively new and unfamiliar topic, we shall begin by describing some of the representative tasks and environments that can benefit from these new systems. This will be followed by a brief description of current hardware and probable future systems. The final section will describe the Boeing 757 application in more depth, including some design guidelines and lessons learned from the field tests.

2 WHY A WEARABLE COMPUTER?

One of the most basic characteristics is that these systems are used for work that cannot be done in a conventional office environment. This includes tasks like aircraft maintenance and vehicle inspections but excludes tasks like using a laptop on an airplane for text editing.

Another characteristic is that the user needs real-time read/write access to more information while doing the work than can be conveniently manually carried to and organized at the worksite. Examples of this may be as simple as an inspection checklist scenario, or a more complex one such as a troubleshooting and repair scenario that requires access to stored diagnostic procedures and schematics, along with keeping a record of what maintenance was actually done. This second scenario raises the issue of just how much data may need to be available at the worksite. For example, the total amount of operational, diagnostic, and repair data for most modern aircraft easily runs to tens of thousands of pages, but only a dozen of these pages may be needed to identify and repair the problem. Identifying and providing timely access to the right subset of the database raises issues that are broader than the scope of this paper, but to the extent that a wearable is seen as a potential delivery vehicle for this data, there is considerable overlap in these issues.

A third characteristic has to do with the manual nature of the application tasks. If the users' hands are already busy, then a system that provides access to this information in a substantially hands-free fashion can provide significant operational advantages over a system that requires frequent physical context switches between doing the work and using a paper manual or a laptop.

The physical environments these systems are used in provide their own unique set of challenges. These include extremes of temperature and humidity, and lighting conditions that can range from complete darkness to full sunlight. In some cases, dirt, grease, oil, sand, hydraulic fluid and various fuels may be common in the environment, all items seen as serious contaminants in the office world. The users may be required to wear gloves, hard hats, safety shields and uniforms, and wearable systems must integrate into this setup as comfortably and non-intrusively as possible or there is likely to be significant resistance from users. In some specialized and exotic applications currently being developed, the wearable systems are woven directly into the clothing worn by a soldier and are designed to run without any intervention from him; in this case they are designed to constantly monitor life signs (heart rate, blood pressure, etc.) and report any deviations due to hazards such as gunshot or knife wounds. As we shall see in later sections, all of these task and environmental characteristics significantly affect the design of the wearable hardware and the applications that run on them.

3 WHAT IS A WEARABLE COMPUTER?

In this section we will describe the hardware and software components used to develop our field-tested systems and how they are likely to evolve in the near future; a more detailed description of the application will be in the next section.

3.1 Current Hardware

Our current prototype systems are from the Flexible PC Company, and consist of:

- 100 MHz 486 CPU with power management
- 32 MB memory
- standard set of I/O connectors
- standard 15-pin VGA connector
- 4 Type III PCMCIA slots

These slots are currently filled with 2 340 MB PCMCIA hard drive cards and 1 PCMCIA speech recognition / sound card from Speech Systems, Inc.

The remaining slot may be filled by a wireless networking card or other device as needed by the task. The standard serial ports mean that we can attach a variety of input devices, such as touchpads, dials, or devices as required.

Our current HMD prototypes are from Virtual Vision, Inc. and feature:

- 8-bit grayscale VGA monocular display
- adjustments for display position and orientation
- adjustments for focus and brightness
- eye safety shield
- optics that focus the image at arms length
- integrated earphone
- integrated noise-cancelling microphone

The dial input device that was used in the final field evaluations was constructed by Carnegie-Mellon from one of their cannibalized Vu-Man 3 wearable computers.

The 777 and 757 applications were built by the author using Visual Basic 3.0 and the PE500 Speech Application Development Environment from Speech Systems, Inc., to run under Microsoft Windows for Workgroups 3.11. The Flexi-PC has a flexible system backplane which allows it to wrap around the user's waist, and physically resembles a money belt. The dial was put into a modified camera case, and was also worn on the waist strap of the Flexi-PC.

3.1 Future Development

Many of the developments in wearables are mirrored by (and often driven by) changes in the enormously larger laptop market but as time goes on and wearables grow into their own category of significant size, we can expect some divergence to take place. The probable future directions for most wearable systems can be divided up into the following four categories:

- CPU, System Bus, and Graphics
- Storage
- Power
- Packaging, Expandability, and Connectivity

3.1.1 CPU, System Bus, and Graphics

Now that most Pentium power management has evolved so that laptops have these CPUs, we can expect to see wearables with 133-150 MHz Pentium CPUs in early 1997, with continued evolution towards Pentium Pro and PowerPC CPUs in the later part of 1997 and into 1998. The current combination of slow and narrow ISA and PCMCIA buses is expected to give way to some combination of CardBus (a laptop version of the desktop PCI bus), Accelerated Graphics Port, Zoomed Video, and Universal Serial Bus by late 1997 or early 1998. This will obviously and dramatically improve system performance in almost all areas. In addition to the dedicated buses for graphics and video just named, we can expect the integration of low-cost 3D graphics hardware (approximately equal to an SGI Reality Engine 1) onto the motherboard by sometime in 1998. This is a specific instance of a more general trend towards greater integration of functions that used to require separate add-in cards (wireless networking, cellular phone, speech recognition) onto the motherboard.

3.1.2 Storage

Current storage solutions such as PCMCIA disks are a focus for a disproportionately large number of problems associated with current wearables. Some of these, such as a difficulty in continuing to function after being dropped 6 feet onto a concrete floor, are intrinsic to the design of the devices themselves. Disks are also one of the two largest consumers of battery power (the display is the other), although the disk hardware is not entirely to blame here. Several modern popular operating systems, such as Windows 95, Windows NT and most flavors of UNIX, are constantly accessing the disk for virtual memory support regardless of the amount of physical memory in the system. Whether or not current capacities are adequate is obviously application-dependent, although for our applications the more storage the better. One solution may be to rely more on dynamically downloading and caching data when it is needed over a wireless link instead of preloading it to begin with. Another expected solution is the development of large (500MB - 1GB) flash memory modules that can serve as solid state disks. These devices would be much faster, have much lower

power consumption, and be far more shock resistant than current drives. Intel has promised such devices by 1998.

3.1.4 Packaging, Expandability, and Connectivity

There is already something of a split in the nascent wearables market between fixed-function units that are not easily (or at all) reconfigurable or expandable and more general purpose wearables with available expansion slots that are accessible by end-users. We expect that the smallest of the fixed function units will soon fit in a shirt pocket, and that some sort of convergence with the current handheld systems and PDAs will then happen. The current tangle of cables and physical connectors that link the central system with the various peripherals will give way to a scheme that takes local advantage of the wireless communications that are also expected to be built in. The notion of a 'personal area network', or a human-sized wireless network that connects all of the body-worn devices is in the early stages of development.

4 FIELD TEST RESULTS AND DESIGN GUIDELINES

In this section we will discuss some of the lessons learned during the development of the 757 application, describe the field tests and discuss some of the early results. The lessons and results have to do with input devices, head-mounted display issues, and the software portion of the system user interface, so the discussion is divided into these three sections following the description of how the test evaluations were organized.

4.1 Field Test Setup

The 757 field tests were done at Renton Field, in Renton, Washington, using flight test mechanics with between 8 and 15 years of experience. Their job is to make sure that the plane is ready to fly before its first flight test. Because these planes have not yet flown, the problems the mechanics encounter can be in any system or location on the plane, so their

troubleshooting experience is substantially broader and deeper than most aviation mechanics. Each subject first filled out a demographic questionnaire that gathered data about work history and previous computer usage. They were then given a brief (2-5 minute) demo and hands-on training on how to use the user interface for the fuel system troubleshooting application.

Each subject was timed and videotaped while using the system - this meant moving around the aircraft out on the tarmac while performing the tests or measurements required by the scenario. Each was asked to provide a stream-of-consciousness think-aloud verbal report of goals, problems, and impressions as they did this. After the scenario was complete, they filled out a post-test survey about the usage of the system and we conducted a videotaped debriefing. The problem scenario itself involved telling the mechanic what diagnostic messages had been reported, and asking them to follow the standard procedures and use the standard manual pages and fault isolation logic trees for the reported problem. The data loaded into the prototype application included all of the manual pages and fault isolation logic trees that a sample of mechanics had asked to see during a documentation of the baseline process for solving this particular problem. This set of pages exceeded what any single mechanic asked for, and at no time during any of the trials did a mechanic ask to see information not available on the wearable at the time of the test.

4.2 Input Device Issues

A general observation about developing applications for use on a wearable is that although some of the user interface issues are the same as for conventional desktop systems, going from a mouse and keyboard in an office to body-worn devices, displays, and voice input in a wide variety of environments challenges a great many assumptions and conventions that have developed over the last 15 years or so.

Recognizing that a conventional 2-button mouse was probably going to be a thoroughly unsuitable input device, we bought, built, or borrowed a wide variety of input devices to try out with the

applications we were developing. The list of devices is shown below.

- Trackballs
- Touchpads
- Rotary dials with buttons
- Ring mice
- Force-sensitive joysticks
- 3D mice
- Speech Recognition

Speech recognition is included for completeness in this list, although it has its own set of associated issues that will be dealt with separately in a later section.

After several months of experimentation, it became clear that many of the observations, problems, and principles were clustering around four larger and interacting groups of ideas

1. The mapping between physical device and interaction task
2. The physical environment the system was to be used in
3. Device orientation
4. Device location and size

4.2.1 Device/Task Mapping

The ubiquity of the mouse is something of a mixed blessing. It is well suited for some types of 2D tasks and has undoubtedly contributed to the popularization of what we now regard as conventional graphical user interfaces. On the other hand, its ubiquity has also led to the development of interface techniques that are varying mixtures of cumbersome and clever solely for the purpose of allowing a mouse to be used for tasks that it is relatively unsuited for - interactive control of many-degree-of-freedom human figures in animation packages, for example (Esposito, 1995).

The absence of a ubiquitous or default input device for wearables squarely raises the question of what device ought to be used for a given task or application. The basic considerations of device/task mapping that arise in conventional user interface (UI) design, as described in (Jacobs, 1992), also arise here. However, for many desktop applications the analysis need not go too much further than this because it can be reasonably assumed that the user is

indoors, sitting down facing the monitor and has a mouse and keyboard available. The next three sections discuss additional factors that need to be considered when designing systems for use on a wearable out in the field.

4.2.2 Physical Environment

Many of the devices in the list just above were not designed to work or tested in the environmental conditions described earlier in this paper. Some fail immediately and some fail soon after becoming dirty or clogged, while a few stand up reasonably well to what we have subjected them to so far. Many touchpads, for example, are capacitive rather than resistive and depend on the electrical characteristics of a bare finger to complete a circuit in order to function properly. Simply putting on a pair of gloves renders these devices useless; the presence of grease or oil usually has the same effect. Pressure-sensitive touch pads depend only on a pair of membranes coming into contact at a point underneath the finger, and so will continue to work where capacitive ones will not. Unfortunately, this capacitive/resistive distinction is rarely noted in product packaging or documentation. In much the same way that the chamber on a mechanical mouse can get dirty and need cleaning, a trackball used by oily dirty gloves can require almost constant cleaning. A force-sensitive pointing device like a Trackpoint that can be easily used with bare hands becomes a lot more difficult to precisely control if the user is wearing a pair of thick gloves that muffle the force cues being sensed by the finger.

A device that has worked well for us was a dial device built by the Wearable Computers Group at Carnegie-Mellon University (CMU). It consists of a large central dial that clicks as it rotates to provide tactile and audible feedback, with 2 small buttons (which act as duplicate right mouse buttons) and 1 large (left) button around the perimeter of the dial. It is well sealed against the set of contaminants already described and can be easily used with or without gloves.

4.2.3 Device Orientation

This is one of those areas in which, perversely enough, the more experience a user has with desktop

devices, the more this experience can work against them. The standard 2D pointing devices provide a mapping between the physical motion of a mouse or finger and the virtual motion of a cursor on the screen. The fixed horizontal orientation of these devices and constant use of these devices in this orientation mean that this particular mapping is so ingrained that interaction techniques like Pie Menus can be very successful relying only on the muscle memory that this constant use provides.

The relative absence of desks in most outdoor environments and the desire for more hands-free operation mean that physical input devices will more likely be worn somewhere on the body rather than constantly held, and worn at different orientations than they had on the desktop. This change in orientation changes the physical - virtual motion mapping that the user may have learned. For example, wearing a resistive touchpad vertically on a belt so that the left and right buttons are still on the left and right makes moving the cursor up and down somewhat awkward because in order to move the cursor up, the finger must physically move down, and vice versa. Rotating the device 180 degrees so those up and down motions are correct now reverses left and right. In some cases the device driver can adjust for these reversals in software, in some cases it can not. In all cases it is an issue that needs to be explicitly dealt with. Rotating the central dial part of the dial device we have been using is orientation independent, but locating the buttons is not.

4.2.4 Device Location and Size

The notion of wearing devices on one's body brings up the obvious question of where one is to wear them. Application requirements may rule out some locations as unsuitable - in a car/truck inspection task, lying on your back underneath the vehicle is common, so the front of the body is a preferable location. During part of our 757 fuel system scenario the mechanic is on his stomach behind a seat in the flight deck checking a circuit breaker near the floor, so we mounted the dial device on the side of one hip or another, based on user preference. Using these devices in tight spaces suggests that devices that are small or at least conform to the body will cause fewer problems than

larger devices or ones that protrude more. One of the frequent problems we had with the dial was that subjects kept bumping it against equipment in the wheel wells as they were climbing up or climbing down; this caused spurious input to the system that they then had to detect and undo. Unlike a desktop mouse that is always in the users peripheral field of view, the body-worn device may not be visible at all for part or all of the time it is being used. Not only is the potential for eye-hand coordination reduced or lost, but now such things as device orientation (if that matters) and button locations may need to be discoverable by touch alone. CMU achieved this with the dial by giving it an asymmetric rounded trapezoid shape, and giving the buttons different sizes, shapes, and numbers of raised plastic dots on each.

4.3 Head-Mounted Display Issues

As described earlier, our HMD had a right-eye monocular grayscale VGA display, with the virtual image focused at about arm's length from the eye. The optics were sufficiently opaque that the outside world could not be seen through them. Two ball joints back near the ears and two near the temple allowed the front visor and eyepiece either to be flipped up out of the way or raised slightly and angled downward for a glance-up style of viewing.

The strongest reaction we received was to the notion of wearing an HMD at all. Some subjects in both evaluations were either positive or neutral about wearing one, while others were adamantly opposed to wearing one. Upon closer examination, the distinguishing factor in most cases was the extent to which the subjects' work was primarily outside the aircraft (generally opposed) or inside the aircraft (neutral or positive). For safety reasons, those subjects that worked on the outside of the plane place an extremely high priority on a completely unobstructed field of view so that they can see anything that might be coming towards them, especially from above.

There were also a number of age-related vision issues to deal with. As already mentioned, the optics focused the image at about arms' length away from the user. In a 'baseline' condition (office light, eyepiece fully down in front of the eye), all subjects

had no difficulty reading the 14-point text used in the HMD, or in switching back and forth between looking at a monitor at arms length and the HMD image. However, in full sunlight we had to cover the frontmost surface of the display unit (the side away from the eye) to prevent ambient light from washing out the display. In addition, there are several places in the scenario when the point of interest was nearer or farther away than arms length, and switching back and forth to read the text of the diagnostic procedure or work on the part required frequent refocusing. The effort required for this repeated refocusing was too noticeable to several subjects over 40. An unexpected interacting factor was the presence of bifocals on many of the older subjects. The near-focus part of their lens was in the lower portion, and the far-focus in the upper portion. Those that used the glance-up style of wearing were then looking at some or the entire arms-length image through a far-focus lens, which contributed to some of the difficulty in reading the text clearly. We are currently in the very early stages of considering a new display design in which the display elements sit beneath the eyes, angled upward as way of dealing with this problem.

4.4 Software Issues

4.4.1 Screen Real Estate

One of the most challenging (but temporary) problems in designing these applications was coping with grayscale VGA screen real estate during a time when 19" or 21" monitors running at least 24 bit color at 1280 x 1024 has become common. Having only one fourth of the screen real estate meant that pixels were doled out with an eyedropper, metaphorically speaking. Since the dial device did not support general 2D positioning and speech input generally makes a poor and imprecise pointer; the user could not manually move or resize the application windows. There was quite a bit of experimentation juggling the variables of window size, window location, font size, and font style to make sure that enough information was visible and that the visible information was readable. We decided against trying to display wiring diagrams and schematics since these images were several feet

on a side and plotted at 300 dpi; given the limited resolution currently available this would have been a bit like looking at a map through a soda straw. Since we had no graphics, it turned out that for these applications the lack of color was not a problem. We finally settled on 14-point type and used a mix of regular and boldface type, using the boldface to indicate selectable hyperlinks in displayed manual pages. Since managing screen real estate was primarily the responsibility of the application, it maintained a record of which windows were open and what their sizes were. As a window opened or closed, the surrounding open windows automatically resized to make best use of the available screen space. Higher resolution (1280 x 1024) displays and color VGA displays are expected in 1997 which will reduce some of these design difficulties.

4.4.2 Speech Input

The speech recognition system that we used, after considerable evaluation of alternatives, was from Speech Systems Inc., and supported speaker-independent continuous speech that is mapped into the best fit from a context-free grammar that we built. There was considerable concern early in the development that speech would not function acceptably well (as reliably as a physical input device) in the noisy environment of an airstrip. There was a deliberate design decision to provide a bit of redundancy by making speech and the dial parallel and equivalent input channels so that if either failed the application could still be used with the other. This had the side benefit of allowing us to look at user preferences for speech or the dial for giving commands to the system. Some early results from the evaluation data suggest that while both devices worked well, users preferred speech over the dial because they viewed it as a more direct way of giving a command; in the rare case that speech did not work they fell back to using the dial.

Early recognition rates were not acceptable, so we set about tuning the vocabulary. We had 10 subjects say every word in the vocabulary and looked for words that were repeatedly misrecognized. Ironically, most of the word and phrase recognition errors came from the presence of 'yes' and 'OK' in the vocabulary; two words that are common in

Windows dialog boxes. After removing these words and substituting 'confirm', recognition rates rose so that on average we had 1 misrecognition per subject.

The last issues to deal with have to do with providing feedback to the user. More specifically, how does the user know what he can say, and how does he know what the system thinks he has said? The second of these was easy - a fixed area of the screen labeled 'I heard' always contained the text of the last utterance. Solving the first was a bit more involved and driven by two goals. The first was the notion of parallel and functionally equivalent input channels. The second goal was to minimize the required training time and make this system as 'walk-up-and-use' as possible. Since the subjects were able to successfully use this system with at most 5 minutes of instruction, we believe we have largely succeeded. After examining quite a few variations, the solution we settled on was to have all of the voice commands visible in one portion of the screen and 1 standard Windows button corresponding to each voice command visible in another portion of the screen. The sets of buttons and voice commands were shown side by side, each in 2 columns of 4, to suggest the parallel and equivalent nature of these channels. Rotating the dial caused the Windows input focus to change from one button to the next, skipping over any buttons that were grayed out because those commands were not available at this point in the application. The post-test usability scores for both of these features were both quite positive - 9 out of 10 for the last utterance, and 8.5 out of 10 for command availability. The application as a whole received an 8 out of 10.

5 FUTURE WORK

Based on the debriefings with the mechanics, two requirements consistently came through. The first was that of all the graphics data that is available to a mechanic, the single most important piece is the wiring diagram. It is not often needed, but when it is, it is essential and there is no real substitute. Higher resolution HMDs and sourceless head orientation trackers will allow us to experiment next year with some ideas for using head orientation as a 'head cursor' to control panning and zooming around graphic images. We also plan on extending

our work to look at collaborative work, by enabling groups of mechanics, either all on the plane or some back at the central hangar, to share voice, video, and application data (like wiring diagrams) as they collectively try to do troubleshooting and repair tasks.

REFERENCES

- Caudell, T., Mizell, D. (1992) Applications of Augmented Reality. Proceedings of HICSS 92.
- Esposito, C. (1995) Of Mice and Monkeys: A Specialized Input Device for Virtual Body Animation. Proceedings of the ACM 1995 Symposium on Interactive 3D Graphics.
- Jacob, R. and Sibert, L. (1992) The Perceptual Structure of Multidimensional Input Device Selection. CHI 92 Conference Proceedings.
- Janin, A. (1994) A Videometric head tracker for augmented reality applications, Proceedings SPIE East.