

Wearable Computers as Intelligent Agents

Thad Starner
GVU Center, College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280 USA

Forty years ago, pioneers such as J.C.R. Licklider and Douglas Englebart championed the idea of interactive computing as a way of creating a “man-machine symbiosis” where mankind would be enabled to think in ways that were previously impossible [3, 5]. Unfortunately, in those times, such a mental coupling was limited to sitting in front of a terminal and keying in requests. Today, wearable computers, through their small size, proximity to the body, and usability in almost any situation, may enable a more intimate form of cognition suggested by these early visions. Specifically, wearable computers may begin to act as intelligent agents during everyday life, assisting in a variety of tasks depending on the user’s context.

1 Wearable computing

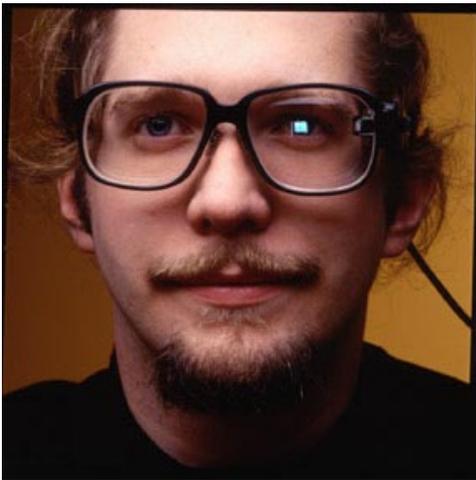


Figure 1: MicroOptical Corporation’s head-up display embedded in prescription eyeglasses.

Wearable computers are generally equated with head-up, wearable displays, one-handed keyboards, and custom computers worn in satchels or belt packs (see Figures 1-2). Often, the public focuses more on the gadgetry involved in wearable computing than on the field’s goals. Wearable computing may be described as the pursuit of an interface ideal – that of a continuously-worn intelligent assistant that augments the wearer’s memory, intellect, creativity, communication skills, physical senses, and physical abilities. Unlike desktop applications such as spreadsheets or graphics drawing programs which monopolize the user’s attention, wearable computer applications should be designed to take a secondary role.



Figure 2: Wearable computing platforms worn at MIT and Georgia Tech.

Wearable computers often perform background tasks such as providing reminders, capturing information or experience, and retrieving time-sensitive information in support of the user. In order to perform such tasks, the interface must take on a form of “symbiosis” between the user and the computational assistant.

J.C.R. Licklider, an early Advanced Research Projects Agency (ARPA) director for the United States whose research agenda is often credited for laying the foundations of the Internet and interactive computing, hinted at such a symbiosis in his writings in 1960 [5]:

“Man-computer symbiosis” is a sub-

class of man-machine systems. There are many man-machine systems. At present however, there are no man-computer symbioses. ... The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today.

To achieve the tight partnership suggested by Licklider in daily life, the computer must be a constant companion for the user. The computer must share in the experience of the user's life, drawing input from the user's environment to learn how the user reasons and communicates about the world. As it learns, the computer can provide increasingly useful and graceful assistance. Ideally, a wearable computer

1. Persists and provides constant access to information services: Designed for everyday and continuous use, the wearable can interact with the user at any given time, interrupting when necessary and appropriate. Correspondingly, the wearable can be accessed by the user quickly and with very little effort.
2. Senses and models context: In order to provide the best cognitive support for the user, the wearable must try to observe and model the user's environment, the user's physical and mental state, and the wearable's own internal state.
3. Adapts interaction modalities: The wearable should adapt its input and output modalities automatically to those which are most appropriate and socially graceful at the time. In many instances, the computer interface will be secondary to the user's primary task and should take the minimal necessary amount of the user's attention. In addition, the interface should guarantee privacy when appropriate, adapt to its user over time, and encourage personalization.
4. Augments and mediates interactions with the environment: The wearable should provide universal information support for the user in both the physical and virtual realms. The wearable should mediate between automation or computation in the environment and the user to present a interface consistent with the user's preferences and abilities. In addition, the wearable should manage potential interruptions, such as phone calls or e-mail, to best serve its user.

Obviously, this is an ambitious list of desired attributes, requiring many years of research. However, the effort may provide insights into human intelligence and everyday human-computer interfaces.

2 Why use wearable computers?

2.1 Portable consumer electronics

Some people wear too many computers already. A businessman might be seen carrying a personal digital assistant (PDA) to keep his contacts, a cellular telephone to make calls, a pager to help screen incoming phone calls and messages, a laptop for business applications, an electronic translator for practicing a new language, and a calculator wristwatch. These devices contain very similar components: a microprocessor, memory, screen, keyboard or input stylus, battery, and, in the case of the first three, a wireless modem [12].

Many other portable consumer electronic devices are evolving to have similar components as well: the compact disc player is becoming the computer-driven MP3 player, the portable dictaphone is being replaced by solid state audio digitizers, digital cameras are beginning to supplant the film camera markets, and digital camcorders have outmoded their analog counterparts. These devices share many of the same components; their main distinctions are determined by their sensors, interfaces, and application software. Why not take advantage of commonality in components to eliminate cost, weight, and redundancy and improve connectivity and services?

One could imagine a box the size of a deck of cards enclosing a powerful yet energy-conserving CPU and a large capacity data storage device. This pocket-sized wearable computer has one output display – an LED to indicate that it is on and that its wireless network is functioning. The wireless network connects peripherals to the wearable computer in a radius of about two to three meters centered at the body.

The user defines the functionality of the wearable computer by the peripherals chosen. For example, if the user likes to listen to music, wireless earphones allow access to the MP3's stored on the wearable computer's hard drive. Adding a walnut-sized camera and the appropriate software transforms the wearable computer into a camcorder. With an Internet modem, the wearable becomes a pager, cellular phone, web-browser, and e-mail reader. Connecting medical sensors to the wearable concentrates many diagnostic and recording devices into one unit.

Thus, with wearable computing and a wireless body-centered network, redundancy in components

can be reduced. In addition, functionality can be increased as derivative services are created leveraging the convergence of media on the body. The portable consumer electronics market will become fertile ground for rapid innovation. Companies can respond to a new need or niche market quickly with an appropriate peripheral or software without having to re-design all the subsystems on each iteration. Sophisticated portable electronics will become cheaper and more powerful for the consumer, and the computer industry will get a new, attractive upgrade path to pursue.

2.2 Mediating interactions in the world

The introduction of the windows, icon, menu, and pointer (WIMP) desktop metaphor in graphical user interfaces provided a means by which a user interface could mediate between applications and the user. Similarly, wearable computers will help provide a consistent interface to computationally-augmented objects in the physical world. For example, the Gesture Pendant (Figures reffig:pendant2-4) is a computer vision-based gesture recognizer worn as a piece of jewelry (the design is similar to the desk-based Toshiba “Motion Processor” project [15]). The device acts as a universal remote control to stereos, televisions, VCRs, home thermostats, and lighting systems [14]. Remote controls may be difficult for older adults to use due to dexterity, vision, or cognitive impairments; instead, the Gesture Pendant recognizes broad gestures to perform the same task. Since many functions are similar for many devices, one gesture may provide an intuitive command across the devices. For example, raising or lowering a finger in front of the pendant might raise or lower the volume of a stereo, increase or decrease the temperature setting on a thermostat, and brighten or dim the lights depending on which object the wearer faces. Thus, the user may define or learn one interface and then apply it repeatedly to many different devices depending on the context. One can imagine a more general purpose wearable computer that, through software, adapts itself to provide a consistent interface to whatever electronics or computer systems are available in the environment. No matter where in the world the wearable user travels, the wearable provides an appropriate interface in the user’s native language to home automation that may be available in the user’s environment. Such infrastructure will be a boon to the handicapped. For example, imagine a tactile display for a blind wearable user that renders simple graphics on the user’s back [2, 7]. When provided with a wireless interface and the appropriate software hooks, the wearable computer can translate images from graphical systems,

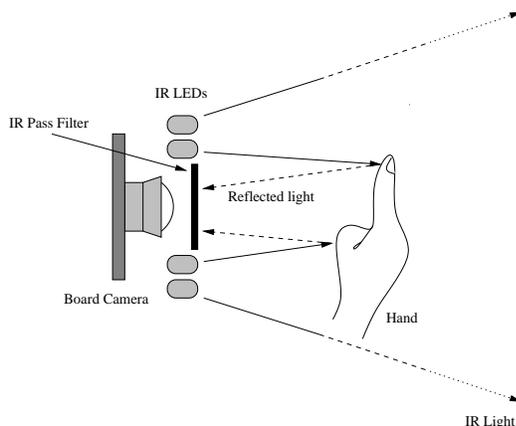


Figure 3: The Gesture Pendant’s infrared light emitting diodes invisibly illuminate the area in front of the user’s body so that the camera embedded in the pendant can recognize the user’s gestures. A filter that blocks visible light helps the vision system ignore lights in the environment

such as those used on interactive map kiosks, into tactile images with which the user can interact. Already wearable computer systems are being investigated for assisting blind users crossing the wide, busy streets of Atlanta. These systems “watch” for infrared beacons embedded in walk street signals and translate this into a tactile wayfinding system for the user [10].

2.3 Aiding human to human communication

The wearable can also assist in human to human communication. For example, Telesign, an ongoing research project at Georgia Tech, uses a hat-mounted, downward looking camera in an effort to recognize sign language (Figure 5). A goal of Telesign is to demonstrate a limited working prototype of a cellular phone for the Deaf as shown in Figure 6. Such a system could also be used as part of an interpreting system when a deaf person needs to communicate with a hearing person who does not sign. In carefully controlled laboratory conditions with user dependent training, the recognition system performs sentence-level American Sign Language recognition at 97% accuracy with signs chosen from a forty word vocabulary. Currently, we are trying to design computer vision and interface techniques that will allow the creation of a more mobile system.

Similar systems might be used to translate between spoken languages. Carnegie Mellon researchers have already demonstrated a English to Croatian speech translator for use in the Bosnian

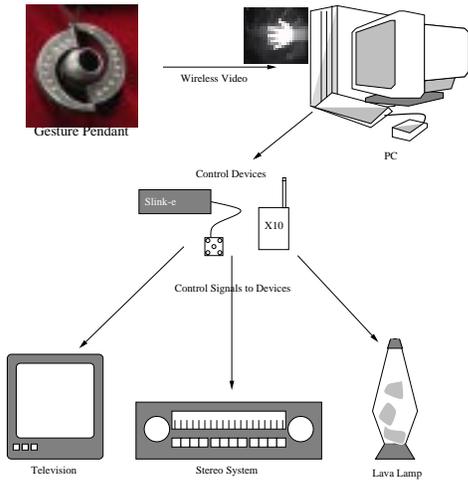


Figure 4: The Gesture Pendant allows users to control appliances in their home through gesture.

peace-keeping mission [13].



Figure 5: The hat-mounted camera, pointed downward to the hands and the corresponding view.

2.4 Providing context sensitive reminders and intelligent assistants

Context sensitive interfaces may become very powerful with wearable computers. The wearable might listen to the wearer's conversations and provide reminders as appropriate to the current topic [8, 16]. Such reminders might range from the text of the last e-mail message exchanged between the participants of an ongoing conversation to recommendations of local restaurants that both participants would enjoy.

2.4.1 Perception and symbiosis

In one approach to creating intelligent assistants, the wearable computer may be a first-class participant in the user's world (see Figure 7). The wearable senses the world from the same "first-person" perspective as the human. Both the wearable and the human have the potential to directly influence the sensors of each other. For example, the human, through his head movement, could direct the wearable's video

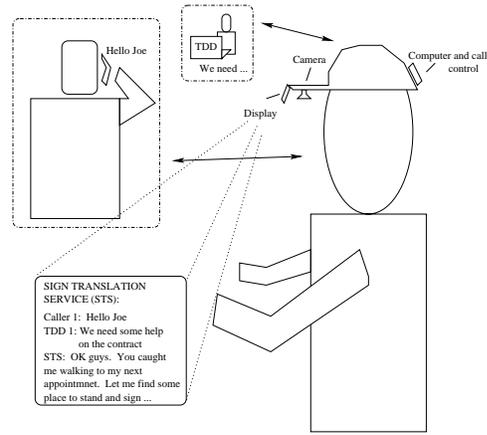


Figure 6: Telesign: A cellular phone system for the hearing-impaired.

camera to look at a particular object. Correspondingly, the wearable could mediate reality [6] for the user by displaying what it is seeing with its camera on the user's head-up display. Such a system may be valuable for augmenting what the user could normally perceive, similar to providing range information while using a telephoto lens. This mediation could also provide protection against distracting or dangerous visual input, such as directly looking at the sun while scanning the sky.

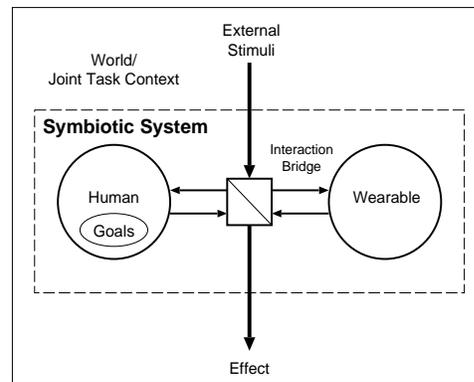


Figure 7: Wearable/user symbiosis.

In addition, both the wearable and the user can manipulate the environment. In most cases, the wearable will effect the environment by communicating with other pieces of automation, such as the user's cellular phone, vehicle, or personal robotics, while the user will effect the environment by performing a physical act, such as a making gesture or utterance. Both the wearable and the human have the ability to observe each other's manipulations. The wearable exploits its on-body cameras and microphones to observe the user's gestures and speech,

and the user may use the wearable interface itself to monitor the traffic history between the wearable and other automation. In addition, the user should be able to query the wearable as to the reasoning for each action.

One of the strengths of the symbiotic system is the cross-monitoring of the agents' (human and wearable) actions. The wearable continuously tries to create and update a user model of the human's goals and abilities. The wearable can passively monitor the human's reactions to external stimuli and learn to recognize meaningful events in the human's life [1]. One can imagine this behavioristic approach being used as research direction for artificial intelligence; such an approach would try to emulate not just human behavior, but a particular human's behavior in a reasonable manner given novel stimuli. From cognitive science we know that such a "user-model complete" approach has fundamental limits due to a lack of knowledge of the internal state of the user; however, a wearable intelligent assistant can take a more active approach. The wearable assistant may gain some insight into the user's goals by how the wearer uses the computer. By maintaining a record of all its interactions and the corresponding external conditions during those interactions, the assistant might learn to associate a certain set of external stimuli with user desires. For example, if the user keys "Joe Snyder" while alone in a shopping mall, the wearable might learn to bring up a map showing Joe's current position on a map relative to the user. However, if the user keys "Joe..." while Joe is in the user's immediate vicinity, the wearable displays the last e-mail sent by Joe so that the users may proceed with an informed conversation.

In addition to learning through interaction, the wearable may form a user model by directly asking the user for help. The assistant may associate a certain user behavior (for example, beginning a computer application for inventory management) with a given context (like exiting a building at a geographical location indicating a supply warehouse). The assistant may then alert the user that it believes it can recognize this behavior, ask whether or not it is a significant event, what the event should be called, and if it should begin the computer application automatically when it next observes the event.

Of course, such intelligent assistants based on real-world, first-person perception are very difficult to create. Most computer gesture and speech recognition systems are designed for use in a laboratory and observe the user from sensors mounted in the environment instead of on the body. In most cases, the variable noise conditions of mobility pose a significant barrier to the adoption of such systems.

2.4.2 Just-in-time Information: The Remembrance Agent (RA)

Computers perform well at storing data and executing repetitive functions quickly. Humans, on the other hand, make intuitive leaps and recognize patterns and structure, even when passive. Thus, an interface in which the wearable computer helps the user remember and access information seems profitable. While word processing may comprise the majority of computer use, it requires only a small fraction of the computer's processing power. Instead of wasting the remaining power, an information agent can use the time to search the user's personal text database for information relevant to the current task. The names and short excerpts of the closest matching files can then be displayed. If the search engine is fast enough, a continuously changing list of matches can be maintained, increasing the probability that a useful piece of information will be recovered. Thus, the agent can act as a memory aid. Even if the user mostly ignores the agent, he will still tend to glance at it whenever there is a short break in his work. In order to explore such a work environment, the Remembrance Agent [9] was created.

The benefits of the Remembrance Agent (RA) are many. First, the RA provides timely information. If the user is writing a paper, the RA might suggest relevant references. If the user is reading email and scheduling an appointment, the RA may happen to suggest relevant constraints. If the user is holding a conversation with a colleague at a conference, the RA might bring up relevant associations based on the notes the user is taking. Since the RA "thinks" differently than its user, it often suggests combinations that the user might never assemble himself. Thus, the RA can act as a constant "brain-storming" system.

The Remembrance Agent can also help with personal organization. As new information arrives, the RA, by its nature, suggests files with similar information. Thus, the user gets suggestions on where to store the new information, avoiding the common phenomenon of multiple files with similar content (e.g. archives-linux and linux-archives). The first trial of the prototype RA revealed many such inconsistencies within the sample database and even suggested a new research project by its groupings.

As a user collects a large database of private knowledge, his RA becomes an expert on that knowledge through constant re-training. A goal of the RA is to allow co-workers to access the "public" portions of this database conveniently without interrupting the user. Thus, if a colleague wants to know about augmented reality, he simply sends a message to the user's Remembrance Agent (e.g. thadra@cc.gatech.edu). The RA can then return its best

guess at an appropriate file. Thus, the user is not interrupted by the query, and he never has to format his knowledge explicitly, as with HTML. Knowledge transfer may occur in a similar fashion. When an engineer trains his successor, he can also transfer his RA's database of knowledge on the subject so that his replacement may continually receive the benefit of his experience even after he has left. Finally, if a large collective of people use Remembrance Agents, queries can be sent to communities, not just individuals. This allows questions of the form "How do I reboot a Sun workstation?" to be sent to 1000 co-workers whose systems, in their spare cycles, may send a response. The questioner's RA, who knows how the user "thinks," can then organize the responses into a top 10 list for convenience.

The Remembrance Agent (RA) is comprised of two parts, the user interface and the search engine. The user interface continuously watches what the user types and reads, and it sends this information to the search engine. The search engine finds old email, notes files, and on-line documents that are relevant to the user's context. This information is then unobtrusively displayed in the user's field of view.

The Remembrance Agent displays one-line suggestions at the bottom of the emacs display buffer, along with a numeric rating indicating the relevancy of the document. These items contain just enough information to represent the contents of the full document being suggested. For example, the suggestion line for a piece of email includes the subject line, the sender, and the date received. The suggestion line for a notes file contains the file name, owner, date last modified, and a summary the file. With a simple key combination, the user can display the full text of a suggested document. The current information retrieval implementation uses a term frequency, inverse document frequency algorithm for comparing documents [11].

Figure 8 shows the output of the Remembrance Agent (in the bottom buffer) while editing a document (top buffer). The reference database for the RA was e-mail archives. Personal email and notes files are good sources for an RA to reference, because these files are already personalized and automatically change with the user's interests.

The Remembrance Agent can also be a method for casual, low-overhead collaboration. Current everyday wearable computer users take copious notes during their normal day to assist in their work and recreation. As an informal experiment, three wearable computer users from the same laboratory combined their personal databases used in conjunction with the RA so that they could gain the benefit of each other's past experiences. Interestingly, RA sug-

gestions from a colleague's database can be quite disquieting. The user recognizes the significance of the suggestion and can almost claim the memory as his own due to the similarity with his own experiences, but he knows that it isn't his entry. These "shadow memories" provide an asynchronous form of collaboration where the collaborators are passive and their wearable agents share information as a matter of the normal use of the wearable.

While the Remembrance Agent gains most of its contextual information from typed text, wearable computers have the potential to provide a wealth of contextual features [4]. Additional sources of user context information may include time of day, location, biosignals, face recognition, and object recognition. Currently the RA is being coupled with voice recognition to provide information support during conversations. Thus, the RA begins to show the advantages of wearable, context-driven augmented reality. However, with an even more comprehensive agent, the wearable computer may be able to uncover trends in the user's everyday life, predict the user's needs, and pre-emptively gather resources for upcoming tasks [1].

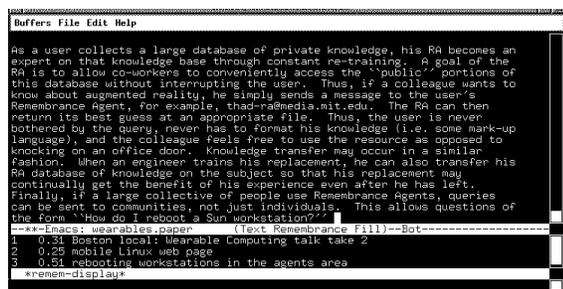


Figure 8: The Remembrance Agent in use while writing a paper. The bottom buffer shows the RA's suggestions for files relevant to the user's current work.

3 Conclusion

Wearable computing enables significant new research opportunities in interface, artificial intelligence, and perception. As research into perception and user modeling through devices carried on the body progresses, new intelligent interfaces will result that will reduce work and complexity and lead to new capabilities.

4 Acknowledgements

This paper reflects the hard work of many members of the Georgia Tech Contextual Computing Group

and the MIT Media Laboratory Wearable Computing Project. In particular, Brad Rhodes and Brad Singletary have influenced the discussions herein.

[16] B. Wong and T. Starner. Conversational speech recognition for creating intelligent agents on wearables. In *Human Computer Interface International (HCII)*, New Orleans, LA, 2001.

References

- [1] B. Clarkson and A. Pentland. Unsupervised clustering of ambulatory audio and video. In *ICASSP*, 1999.
- [2] C. Collins, L. Scadden, and A. Alden. Mobility studies with a tactile imaging device. In *Fourth Conf. on Systems and Devices for the disabled*, Seattle, WA, June 1977.
- [3] D. Engelbart. Augmenting human intellect: A conceptual framework. Technical Report AFOSR-3233, Stanford Research Institute, Menlo Park, CA, October 1962.
- [4] M. Lamming and M. Flynn. Forget-me-not: Intimate computing in support of human memory. In *FRIEND21: Inter. Symp. on Next Generation Human Interface*, pages 125–128, Meguro Gajoen, Japan, 1994.
- [5] J. Licklider. Man-computer symbiosis. *IRE Transactions of Human Factors in Electronics*, HFE-1(1):4–11, March 1960.
- [6] S. Mann. On the bandwagon or beyond wearable computing. *Personal Technologies*, 1(4), 1997.
- [7] J. Nissen. Cupid and tactile exploration of virtual maps. In *IEEE VRAIS98 Workshop on Interfaces for Wearable Computers*, March 1998.
- [8] B. Rhodes. *Just-In-Time Information Retrieval*. PhD thesis, MIT Media Laboratory, Cambridge, MA, May 2000.
- [9] B. Rhodes and T. Starner. Remembrance agent: A continuously running automated information retrieval system. In *Proc. of Pract. App. of Intelligent Agents and Multi-Agent Tech. (PAAM)*, London, April 1996.
- [10] D. Ross and B. Blasch. Wearable interfaces for orientation and wayfinding. In *ACM conference on Assistive Technologies*, pages 193–200, 2000.
- [11] G. Salton. *The SMART Retrieval System - experiments in automatic document processing*. Perntice-Hall, Inc., Englewood Cliffs, NJ, 1971.
- [12] O. Shivers. Bodytalk and the bodynet: A personal information infrastructure. Technical report, MIT Laboratory for Computer Science, Personal Information Architecture Note 1, Cambridge, MA, 1993.
- [13] A. Smailagic, D. Siewiorek, R. Martin, and D. Reilly. Cmu wearable computers for real-time speech translation. In *IEEE Intl. Symp. on Wearable Computers*, San Francisco, CA, 1999.
- [14] T. Starner, J. Auxier, D. Ashbrook, and M. Gandy. The gesture pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. In *IEEE Intl. Symp. on Wearable Computers*, Atlanta, GA, 2000.
- [15] Toshiba. Toshiba's motion processor recognizes gestures in real time. Available at: <http://www.toshiba.com/news/980715.htm>, July 1998.