7

Synergies between Head-Mounted Displays and Head-Mounted Eye Tracking: The Trajectory of Development and Its Social Consequences

by

Rich Ling, Professor, IT University of Copenhagen
(rili@itu.dk)
Diako Mardanbeigi, PhD Fellow, IT University of Copenhagen
(dima@itu.com)
Dan Witzner Hansen, IT University of Copenhagen
(witzner@itu.dk)

Abstract

In this paper we examine the potentials and limitations of fusing head-mounted computing devices such as Google Glass with head-mounted gaze tracking (HMGT). In the current version of Google Glass, for example, there is no gaze-tracking functionality—instead there is "point of vision" video functionality that captures the broader scene in front of the user. Gaze tracking adds a high degree of specificity to head-mounted computing devices that follows the precise gazing point of the user. We suggest that because of technical developments, HMGT is becoming increasingly mobile and that a likely path of adoption for HMGT will be as a feature

of head-mounted computing devices. We suggest several general applications for this technology. Further, we discuss the potential for head-mounted gaze tracking to become a widely used technology. However, there are several issues that hinder this trajectory. These include issues of privacy (both in the legal and in the ethical sense), the idea of how we present ourselves in the Goffmanian sense, and the difficulties of developing reciprocal expectations for the technology. For these reasons, both head-mounted computing devices and HMGT are likely to remain niche technologies.

Introduction

Gaze tracking has moved from being unwieldy and intrusive to simple and discreet. It has moved from being a technology that is complex to use and reliant on the care and prodding of highly trained engineers and scientists to becoming non-invasive and relatively straightforward to use. It has also moved to become a technology with an increasing number of use areas. This is not to say that head-mounted gaze tracking (HMGT) is a mainline technology. There are, however, significant areas where the technology can enhance data collection and can assist in the execution of important tasks. In this paper we are interested in looking at this technology in the context of head-mounted displays and consider the likely trajectory of development.

As with many other electronic devices, HMGT technology has become smaller and more agile (Hansen et al. 2005). Early in their history, eye-tracking devices often involved elements attached directly to the eye and had the need to stabilize the head by fixing it into place with various frames and straps. By contrast, contemporary eye tracking technology can disappear into simple, lightweight, mobile devices; this development has been seen on many technical fronts. Indeed we are on the cusp of another transition; namely, mobile head-mounted displays that will have the ability to retrieve information and to help us mediate our communication.

It is likely that in the near future HMGT functionality will be compact enough to fit into wearable displays such as point-of-vision (POV) devices, including Google Glass, that replicate an individual's field of vision. The current crop of these devices allows the user to capture video of, for example, a person as they parachute out of a plane or a law enforcement officer as they go about their rounds. The image captured, however, replicates the broad field of vision and not a particular point of gaze. In many cases, this broader image is what is best; however, we contend that there are also situations where a more specific focal point is also of interest.

There has been limited discussion of HMGT and heads-up displays in the literature (Hansen and Ji 2010). In the work that exists they have been examined as extensions virtual reality and immersive computing (Park, Lee, and Choi 2008) and as a way of apportioning attention (Kurauchi and Morimoto 2013). HMGT has also been examined in terms of its impact on social interaction in a laboratory context (McAtamney and Parker 2006). Thinking somewhat more broadly HMGT-enhanced head-mounted computing devices, such as what we see in the Google Glass project, we will have the ability to further indicate our point of attention and eventually transmit this to others or make it available for later examination. HMGT will tell us, for example, that a user is looking at a specific individual and not a crowd; a certain product in the shelf in the grocery store and not the whole shelf; or a particular part of the PC screen and not the whole screen. This can change the way that we can interact with our environment. In this paper we consider how head-mounted computing devices and HMGT can fuse into a single platform. Because of this development it is likely that HMGT will find new applications. In this process, we also see that it there are consequences in relation to privacy and power relationships (Katz 2013).

We will first go through the development and application of wearable computing, and follow this with a short account of the history and functionality of gaze tracking. We next discuss the melding of HMGT and heads-up display technologies and the potential for using this when it facilitates interacting with information that is embedded in the local context. This touches on issues such as the so-called internet of things. Finally, we look into the eventual applications for HMGT-enhanced wearable displays both in terms of the possibilities and the threats that they represent for at the personal and the social levels.

Head-Mounted Display and Wearable Computing Technology
Technical Development of HMD and Wearable Computing

Wearable devices that enhance our interaction with the world might be traced back to the development of glasses (Kriss and Kriss 1998). Following this line of thought, the watch, for example, was carried on the body (often in a well-protected pocket) from the 1600s (Landes 1983) and in the case of women, on the wrist, often as a piece of jewelry. The wristwatch made its appearance with males during the First World War since it was awkward for pilots to dig out pocket watches (Kahlert, Mühe, and Brunner 1986). Moving to head-mounted electronic devices, earphones have been a part of the technical landscape since the early period of the radio (Howeth 1963) and the idea of a head-mounted display (HMD) was first was patented by Thelma McCollum (1945) and as a stereoscopic television HMD by Morten Heilig (1960). Because of the technical limitations at that time, the idea of HMD was more focused on giving the user a virtual experience by showing a video, not as a "see-through" device that augmented vision. The first video "see-through" augmented reality system was made in the 1960s by the Bell Helicopter Company, which was a servo-controlled, camera-based HMD (Azuma et al. 2001). This provided the pilot with an augmented view captured by an infrared camera under the helicopter that was useful for landing at night. Since the early 1970s, the US Air Force has carried out research on HMD systems as a way of providing their aircrew with a variety of flight information and also a method for interacting with the airplane and user interfaces (Kiyokawa 2007). In the 1980s we began to see the use of HMDs where the user is able to "see through" the device, either optically or based on a video image. The user can see, for example, 3D computer- generated objects superimposed on his/her real-world view. The optical and the video approaches for HMD hardware design merge and superimpose the virtual view onto the real views of the world either via a semi-transparent mirror as with optical see-through HMDs (Berman and Melzer 1989; Buchroeder, Seeley, and Vukobratovich 1981; Droessler and Rotier 1990; Rolland et al. 1995), or via video cameras mounted on the head as with video see-through HMDs (Bajura, Fuchs, and Ohbuchi 1992; Edwards, Rolland, and Keller 1993).

The most recent of HMD project, what we refer to as head-mounted computing devices, and the one that seems to have garnered the greatest general interest, is the Google Glass project that includes an augmented reality head-mounted display. As of this point, Google Glass includes a heads-up display in addition to an embedded POV scene camera, microphone, different types of radio-based communication (Wi-Fi - 802.11b/g and Bluetooth), GPS functionality, an accelerometer and "bone conduction" in lieu of speakers. Voice control is used to operate the device including taking pictures/video, sending messages, getting directions, etc. Google Glass, as well as other smart glasses (e.g., Vuzix M100), show that head-mounted computing devices can potentially be used as a visual interface for mobile appliances; they can become a common display for various tools that we use such as mobile phones, tablets and even laptops.

Applications of the Head-Mounted Computing Technology

Head-mounted computing devices have been used in many different application fields such as: military, law enforcement (e.g., police), civilian (e.g., engineering, medicine, and computer-guided surgery), video gaming, sports, and simulation (e.g., driving and flight). Perhaps the most promising future uses of these tools are those in which the display allows for enhanced virtual environments (e.g., enhanced reality) rather than replacing real environments as in virtual reality (Bajura et al. 1992).

Head mounted displays provide the ability to use context-aware information such as weather reports, incoming text messages, public transportation schedules, route finding, information sharing with others, etc. Additional functionality will likely include pattern recognition perhaps similar to that in Google Goggles that references libraries of photos taken by others in addition to GPS data to search for further information on the item in question.

Gaze-Tracking Technology

Parallel with the development of wearable computing and head-mounted displays, there is also a development in the area of gaze tracking. Gaze tracking monitors and records the point of regard (i.e., where a person is looking as well as a direction in space) (Witzner and Ji 2010). In this section, we give a short history of the gaze-tracking technology in terms of technical development; next, different application areas of this technology are briefly described. At the end of this section, some of the limitations of the gaze trackers are described.

A Short History of Gaze Tracking

The functioning of the eyes and the interaction between gaze and cognition has long been the subject of interest. The people who have contributed to our understanding of vision include some of the luminaries of science such as Kepler and Descartes (Wade and Tatler 2005), and people have been developing ways of mechanically tracking eye movement for over 100 years (Jacob and Karn 2003). Seen from our remove, many of the early systems were quite draconian. The earliest devices were physical "contact lenses" that were attached to the eye using either an adhesive or suction to hold them in place. These lenses were sometimes attached to a mechanical lever in order to track the movement of the eye; it goes without saying that this hindered natural observations. As Jacob notes, "This method is obviously practical only for laboratory studies, as it is awkward, uncomfortable and interferes with blinking" (1995, 267). An early researcher, Edmund Huey, described his approach to recording the movement of a subject's eyes:

> I arranged apparatus as follows: A plaster of Paris cup was molded to fit the cornea accurately and smoothly, sand-papered until it was very light and thin, and placed upon the front surface of the eye, the cup adhering tightly to the moist cornea. No inconvenience was felt, as the corneal surface was made insensitive by the use of a little holocain, or sometimes cocaine. A round hole in the cup permitted the observer to read with this eye, the other eye was left free. A light tubular level of celloidin and glass connected the cup to the aluminum pointer, flat and thin, which responded instantly to the slightest movement of the system; and, suspended over the smoked-paper surface of a

> moving drum cylinder, the aluminum point traced a record of the eye's movement as the observer read. (*The Psychology and Pedagogy of Reading* pp. 17)

The system of tracking eye movement became progressively less invasive as the technology for observation developed. The use of film cameras eased the burden on (and presumably the irritation of) subjects. Shortly after the turn of the last century, researchers attached a simple "white speck of material" to "the eye of a subject and filmed it as the individual read" (Jacob and Karn 2003, 574). Researchers began to photograph the light reflected from the cornea (Majaranta 2009; Wade and Tatler 2005). In 1901, Dodge and Cline developed what they called the "Dodge Photochronograph" that is seen as the progenitor of today's eye reflection tracking systems; these have since dispensed with attaching anything to the eye (Judd et al. 1905). This is not to say, however, that the gaze-tracking systems were not bulky: they might take up whole sections of the laboratory. Buswell's 1935 device, for example, was a rambling collection of tubes, monitors, electronics, struts, lights, and frames with which to stabilize the subject's head. It filled a large desk and spilled over onto area behind, and it was nothing if not voluminous (1935).

As with many other areas of research, the rise of computerization dramatically changed the way we were able to gather and analyze gaze information. The equipment for tracking eye movement has undergone a radical reduction in size and devices have seen a similarly radical increase in processing power, accuracy, price and responsivity. With time, researchers developed head-mounted devices that allowed the subjects greater freedom of movement (Jacob and Karn 2003).

Early eye tracking systems used retrospective analysis of film or other recording material. Starting in the 1960s computers gave researchers the ability to digitally gather gaze tracking information, process the data, and provide feedback in real time (Jacob and Karn 2003). These developments mean that gaze trackers can be used as a computer-pointing device; they can also be used for sending commands (e.g., making selections on a screen).[1]

1 Dwell-time selection, eye blinks, gaze-gestures, and context switching have been typical ways of extending the capabilities of eye trackers for gaze-based interaction. Gaze as a pointing modality can also be used together with some other

Gaze interaction with computers has, until now, mostly been applied to the situation of a single, stationary user is sitting in front of a screen. It has used a camera, often mounted on or near the PC screen, to first calibrate and then to track the user's gaze (a remote eye tracker). Recent work has moved in the direction of head-mounted devices that are increasingly mobile and where, as the name suggests, the camera that captures the individual's eye movement is mounted on the person's head using either a helmet, a headband, or glasses (Ishiguro et al. 2010; Mardanbegi and Hansen 2011; Toyama et al. 2012). This has extended the domain of gaze-based interaction into the mobile situations that allow the user almost complete freedom of head movement as well as mobility.

Compared to the previous generation of gaze trackers, HMGT devices afford an unheard of degree of mobility. The developments in camera technology and miniaturization mean that it is now possible to move away from the desk-bound notion of eye tracking. Indeed, we are entering a period where head-mounted eye trackers have become much smaller, lighter, and thus easier to integrate with other mobile devices. Further, the integration of a variety of input possibilities (gaze, haptics, gestures, etc.) means that HMGT is becoming more flexible and more suitable for mobile, gaze-based interactive applications.

HMGT is currently at a stage where size and quality allow seamless integration of eye trackers into normal glasses. HMGT software is, to a large extent, also equal to an increasing number of gaze tracking tasks.[2] As we will discuss below, this also expanded the areas of use of gaze tracking.

---

interaction modalities such as body gestures and speech. Eye-based head gesture is a novel technique for enhancing gaze-based interaction through voluntary head movements. Gaze and head gestures measured by these trackers provide a gaze-based method for interacting with computers and objects in the environments.

2 We currently have cameras that are only several millimeters in size. In addition, the use of infrared light sources in glasses mean that glasses-mounted eye trackers are not a significant technological challenge. Clearly, several issues remain that will improve eye trackers even further (e.g., the ability to handle large and rapid light changes). A general problem for most current trackers is their need to be calibrated to the individual. While this is a current problem with most commercial eye trackers, there exist several possible techniques that could limit explicit per session calibration (Witzner, Hansen, and Ji 2010).

Gaze Tracking Applications

Gaze tracking applications can broadly be divided into two categories: *diagnostic applications*, where the eye tracker provides objective and quantitative evidence of the user's visual and attentional processes or neurological disorders (e.g., identification of neurological disorders by studying the diagnostic data provided by properties of saccades and fixations, and applications in psychology, cognitive linguistics, and product design), and *interactive applications*, where the eye tracker is used as an input device of an interactive system, and the system responds to the user's gaze (Duchowski 2007).

Diagnostic Applications

The earliest questions that used gaze tracking considered the interaction between gaze and tasks such as reading and looking at a picture. The research questions revolved around the interaction between vision and comprehension. Yarbus and Riggs (1967; see also Buswell 1935), for example, recorded people's gazes as they looked at an image when there was no particular task required of the viewer, and then when the viewer was asked to retrieve different types of information from the image (i.e., the number of people in the image or the type of clothes they are wearing). In other cases, gaze was recorded when people were asked to synthesize information from the image such as the class status of the people. In each case, Yarbus and Riggs recorded different patterns of eye movement.

Eye tracking has also been used when examining how people read (Rayner 1998). Just and Carpenter, for example, have used eye tracking to measure the time (in milliseconds) that subjects looked at words in sentences (1980). They suggest that the time to integrate gaze and comprehension depends on the frequency of a word's general use and its thematic importance; there is also a pause at the end of a sentence. The research also shows that eye movement differs when a person is reading aloud or silently. In addition, the research has indicated that as the complexity of the material becomes more difficult, we spend a longer time on each word and have a narrower field of focus (Duchowski 2007).

A similar application has been to study the use of gaze in how people carry out everyday tasks, such as simple food preparation, and how people handle different situations that arise in driving in traffic (often examined using driving simulators). In the case of the simple tasks, the research has been concerned with the role of gaze when going through a sequence of actions. The findings show that the subjects gaze will often presage the next physical action: when we are making a sandwich we look at the butter immediately before we move our hands to retrieve it. In the case of driving, while this is a dynamic situation as compared to the static analysis of reading or viewing a photograph, it has a common thread in that gaze tracking is used to understand the how the eyes focus on certain things and perhaps ignore or overlook other items that may also have importance.

Gaze tracking has also been applied to usability studies. In a classic study, Fitts et al. (1949) used a film camera to record the gaze of pilots as they landed airplanes. This has been extended later with other dimensions of flying (Duchowski 2007) to better understand where to place the instruments. This type of research has been applied to other arenas as well. Researchers have been interested in understanding, for example, the best arrangement of items on a web page or in printed material. It is often the case that the diagnostic applications have not relied on real-time feedback. Rather, the data is captured and analyzed later.

Another area of research has been to control how people carry out various types of visual analysis. This includes questions of, for example, X-ray inspection by doctors, production control inspection, and photo interpretation (e.g., in the case of astronomy or national security).

A question that has been broached in this context is the connection between seeing and cognition. According to Jacob and Karn:

> Psychologists who studied eye movements and fixations prior to the 1970s generally attempted to avoid cognitive factors such as learning, memory, workload, and deployment of attention. Instead their focus was on relationships between eye movements and simple visual stimulus properties such

> as target movement, contrast, and location. Their solution to the problem of higher-level cognitive factors had been to ignore, minimize or postpone their consideration in an attempt to develop models of the supposedly simpler lower-level processes, namely, sensorimotor relationships and their underlying physiology. (2003, 575)

Perhaps as an attempt to address this issue, the next step in this line of research was to combine eye tracking with brain activity as recorded with Functional Magnetic Resonance Imaging (fMRI). This development has provided us with a new tool with which to study the interaction between reading/looking and cognition. The research generally shows the correlation between eye fixation and brain activity (Duchowski 2007). This approach allows us to better understand the way that cognition works as we access different types of information in our brains. A related question is the interaction between vision and cognition for populations that are not able to communicate or have only fundamental communication capacity,: newborn babies, for example (Johnson et al. 1991). The research has investigated how newborns fixate on various shapes such as images with faces vs. more abstract images, providing insight into the bonding process.

Interactive Applications

As noted above the development of computing capacity meant that gaze tracking provided for immediate feedback. This led to the use eye movement as a pointing device for computer-based user interfaces. The most common application of this capability has been to allow disabled persons who cannot use their hands to control a mouse or keyboard (Handa and Ebisawa 2008; Hutchinson et al. 1989; Jacob and Karn 2003; Majaranta 2009). Indeed, this has been one of the most central applications for gaze tracking heretofore. The coming development of more compact HMGT devices will likely see its application in other interactive situations as noted below.

We are now seeing that the gaze tracking devices are becoming smaller, priceless expensive, more robust, and less in need of the careful goading and maintenance of engineers and scientists.

Further, they are no longer leashed to large computing devices. This means that the uses of gaze tracking can move into more natural settings, and thus we can begin to consider a broader range of applications. In addition to the traditional uses of cognition research, usability studies, and as aids for disabled persons, it is possible to develop gaze-tracking applications for more quotidian purposes. This is a discussion to which we will return below.

The Synergies of HMGT and Wearable Computers
Limitations of the Head-Mounted Display/Computer

The current implementation of the Google Glass, as well as various POV "action video cameras,"[3] have the ability to capture, in a broad sense, what the individual is looking at. Many of the head-mounted devices replicate the users' field of vision. However, the field of view for these video-based applications (often about 170 degrees) is broader than our active field of vision (which is about 135 degrees vertically and 160 degrees horizontally) (Wandell 1995). However, the most sensitive part of the eye is actually a small part of the total organ, and the field of vision is divided into three different areas of differing sensitivity and clarity. In order of decreasing clarity there is the fovea (about one to two degrees of vision), parafoveal (about three to five degrees), and peripheral region (everything beyond about six degrees). The foveal area is our major source of visual information, as the peripheral area is only able to register movements and contrasts as it has very poor visual acuity.

When we are looking at a scene before us, we focus on only a small portion of the total information; we continually scan a scene in order to gather further information. In some cases we can move our attention to the peripheral areas of vision, albeit not with the same natural ease. Within the brain a large portion of the cerebral cortex is devoted to processing the visual information from the foveal area. Thus, the wide frame captured by a many POV video system does not map onto our foveal-intensive vision.

3 These include for example the GoPro, Contour+, Ion Air Pro Drift HD, Panasonic HX-A100, AXON flex, and car-mounted video devices. An increasing number of other devices are moving into this space.

The Affordances of the Current HMGT

As we have noted, it is technically possible to have head-mounted eye trackers integrated with a POV scene camera can indicate the point of gaze. Additionally, we can use computer vision techniques for recognizing the objects in the scene and also for reconstructing the environment around the user. When the apparatus is attached to the user's head, it is also possible to know the direction and the speed of the movements of the head.

Gaze tracking can provide an abundance of information about the subject and their environment. This can include personal information (such as their focus, reading capabilities/content) as well as the general insight into the things and images that they visually dwell upon. The eye image recorded by today's gaze trackers can be used for measuring the eye movements and fixations[4] (Jacob and Karn 2003). In addition, the technology can also provide other types of eye-based information such as the pupil diameter (e.g., as an indicator of the cognitive load), different eye features like iris pattern (e.g., used as a biometric), the frequency of blinking, the behavior of the eye muscles (e.g., as one of the indicators of the user's fatigue) (Singh, Bhatia, and Kaur 2011), and the reflection of the environment on the surface of the cornea. In addition, the vestibulo-ocular reflex that coordinates eye movements relative to head movements makes it possible to even measure changes in head rotations (roll, tilt, and pan) through the eye movements (Mardanbegi, Hansen, and Pederson 2012).

By looking at the future interactive applications of wearable computers, and different ways of interaction with the head-mounted graphical user interfaces, we see that gaze as a pointing mechanism will likely be an early functionality to head-mounted computing devices. In addition, speech and gestures will also likely be added as mechanisms for sending commands (e.g., doing selection) and enhancing communication. Other technologies such as haptic, accelerometers, electroencephalography (EEG), and perhaps other biosensors may also be used to give more functionality to wearable computers.

4 For example, the number of fixations, the amount of time in each area, the number of times returned to a point, etc.

Applications of Gaze-Enhanced Head Mounted Computing Devices

There is a wide range of applications that are possible with gaze -enhanced, head-mounted computing devices that would allow for extremely detailed interaction between users. Indeed, when the gaze of one person is transmitted to another, the second person could specifically understand what the first person is looking at and, by inference, where their attention is directed.

Using this functionality a technician, for example, could call to a remote expert and be "talked through" exceedingly detailed procedures. Gaze-enhanced devices could be used when teaching people to react to visually specific clues (e.g. the investigation of X-ray images or when learning to drive). It is also possible to conceive of these technologies being used to deploy and direct remotely located workers across a broader geographical area. Gaze tracking could facilitate the logistical systems of delivery people, who could visually check the stocks of items on the shelves. Gaze-tracking systems could "check off" the QR codes of the existing stock and compile a list of needed items and flag those that are out of date. Shared gaze tracking could help us assist one another in focusing in on relevant (and very detailed) information when navigating in unfamiliar areas. Alternatively, if an individual were lost he/she could track on a sign showing the name of the street (or perhaps another sign such as a local restaurant) and this would help the system locate the individual.

HMGT and heads-up display technology has many applications for individuals. Further, combining head-mounted computing devices and HMGT, we also move beyond applications for single individuals. As with many other technologies, we suggest that the first users will likely be larger institutions, particularly those where there is a need for central coordination and mutual understanding of one another's situation. With time, we suggest that the technology will be further diffused for use by less formal social clusters, such as families or groups of friends. The technology will allow us to enhance the interaction between individuals since it provides for real-time updates of our social situation. That said, the likely areas of adoption will be niche applications in the near future. This is a theme to which we will return below.

Social Consequences of HMGT and Digital Artifacts

As noted above there have been several phases in the development of gaze tracking. These have included the basic understanding of eye movement, the application of this basic understanding to both the study of cognition and to usability and, most recently, the use of gaze tracking with live video and sophisticated computing power to control computers. We are now entering a phase when gaze tracking is moving out of the sheltered environment of the laboratory and moving "into the wild" (D. W. Hansen and Pece 2005). As noted above, the devices are becoming easy enough to use that they can be imbedded in other head-mounted gadgets, such as POV video devices and heads-up display units. The technology is available—this means that HMGT is becoming available for the development of a variety of applications that were not possible when it was bound to specific locations by the bulkiness of the equipment.

However, the very mobility of the equipment also means that there are several new uncertainties that arise. These include the qualms of privacy and the issues of recording the social interactions. In addition, there are questions focusing on the degree to which HMGT will become embedded in the structure of social interaction.

Privacy and Legal Issues of HMGT

The head-mounted POV scene cameras are a common element in computing glasses with HMD (e.g., Google Glass) just as they are common in head-mounted gaze trackers. The privacy issues of the HMGT are, on the one hand, associated with the scene camera and on the other hand, related to the gaze data and the information that the eyes can reveal (e.g., of a personal nature).

Use of video equipment raises question with regards our rights to gather photographic information and being photographed (Mann 2013). The use of photographic equipment is well trod: as soon as photography became common the question of our right "to be let alone" was an issue (Warren and Brandeis 1890). Warren and Brandeis wrote in 1890 that "Instantaneous photographs . . . have invaded the sacred precincts of private and domestic life; and numerous mechanical devices threaten to make good the prediction 'what is whispered in the closet shall be proclaimed from the house-tops'". The context in which Warren and Brandeis were discussing privacy was an era when photography was largely practiced by professional news photographers, previous to the popularization of smaller personal cameras, and more than a century before digital photography became standard. With time, the development of closed-circuit television and a variety of other digital recording systems adds unheard of dimensions to "shouting from the house-tops." In many cases, however, there has been and continues to be a power differential between those who record and those who are recorded. It is the local convenience store and gas station that has the security cameras, and these were used in the context of protecting their private property. The ability of individuals to record material in these private settings is different from the right of the property owner to do the same. This question has been brought into the public discussion by the so-called McDonalds incident with Steve Mann. In short, Steve Mann entered a McDonald's in France wearing his "eye tap" (Mann et al. 2005). The eye tap is, among other things, a forward-mounted video camera set in a glasses frame, wherein the video camera covers one eye. According to Mann's version of the incident one of the employees tried to tear the glasses off his face and Mann was eventually pushed out the door.[5]

Among the other issues that the incident touches on, there are issues are associated with who is allowed to capture video in a particular situation. In the case of the commercial establishments, they often have the right to possess surveillance. Also, since it is considered their domain, they can, to some degree, set other conditions with regards who they will serve. Clearly the incident raises the question of the conditions for video capture both on the part of establishments as well as with customers. The incident has been couched in terms of power to surveil and be surveilled as a function of power. A somewhat parallel query arises with the equipping of police with eye-mounted video cameras as in Rialto, California.[6] In this case, the local police department realized a major

5 http://www.slashgear.com/broken-glass-father-of-wearable-computing-allegedly-assaulted-17238802/

6 http://www.nytimes.com/2013/04/07/business/wearable-video-cameras-for-police-officers.html?_r=0

reduction in the number of complaints against officers. There is the idea that words and comments are no longer ephemeral, but now they have become a digital artifact. How is this data being collected, stored, and used? HMGT in public situation adds a new and untested dimension to this issue.

There is, however, another issue associated with eventually wearing a digital recording device in the normal flux of daily life (as seen in, for example, the idea of Memex, MyLifeBits and in so called lifecasting) (Gemmell et al. 2002; Mann 2013); namely, it imposes a dimension on the situation that has not hereto been a part of our understanding of a social situation. A tacit idea associated with social repartee is the idea that the interaction is not recorded, it is ephemeral. The imposition of a record on the interaction eventually changes the way that we are willing to commit ourselves to the situation, and raises the specter of being accountable for our comments and our actions in a way that we are not accountable when they are fleeting.

HMGT in natural settings ratchets up the issue of privacy to yet another level since the technology not only records what is happening in a particular situation, but where the gaze of one of the actors in the situation is resting at any given moment. To be the subject of others' digital gaze and to know that it is recorded means that the scene takes on a different social character. Even though photographing (or tracking gaze) in public places is not illegal, it raises ethical issue and it challenges our notions of privacy. In addition, the ability to capture gaze changes our heretofore-private behavior into a documented event. I could eventually be held responsible for my comments, actions and gaze in a way that was not possible before.

In some ways this might be simple embarrassment that we are caught looking at things better left undisturbed by our glance. However, our use of HMGT record could incriminate us. If, for example, the gaze tracker recorded a car accident it might show that I was adjusting the radio or texting at the time of the crash. HMGT can provide important feedback to a driver such as monitoring eye activity and sensing when they are in need of a rest stop. However, gaze tracking could also be used against a driver if it finds that their gaze was not on the appropriate place when they were involved in an accident. Further, if they showed that I continued to drive even after my HMGT device found I was drowsy; it might also make me culpable. Thus the technology has implications for the apportionment of responsibility.

Following the work of Goffman, significant parts of social interaction take place in guarded settings (Goffman 1959). The documentation of these would violate our sense of the situation at many levels. It would, in a sense, formalize that which up until now had been informal. The resistance to this development would likely hinder the eventual adoption of HMGT and, for that matter, head-mounted computing devices.

There are yet other dimensions to this issue: HMGT could eventually record the individuals we see or the items we look at in a store. In this latter case, the collection of QR codes that we gaze at can be valuable information for marketing purposes. The question then arises as to ownership of that data and how that data might be used by marketers to form a profile of the individual. Since HMGT is far more specific than simple POV devices (or GPS information) ownership and use of the information presents an important unsolved issue. Thus there are potentially some difficult unresolved questions that need to be settled.

HMGT as a Social Mediation Technology

Another issue associated with the eventual development of HMGT is the degree to which it can become embedded in the flux of social interaction. There are a range of technologies and systems that take on dimensions of being Durkheimian social facts (Ling 2012). Mechanical timekeeping, telecommunication, and dimensions of the internet can been seen in this context on a broad social level. In addition, in more restricted groups, technologies such as calendaring systems and, in its time, the network of fax machines, are examples of social mediation technologies (Ling and Canright 2013).

There are many characteristics that are common for these technologies, including their critically large number of users, their supported adoption by an ideology that legitimates their position in society (we feel safer by having a mobile phone with us), their arranging the social landscape to the exclusion of alternative systems that provide approximate the same function (e.g., the clock displacing the sun dial) and perhaps most importantly, the reciprocal expectation that others will also either operate based on the edicts of the system (everyone needs to respect time and timekeeping) or be mutually available via a particular mediation form. This is not to say that all technical developments become social mediation technologies—there are many that have become thoroughly embedded in society in spite of not being used for social mediation. Refrigeration is an example of a technology that has made dramatic changes in the social ecology. It is not, however, used for the mediation of social interaction.

The question here is whether HMGT (or for that matter, head-mounted computing devices), will become a technology of social mediation. It is indeed difficult to make the case that this will happen. As we have noted there is undeniable functionality that is provided by HMGT, and the trajectory that is perhaps most likely is that HMGT will be implemented in a future heads-up devices. In this course it will be developed for special applications such as remotely mediated group work where the detailed knowledge of one another's focus is important (i.e., coaching of detailed repairs). It might be that teams of repair personnel could be linked to one another as they carry out a distributed repair task and can thereby interact with one another to facilitate their common work. It might be that we use gaze tracking when discussing detailed co-editing of documents with one another so that we can tacitly see where our co-authors are looking. Other applications might extensions of the inspection functionality noted above—where, for example, delivery people will need to gaze at particular points in a store where they deliver products to insure that they are displayed properly.

This suggests, however, that video recording (and also the more specific use of gaze tracking) may find a niche when used in formalized settings for well-defined purposes. When thinking of personal uses of HMGT it is possible to imagine people using gaze to access specific types of information in specific setting. It might, for example, be useful to have detailed gaze tracking while shopping so that we can read in barcodes or QR codes to gather information about products like their nutritional value as compared to our favorite diet or, eventually, that the item is on sale at a store down the street. As noted above, however, there are a variety of questions that need to be addressed before this is universally accepted.

Still, it is more difficult to understand how either HMGT or head-mounted computing devices will quickly become a part of the general flux of social interaction. While there is a begrudging acceptance of surveillance in society and there has been the development of sousveillance (i.e., people below observing those above), there is not a major discussion of what is termed "veillance" where there is not a power differential between the individuals involved. This has been a sphere based on trust and forgiveness. The insertion of digital recording and, more specifically, gaze tracking into this context will likely not be as simple as it raises a broader set of questions (McAtamney and Parker 2006). The point here is that HMGT *can*, and likely *will*, become a part of the broader digital landscape, but that the first applications will not be associated with social interaction but with commercial situations.

In a similar way, we will also likely develop norms of when we are explicitly NOT looking at the activities of others. These types of processes were seen with the adoption of the mobile phone (Ling 1997). We will develop the sense that it is not appropriate to have on our HMGT unit when another person is using their PIN code. We may need to have a function that shows the recorder is not on, or we will take off the HMGT device much as we take off sunglasses, as a sign of courtesy.

Conclusion

In this paper we have considered the eventual melding of HMGT with heads-up display technology, and we see that heads-up devices are moving into the diffusion process. The commercialization of devices such as Google Glass indicates that there is a certain interest in this direction. At this point, HMGT and heads-up technology are two separate threads of development.

HMGT technology is technically available. The cameras that will provide for gaze tracking, the computing capacity, and the batteries are already available. It is very possible that gaze tracking will become a feature of head-mounted computing devices such as the Google Glass. This may well come as a part of the "feature creep" that is often associated with these types of gadgets. Thus, rather than being seen as a separate technology with its own trajectory, we suggest that it will be included in the eventual development of wearable computing.

We will certainly see that it is applied to various types of "niche" applications such as those noted above. We suggest that the possibilities afforded by the integration HMGT and HMD will allow efficiencies in various use situations. In a variety of commercial settings the functionality provided by exact gaze tracking will be able to make a contribution. It is also possible to imagine implementations that integrate the gaze point of the individual inspector or worker into a larger system of quality control. In addition, it is clear that HMGT and HMD can be useful in situations where careful inspection is necessary. In addition, they have the ability to make a contribution to different types of research.

That said, the technologies must face a significant social threshold. As noted above, the introduction of recording technology to what is largely seen as ephemeral social interaction violates what Goffman saw as the guarded nature of social settings. It would lead to more caution in our willingness to commit ourselves to the setting and it would also, perhaps, provide the raw materials for others to parody the ways we present ourselves.

Because of these considerations we suggest that the maturity of the HMGT technology will mean that it is easily integrated into head-mounted computing devices; these will likely find a variety of innovative applications. However, we must be sober in our suggestion that these technologies will be used in a wide range of informal social settings.

References

Azuma, Ronald, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. "Recent Advances in Augmented Reality." *IEEE Computer Graphics and Applications* 21, no. 6 (2001): 34–47. doi:10.1109/38.963459

Bajura, Michael, Henry Fuchs, and Ryutarou Ohbuchi. (1992). "Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient." *ACM SIGGRAPH Computer Graphics* 26 (1992): 203–210. [Accessed 26.2.14] http://dl.acm.org/citation.cfm?id=134061.

Berman, Arthur L., and James E. Melzer. Optical collimating apparatus. US Patent 4,859,031, filed August 3, 1987, and issued August 22, 1989.

Buchroeder, R. A., G. W. Seeley, and D. Vukobratovich. "Design of a Catadioptric VCASS Helmet-Mounted Display." Defense Technical Information Center Document, 1981. http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA109431Buswell, Guy T. *How People Look at Pictures: A Study of the Psychology of Perception in Art*. Chicago: University of Chicago Press, 1935. http://psych.wfu.edu/art_schirillo/articles/Buswell,%201935.pdf

Droessler, Justin G., and Donald J. Rotier. "Tilted Cat Helmet-Mounted Display." *Optical Engineering* 29, no. 8 (1990): 849–854.

Duchowski, Andrew T. *Eye Tracking Methodology: Theory and Practice*. London: Springer, 2007.

Edwards, Emily. K., Jannick P. Rolland, and Kurtis P. Keller. "Video See-through Design for Merging of Real and Virtual Environments." In *IEEE Virtual Reality Annual International Symposium, Septemer 18–22, 1993 Seattle, Washington/93Ch3336-5,* 223–33. Seattle: IEEE, 1993.doi:10.1109/VRAIS. 1993. 380774.

Fitts, Paul M., Richard E. Jones, and John L. Milton. "Eye Fixations of Aircraft Pilots. III. Frequency, Duration, and Sequence Fixations When Flying Air Force Ground-Controlled Approach System (GCA)". Defense Technical Information Center Document, 1949. http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA329371

Gemmell, Jim, Gordon Bell, Roger Lueder, Steven Drucker, and Curtis Wong. "MyLifeBits: Fulfilling the Memex Vision." In

Proceedings of the tenth ACM international Conference on Multimedia, 235–38. New York: ACM, 2002. doi:10.1145/641007.641053.

Goffman, Erving. *The Presentation of Self in Everyday Life*. New York: Doubleday Anchor Books, 1959.

Handa, S., and Y. Ebisawa. “Development of Head-Mounted Display with Eye-Gaze Detection Function for the Severely Disabled.” In *2008 IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems,* 140–44. IEEE, 2008. doi:10.1109/VECIMS.2008.4592769.

Hansen, Dan W., and Arthur E.C. Pece. “Eye Tracking in the Wild.” *Computer Vision and Image Understanding* 98, no. 1 (2005): 155–181.

Hansen, Dan W., and Qiang Ji. ”In the Eye of the Beholder: A Survey of Models for Eyes and Gaze.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, no. 3 (2010): 478–500. doi:10.1109/TPAMI.2009.30.

Hansen, John P., Dan W. Hansen, Anders S. Johansen, and John Elvesjö. “Mainstreaming Gaze Interaction Towards a Mass Market for the Benefit of All.” *Human Computer International 2005*. London: Springer

Heilig, Morton L. Stereoscopic-television apparatus for individual use. US Patent 2,955,156, filed May 24, 1957, issued October 4, 1960.

Howeth, Linwood S. *History of Communications-Electronics in the United States Navy*. Washington: US Government Printing Office, 1963.

Huey, Edmund B. *The Psychology and Pedagogy of Reading: With a Review of the History of Reading and Writing and of Methods, Texts, and Hygiene in Reading*. Cambridge, MA: MIT Press, 2009.

Hutchinson, T. E., K. P. White, Jr., W. N. Martin, K. C. Reichert, and L. A. Frey. “Human-computer interaction using eye-gaze input.” *IEEE Transactions on Systems, Man and Cybernetics* 19, no. 6 (1989): 1527–34.

Ishiguro, Yoshirto, Adiyan Mujibiya, Takashi Miyaki, and Jun Rekimoto. “Aided Eyes: Eye Activity Sensing for Daily Life.” In *Proceedings of the 1st Augmented Human International Conference*. New York: ACM, 2010. doi:10.1145/1785455.1785480.

Jacob, Robert J.K. “Eye tracking in advanced interface design.” In *Virtual Environments and Advanced Interface Design*, edited by

Woodrow Barfield and Thomas A. Furness, 258–88. New York, NY: Oxford University Press, 1995.

———, and Keith S. Karn. "Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises." *Mind* 2, no. 3 (2003): 573–604.

Johnson, Mark H., Suzanne Dziurawiec, Hadyn Ellis, and John Morton. "Newborns' Preferential Tracking of Face-Like Stimuli and its Subsequent Decline." *Cognition* 40, no. 1 (1991): 1–19.

Judd, Charles H., Cloyd N. McAllister, N. Cloyd, and W. M. Steele. "General Introduction to a Series of Studies of Eye Movements by Means of Kinetoscopic Photographs." *Psychological Monographs*. 1905. http://www.citeulike.org/group/2640/article/420454.

Just, Marcel Adam, and Patricia A. Carpenter. "A Theory of Reading: From Eye Fixations to Comprehension." *Psychological Review* 87, no. 4 (1980): 329–54.

Kahlert, Helmut, Richard Mühe, and Gisbert L. Brunner. *Wristwatches: History of a Century's Development*. West Chester, PA: Schiffer, 1986.

Katz, James. "Mobile Gazing Two-Ways: Visual Layering as an Emerging Mobile Communication Service." *Mobile Media & Communication* 1, no. 1 (2013): 129–33.

Kiyokawa, Kiyoshi. "An Introduction to Head Mounted Displays for Augmented Reality." In *Emerging Technologies of Augmented Reality Interfaces and Design*, edited by Michael Haller, Mark Billinghurst, and Bruce H. Thomas, 43–63. London: Idea Group Publishing, 2007. doi:10.4018/978-1-59904-945-8.ch022.

Kriss, Timothy C., and Vesna M. Kriss. "History of the Operating Microscope: From Magnifying Glass to Microneurosurgery." *Neurosurgery* 42, no. 4 (1998): 899–907.

Kurauchi, Andrew T., Carlos H. Morimoto, Diako Mardanbegi, and Dan Witzner Hansen. "Towards Wearable Gaze Supported Augmented Cognition." Paper presented at the CHI 2013 Workshop on "Gaze-Interaction in the Post-WIMP World," Paris, France, April 2013. http://gaze-interaction.net/wp-system/wp-content/uploads/2013/04/KMM+13.pdf

Landes, David S. *Revolution in time: Clocks and the Making of the Modern World.* Cambridge: Belknap, 1983.

Ling, Rich. " 'One Can Talk about Common Manners!': The Use of Mobile Telephones in Inappropriate Situations." In *Themes in mobile telephony Final Report of the COST 248 Home and Work group*, edited by L. Haddon, Stockholm: Telia, 1997.

———. *Taken for Grantedness: The Embedding of Mobile Communication into Society*. Cambridge, MA: MIT Press, 2012.

———, and Geoff Canright. "Perceived Critical Adoption Transitions and Technologies of Social Mediation." Paper presented at the Cell and Self Conference, Ann Arbor, MI, 2013.

Majaranta, Päivi. "Text Entry by Eye Gaze." PhD diss., University of Tampere, Finland, 2009.

Mann, Steve. "Wearable computing: A First Step Toward Personal Imaging." In M. Soegaard & R. F. Dam (Eds.), *Encyclopedia of Human-Computer Interaction*. 2nd ed, 25–32. Aarhus, Denmark: The Interaction Design Foundation, 2013. doi:10.1109/2.566147.

Mann, Steve, James Fung, Chris Aimone, Anurag Sehgal, and Daniel Chen. "Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences." Paper presented at the Conference on Human Factors in Computing Systems, Portland, OR, April 2005. http://americasbesteyeglasses.org/goto/http://eyetap.org/papers/docs/Eyetap.pdf.

Mardanbegi, Diako, Dan Witzner Hansen, and Thomas Pederson. "Eye-Based Head Gestures." In *Proceedings of the Symposium on Eye Tracking Research and Applications*, 139–46. New York: ACM, 2012. doi:10.1145/2168556.2168578.

———, and Dan Witzner Hansen. "Mobile gaze-based screen interaction in 3D environments." In *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications*. New York: ACM, 2011. doi:10.1145/1983302.1983304

McAtamney, Gerard and Caroline Parker. "An Examination of the Effects of a Wearable Display on Informal Face-to-Face Communication." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 45–54. New York: ACM, 2006. doi:10.1145/1124772.1124780.

McCollum, Thelma. Stereoscopic television apparatus. US Patent 2,388,170, filed April 15, 1943, and issued October 1945.

Park, Hyung Min, Seok Han Lee, and Jong Soo Choi. "Wearable Augmented Reality System Using Gaze Interaction." In *Proceedings of the 7th IEEE/ACM international Symposium on Mixed and*

Augmented Reality, 175–76. Washington, DC: IEEE Computer Society, 2008. doi:10.1109/ISMAR.2008.4637353.

Rayner, Keith. "Eye Movements in Reading and Information Processing: 20 Years of Research." *Psychological Bulletin* 124, no. 3 (1998): 372.

Rolland, Jannick P., Christiana A. Burbeck, William Gibson, and Dan Ariely. "Towards Quantifying Depth and Size Perception in 3D Virtual Environments." *Presence: Teleoperators and Virtual Environments*, 4, no. 1 (1995): 24–48.

Singh, Hardeep, J. S. Bhatia, and Jasbir Kaur. "Eye Tracking Based Driver Fatigue Monitoring and Warning System." In *2010 India International Conference on Power Electronics (IICPE)*, 1–6. New Delhi, India: 2001. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5728062.

Toyama, Takumi, Thomas Kieninger, Faisal Shafait, and Andreas Dengel. "Gaze Guided Object Recognition Using a Head-Mounted Eye Tracker." In *Proceedings of the Symposium on Eye Tracking Research and Applications*, 91–8.New York: ACM, 2012. doi:10.1145/2168556.2168570.

Wade, Nicholas J., and Benjamin W. Tatler. *The Moving Tablet of the Eye: The Origins of Modern Eye Movement Research*. New York, NY: Oxford University Press, 2005. http://books.google.com/books?hl=en&lr=&id=hJg2xhz7XKUC&oi=fnd&pg=PR7&dq=wade+and+tatler+2005&ots=FrHfsV0T30&sig=ZLFlwzQuzkkT788vu03Wkc2TSmY

Wandell, Brian A. *Foundations of Vision* Vol. 93. Sunderland, MA: Sinauer Associates, 1995. http://www.getcited.org/pub/103207005

Warren, Samuel D. and Louis D. Brandeis. "The Right to Privacy." *Harvard Law Review* 4, no. 5 (December 1890): 193–220. doi:10.2307/1321160.

Yarbus, Alfred L., and L. A. Riggs. *Eye Movements and Vision* Vol. 2. New York: Plenum Press, 1967.

Author Profiles

Rich Ling is currently a professor at the IT University of Copenhagen. He has a PhD from the University of Colorado in Sociology and an adjunct position at the University of Michigan. In addition, he holds a position with the research organization of the mobile communications concern Telenor, and for the past two decades he has studied the social consequences of mobile communication. He has written several books in this area including *The Mobile Connection* (Morgan Kaufmann, 2004), *New Tech, New Ties* (MIT Press, 2008) and most recently *Taken for Grantedness* (MIT Press, 2012). He is also a founding co-editor of the Sage Journal *Mobile Media and Communication.*

Diako Mardanbegi is a post-doc at IT University of Copenhagen. He has a BS in Mechanical Engineering at Amirkabir University of Tehran and a MS in Biomechanical Engineering at Iran University of Science and Technology. His current research is within mobile gaze tracking for control of home appliances and investigating how mobile eye trackers can be used for interaction with the environment. He has been doing research in the field of eye tracking since 2005 and as part of his work at the IT University of Copenhagen he developed the open-source Haytham Gaze Tracker.

Dan Witzner Hansen is associate professor at the IT University of Copenhagen. He received his PhD from that institution; his thesis was about low-cost eye tracking. His research interests are within image analysis and machine learning with a particular emphasis on robust and low-cost eye tracking and their applications. Within eye tracking, his interest spans gaze estimation (fully, partially, and uncalibrated setups), robust methods for image processing, and noise-tolerant gaze interaction. Within eye tracking research he is aiming towards "mobile eye tracking in the wild." A review paper on this topic is available in his publications.