



Wearable Computing: Toward Humanistic Intelligence

Steve Mann, *University of Toronto*

Over the past 20 years, wearable computing has emerged as the perfect tool for embodying *humanistic intelligence*. HI is intelligence that arises when a human is part of the feedback loop of a computational process in which the human and computer are inextricably intertwined.

It is common in the field of human–computer interaction to think of the human and computer as separate entities. (Indeed, the term “HCI” emphasizes this separateness by treating the human and computer as different entities that interact.) However, in HI theory, we prefer not to think of the wearer and the computer with its associated I/O apparatus as separate entities. Instead, we regard the computer as a second brain and its sensory modalities as additional senses, which synthetic synesthesia merges with the wearer’s senses.

When a wearable computer functions in a successful embodiment of HI, the computer uses the human’s mind and body as one of its peripherals, just as the human uses the computer as a peripheral. This reciprocal relationship is at the heart of HI.

Assisting human intelligence

HI also suggests a new goal for signal-processing hardware—that is, in a truly personal way, to directly assist, rather than replace or emulate, human intelligence. To facilitate this vision, we need a simple and truly personal computational signal-processing framework that empowers the human intellect.

The HI framework, which arose in Canada in the 1970s and early 1980s, is in many ways similar to Douglas Engelbart’s vision that arose in the 1940s while he was a radar engineer. Engelbart, while seeing images on a radar screen, realized that the cathode ray screen could also display letters of the alphabet and computer-generated pictures and graphical

content. Thus, computing could be an interactive experience for manipulating words and pictures. Engelbart envisioned the mainframe computer as a tool for augmented intelligence and communication, which many people in a large amphitheater could use to interact.^{1,2}

Although Engelbart did not foresee the personal computer’s significance, modern personal computing certainly embodies his ideas. This special issue presents a variety of attempts at realizing a similar vision, but with the computing resituated in the context of the user’s personal space. The idea is to move the tools of augmented intelligence and communication directly onto the body. This will give rise not only to a new genre of truly personal computing but also to some new capabilities and *affordances* arising from direct physical proximity to the human body, allowing the HI feedback loop to develop. (Affordances are what an environment offers to an organism.³) Moreover, a new family of applications will arise, in which the body-worn apparatus augments and mediates the human senses.

HI theory

HI’s goals are to work in extremely close synergy with the human user and, more important, to arise partly because of the very existence of the human user.⁴ HI achieves this synergy through a user interface to signal-processing hardware that is in close physical proximity to the user and is continuously accessible.

Operational modes

An embodiment of HI has three fundamental operational modes: *constancy*, *augmentation*, and *mediation*.

Constancy. An embodiment of HI is *operationally constant*; that is, although it might have power-saving (sleep) modes, it is never completely shut down (as is typically a calculator worn in a shirt pocket but turned off most of the time). More important, it is also *interactionally constant*—that is, the device’s inputs and outputs are always potentially active. Interactionally constant implies operationally constant, but operationally constant does not necessarily imply interactionally constant.

So, for example, a pocket calculator kept in your pocket but left on all the time is still not interactionally constant, because you cannot use it in this state (you still have to pull it out of your pocket to see the display or enter numbers). A wristwatch is a borderline case. Although it operates constantly to keep proper time and is conveniently worn on the body, you must make a conscious effort to orient it within your field of vision to interact with it.

Wearable computers are unique in their ability to provide this always-ready condition, which might, for example, include retroactive video capture for a face-recognizing reminder system. After-the-fact devices such as traditional cameras and palmtop organizers cannot provide such retroactive computing.

Figure 1a depicts the signal flow from human to computer, and computer to human, for the constancy mode.

Once, people did not see why devices should be operationally and interactionally constant; this shortsighted view led to the development of many handheld or so-called “portable” devices. In this special issue, however, we will see why it is desirable to have certain personal-electronics devices, such as cameras and signal-processing hardware, always on—for example, to facilitate new forms of intelligence that assist the user in new ways.

Augmentation. Traditional computing paradigms rest on the notion that computing is the primary task. Intelligent systems embodying HI, however, rest on the notion that computing is not the primary task. HI assumes that the user will be doing something else while computing, such as navigating through a corridor or walking down stairs. So, the computer should augment the intellect or the senses, without distracting a primary task. Implicit in this mode is a spatiotemporal con-

textual awareness from sensors (wearable cameras, microphones, and so on).

Figure 1b depicts the signal flow between the human and computer in this mode.

Mediation. Unlike handheld devices, laptop computers, and PDAs, good embodiments of HI can *encapsulate* the user (see Figure 1c). Such an apparatus doesn’t necessarily need to completely enclose us. However, the basic concept of mediation allows for whatever degree of encapsulation is desired (within the limits of the apparatus), because it affords us the possibility of a greater degree of encapsulation than traditional portable computers. As with the augmentation mode, a spatio-temporal contextual awareness from sensors is implicit in this mode.

The encapsulation that mediation provides has two aspects, one or both of which can be implemented in varying degrees, as desired.

The first aspect is *solitude*. The ability to mediate our perception lets an embodiment of HI act as an information filter. For example, we can block out material we might not wish to experience (such as offensive advertising) or replace existing media with different media (for example, see the “Filtering Out Unwanted Information” sidebar). In less extreme manifestations, it might simply let us moderately alter aspects of our perception of reality. Moreover, it could let us amplify or enhance desired inputs. This control over the input space contributes considerably to the most fundamental HI issue: user empowerment.

The second aspect is *privacy*. Mediation lets us block or modify information leaving our encapsulated space. In the same way that ordinary clothing prevents others from seeing our naked bodies, an embodiment of HI might, for example, serve as an intermediary for interacting with untrusted systems, such as third-party implementations of digital anonymous cash. In the same way that martial artists, especially stick fighters, wear a long black robe or skirt that reaches the ground to hide the placement of their feet from their opponent, a good embodiment of HI can clothe our otherwise transparent movements in cyberspace and the real world.

Other technologies such as desktop computers can, to a limited degree, help us protect our privacy with programs such as Pretty Good Privacy. However, the primary weakness of these systems is the space between them and their user. Compromising the link between the human and the computer (perhaps through a Trojan horse or other planted

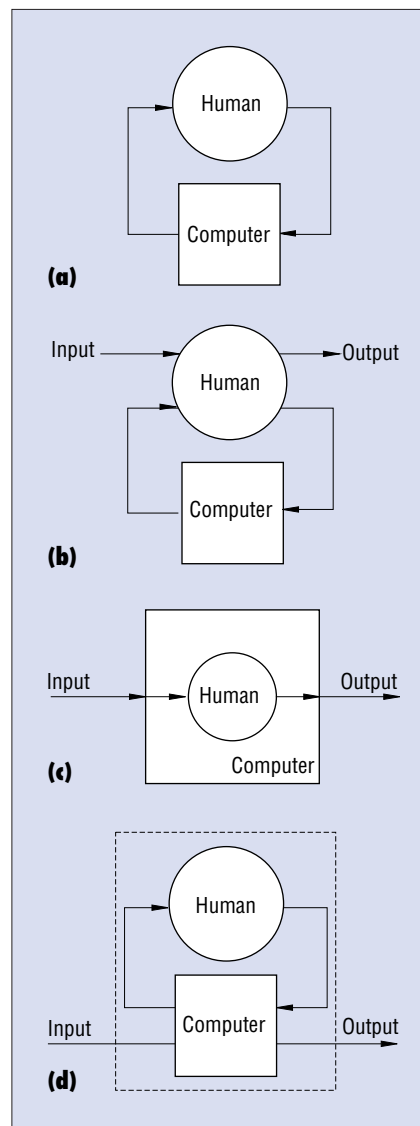


Figure 1. Signal flow paths for the three basic operational modes of devices that embody HI: (a) constancy; (b) augmentation; (c) mediation; (d) mediation (redrawn to resemble Figures 1a and 1b) emphasizing the separate protective shell that encapsulation can provide.

virus) is generally far easier when they are separate entities.

A personal information system that the wearer owns, operates, and controls can provide a much greater level of personal privacy. For example, if the user always wears it (except perhaps during showering), the hardware is less likely to fall prey to attacks. Moreover, the close synergy between the human and computer makes the system less vulnerable to direct attacks, such as someone

Filtering Out Unwanted Information

The owner of a building or other real estate can benefit financially from placing advertising signs in the line of sight of all who pass by the property (see Figure A1). These signs can be distracting and unpleasant. Such theft of solitude benefits the owner at the expense of the passersby.

Legislation is one possible solution to this problem. Instead, I propose a *diffusionist*¹ approach in the form of a simple engineering solution that lets the individual filter out unwanted real-world spam. Such a wearable computer, when functioning as a reality mediator, can create a modified perception of visual reality (see the coordinate-transformed images in Figure A2). So, it can function as a visual filter to filter out the advertising in Figure A1 and replace it with useful subject matter, as in Figure A3. Such a computer-mediated intelligent-signal-processing

system is an example application of humanistic intelligence.

Reference

1. S. Mann, "Reflectionism and Diffusionism," *Leonardo*, vol. 31, no. 2, 1998, pp. 93–102; <http://wearcam.org/leonardo/index.htm> (current 5 June 2001).

Figure A. Filtering out unwanted advertising messages (each row shows frames from a movie): (1) Advertising can be distracting and annoying. (2) A wearable computing device together with an EyeTap system (see the other sidebar) creates a modified perception of the advertising. (3) It then replaces the advertising with subject matter useful to the user.



looking over your shoulder while you're typing or hiding a video camera in the ceiling above your keyboard.

For the purposes of this special issue, we define privacy not so much as the absolute blocking or concealment of personal information, but as the ability to control or modulate this outbound information channel. So, for example, you might wish members of your immediate family to have greater access

to personal information than the general public does. Such a family-area network might feature an appropriate access control list and a cryptographic communications protocol.

In addition, because an embodiment of HI can encapsulate us—for example, as clothing directly touching our skin—it might be able to measure various physiological quantities.

Thus, the encapsulation shown in Figure 1c enhances the signal flow in Figure 1a. Fig-

ure 1d makes this enhanced signal flow more explicit. It depicts the computer and human as two separate entities within an optional protective shell, which the user can fully or partially open if he or she desires a mixture of augmented and mediated interaction.

Combining modes. The three modes are not necessarily mutually exclusive; constancy is embodied in augmentation and mediation.

These last two are also not necessarily meant to be implemented in isolation. Actual embodiments of HI typically incorporate aspects of augmentation and mediation. So, HI is a framework for enabling and combining various aspects of each of these modes.

Basic signal flow paths

Figure 2 depicts the six basic signal flow paths for intelligent systems embodying HI. The paths typically comprise vector quantities. So, the figure depicts each basic path as multiple parallel paths to remind you of the vector nature of the signals.

Each path defines an HI attribute:

1. *Unmonopolizing*. The device does not necessarily cut you off from the outside world as a virtual reality game or the like does.
2. *Unrestrictive*. You can do other things while using the device—for example, you can input text while jogging or running down stairs.
3. *Observable*. The device can get your attention continuously if you want it to. The output medium is constantly perceptible. It is sufficient that the device is almost always observable, within reasonable limitations—for example, as when a camera viewfinder or computer screen is not visible when you blink your eye.
4. *Controllable*. The device is responsive. You can take control of it at any time. Even in automated processes, you should be able to manually override the automation to break open the control loop and become part of the loop. Examples of this controllability might include a Halt button you can invoke when an application mindlessly opens all 50 documents that were highlighted when you accidentally pressed Enter.
5. *Attentive*. The device is environmentally aware, multimodal, and multisensory. This ultimately gives you increased situational awareness.
6. *Communicative*. You can use the device as a communications medium when you wish. It lets you communicate directly to others or helps you produce expressive or communicative media.

Adapting to HI

Because devices embodying HI often require that the user learn a new skill set, adapting to them is not necessarily easy. Just as a young child takes many years to become

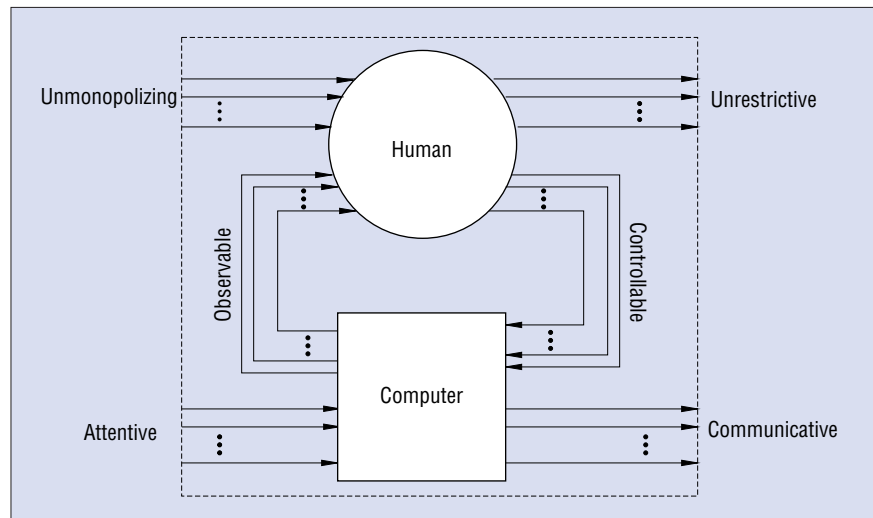


Figure 2. The six signal flow paths for intelligent systems embodying HI. Each path defines an HI attribute.

proficient at using his or her hands, some devices that implement HI have taken years of use before they begin to behave like natural extensions of the mind and body. So, in terms of human-computer interaction,⁵ the goal is not just to construct a device that can model (and learn from) the user, but, more important, to construct a device from which the user also must learn. Therefore, to facilitate the latter, devices embodying HI should provide a constant user interface that is not so sophisticated and intelligent that it confuses the user. Although the device might implement sophisticated signal-processing algorithms, the cause-and-effect relationship of the input (typically from the environment or the user's actions) to this processing should be clearly and continuously visible to the user.

Accordingly, the most successful examples of HI afford the user a very tight feedback loop of system observability. A simple example is the viewfinder of an EyeTap imaging system (see the related sidebar). In effect, this viewfinder continuously endows the eye with framing, a photographic point of view, and an intimate awareness of the visual effects of the eye's own image-processing capabilities.

A more sophisticated example of HI is a biofeedback-controlled EyeTap system, in which the biofeedback process happens continuously, whether or not the system is taking a picture. Over a long period of time, the user will become one with the machine, constantly adapting to the machine intelligence, even if he or she only occasionally deliberately uses the machine.

This special issue

In their profound and visionary article, Joshua Anhalt and his colleagues provide a background for context-aware computing, along with some practical examples of HI implemented in such forms as a portable help desk. This work comes from Carnegie Mellon University's Software Engineering Institute and IBM's T.J. Watson Research Center. The SEI is under the direction of Daniel Siewiorek, who has been working on wearable computing for many years.

This article marks an interesting departure from their previous work in military equipment maintenance applications, and suggests a branching out into applications more suitable for mainstream culture. Wearable computing has gone beyond the military-industrial complex; we are at a pivotal era where it will emerge to affect our daily lives.

Recognizing the importance of privacy and solitude issues, the authors formulate the notion of a *distraction matrix* to characterize human attentional resource allocation.

Li-Te Cheng and John Robinson also look at an application targeted for mainstream consumer culture. They report on context awareness through visual focus, emphasizing recognition of visual body cues, from the first-person perspective of a personal imaging system. They provide two concrete examples: a memory system for playing the piano and a system for assisting ballroom dancing. This work shows us further examples of how wearable computers have become powerful enough to perform vision-based intelligent signal processing.

EyeTap

One application of humanistic intelligence is an EyeTap.¹ An EyeTap is a nearly invisible miniature apparatus that causes the human eye to behave as if it were both a camera and a display. This device can facilitate lifelong video capture and can determine the presence of an opportunity or a threat, based on previously captured material.

One practical application of an EyeTap is in assisting the visually impaired. In the same way that a hearing aid contains a microphone and speaker with signal processing in between, the EyeTap causes the eye itself to, in effect, contain an image sensor and light synthesizer, with processing in between the two.

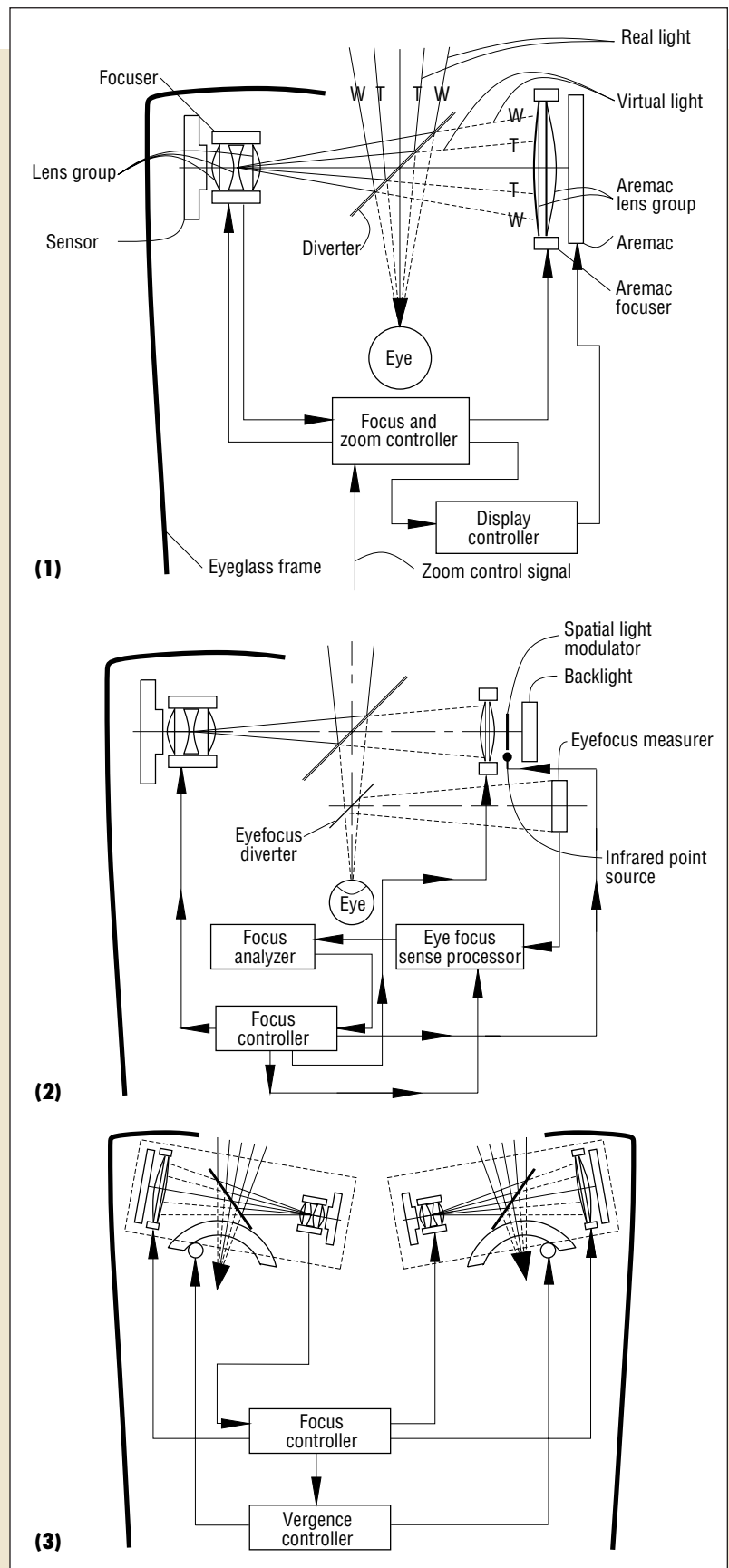
The EyeTap tracks depth by using a single control input to manually or automatically focus a camera and an *aremac* together.¹ The *aremac* ("camera" spelled backwards) is a device that resynthesizes light that was absorbed and quantified by the camera. Figure B diagrams three approaches to depth tracking. Solid lines denote real light from the subject matter, and dashed lines denote virtual light synthesized by the *aremac*.

Figure B1 shows an autofocus camera controlling the *aremac*'s focus. When the camera focuses to infinity, the *aremac* focuses so that it presents subject matter that appears as if it is infinitely far. When the camera focuses closely, the *aremac* presents subject matter that appears to be at the same close distance. A zoom input controls both the camera and *aremac* to negate any image magnification and thus maintain the EyeTap condition. W denotes rays of light defining the widest field of view. T (for tele) denotes rays of light defining the narrowest field of view. The camera and *aremac* fields of view correspond.

Figure B2 shows eye focus controlling both the camera and *aremac*. An eye focus measurer (via the *eye focus diverter*, a beamsplitter) estimates the eye's approximate focal distance. Both the camera and *aremac* then focus to approximately this same distance.

The mathematical-coordinate trans-

Figure B. Depth tracking with the EyeTap: (a) An autofocus camera controls focus of the *aremac*, which resynthesizes light that was absorbed and quantified by the camera. Solid lines denote real light from the subject matter; dashed lines denote virtual light synthesized by the *aremac*. W denotes rays of light defining the widest field of view. T (for tele) denotes rays of light defining the narrowest field of view. (b) Eye focus controls both the camera and the *aremac*. (c) An autofocus camera on the left controls focus of the right camera and both *aremac*s (as well as vergence).



formations in Figure B2 arise from the system's awareness of the wearer's gaze pattern, such that this intelligent system is *activity driven*. Areas of interest in the scene will attract the human operator's attention, so that he or she will spend more time looking at those areas. In this way, those parts of the scene of greatest interest will be observed with the greatest variety of quantization steps (for example, with the richest collection of differently quantized measurements). So, the EyeTap will automatically emphasize these parts in its composite representation.¹

This natural *foveation* process arises, not because the EyeTap itself has figured out what is important, but simply because it is using the operator's brain as its guide to visual saliency. Because operating the EyeTap does not require any conscious thought or effort, it resides on the human host without presenting any burden. However, it still benefits greatly from this form of humanistic intelligence.

In Figure B3, an autofocus camera on the left controls the focus of the right camera and both aremacs (as well as the vergence). In a two-eye system, both cameras and both aremacs should focus to the same distance. So, one camera is a focus master, and the other is a focus slave. Alternatively, a focus combiner can average the focus distance of both cameras and then make the two cameras focus at an equal distance. The two aremacs and the vergence controllers for both eyes track this same depth plane as defined by the camera autofocus.

Computing such as the EyeTap provides blurs the line between remembering and recording, as well as the line between thinking and computing. So, we will need a whole new way of studying these new human-based intelligent systems. Such an apparatus has already raised various interesting privacy and accountability issues. Thus, HI necessarily raises a set of humanistic issues not previously encountered in the intelligent systems field.

Reference

1. S. Mann, "Humanistic Intelligence/ Humanistic Computing: 'Wearcomp' as a New Framework for Intelligent Signal Processing," *Proc. IEEE*, vol. 86, no. 11, Nov. 1998, pp. 2123–2151; <http://wearcam.org/procieee.htm> (current 5 June 2001).

Kaoru Sumi and Toyoaki Nishida put context awareness in a spatiotemporal global framework, with computer-based human communication. In the context of conversation, the system illustrates how HI can serve as a human–human communications medium, mediated by wearable computer systems.

David Ross provides an application of HI for assistive technology. Besides the military-industrial complex, early HI adopters might well be those with a visual or other impairment. For this sector of the population, wearable computing can make a major difference in their lives.

Ömer Faruk Özer, Oguz Özüin, C. Öncel Tüzel, Volkan Atalay, and A. Enis Çetin describe a personal-imaging system (wearable camera system) for character recognition. Chain-coded character representations in a finite-state machine are determined by way of personal imaging as a user interface.

Soichiro Matsushita describes a wireless sensing headset. Indeed, it has often been said that a good embodiment of HI will replace all the devices we normally carry with us, such as pagers, PDAs, and, of course, cellular telephones. Thus, a context-awareness-enhancing headset is a good example of how HI will improve our daily lives.

Although I have formulated a theoretical framework for humanistic intelligence, the examples I've described in this introduction are not merely hypothetical; they have been reduced to practice. Having formulated these ideas some 30 years ago, I have been inventing, designing, building, and wearing computers with personal-imaging capability for more than 20 years. Actual experience of this sort has grounded my insights in this theory in a strong ecological foundation, tied directly to everyday life.

We are at a pivotal era in which the convergence of measurement, communications, and computation, in the intersecting domains of wireless communications, mobile computing, and personal imaging, will give rise to a simple device we wear that replaces all the separate informatic items we normally carry.

Although I might well be (apart from not more than a dozen or so of my students) the only person to be continuously connected to, and living in, a computer-mediated reality, devices such as EyeTaps and wearable computers doubtlessly will enjoy widespread use in the near future.

Twenty years ago, people laughed at this

The Author



Steve Mann is a faculty member at the University of Toronto's Department of Electrical and Computer Engineering. He built the world's first covert fully functional wearable image processor with computer display and camera concealed in ordinary eyeglasses and was the first person

to put his day-to-day life on the Web as a sequence of images. He received his PhD in personal imaging from MIT. Contact him at the Dept. of Electrical and Computer Eng., Univ. of Toronto, 10 King's College Rd., S.F. 2001, Canada, M5S 3G4. He can be reached via e-mail at mann@eecg.toronto.edu or by tapping into his right eye, <http://eyetap.org>.

idea. Now I simply think of Alexander Graham Bell's prediction that the day would come when there would be a telephone in every major city of this country.

Thus, there is perhaps no better time to introduce HI by way of a collection of articles showing how these ideas can be actually reduced to practice. ■

References

1. D.C. Engelbart, *Augmenting Human Intellect: A Conceptual Framework*, research report AFOSR-3223, Stanford Research Inst., Menlo Park, Calif., 1962; www.histech.rwth-aachen.de/www/quellen/engelbart/ahi62index.html (current 5 June 2001).
2. D.C. Engelbart, "A Conceptual Framework for the Augmentation of Man's Intellect," *Visitas in Information Handling*, P.D. Howerton and D.C. Weeks, eds., Spartan Books, Washington, D.C., 1963, pp. 1–29.
3. J. Zhang, "Categorization of Affordances," Dept. of Health Informatics, Univ. of Texas at Houston, <http://acad88.sahs.uth.tmc.edu/courses/hi6301/affordance.html> (current 3 July 2001).
4. S. Mann, "Humanistic Intelligence/Humanistic Computing: 'Wearcomp' as a New Framework for Intelligent Signal Processing," *Proc. IEEE*, vol. 86, no. 11, Nov. 1998, pp. 2123–2151; <http://wearcam.org/procieee.htm> (current 5 June 2001).
5. W.A.S. Buxton and R.M. Baecker, *Readings in Human-Computer Interaction: A Multidisciplinary Approach*, Morgan Kaufmann, San Francisco, 1987, chapters 1 and 2.
6. M. Weiser, *Ubiquitous Computing*, <http://sandbox.parc.xerox.com/ubicomp> (current 5 June 2001).
7. J. Cooperstock, *Reactive Room*, www.dgp.toronto.edu/~rroom/research/papers (current 5 June 2001).