

Envisioning, designing, and implementing the user interface require a comprehensive understanding of interaction technologies. In this forum we scout trends and discuss new technologies with the potential to influence interaction design. — **Albrecht Schmidt, Editor**

# Eyewear Computers for Human-Computer Interaction

**Andreas Bulling**, Max Planck Institute for Informatics, **Kai Kunze**, Keio University

**H**ead-worn displays and eye trackers, augmented and virtual reality glasses, egocentric cameras, and other “smart eyewear” have recently emerged as a research platform in fields such as ubiquitous computing, computer vision, and cognitive and social science. While earlier generations of devices were too bulky to be worn regularly, recent technological advances have made eyewear unobtrusive and lightweight, and therefore more suitable for daily use. Given that many human senses are located on the head, smart eyewear provides opportunities for types of interaction that were impossible before now. In this article, we highlight the potential of eyewear computing for HCI, discuss available input and output modalities, and suggest the most promising future directions for eyewear computing research, namely multimodal user modeling, lifelong learning, and large-scale (collective) human-behavior sensing and analysis.

## HEAD-BASED INPUT MODALITIES

The latest eyewear computers offer a wide range of sensing and input modalities, including eye tracking, egocentric cameras, microphones and inertial sensors, and sensors for measuring brain activity.

Head-worn eye tracking has seen significant advances in recent years and has therefore been a particularly active area of research in eyewear

computing. Eye movements are a rich source of information about the user, indicating their activities and cognitive processes [1,2,3]. The two most promising eye-tracking methods for measurements in daily-life settings are Electrooculography (EOG) and computer-vision-based eye tracking. EOG requires electrodes positioned on the skin around the eyes and measures changes in the electric potential fields caused by eye movement. EOG is computationally lightweight and can therefore be unobtrusively integrated into eye glasses, such as J!NS MEME (Figure 1; jins-meme.com). However, EOG can be subject to significant signal drift and noise and provides information only about relative eye movements; it does not provide accurate gaze direction.

In contrast, vision-based eye tracking relies on special-purpose eye cameras as well as infrared

illumination, and provides gaze direction in either 2D-scene-camera or 3D-world coordinates. The scene camera, readily integrated into head-mounted eye trackers, records high-resolution videos from the user’s perspective and therefore enables a variety of egocentric vision applications. For example, the open source PUPIL head-mounted eye tracker from Pupil Labs ([pupil-labs.com/pupil](http://pupil-labs.com/pupil)) is affordable and fully customizable, and can be implemented for long-term use (Figure 2). In addition, given that the eye cameras are typically high-resolution, they can be used for corneal imaging, which analyzes the scene reflected on the wet surface of the user’s cornea (Figure 3). Corneal imaging can provide contextual information such as details on the user’s attention, the objects and interlocutors with whom they interact, and their surroundings. Taken together, head-worn eye tracking has significant potential to enable a new class of mobile HCI applications, in particular when combined with other smart eyewear such as head-mounted displays (HMDs).

Egocentric vision has seen a similar increase in research activity. Egocentric vision can provide additional user context, for example, the objects users are interacting with as well as their activities and social interactions [4]. It can also be used for direct interactions such as detecting hand gestures, as well as for security applications such as secure identification, document verification,

### Insights

- As eyewear computers become more lightweight and unobtrusive, they become more suitable for everyday use.
- Given that many human senses are located on the head, eyewear computing will enable types of interaction between humans and machines never before possible.
- Smart eyewear promises new applications in multimodal user modeling, lifelong learning, and collective human-behavior sensing and analysis.



Figure 1. From fNIRS and mobile eye tracking glasses to Google Glass to unobtrusive smart glasses (J!NS MEME prototype recording eye and head movements).

and content hiding [5] (see framework in Figure 4). Embedded computers have reached a level of computing power and miniaturization that make it possible to use many state-of-the-art computer-vision algorithms on head-worn devices, such as those for activity, gesture, and text recognition. However, many algorithms need to be adapted or even rewritten from scratch because egocentric vision is fundamentally different from established computer-vision setups. In particular, egocentric cameras move with the users and therefore are not only subject to blur but also observe the world from viewpoints for which little training data is available.

Given they are relatively cheap, small sensors, microphones can be easily integrated into eyewear computers. Microphones have been used for tasks such as speech recognition or, when combined with bone-conduction speakers, biometric



Figure 2. PUPIL from Pupil Labs is an accessible, affordable, and extensible open source platform for head-mounted eye tracking, mobile gaze-based interaction, and egocentric vision research.

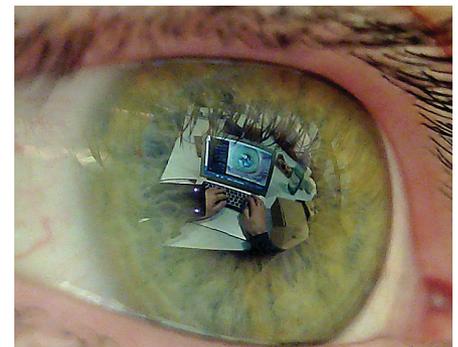


Figure 3. Current head-mounted eye trackers feature high-resolution eye cameras that can be used for corneal-imaging applications, such as gaze estimation, hand-gesture recognition, or semantic scene analysis.

user identification [6]. Ambient sound in particular is a rich source of contextual information about the user's environment and the people and objects with which they are interacting. Voice input is also a viable interaction modality for head-mounted computers, as seen in Google Glass. Yet, when used in public, voice

poses privacy and social challenges that are interesting from an HCI research perspective.

Recent work has investigated the use of other lightweight sensors. For example, many projects have focused on physical-activity recognition using inertial motion sensors, including sensors attached to the head [3].

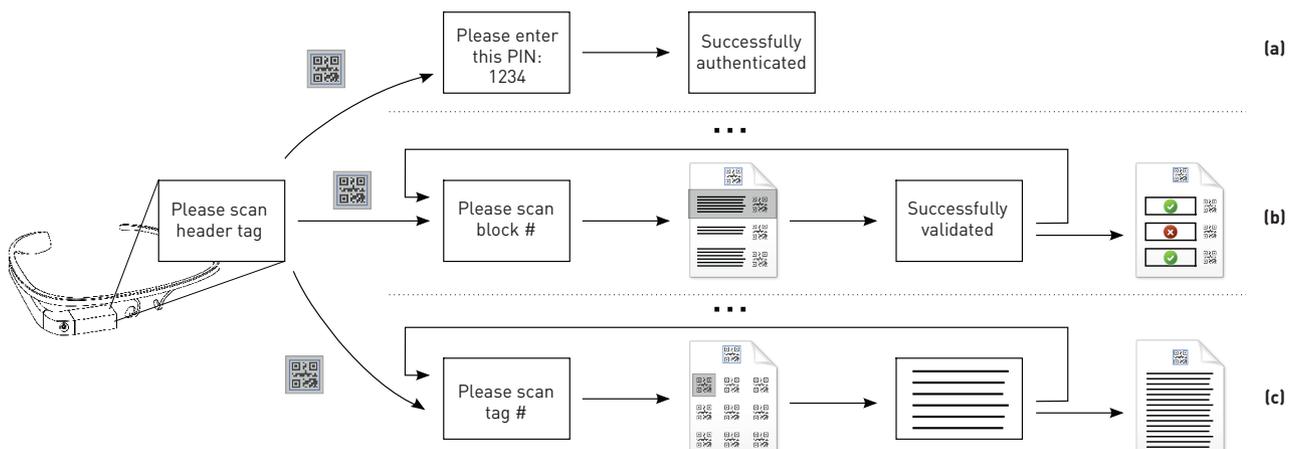


Figure 4. The Ubic framework combines a head-mounted display, such as Google Glass, with egocentric vision methods to bridge the gap between digital cryptography and the physical world. The framework covers key cryptographic primitives, such as secure user identification, document verification using a special secure physical document format, and content hiding.

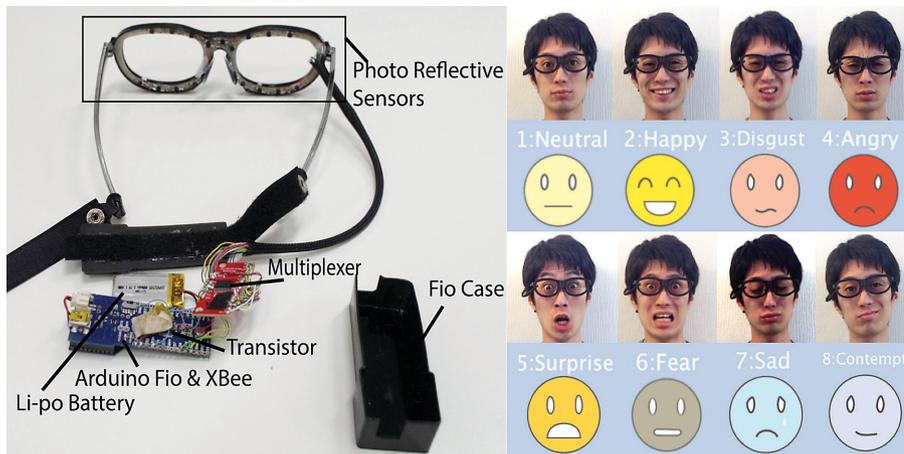


Figure 5. Affective Wear detecting facial expressions [7].

Similar to other modalities, head-worn sensors can provide additional information about the wearer, such as their cognitive tasks and attention [1,2]. Head motion can also indicate social activities and be used as an input modality for HMDs. Katsutoshi et al. investigated embedded photo-reflective sensors to recognize a wearer’s facial expressions in daily life [7]. The system leveraged the skin deformations created when wearers make facial expressions by measuring the proximity between skin surface and glasses frame (Figure 5).

Finally, brain activity is particularly interesting in the context of head-worn HCI because it can provide additional insights into users’ cognitive processes and can be used in brain-computer interfaces. The two sensing modalities most promising for wearable measurement of brain activity are electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS). Recent devices such as the Emotiv Insight headset (emotivinsight.com) or the Hitachi HOT-1000, a consumer fNIRS device, point the way toward unobtrusive brain-activity sensing as well as mobile brain-computer interfaces.

### HEAD-WORN OUTPUT MODALITIES

Our impressions of the world and our memory of it are constructed in large part by the human visual system. There are more than 30 areas concerned with vision in the brain, which occupy more than half of the total surface area of the cortex [8]. Head-worn displays are therefore positioned to play a key role in potentially extending our visual sense by overlaying the wavelengths our eyes do not naturally see (such as infrared), potentially overlaying zoom-lens output, and overlaying 2D or 3D computer-generated imagery. Optical design of head-worn displays involves a rich space of design parameters and trade-offs (see Figure 6 for some example displays). Main modes of operation include optical see-through, video see-through, and opaque displays. Optics can be designed for one eye or both. Important optical design parameters include eyebox, eye clearance, field of view, virtual-image location and distance, wavelength band, and image quality. For augmented reality (AR) applications, a see-through display has several advantages, including minimal obstructions of the visual

field, making them safer to wear. In addition, the resolution, distortion, and depth perception of the real scene can be as good as that which is seen with the naked eye. In the consumer space, aesthetics and fashion make up the single most important design parameter.

At the moment, there is no publicly known optical architecture that can satisfy all of the desired constraints (see-through, aesthetics and fashion, full color, large field of view, and so on) simultaneously [9]. On the positive side, the increasing popularity of smart watches indicates that there are various microinteraction-based AR applications that can be realized with a small field of view. There are smart watches in the market with approximately 30 mm<sup>2</sup> displays, which, when viewed at 12 inches, would correspond roughly to a 6x6-degree field of view. One challenge is that the display has no access to the user’s exact view as rendered on the retina. Corneal-feedback augmented reality is a novel technology with the potential to address this problem by providing a closed-feedback loop by continuously inspecting the reflections in the eye. More futuristic are contact-lens-based displays. A main drawback is the more invasive nature of this approach from a user-acceptance point of view, especially for users who do not need corrective lenses.

Other feedback modalities potentially available on eyewear computers are sound, tactile, and smell. Sound has long been used to provide information as a secondary and reinforcement channel, when the user needs to focus visually on the task at hand. More interesting, however, are newer research works that explore how to alter what we hear. This even led to some pioneering consumer products such as Active Listening by Doppler Labs (<https://www.dopplerlabs.com>).

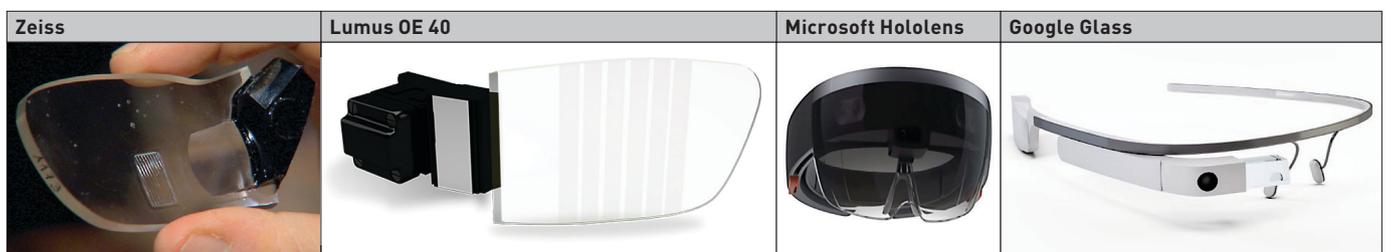


Figure 6. Different head-up displays.

In this case, sound is not used as mere notification; rather, the devices alter the sound perception to filter given frequencies to be able to hear voices more clearly, for example. Sonic interface design in AR environments seems to be a promising new research direction, especially when combined with other eyewear sensing and feedback modalities.

Research in tactile feedback so far has focused on people with disabilities (mostly people who are vision- and hearing-impaired). A major advantage of using haptics over sound: It is privacy conserving. Haptics on the head for non-disabled people is not very well explored. Early work investigates haptic feedback for navigation on smart glasses combining eye gaze and tactile output [10].

Smell as feedback might seem unconventional at first. Yet there are interesting links between memory retrieval processes and smell. A specific odor might enable a user to easier recall specific facts. And, especially on eyewear computers, it is easy to implement them in a personalized fashion (so that other people aren't disturbed).

## WHAT IS DOWN THE ROAD?

The input and output modalities available on current smart eyewear already enable a wide range of mobile applications and, at the same time, pose interesting challenges for human-computer interaction and user interface design. Smart eyewear also points toward completely new applications, the three most promising of which are multimodal user modeling, lifelong learning, and large-scale (collective) human-behavior sensing and analysis.

While current smart eyewear typically relies on a single input and output modality, future head-worn devices can exploit multiple modalities at the same time. Multimodal sensing will enable researchers to model user behavior not only more holistically but also more robustly and with improved performance compared with when using only a single modality [1,2,3]. Multimodal user modeling requires us to establish which modalities are best

suited for a given task, and in addition, these modalities need to be combined in the right way to allow for seamless sensing and natural interactions and user feedback.

With further advances in miniaturization and improvements in signal processing, egocentric vision methods, and battery lifetime, future smart eyewear has significant potential for long-term sensing, modeling, and analysis of user behavior. Current systems can monitor users only over relatively short amounts of time and therefore have a temporally limited view on their behavior. Methods for lifelong learning, however, can leverage the full information content available in all interactions that users perform explicitly or implicitly with computing systems and with each other over the years. Lifelong learning therefore represents both a paradigm shift and key characteristics to realize the vision of collaborative or even symbiotic human-machine systems.

Finally, while all of these developments will be beneficial for individual users, eyewear computing promises multimodal user-behavior modeling and lifelong learning at large, in other words, for collectives of large numbers of users in daily life. Mobile phones are already starting to be employed for collective human behavior analysis, for example to analyze movements of large groups of people at large-scale events such as music festivals. Eyewear computers show even bigger promise for such analyses, given their close proximity to the user and their location on the head.

Given the increasing number of consumer devices and recent advances in head-mounted eye tracking, egocentric vision, as well as head-mounted augmented and virtual reality displays, it is a perfect time for HCI researchers and user interface designers to take up both the opportunities and the challenges of smart eyewear.

## ACKNOWLEDGMENTS

This article is based, in part, on discussions with participants of the

Dagstuhl Seminar 16042 “EyeWear Computing - Augmenting the Human with Head-Mounted Wearable Assistants” (<http://www.dagstuhl.de/16042>).

## ENDNOTES

1. Amft, O. et al. Making regular eyeglasses smart. *IEEE Pervasive Computing* 14, 3 (2015), 32–43, 2015.
2. Bulling, A. and Zander, T.O. Cognition-aware computing. *IEEE Pervasive Computing* 13, 3 (2014), 80–83, 2014; <http://dx.doi.org/10.1109/mprv.2014.42>
3. Ishimaru, S. et al. In the blink of an eye: Combining head motion and eye blink frequency for activity recognition with google glass. *Proc. of Augmented Human 2014*; <http://dx.doi.org/10.1145/2582051.2582066>
4. Fathi, A., Farhadi, A., and Rehg, J.M. Understanding egocentric activities. *Proc. of ICCV 2011*.
5. Simkin, M. et al. Ubic: Bridging the gap between digital cryptography and the physical world. *Proc. of ESORICS 2014*, 56–75; [http://dx.doi.org/10.1007/978-3-319-11203-9\\_4](http://dx.doi.org/10.1007/978-3-319-11203-9_4)
6. Schneegass, S. et al. SkullConduct: Biometric user identification on eyewear computers using bone conduction through the skull. *Proc. of CHI 2016*; <http://dx.doi.org/10.1145/2858036.2858152>
7. Katsutoshi, M. et al. Facial expression recognition in daily life by embedded photo reflective sensors on smart eyewear. *Proc. of IUI 2016*.
8. Kandel, E. et al. *Essentials of Neural Science and Behavior*. McGraw Hill Professional, 1995.
9. Cakmakci, O. and Rolland, J. Head-worn displays: a review. *Journal of Display Technology* 2, 3 (2006), 199–216.
10. Kangas, J. et al. Using gaze gestures with haptic feedback on glasses. *Proc. of NordiCHI 2014*.

Andreas Bulling is head of the Perceptual User Interfaces Group at the Max Planck Institute for Informatics as well as the Cluster of Excellence on Multimodal Computing and Interaction at Saarland University, Germany. His research interests are in human-centered computing, human-computer interaction, and computer vision.  
→ [bulling@mpi-inf.mpg.de](mailto:bulling@mpi-inf.mpg.de)

Kai Kunze is an associate professor at the Graduate School of Media Design, Keio University, Yokohama, Japan. His research interests include eyewear computing, human-computer interaction, and wearable computing with a focus on amplifying skills and augmenting the embodied mind.  
→ [kai@kmd.keio.ac.jp](mailto:kai@kmd.keio.ac.jp)