Eye Tracking: A Brief Introduction

Vidhya Navalpakkam and Elizabeth F. Churchill

The countenance is the portrait of the soul, and the eyes mark its intentions.

Markus Tullius Cicero

When the arm is stretched in front of one's face, the size of the thumb is approximately what we see in high resolution. Visual acuity drops as we move toward the periphery. It is remarkable that despite this drop in acuity, we perceive, scan, recognize, and navigate visual information in the world around us—apparently effortlessly. This is largely due to eye movements. To counter limitations in peripheral resolution, our eyes rapidly shift from one position to another about three to four times per second, to sample visual information from the interesting areas of the world. The brain stitches together these different pieces of information in real time to present a picture of the world around us in good visual resolution. These sudden jumps in eye position that occur through fast eye movements are known as saccades.

Information processing is thought to occur during fixations,[1] when the eye position is relatively static. Therefore, eye fixations are taken to be a good proxy for cognitive

[1] Although saccades and fixations are most commonly analyzed for information processing tasks, there exist other types of eye movements such as pursuit, vergence, and vestibular eye movements. Pursuit eye movements have lower velocity than saccades and occur when the eyes follow a moving object. Vergence eye movements occur when the eyes move toward each other, to fixate on a nearby object. Vestibular eye movements occur when the eyes rotate to compensate for head and body movements in order to maintain the same direction of vision. Other smaller movements of the eyes include drifts and microsaccades.

V. Navalpakkam
Google, Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA
e-mail: vidhyan@google.com

E.F. Churchill (✉)
eBay Research Labs, eBay Inc., 2065 Hamilton Ave, San Jose, CA 95125, USA
e-mail: churchill@acm.org

J.S. Olson and W.A. Kellogg (eds.), *Ways of Knowing in HCI*,
DOI 10.1007/978-1-4939-0378-8_13,

attention[2] and focused problem solving, and have been of interest to communities studying the perceptual, cognitive, and social processing of information.

In this chapter we briefly introduce eye tracking as a way of knowing in Human Computer Interaction. Specifically, we look at how eye tracking is used as a method for assessing how people perceive, process, and interact with images and interfaces to digital, computer-based technologies. In addition, we consider how eye tracking can facilitate understanding and supporting human-to-human communication and collaboration in technologically mediated environments. We offer a brief grounding with some examples and ask how can eye tracking help us understand the way humans interact with each other, and with displays.

We briefly consider the anatomy of the eye, and list eye tracking measurements using popular and contemporary technologies. We also discuss the benefits of eye tracking (including strengths and limitations), and illustrate with examples on how to effectively apply this method for research in Human Computer Interaction.

What Is Eye Tracking?

Eye tracking is the process of measuring either the point of gaze ("where we are looking") or the movement of the eye relative to the head. Although eye tracking as a method has gained a lot of press in recent years, eye tracking has been a method for understanding conscious and unconscious information processing since the 1800s (for example, Javal, 1990). Much of the early work into eye tracking was conducted through direct observation of people's gaze. However, today an array of sophisticated eye tracking technologies is readily available from trusted vendors who offer services as well as software and hardware products. An *eye tracker* is a device for measuring eye positions and eye movement. Eye trackers are used for research on the human (primate) visual system, in a number of research areas including psychology, cognitive science, marketing, and product design. There are a number of methods for measuring eye movement using eye trackers. Before we discuss the details of how eye tracking data is measured, gathered, and interpreted, we briefly discuss the anatomy of the eye and the various ways to obtain eye tracking data.

[2] Attention can be of two types: overt (the focus of attention matches where the eyes look) and covert (the focus of attention is different from where the eyes look). For example, when one is looking up to concentrate, where their eyes look is not correlated with they are thinking. This is a case of covert attention. It has been argued that for most natural viewing conditions, the focus of attention correlates with where the eyes look. In the rest of this article, we refer to overt attention as simply attention.

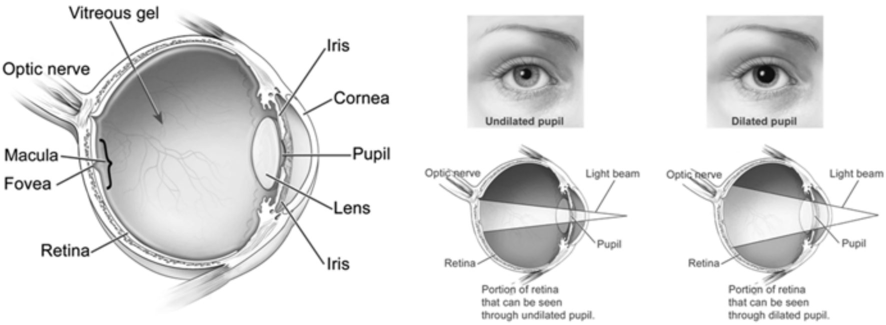

Fig. 1 Anatomy of the eye. Courtesy: National Eye Institute, National Institutes of Health (NEI/NIH)

The Anatomy of the Eye

The human eye, a slightly asymmetrical globe, is filled with a clear gel called the vitreous humor. There are seven parts of the eye that are worth knowing about for the purposes of understanding how eye tracking works (see Fig. 1). These are the iris (the pigmented part which gives us our eye color), the cornea (a clear dome over the iris), the pupil (the open circle at the center of the iris where light comes into the eye, appearing black), the sclera (the white part of the eye), the conjunctiva (a clear layer of tissue that is not visible, but which covers all but the cornea of the eye), the lens (which lies behind the pupil and the iris and which helps to focus incoming light on the back of the eye), and the retina (which is comprised of light-sensing cells the inside lining of the eye).

In the center of the retina is an area called the macula; at *its* center is the fovea, a slight depression which is responsible for high-resolution vision. Light waves enter the eye through the cornea, and pass through the pupil. As light intensity changes, so does the size of the pupil: Brighter light constricts the pupil; less light causes the pupil to dilate. Light is converted into electrical impulses by the retina, and the optic nerve transmits these impulses to the brain via the visual pathway, to the occipital cortex at the back of the brain.

There are two kinds of light receptor cells found in the retina: the cone cells and the rod cells. Foveal vision is created by tightly packed cone cells; these only account for 6 % of the total retinal light receptors. Cone cells require the most light for creating a clear, detailed image. Rod cells account for the other 94 % of light receptors in the retina. They require less light but create the blurry, less colorful qualities of peripheral vision.

Once the electrical signals get to the brain they are interpreted or "seen" by the brain as a visual image, sometimes called the "visual field." The visual field is a combination of the two primary types of vision mentioned in the introduction: *foveal* (high resolution and colorful) and *peripheral* vision (blurry and less colorful).

Peripheral vision also permits vision under low light conditions. Gross movements, color, and shape contrasts in the periphery are processed and, if they warrant further examination by the foveal vision we typically move our eyes and/or head to bring the objects of interest into the fovea for further recognition and action.

Deeper discussions regarding the anatomy of the eye can be found in “Eye and Brain: The Psychology of Seeing” by Richard L. Gregory and “Eye Tracking Methodology” by Andrew T. Duchowski, especially chapter 2 on neurological substrates.

Eye Tracking Methods

Since eye tracking is considered to provide a window onto the user’s attention, we begin by addressing the strengths and current limitations of current methods. Eye gaze reveals a lot about the user that is otherwise hard to know: It tells us where the user looked, for how long, and in what order. Eye trackers are easy to use; most commercial eye trackers come with user-friendly data collection capabilities and data management interfaces, and some offer built-in data analysis software.[3] Eye tracking as a method enables tracking user eye gaze at a fine temporal resolution (~2–20 ms/sample) and high spatial resolution (<0.5° error in accuracy). The high temporal and spatial resolution can be valuable for a variety of applications. The uses range from diagnosis of medical disorders to determining user examination strategies on web search pages. In the latter, they examine the order in which people examine the search results, how much time do they spend on titles, urls, snippets, that are critical for inferring document relevance, and for applications such as ranking and search optimization. The current limitations of the method are that eye trackers are currently expensive (commercial equipment ranges from $10 K upwards), that studies tend to be small scale (involving 10–30 users), and that studies are usually conducted in controlled lab settings (raising questions about ecological validity—that is, whether results generalize to natural settings). As we discuss toward the end of this chapter, recent work has started addressing some of these limitations.

Eye tracking methods have come a long way since the method was first proposed. Early studies of eye gaze—“looking behavior”—involved simply filming subjects while they looked at a picture or watched a video clip. Researchers hand scored

[3] There are many companies offering hardware and software for eye tracking studies both in the laboratory or in controlled desktop settings and also for mobile contexts. Well-known companies include SMI (SensoMotoric Instruments) a spin-off from led by Dr. Winfried Teiwes and his academic mentors in 1991 (http://www.smivision.com/), Tobii Technology established in 2001 by John Elvesjö, Henrik Eskilsson, and Mårten Skogö (http://www.tobii.com/), and Arrington Research which was founded in 1995 by Dr. Karl Frederick Arrington as part of a technology transfer initiative at the Massachusetts Institute of Technology (http://www.arringtonresearch.com/). Other companies include Applied Science Laboratories (ASL), EyeTech, Mirametrix, Seeing Machines and SR. Webcam-based eye tracking solutions include GazeHawk and eye-trackShop.

Fig. 2 Some examples of contemporary eye tracking equipment. From *top* to *bottom*: (**a**) Head-mounted mobile eye tracker in the form of eye tracking glasses (Courtesy: Tobii Technology); (**b**) table-based remote eye tracker (Courtesy: Tobii Technology); (**c**) eye tracking setup for mobile and personal devices (Courtesy: Tobii Technology); and (**d**) EOG (*Source*: Utah Medical School, http://webvision.med.utah.edu/)

recorded material to obtain a crude indication of their gaze direction. Since these early beginnings, a variety of eye tracking methods have evolved to determine the direction of gaze more accurately. Techniques include:

(a) Surface electrodes, electrooculogram (EOG)
(b) Infrared corneal reflections
(c) Video-based pupil monitoring
(d) Scleral search coils

These methods differ in their utility and in their invasiveness: EOG techniques are helpful in measuring saccade latency, but not good at measuring location; while scleral coils offer high spatial resolution (0.01°) and high temporal resolution (1,000 Hz), they are invasive and uncomfortable for participants, hence less preferred, except in clinical settings. These methods also differ depending on whether the head is free to move or not. For some applications, the head position is fixed using a forehead support or a bite bar or some other restraining mechanism that holds the eye position steady. In other cases, the head is free to move; here, head movement is accounted for with magnetic or video-based head trackers. Examples of these are illustrated in Fig. 2.

Methods are constantly evolving thanks to new technologies that are appearing—lighter weight, mobile, high resolution, infrared-enabled webcams—and due to advances in computer vision. Well-designed, lightweight and comfortable desktop and laboratory-based eye tracking equipment is nowadays standard fare for usability labs, psychology and vision science laboratories. Further, improvements in cameras and in recording technologies mean that mobile eye trackers for following gaze as people navigate the physical environment can now help researchers investigate the effect of complex environments on eye gaze, and in understanding the role of eye gaze in face to face and technology-mediated human–human interactions. An example of using mobile eye tracking to study coordination and communication in a real-world space can be found in work carried out by Gergle and Clark (2011); the researchers report that in collocated conversations, two people who are moving around and conversing tend to use more local deictic references to point to objects (e.g., “this,” “these”) and have lower gaze overlap than two people who are seated and thus stationary when conversing.

Finally, the possibility of webcam-based eye tracking offers hope of studies at large scale, where eye gaze patterns of hundreds and thousands of people in natural settings can be tracked using webcams. Devices increasingly come fitted with webcams, and software for data collection and analysis is readily available. In fact, this offers the possibility of eye tracking as an easy-to-use web service, where the video of the eye captured by the webcam is sent to the server in a cloud infrastructure, which then extracts eye positions from the video, analyzes eye tracking data, and sends the results back to the user’s device, thereby enabling the use of eye tracking on mobile devices and phones with low computing power. Eye tracking as a cheap web service could lead to several interesting opt-in applications such as hands-free, eye-controlled scrolling, swiping, navigating, typing, and gaming on any web and webcam-enabled device (big or small) including smart phones, tablets, laptops, and desktop computers, and more importantly, can enable patients with motor disorders to interact with computers and mobile devices. A current limitation is that the accuracy of these methods is still low and a rigorous comparison against high accuracy commercial eye trackers is lacking. Thus, there exists a tradeoff between accuracy and scalability of eye tracking. However, it remains the case that the increasing availability and accuracy of high-resolution, inexpensive, lightweight, and highly configurable sensors like cameras means that eye tracking equipment is becoming readily available for researchers cheaply.

The resolution in terms of temporal sampling and point of gaze in the collection of eye tracking data varies according to the type and model of eye tracker used. Today, reduced price means reduced accuracy. Boraston and Blakemore (2007) offer examples of the variations in temporal sampling for various methods, reporting that technologies are improving all the time. At the time of their paper, pupil-only and pupil-CR eye trackers typically operated at sampling rates of between 50 Hz and 2 kHz (i.e., sampled eye position at 0.5–20 ms/sample). Direct tracking of the fovea was being accomplished at speeds of up to 200 Hz (5 ms/sample). Spatial resolution varies from 0.005° of visual angle (Clarke, Ditterich, Druen, Schonfeld, & Steineke, 2002) to 0.5°, or approximately 0.1° for methods that involve direct detection of the fovea (Gramatikov, Zalloutm, Wu, Hunter, & Guyton, 2007). Indeed, more recently, for the purposes of evaluating equipment for usability studies where considerable accuracy at the pixel-on-screen level can be needed, Johansen et al. (2011) conducted a series of tests of eye trackers, comparing an open source remote eye tracking system with a state-of-the-art commercial eye tracker. While both devices were fairly stable over time, the commercial tracker was clearly more accurate at the pixel level. They concluded that low cost eye tracking is a viable alternative to expensive equipment only when usability studies do not need to distinguish between particular words or menu items that participants are looking at. If the research is focused on larger areas of interest, e.g., whether a person is looking at an object or another person in the room, cheaper solutions are adequate.

Video-based eye trackers are the most commonly used method today; indeed, when people talk about eye tracking, they usually mean video-based methods. Most commercial eye trackers use infrared cameras with high zoom to capture high-resolution images of the eye. Points of interest such as the center of the pupil and corneal reflection are extracted from these images to determine the point of regard of the user, or simply the eye position (Goldberg & Wichansky, 2003). In order to learn how the eye position (on the image of eye) maps to what the user is looking at (on a screen for example), a short procedure known as "calibration" is performed where the user is asked to look at various points (usually in a 3×3 grid) on the display, and the relationship between the two coordinate systems (pupil-center/corneal-reflection on the image of the eye, and the x,y coordinate on the display being viewed) is established. Once good calibration is achieved (high accuracy is $<1/2°$ error) the study can commence.

Below are a few general practices for accurate and reliable calibration. One must ensure a good initial view of the eye that is robust to wide-angle glances (especially for participants wearing glasses), and use a calibration grid that is at the approximate distance of the testing stimuli (e.g., don't calibrate on a wall then test on a nearby display or vice versa). Next, one must ensure that the calibration grid covers just outside the boundaries that will be used by the participants. Participants must be requested to move their eyes then their head when performing calibration and during the study.

Once the raw eye tracking data are obtained using one of the methods above, the data can be parsed to obtain various measures such as eye fixations (brief pauses in eye position lasting around 200–250 ms each on average) and saccades (fast eye movements). Most commercial eye trackers come with built-in software for extraction of fixations and saccades, and provide output in the form of a sequence of fixations (with timestamp, x,y position, duration, and link to the display viewed). Because these are now provided automatically, we skip the discussion of these computations and refer the reader to Salvucci and Goldberg for an overview of algorithms to extract these measures (Salvucci & Goldberg, 2000).

For the remainder of the chapter, we focus on the process of inferring useful information from eye movement recordings, which involves the researcher defining "areas of interest" over certain parts of a display or interface under evaluation, and analyzing the eye movements that fall within such areas. Commonly used measures for areas of interest are fixation duration (how long do users notice as measured by dwell-time on a part of the visual scene), number of fixations (how often do users

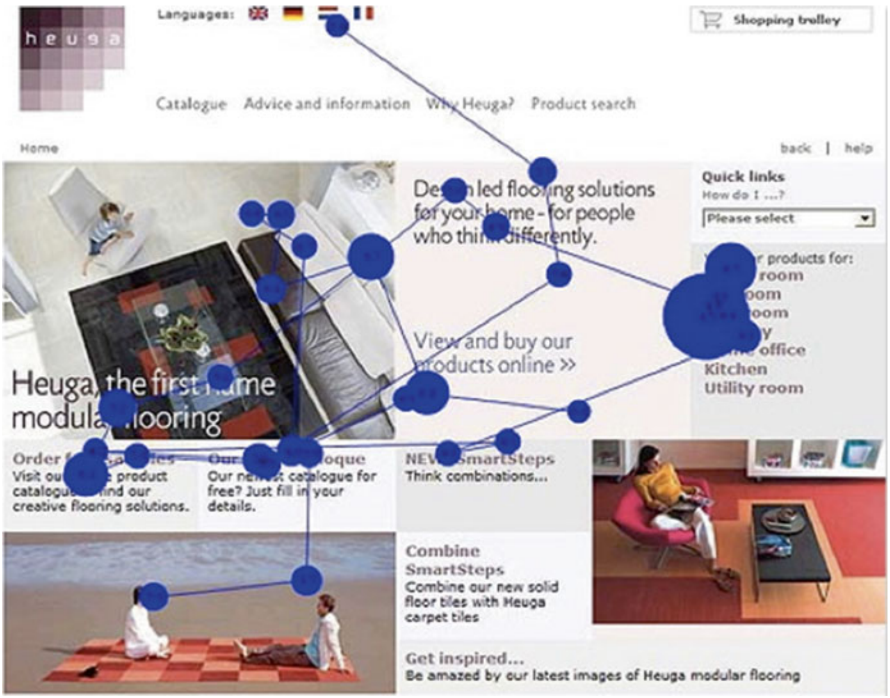

Fig. 3 An example of a sequence of eye fixations obtained from a single user. The *lines* indicate saccades and the *circles* indicate fixations. The size of the circle is proportional to the duration of the fixation. *Source*: Nielsen & Pernice, 2010

notice a part of the visual scene), sequence of fixations (the order in which users notice different parts of the visual scene), and transitions between pairs of areas of interest (how frequently users visit one area of interest from another). Figure 3 shows an example of a sequence of eye fixations interspersed by saccades.

What Is Measured with Eye Tracking? What Questions the Method Can Answer?

Eye fixations are known to be driven by perceptual salience and relevance as determined from prior experience to be important or informative (Loftus & Mackworth, 1978). A strong hypothesis is the "eye-mind" hypothesis (Just & Carpenter, 1976), according to which the eye provides a window to the user's mind, i.e., it provides a "dynamic trace of where a person's attention is being directed in relation to a visual scene." Although several exceptions to this hypothesis have been reported (for example, in covert attention, where the focus of user attention is different from where the eye is looking) for most natural viewing scenarios, eye fixations are thought to reflect

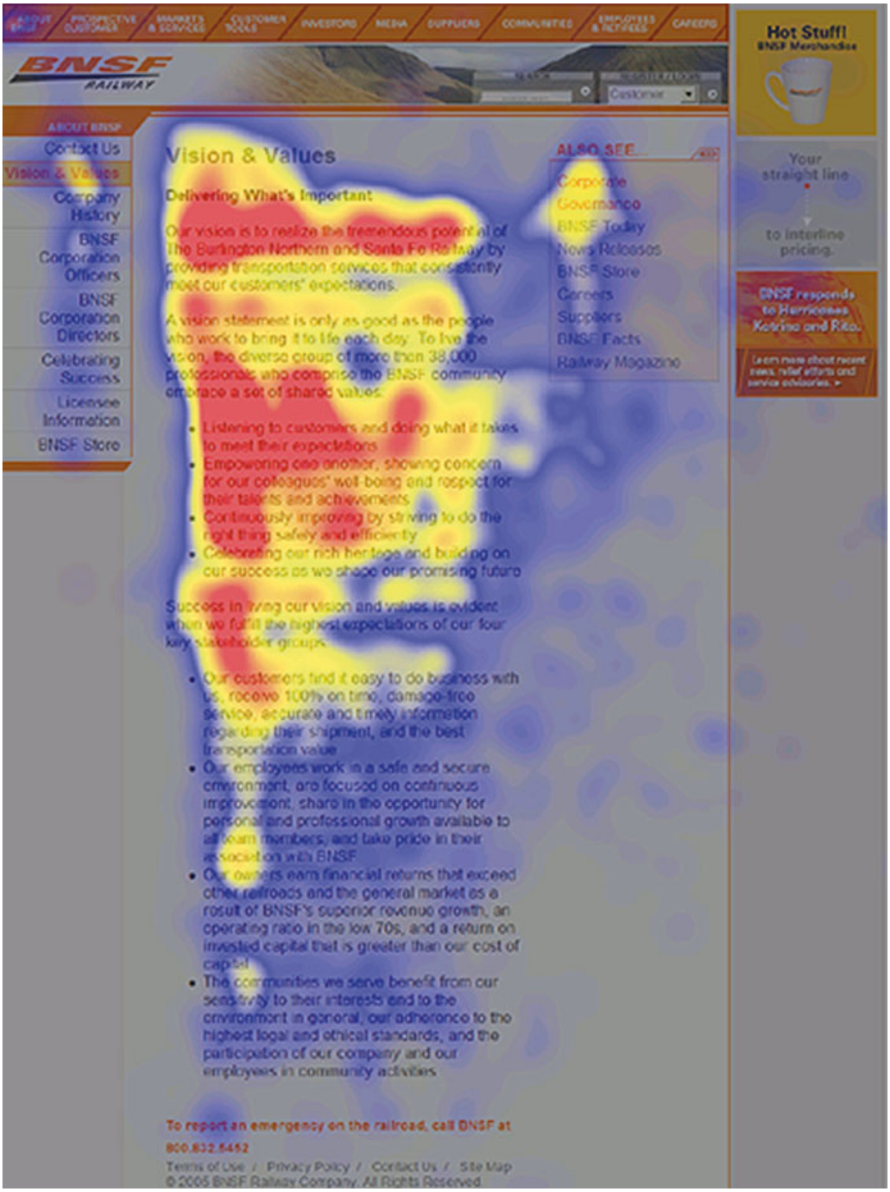

Fig. 4 Example F-pattern heatmap (*Source*: Nielsen Norman Group, http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/)

the current focus of the user's attention and the amount of cognitive processing on the fixated object(s).

In contemporary studies, a typical representation of results from an eye tracking study aggregated over several users is the "heatmap." Heatmaps use different colors to visualize the distribution and intensity of user attention on the display (see Fig. 4).

This contrasts with Fig. 3, which shows the data from a single user. Areas where users looked the most are colored red, yellow areas indicate fewer fixations, followed by the least-viewed areas in blue. Gray areas did not attract any fixations. The example in Fig. 4 is from a website's "About Us" pages. The heatmap clearly shows users' tendency to read in an "F" pattern, and their focus on information that is presented in bulleted lists.

While heatmaps are in common use these days, with colors depicting levels of attention/interest, there have been a number of different ways of representing what holds peoples' gaze and how to interpret that gaze. Duchowski (2002) characterizes eye tracking research as having developed along three historical periods. The first from 1879 to 1929 focused on psychophysiological characteristics of eye movements. Movement characteristics of the eyes, such as latency of saccadic eye movements, were studied in this period. The second came under the auspices of the behaviorism movement in understanding human behavior (1930–1958). The third era has been focused on technological developments and the production of increasingly accurate and reliable systems. Mele and Federici (2012) offer a thorough review of papers about eye tracking and discuss what they call the "fourth era" of eye tracking research: here, we see a greater emphasis on multidisciplinary contributions to the understanding of the significance and relevance of eye movements to the situational context in which the observed participant is engaged, including the disciplines of neuro-, cognitive-, and social-psychology and sociology.

Below we offer a further breakdown of the kinds of questions to which eye tracking as a method has been applied across different disciplines. While many of these may not be directly reported in the Human Computer Interaction (HCI) literature, they lead to new possibilities for understanding the power of this technique.

Vision Science (Neuroscience/Psychology)

Eye tracking has been widely used to study perceptual and cognitive processes in attention (e.g., visual search, memory, scene perception). We can think of fixation as either being "pulled" to a particular scene location by the visual properties at that location, or "pushed" to a particular location by cognitive factors related to what we know and what we are trying to accomplish (Henderson, 2003). For example, a bright or colorful area of a scene might attract the eyes simply because of its visual properties, where gaze reflects low-level processing of the human brain. Indeed, analysis of image statistics at fixated vs. non-fixated locations has shown that fixations tend to have higher density of edges, higher brightness contrast, and more generally, higher image saliency (Reinagel & Zador, 1999; Parkhurst, Law, & Niebur, 2002; reviewed in Itti & Koch, 2001). However, at a higher processing level, a viewer might want to look at scene regions that are relevant given current tasks and goals, whether or not those regions are visually prominent. Fixated regions differed in both their image statistics and their semantic content compared to regions that were not fixated (Henderson, Brockmole, Castelhano, & Mack, 2007). An ongoing focus of research is understanding the extent to which fixations are pulled by the stimulus or pushed by cognitive processes. Hayhoe and Ballard (2005) provide a good review of eye movements in natural behavior.

Computer Vision: Perceptual Models of Eye Gaze

Based on insights from eye tracking, several computational models have been developed to predict eye gaze on images and videos. These models compute visual saliency or eye-catchiness of image regions based on differences in visual properties such as color, orientation, size, brightness, motion, etc. A popular such model, which is inspired by the functioning of primate visual cortex, is the saliency model of Itti and Koch (2000). Other approaches include:

- Computing visual "surprise" in a Bayesian sense, the difference between prior and posterior belief in distribution of visual features in the world (Itti & Baldi, 2009)
- Use information theoretic measures (Bruce & Tsotsos, 2009; Zhang, Tong, Marks, Shan, & Cottrell, 2008)
- Machine learning classifiers to differentiate between fixated and non-fixated locations in images

These models perform reasonably accurately in predicting eye gaze on images and videos, especially for the first few seconds of viewing. They have potential applications in evaluating the visual catchiness of web page designs and advertisements.

In addition to visual attention models that are solely driven by image properties, there have also been attempts to model the role of user knowledge about the world. For example, when searching for cars or pedestrians, people tend to look at salient objects in the bottom half of the image where the street is most likely to appear. Thus, eye gaze is driven as a combination of saliency and knowledge of scene context. Torralba and colleagues have developed models of eye gaze that take both factors into account, and provide better prediction of user eye gaze on natural scenes (Torralba, Oliva, Castelhano, & Henderson, 2006).

An ongoing challenge is to develop predictive models of eye gaze that combine low-level image saliency with high-level semantics of the image and user intent. Examples of initial attempts in this direction and further refinements in the context of visual search task can be found in Navalpakkam and Itti (2002, 2005, 2007). In addition, there is an increasing number of models based on the emerging notion that attention and eye movement strategies serve to optimize human visual task performance (Najemnik & Geisler, 2005; Navalpakkam & Itti, 2007; Renninger, Verghese, & Coughlan, 2007; Stritzke, Trommershäuser, & Gegenfurtner, 2009).

Psychology: Reading Behavior

The study of eye movements during reading has a long and rich history dating back to the latter part of the nineteenth century (see Rayner (1998) for a good review). Eye tracking has been used for a critical examination of the cognitive processes underlying reading. For example, when reading English, eye fixations last about 200–250 ms and the mean saccade size is 7–9 letter spaces. Interestingly, many words are skipped so that foveal processing of each word is not necessary. As word length increases, the probability of fixating a word increases (Rayner & McConkie, 1976). As text becomes conceptually more difficult, fixation duration increases, saccade length decreases, and the frequency of backward saccades increases (Jacobson & Dodwell, 1979; Rayner & Pollatsek, 1994).

Language Processing

Methodologically, the drive to look at objects as they are mentioned provides an important tool for studying online language processing. For example, viewers will typically look to a scene area that contains an object when that object is mentioned by a speaker. Thus, eye fixation provides a pointer or index (Ballard, Hayhoe, Pook, & Rao, 1997) that anchors cognitive processes such as language understanding to entities in the world (Henderson & Ferreira, 2004; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

Neuroscience: Medical Conditions/Disorders

Eye tracking has also been used as a method for detecting medical conditions and/or disorders such as autism and attention deficit disorders. Kanner's original description of autism highlighted the social and emotional aspects of this disorder and demonstrated it with eye tracking data (Kanner, 1943). The most commonly used stimuli are pictures of human faces, but videotapes of social interactions, human voices, and abstract animations have also been employed. Normal adults show a very specific pattern of gaze when viewing faces, fixating mainly on the eyes, but also on the nose and mouth, the so-called "core features" (Walker-Smith et al., 1977). People with autism spend less time examining the eyes (Dalton et al., 2005; Pelphrey et al., 2002) and look more frequently at the mouths and bodies, and at other objects in the scene. Eye tracking could therefore be used to diagnose and understand cognitive processing in individuals with autism (Klin, Jones, Schultz, Volkmar, & Cohen, 2002).

Market Research

Eye tracking has been used extensively in Market Research to assess product designs and also the impact of advertisements on salience and memorability of brands, logos, and products. Common use cases include comparative experiments that determine which advertisement designs attract more attention (e.g., Lohse, 1997), as are experiments focused on determining whether Internet users look at banner advertising on websites (they do not) (Burke, Hornof, Nilsen, & Gorman, 2005). An example of a recent study using eye tracking to examine the relative impact of visual salience of the design and its perceived value on user choice can be found in Milosavljevic, Navalpakkam, Koch, and Rangel (2012).

Human Computer Interaction (HCI)

Turning specifically to HCI, eye tracking has been used as a method for several purposes:

- Understanding the perceptual aspects of user attention on displays (what do users notice)
- Cognitive aspects of attention (what do users focus on, or spend time processing)
- Social aspects of attention (e.g., mutual gaze in human–human interactions, explained later)
- As an input method, using gaze as an alternative to the keyboard and mouse

These various use cases are outlined in the next section.

The Variety of Uses of Eye Tracking in Human Computer Interaction

Research over the years has investigated eye gaze patterns while driving, flying, and reviewing X-ray images. More recently, researchers are increasingly using eye tracking as a method to understand *information seeking*, *searching, and browsing* with desktop and handheld devices. The process of inferring useful information from eye movement recordings involves defining “areas of interest” over certain parts of a display or interface under evaluation, and analyzing the eye movements that fall within such areas. Using heatmaps and measures such as fixation duration described in the methods section, the visibility, meaningfulness, and placement of specific interface elements can be objectively evaluated. The findings can be used to improve the design of the interface (Goldberg & Kotval, 1999). For example, usability studies in HCI routinely use eye gaze heatmaps and the sequence of eye fixations to evaluate websites and designs (Nielsen & Pernice, 2010). Dabbish and Kraut

(2004) use it to assess attentional distribution to better understand awareness on collaborative tasks.

Eye tracking has also been used to study people's *web search behaviors* (Cutrell & Guan, 2007; Granka & Rodden, 2006) and the relationship between mouse tracking and eye tracking (Rodden & Fu, 2007; Rodden, Fu, Aula, & Spiro, 2008). For example, Rodden et al. identified different types of eye-mouse coordination behavior, including the mouse moving randomly without any correlation to eye movements, or being parked at some spot on the page, or marking something as important, or following the eye vertically and, to a lesser extent, horizontally. A more recent study by Huang, White, and Buscher (2012) identified how correlations between eye tracks and mouse tracks vary with time from page load, and attempted to model eye position from mouse position. In one of our studies (Navalpakkam & Churchill, 2012), we identified eye gaze and mouse markers that are predictive of when users struggle to read content. Because mouse tracking is scalable (and unlike current forms of eye tracking, it doesn't require the user to wear special equipment), understanding user attention and other behaviors through mouse tracking and in relation to eye tracking is becoming a hot topic of research. Certainly there are newly emerging eye tracking techniques that are more scalable (e.g., using embedded laptop cameras and webcams); however, as mentioned in the method section, these have low accuracy and may have particular challenges if successful, webcam-based eye tracking could have a big impact on HCI and other fields (advertising, web page optimization, marketing research).

Eye tracking can be useful in studying *social interactions*, and *human–human conversations*. Within sociological studies, eye gaze has been used to uncover the role of mutual gaze in the ongoing conduct of social organization and social interactions (Argyle & Cook, 1976; Goffman, 1964; Goodwin, 1984; Kendon, 1967). Goffman, in particular, points out that eye gaze plays a crucial role in the initiation and maintenance of social encounters, describing what he calls an "eye-to-eye ecological huddle which tends to be carefully maintained, maximizing the opportunity to monitor one another's mutual perceivings" (Goffman, 1964, page 95). Goodwin and Kendon both offer detailed accounts with careful notations of the ways in which eye gaze is used as part of conversational turn-taking and to manage embodied mutual orientation within a conversation or toward a shared resource. In the context of human–human communication, eye gaze has been used in two ways:

- A "measure" of interpretation (e.g., to resolve ambiguity in utterances in face-to-face conversations, Henderson & Ferreira, 2004; Tanenhaus et al., 1995)
- To place "constraints" upon interpretation (Hanna & Brennan, 2007) that make eye gaze a powerful cue about attention and intention in face-to-face communication.

Eye tracking has been used to determine mutual gaze achievement in video conferencing and also with synthetic interface characters and on-screen avatars (Roberts et al., 2009; Steptoe et al., 2009).

Eye tracking has also been used as a method in Computer Supported Cooperative Work (CSCW) to understand and support human-to-human communication in mediated environments—both to better understand the process as well as to design and develop systems to support coordination and collaboration. For example, work by Fussell and colleagues show how the transfer of gaze in remote collaboration scenarios (i.e., sharing a speaker's gaze with listener by transferring or projecting it on the listener's system) can be incredibly useful to support communication and coordination (e.g., Fussell, Setlock, Parker, & Yang, 2003; Ou, Oh, Yang, & Fussell, 2005). Further, Ou et al. (2005), Ou, Oh, Fussell, Blum, and Yang (2008), and Ou, Shi, Wong, Fussell, and Yang (2006) use eye gaze to predict focus of attention and attentional distribution in a visual setting.

Researchers have recently started using eye tracking to track two people (dyadic eye tracking) at once, and also in mobile settings to understand things like *initiative*, *lead and follow patterns*, *attention*, *ambiguity resolution*, *coordination failures*, and so on (Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008; Cherubini, Nüssli, & Dillenbourg, 2008; Gergle & Clark, 2011; Jermann & Nüssli, 2012). An ongoing topic of research, some of the methodological challenges that arise for dyadic (more generally, multi-person) eye tracking are discussed in Richardson, Dale, and Kirkham (2007).

Finally, although oriented to more traditional web page interaction, recent work has used eye tracking as an investigative method to uncover patterns in human–device interaction from the perspective of a *social engagement*, rather than purely as cognitive information processing (Moore & Churchill, 2011; Moore et al., 2011).

In addition to being used as a method to evaluate people's interaction with devices and to study mediated communication (e.g., human–human communication through video conferencing), eye tracking has also been used as an input method for adaptive interfaces (Jacob & Karn, 2003a, 2003b). Experiments suggest that selection with eye gaze can be a robust input/selection technique, and in fact, eye selection can be faster than using a mouse. In 2007, Oyekoya demonstrated experimentally that eye gaze can be a viable selection method for visual search tasks, and that prior experience on visual tasks with a mouse can create a "training effect" (Oyekoya, 2007). Focusing on assistive technologies and what they call "psychotechnologies," Mele and Federici (2012) outline several opportunities for eye tracking technologies as the underlying technology for eye gaze to be more fully realized as an input technique.

How to Do It: What Constitutes Good Work?

Like any other method, a clean experiment design makes eye tracking a viable method for understanding user behavior.[4] We describe the main aspects of a good eye tracking study (described in detail in Duchowski, 2007) with an example from our work below (Navalpakkam & Churchill, 2012; Navalpakkam, Rao, & Slaney, 2011).

[4] For more details on experimental design, please see the chapter on Experimental Research in HCI in this volume.

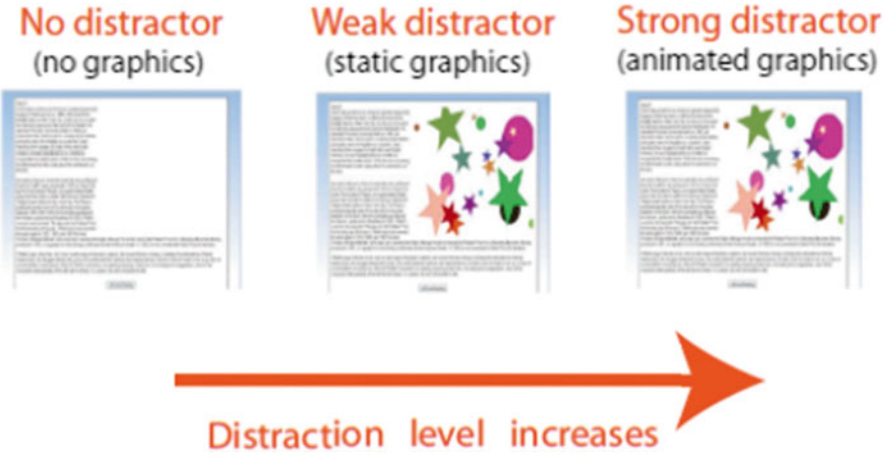

Fig. 5 This figure illustrates the experiment design with control group (no distractor), treatment group 1 (weak distractor), and treatment group 2 (strong distractor). *Source*: Navalpakkam et al., 2011

1. *Hypothesis*: Formulate clear null and alternate hypotheses, motivate them, and state the underlying assumptions.

   For example, in our case study, for *Motivation*, we wished to understand how the presence of distracting elements on a web page affects eye gaze patterns. Our *Null hypothesis H0* was that eye gaze patterns on a web page do not change in the presence of distracting elements; our *alternate hypothesis H1* was that users spend more time looking at the distracting element; our second *alternate hypothesis H2* was that users spend more time looking at the page content.
2. *Design*: Determine whether it is an observational study or an experimental study. If the latter, design control and treatment groups to test the hypotheses. Identify independent and dependent variables; change the independent variable's values, while keeping everything else more or less a constant (thus avoiding confounding factors); and measure the impact on the dependent variable. Also determine whether it is a within-subjects design or between-subjects design. There can be differences in eye movements between participants on identical tasks, thus it may be prudent to use a within-participants design in order to make valid performance comparisons (Goldberg & Wichansky, 2003).

   In the case study, the control group did not see any distracting elements on the web page. Treatment group 1 saw a moderately distracting element in the form of a static, irrelevant graphic on the top right of the page. Treatment group 2 saw a highly distracting element in the form of an animated, irrelevant graphic on the top right of the page, shown in Fig. 5. A within-subject design was used, and the order of experimental conditions was randomized to balance familiarity and fatigue. Each participant saw three essays (from the Test of English Language Fluency) that were randomly paired with one of the graphic types. Each essay consisted of 300–400 words, followed by 5 factual/theme-based multiple-choice questions (to confirm that users indeed read the essays and performed the task as told), and 2 subjective questions where subjects were asked to rate their user experience on a scale of 1–5 for pleasantness.

In this study, the independent variable was the level of distraction (varied from “none” in the control group, to “medium” in treatment 1 to “high” in treatment 2). Dependent variables included the amount of time users spent looking at the page content and distracting element, the corresponding number of eye gaze fixations, the time to first eye fixation, and the user-reported levels of pleasantness of experience.

3. *Task description*: What task is assigned to the participants? Are they freely viewing the displays, or are they performing a task such as searching for a particular object in the display. The task description is a critical part of an eye tracking study. Its importance is highlighted in a classic study by Yarbus (1967) that shows how eye movements on a painting are influenced by the task given to the user, with other things kept constant. The eye tracking data shows that when asked to determine the ages of the people in the painting, eye gaze focused on the faces of people, whereas when asked to assess material circumstances of the family, eye gaze focused on the clothes that people wore, the furniture, and other features in the visual scene (see Fig. 6). In our example study, the task assigned to participants was a reading comprehension task—“Read the article on this web page and answer the questions that follow.” Unknown to the participants, the goal of our study was to test how the presence of distracting elements on the page affects eye gaze. Thus we avoided any potential biasing of participants’ behavior that may result from them knowing the study’s goal.
4. *Participants*, *Apparatus, and Procedure*: A good eye tracking study should include a description of the participants (e.g., number, age group, gender, demographics); compensation or incentive structure (is there a performance bonus?); apparatus used (eye tracker model, monitor resolution, display viewing angle, calibration accuracy); and procedure of the study (instructions before the study, flow of the study, posttask feedback).

In the case study, there were 20 participants (8 female, 12 males; residents of United States), aged 19–60, with normal or corrected vision. Participants were fluent in English (spoken and written), and had either completed or were pursuing undergraduate education. We recorded participants’ gaze patterns during task performance using a Tobii 1750 eye tracker (50 Hz sampling frequency), with a 17″ LCD monitor, set at resolution $1{,}024 \times 768$, at roughly 85 cm viewing distance. We collected a log of eye and mouse movement.

Participants were compensated as follows: a flat payment rate for participation in the study, and in addition, $1 for every correct answer (5 questions per essay $\times$ 3 essays). During data cleaning, three participants were excluded for the following reasons: poor calibration (two did not maintain their head in the correct position), or outliers in fixation duration or number of fixations (3 standard deviations, 1 participant).

The study began with a 5-point calibration procedure followed by the task-instruction screen. This was followed by one practice essay paired with animated

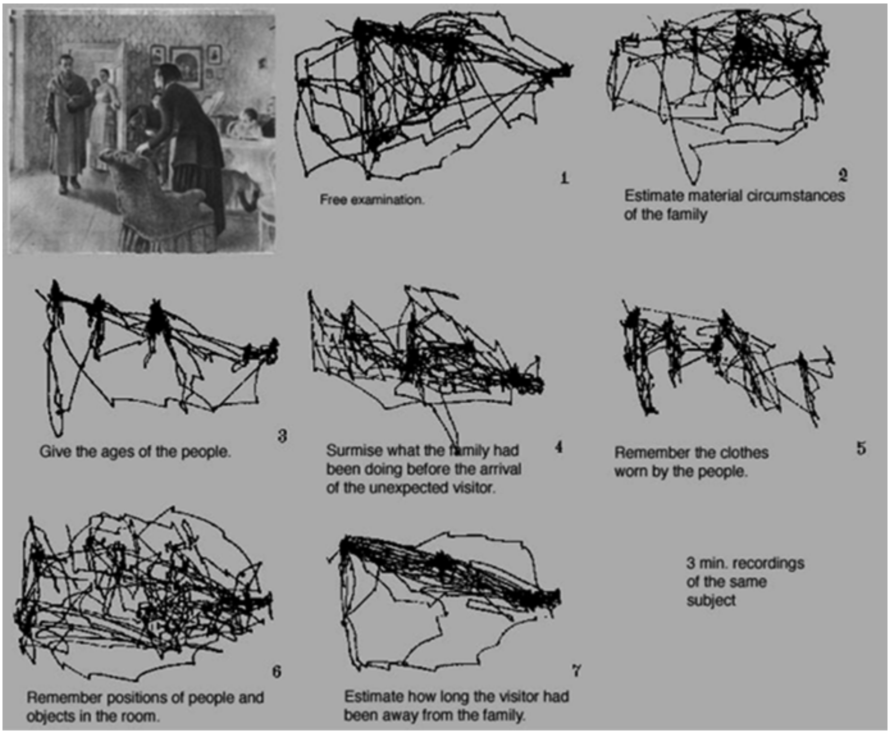

Fig. 6 Eye movements given different tasks—From Yarbus (1967)

graphics to help participants familiarize themselves with the task, types of graphics, and the format of the questions in the reading comprehension test. Following the practice trial, participants saw three essays randomly paired with no, static, or animated graphics (i.e., control, treatment 1 and 2 groups in randomized order). At the conclusion of the study, participants were paid based on task performance.

5. *Analysis*: A critical part of a good eye tracking study is conducting careful and appropriate analysis. Eye tracking data can be analyzed in qualitative ways (e.g., heatmap visualization, observing where people look) and using rigorous quantitative methods. The latter consists of defining areas of interest and extracting measures for each area of interest, such as the number of eye fixations, duration of eye fixations, number of saccades, time to first fixation, and number of backward saccades (called "regressions," suggesting confusion or distraction). As mentioned later under challenges, defining and determining areas of interest (AOIs) for analysis can be highly complex as one moves toward dynamic and/or longer-term tasks. It is much easier for a stable, 2D image, with predefined areas of interest where we want to analyze gaze. If interested in questions such as "where does the user look next," one could also look at the temporal ordering of eye fixations by extracting transition probabilities P(x,y), which describe the

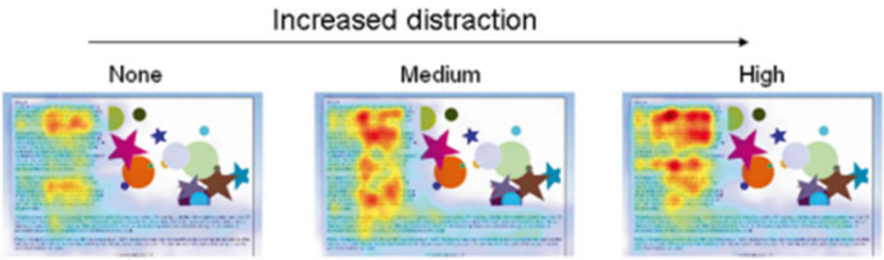

Fig. 7 Example of heatmap from Navalpakkam et al., CHI 2011

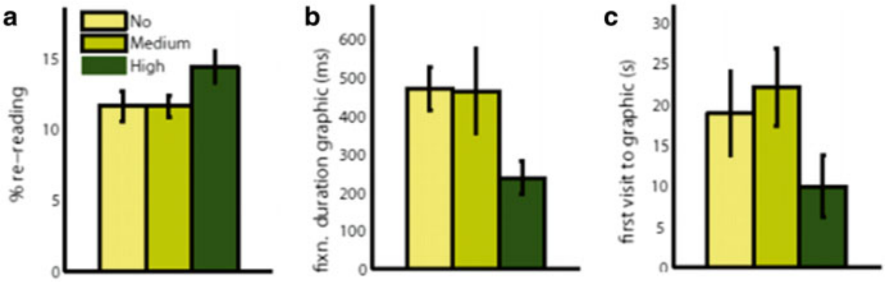

Fig. 8 Example eye tracking measures for the distraction study

probability that the user will look next at display item or location "y," given that she is currently looking at item or location "x." Navalpakkam et al. (2011) offer an example of using the above eye tracking metrics and analysis in the context of understanding and modeling how image presence, position, and user interest drive the way users attend to and select online news content. More common types of analyses such as heatmaps and fixation metrics are illustrated below with an example from the case study.

Figures 7 and 8 show that users spend more time processing the page content as the amount of distraction on the page increases. Similar effects were observed with mouse tracks as well. For easy comparison, we overlay the heatmaps on the same image; however, note that in the "no" distraction condition, there was no graphic on the right-hand side of the page.

A particularly interesting finding in the study was that more time on page is not always good. We found that users spent more time on page in the high distraction condition, but they reported being more annoyed (very low pleasantness scores). Analysis of eye tracking data revealed that the increased time was due to struggle in reading, and increased cognitive effort in processing the content in the presence of a highly distracting graphic. For example, in Fig. 8, although the highly (animated) distracting element was noticed earlier in time (panel C), it was rejected faster (panel B), and users spent more time processing and re-reading the page content (panel A), which is an indication of struggles with reading. The authors further show in the paper that further, these eye tracking patterns can predict subjective assessments, like user frustration with high accuracy. Further details can be found in the paper itself (Navalpakkam et al., 2011).

6. *Clarify assumptions*: Finally, one must clarify the assumptions underlying the conclusions from the data. For example, a common assumption is that the amount of time spent looking at an item on page reflects the amount of user attention and cognitive processing on that item—more time spent is assumed to imply more attention and deeper cognitive processing. While this assumption is reasonable for most scenarios, as has been discussed already eye fixation and gaze duration are sometimes not correlated with the depth of the user's cognitive processing. For example, sometimes people look up when they concentrate or are conducting hard mental operations. Assumptions about the correlations between eye gaze direction, focus, and time must be examined, and we must be cautious about over-generalization.

Today's Challenges

The dynamic nature of modern computer interfaces provides a technical challenge for studying eye fixations. For example, with pop-up messages, animated graphics, and user-initiated object movement and navigation, objects can move around a screen, or move off of 2D screens. As a result, the definition of areas of interest becomes a challenging problem. Knowing that a person was fixating 10° above and 5° to the left of the display's center does not allow us to know what object the person was looking at in the computer interface unless we keep track of the changes in the computer display. Analysts must bear this in mind while considering dynamic displays. Recent software packages sold by the various eye tracking vendors now have definable areas of interests that can change over time, but even those often must be generated post hoc. Gergle and Clark (2011) suggest another solution that couples eye tracking with vision tracking techniques by using "objects of interest" as opposed to static areas of interest.

Conclusions

As modern computer interfaces continue to shrink in size (from desktops to laptops to tablets to phones) and become more mobile, understanding how users process information as they move around in the world becomes important. Upcoming technologies like webcam-enabled eye tracking are exciting and offer hope, but need to deal with challenges in calibration and accuracy (that are rendered difficult due to varying distance between user and device, varying head pose, and lighting conditions).

Finally, as Chi and colleagues (2009) have discussed, ideally, eye tracking methods should possess the following factors:

1. Accuracy
2. Reliability
3. Robustness
4. Non-intrusiveness
5. The possibility for free head movements
6. No prior calibration
7. Real-time response

We add to this list that eye tracking methods should be

8. Work for Dynamic displays
9. Allow for study participants' mobility
10. Be Scalable

Achieving all of these factors in one system is not yet possible as systems still require calibration, and because there are accuracy-intrusiveness and accuracy-scalability trade-offs. However, it is possible to imagine, as cameras improve and mounts for mobile eye trackers become increasingly lightweight, that we will see the emergence of more powerful eye tracking opportunities in the future. With these we will be able to more accurately discern and assess degree of attention, level of interest, and management of cognitive and social interaction.

Exercises

1. Unlike regular experiments, what do experiments with eye tracking have to begin with? How is this done?
2. When is the eye's direction not a good indicator of what the person is looking at or thinking about? How would you separate those from "real" perception?

References and Further Reading[5]

References for getting more expert in this method

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. Secaucus, NJ: Springer.
Duchowski, A. T. (2002). A breadth-first survey of eye tracking applications. *Behavior Research Methods, Instruments, and Computers, 34*, 455–470.
Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movements* (pp. 493–516). Oxford, UK: Elsevier Science.
Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van De Weijer, J. (2011). *Eye tracking. A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

[5] There are many texts that can be drawn on to learn more about eye tracking. We list some below.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Henderson, J. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188–194.

Jacob, R. J. K., & Karn, K. S. (2003a). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573–603). Oxford, England: Elsevier.

Poole, A., & Ball, L. J. (2005). Eye tracking in human-computer interaction and usability research: Current status and future prospects. *Psychology, 10*(5), 211–219.

References to example papers that have used it well, a short commentary about each reference and what questions they answer can be found in the footnotes

Goldberg, J. H. & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics, 24*, 631–645[6]

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*, 766–786[7]

Reinagel, P. & Zador, A. M. (1999). Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems, 10*, 341–350[8]

Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative visual saliency differences induce sizeable bias in consumer choice. *Journal of Consumer Psychology, 22(1)*, 67–74[9]

Navalpakkam, V., Rao, J. M., & Slaney, M. (2011). Using gaze patterns to study and predict reading struggles due to distraction. *CHI 2011*, 7–12 May, 2011, Vancouver, BC, Canada[10]

[6] Comparative assessment of measures of eye movement locations and scanpaths used for evaluation of interface quality. Revealed that well-organized functional grouping of icons result in shorter scanpaths, covering smaller areas. Less well organized interfaces produce less efficient search behavior. However, poorly organized icon groupings do not affect users' ability to interpret/understand icons.

[7] While searching for objects in scenes, eye gaze is affected by the scene context, e.g., one searches for pedestrians on the lower half of a scene, where the street is most likely to be; with this one searches for birds in the upper half of the scene, where the sky is most likely to be.

[8] While freely viewing images, eye gaze is biased towards image locations that have high spatial contrast.

[9] Fast eye movements and choices between food items are driven by perceptual factors such as the visual catchiness or saliency of items, while slower choices are driven by high-level factors such as the value of items.

[10] Eye movements differ when users are distracted compared to when they are not. This study identifies eye gaze markers that are predictive of user reading struggle, or frustration. This study and its companion paper (Navalpakkam & Churchill CHI 2012, focusing mainly on the relationship between eye and mouse tracking) have been used as examples of designing, conducting and analyzing an eye tracking study.

Moore, R. J. & Churchill, E. F. (2011). Computer Interaction Analysis: Toward an Empirical Approach to Understanding User Practice and Eye Gaze in GUI-Based Interaction. *Computer Supported Cooperative Work, 20 (6)*, 497–528[11]

Other references

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*(04), 723–742.

Brennan, S. E., Chen, X., Dickinson, C., Neider, M., & Zelinsky, G. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition, 106*, 1465–1477.

Boraston, Z., & Blakemore, S. (2007). The application of eye tracking technology in the study of autism. *The Journal of Physiology, 581*, 893–898.

Burke, M., Hornof, A., Nilsen, E., & Gorman, N. (2005). High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Transactions on Computer-Human Interaction, 12*(4), 423–445.

Buswell, G. T. (1935). *How people look at pictures*. Chicago, IL: University of Chicago Press.

Bruce, N. D., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision, 9*(3), 5. 1–24.

Cherubini, M., Nüssli, A. M., & Dillenbourg, P. (2008). Deixis and gaze in collaborative work at a distance: A computational model to detect misunderstandings. In *ETRA '08: Proceedings of the 2008 symposium on eye tracking research and applications* (pp. 173–180). New York, NY: ACM.

Chi, J.-n., Zhang, P.-y., Zheng, S.-y., Zhang, C., & Huang, Y. (2009). Key techniques of eye gaze tracking based on pupil corneal reflection. In *WRI global congress on intelligent systems, 2009: GCIS '09, 19–21 May, 2009, Xiamen, CN* (pp. 133–138). Washington, DC: IEEE.

Clarke, A. H., Ditterich, J., Druen, K., Schonfeld, U., & Steineke, C. (2002). Using high frame rate CMOS sensors for three-dimensional eye tracking. *Behavior Research Methods, Instruments, and Computers, 34*, 549–560.

Cooke, L. (2006). Is eye tracking the next step in usability testing? In *IEEE international professional communication conference: IPCC '06, 23–25 October, 2006, Saratoga Springs, NY* (pp. 236–242). Washington, DC: IEEE. doi:10.1109/IPCC.2006.320355.

Cutrell, E., & Guan, Z. (2007). What are you looking for?: An eye tracking study of information usage in web search. In *Proceedings of the SIGCHI conference on human factors in computing systems, Montréal, QC, Canada, 22–27 April, 2006* (pp. 407–416). New York, NY: ACM.

Dabbish, L., & Kraut, R. (2004). Controlling interruptions: Awareness displays and social motivation for coordination. In *Proceedings of the 2004 ACM conference on computer supported cooperative work (CSCW '04), Chicago, IL, 6–10 November, 2004* (pp. 182–191). New York, NY: ACM.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8*, 519–526.

[11] The authors outline a novel methodological approach to understanding how humans interact with interactive applications and services. They call their approach "computer interaction analysis". Computer interaction analysis extends traditional eye tracking approaches by interweaving and drawing out the relationship between people's input actions, system display events and people's eye movements.

Ellis, S., Candrea, R., Misner, J., Craig, C. S., & Lankford, C. P. (1998). Windows to the soul? What eye movements tell us about software usability. *7th annual conference of the Usability Professionals' Association conference: UPA '98*, June 22–26, 1998, Washington, DC

Exline, R. V. (1974). Visual interaction: The glances of power and preference. In S. Weitz (Ed.), *Nonverbal communication* (pp. 65–92). New York, NY: The Oxford University Press.

Exline, R. V., & Winters, L. G. (1965). Affective relations and mutual glances in dyads. In S. Tomkins & C. Izzard (Eds.), *Affect, cognition and personality*. New York, NY: Springer.

Fussell, S. R., Setlock, L. D., Parker, E. M., & Yang, J. (2003). Assessing the value of a cursor pointing device for remote collaboration on physical tasks. In *CHI '03 extended abstracts on human factors in computing systems* (pp. 788–789). New York, NY: ACM.

Garrett, J. J. (2003). *The elements of user experience: User-centered design for the web*. New York, NY: New Riders Press.

Gergle, D., & Clark, A. T. (2011). See what I'm saying? Using dyadic mobile eye tracking to study collaborative reference. In *Proceedings of the ACM 2011 conference on computer supported cooperative work (CSCW '11)* (pp. 435–444). New York, NY: ACM.

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *American Journal of Psychology, 76*, 386–394.

Gramatikov, B. I., Zalloutm, O. H., Wu, Y. K., Hunter, D. G., & Guyton, D. L. (2007). Directional eye fixation sensor using birefringence-based foveal detection. *Applied Optics, 46*, 1809–1818.

Goffman, E. (1964). *Behavior in public places*. Glencoe: The Free Press.

Goodwin, C. (1984). Notes on story structure and the organization of participation. In M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 225–246). Cambridge: Cambridge University Press.

Granka, L. & Rodden, K. (2006). Incorporating eye tracking into user studies at Google. In *Workshop Position paper presented at CHI.*

Hanna, J., & Brennan, S. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language, 57*(4), 596–615.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford: Elsevier.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world*. New York, NY: Psychology Press.

Huang, J., White, R., & Buscher, G. (2012). User see, user point: Gaze and cursor alignment in web search. In *Proceedings of the 2012 ACM annual conference on human factors in computing systems* (pp. 1341–1350). New York, NY: ACM.

Humphrey, K., & Underwood, G. (2009). Domain knowledge moderates the influence of visual saliency in scene recognition. *British Journal of Psychology, 100*, 377–398.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research, 49*(10), 1295–1306.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10–12), 1489–1506.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience, 2*, 194–203.

Jacob, R. J., & Karn, K. S. (2003b). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind, 2*(3), 4.

Jacobson, J., & Dodwell, P. C. (1979). Saccadic eye movements during reading. *Brain and Language, 8*(3), 303–314.

Javal, E. (1990). Essay on the physiology of reading. *Ophthalmic and Physiological Optics, 10*, 381–384.

Jermann, P., & Nüssli, M. A. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In *Proceedings of the ACM 2012 conference on computer supported cooperative work (CSCW '12)* (pp. 1125–1134). New York, NY: ACM.

Johansen, S. A., Agustin, J. S., Skovsgaard, H., Hansen, J. P., & Tall, M. (2011). Low cost vs. high-end eye tracking for usability testing. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1177–1182.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–480.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*(3), 217–250.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psycholoigica, 26*, 22–63.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*, 809–816.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*, 3559–3565.

Latimer, C. R. (1988). Eye-movement data: Cumulative fixation time and cluster analysis. *Research Methods, Instruments and Computers, 20*(5), 437–470.

Loftus, G. P., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 565–572.

Lohse, G. L. (1997). Consumer eye movement patterns on yellow pages advertising. *Journal of Advertising, 26*(1), 61–73.

Mele, M. L., & Federici, S. (2012). A psychotechnical review on eye tracking systems: Toward user experience. *Disability and Rehabilitative Technology, 7*(4), 261–281.

Moore, R. J., Churchill, E. F., & Kantamneni, R. G. P. (2011). Three sequential positions of query repair in interactions with internet search engines. *In Proceedings of CSCW 2011* (pp. 415–424). New York: ACM

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature, 434*(7031), 387–391.

Navalpakkam, V., & Churchill, E. (2012). Mouse tracking: Measuring and predicting users' experience of web-based content. In *Proceedings of the 2012 ACM annual conference on human factors in computing systems* (pp. 2963–2972). New York, NY: ACM.

Navalpakkam, V., & Itti, L. (2002). A goal oriented attention guidance model. In *Biologically motivated computer vision* (pp. 81–118). Berlin: Springer.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research, 45*(2), 205–231.

Navalpakkam, V., & Itti, L. (2007). Search goal tunes visual features optimally. *Neuron, 53*(4), 605–617.

Nielsen, J., & Pernice, K. (2010). *Eye tracking web usability*. Berkeley, CA: New Riders.

Ou, J., Oh, L. M., Yang, J., & Fussell, S. R. (2005). Effects of task properties, partner actions, and message content on eye gaze patterns in a collaborative task. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 231–240). New York, NY: ACM.

Ou, J., Shi, Y., Wong, J., Fussell, S. R., & Yang, J. (2006). Combining audio and video to predict helpers' focus of attention in multiparty remote collaboration on physical tasks. In *Proceedings of the 8th international conference on multimodal interfaces* (pp. 217–224). New York, NY: ACM.

Ou, J., Oh, L. M., Fussell, S. R., Blum, T., & Yang, J. (2008). Predicting visual focus of attention from intention in remote collaborative tasks. *IEEE Transactions on Multimedia, 10*(6), 1034–1045.

Oyekoya, O. (2007). Eye tracking: A perceptual interface for content based image retrieval. Unpublished Ph.D. thesis, Department of Electronic & Electrical Engineering, Adastral Park Campus, University College London.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42*(1), 107–124.

Poole, A., & Ball, L. (2006). Eye tracking in human-computer interaction and usability research: Current status and future prospects. In C. Ghaoui (Ed.), *Encyclopedia of human computer interaction* (pp. 211–219). London, UK: Idea Group Reference.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders, 32*, 249–261.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16*(8), 829–837.

Rayner, K., & Pollatsek, A. (1994). *The psychology of reading*. Hillsdale, NJ: Lawrence Erlbaum.

Reimer, M. D. (1955). Abnormalities of the gaze – A classification. *Psychiatric Quarterly, 29*, 659–672.

Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision, 7*(3), 6. 1–17.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*, 368–373.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science, 18*(5), 407–413.

Roberts, D., Wolff, R., Rae, J., Steed, A., Aspin, R., McIntyre, M., et al. (2009). Communicating eye-gaze across a distance: Comparing an eye-gaze enabled immersive collaborative virtual environment, aligned video conferencing, and being together. In *Virtual reality conference, 2009. VR 2009. IEEE* (pp. 135–142). Washington, DC: IEEE.

Robinson, G. H. (1979). Dynamics of the eye and head during movement between displays: A qualitative and quantitative guide for designers. *Human Factors, 21*(3), 343–352.

Rodden, K. & Fu, X. (2007). Exploring how mouse movements relate to eye movements on web search results pages. In *SIGIR 2007 workshop on web information seeking and interaction (WISI)*, July 27, 2007, Amsterdam, The Netherlands, pp. 29–32.

Rodden, K., Fu, X., Aula, A., & Spiro, I. (2008). Eye-mouse coordination patterns on web search results pages. In *CHI '08 extended abstracts on human factors in computing systems* (pp. 2997–3002). New York, NY: ACM.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the Eye Tracking Research and Applications Symposium (pp. 71–78). New York: ACM Press.

Steptoe, W., Oyekoya, O., Murgia, A., Wolff, R., Rae, J., Guimaraes, E., et al. (2009). Eye tracking for avatar eye gaze control during object-focused multiparty interaction in immersive collaborative virtual environments. In *Virtual reality conference, 2009. VR 2009. IEEE* (pp. 83–90). Washington, DC: IEEE.

Stritzke, M., Trommershäuser, J., & Gegenfurtner, K. R. (2009). Effects of salience and reward information during saccadic decisions under risk. *JOSA, A26*(11), B1–B13.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 632–634.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*, 766–786.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies in face perception. *Perception, 6*, 313–326.

Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W. (2008). SUN: A Bayesian framework for saliency using natural statistics. *Journal of Vision, 8*(7), 32.