# 2

# PERVASIVE EYE-TRACKING FOR REAL-WORLD CONSUMER BEHAVIOR ANALYSIS

*Andreas Bulling and Michel Wedel*

## Introduction

Human gaze has long held a particular fascination among researchers and practitioners alike because of its fundamental importance in human communication and interaction as well as its close links to human perception and cognition. A large body of work in the psychological and social sciences as well as in consumer behavior research has shed light on the many different ways in which a wide range of factors influence or are influenced by gaze behavior. Much of this research has been conducted in controlled laboratory settings where users sit behind a computer monitor and look at carefully designed visual stimuli. Real-world consumer behavior has, for a long time, been beyond the reach of gaze behavior research.

Methods to assess consumers' attention, their point of gaze, or the movement of their eyes over time use either sensors placed in the environment (so-called stationary eye-tracking) or worn on the head (so-called mobile eye-tracking). While early eye trackers were intrusive, cumbersome to use, and restricted data collection to short-term recordings in controlled laboratory settings, two recent developments have started to change this. First, mobile eye trackers can now be implemented as lightweight embedded systems, and therefore have become suitable for recordings in mobile daily life settings (Bulling & Gellersen, 2010; Tonsen, Steil, Sugano, & Bulling, 2017)—facilitating recordings over several hours or days and for large groups of users (Bulling, Weichel, & Gellersen, 2013; Pieters & Wedel, 2004; Steil & Bulling, 2015). Second, methods for stationary eye-tracking that only require a single off-the-shelf camera have recently improved considerably and now promise accurate eye-tracking capabilities on the millions of handheld devices, displays deployed in public, and smart appliances at home that increasingly feature integrated cameras (Sugano, Zhang, & Bulling, 2016; Wood & Bulling, 2014; Zhang, Sugano, & Bulling, 2017). Taken together, both advances give rise to a new class of *pervasive eye-tracking* systems that will enable continuous, robust, and accurate monitoring of gaze in everyday life.

The ramifications of this imminent paradigm shift are far-reaching in the social, psychological, business, and computer sciences, as well as for practice. Gaze has a long history as a modality for explicit human-computer interaction, such as for gaze-based pointing or object selection (Majaranta & Bulling, 2014), as well as in social signal processing (Adams & Kleck, 2003) and in artificial conversational agents (Vertegaal, Slagter, van der Veer, & Nijholt, 2001). Gaze also serves as a source of implicit information about users, including their behavioral context (Bulling, Ward, Gellersen,

& Tröster, 2011; Zhang, Wedel, & Pieters, 2009), intents and goals (Bednarik, Vrzakova, & Hradis, 2012; Pieters & Wedel, 2007), cognitive processes and states (Bulling & Roggen, 2011; Bulling & Zander, 2014), and even personality traits (Hoppe, Loetscher, Morey, & Bulling, 2015). With eye-tracking now moving into everyday life, gaze will also become a rich source of information on the "inner workings" of consumers—information that is difficult if not impossible to obtain from other modalities available today (see Chapters 1, 3–5). This information will become readily available at large scales in real-world environments, for example, while consumers are making purchases in brick-and-mortar and online stores, providing unprecedented insights into consumers' preference formation and decision-making processes (Stüttgen, Boatwright, & Monroe, 2012). As such, *pervasive eye-tracking, defined as the collection and utilization of gaze data in real-world settings in which consumers go about their everyday tasks*, has the potential to become a core technology to passively monitor, analyze, and actively manage consumer attention. Pervasive eye-tracking also has significant potential to facilitate, support, and enhance consumers' interactions, such as with interactive billboards or smart shelves. Figure 2.2 illustrates the spectrum of pervasive eye-tracking applications from offline to real-time.

## Eye-Tracking

One of the first eye-tracking devices was developed by Huey (1898). The device consisted of a lever that was attached to a cup that was placed on the eye and that had a hole for the respondent to see through. Moving the eyes caused a pen attached to the cup to move across the surface of a drum, which recorded the eye movement. This device had obvious mechanical limitations. To alleviate these, Orschansky (1899) attached a mirror to the eye cup, and recorded the reflection of light on the mirror. It soon became clear that it was even better to record light reflected by the surface of the eye itself. Dodge first used this principle when developing his "falling plate" camera (Dodge, 1900). His eye tracker consisted of a photographic plate that was lowered gradually to record the reflection of sunlight off a white piece of cardboard placed in front of the eye. This device produced the first published trace of eye movements (Wade, 2010).

Today, eye trackers typically measure eye movements in one of three ways (see also Chapter 4; Duchowski, 2007): a) video-based infrared pupil-corneal reflection (PCR), b) measurement of the cornea-retinal standing potential between the front and the back of the human eye (Electrooculography, EOG), and c) video-based eye-tracking using head-mounted or stationary visible light video cameras (video-oculography). We call the latter systems *pervasive eye-tracking systems*. They can be further categorized depending on whether they use eye landmarks, such as pupil centers and eye corners, or directly estimate gaze direction from the eye image using machine learning techniques (Hansen & Ji, 2010). Video-based eye trackers can be either used in a stationary or mobile configuration. A typical setup consists of a video camera that records the movements of the eye(s) and a computer and software that processes, saves, and analyzes the gaze data. In stationary systems, this eye camera is often placed below the screen while in mobile systems the camera is mounted on a glasses-like frame. Mobile eye trackers additionally include an egocentric scene camera to map the user's gaze direction to the visual scene and thus facilitate subsequent analysis.

### *PCR-Based Eye-Tracking*

Most commercial stationary eye trackers are PCR-based, and have gained popularity because they typically have high spatial and temporal precision. PCR-based eye trackers emit infrared light and use cameras to detect the reflection of the light on the cornea, the outer layer of the eye. Most devices are binocular, i.e., they detect the reflection on both eyes. The point of focus of the eyes is estimated from the relative distance between the corneal reflection(s) and the pupil center, which has to be determined through a calibration task prior to first use. Because the distance between the pupil and the corneal reflection does not change much during head movements, modern stationary eye trackers allow small head movements. PCR-based eye trackers record the $x$- and $y$-coordinates of the point of gaze with sampling rates of up to 2 kHz and a spatial resolution of 0.5°.

Stationary eye-tracking confines users' body and head movement to a rather small virtual tracking box about half a meter away from the tracker. This requirement has, for a long time, restricted gaze recordings to controlled laboratory settings and carefully selected stimuli that were presented to participants on a computer screen for predefined durations. This approach has been increasingly criticized because principles guiding the eyes when looking at computer screens can be very different from those when engaging in everyday behavior (Foulsham, Walker, & Kingstone, 2011). Findings obtained in controlled settings may thus have limited validity for natural environments (Kingstone, Smilek, & Eastwood, 2008). Compelling evidence for these differences was provided in a study which compared eye movements of participants while exploring different real-world environments and watching videos of these environments (Marius't Hart et al., 2009). The distribution of eye movements obtained in the laboratory predicted the gaze distribution in the real world with around 60% accuracy—indicating significant differences in eye movements between laboratory and real-world situations. Mobile eye trackers (see Figure 2.1) address these challenges by allowing collection of gaze data in everyday settings and during unconstrained head and body movements, including daily activities such as making tea or sandwiches, driving, walking, playing sports, and shopping (Hayhoe & Ballard, 2005).

Mobile PCR-based eye trackers consist of lightweight glasses, in which miniature infrared (IR) cameras pointing at one or both eyes, as well as a scene camera, are built in. The IR cameras record the Purkinje reflections off the cornea as in stationary PCR-based eye trackers, while the scene

**FIGURE 2.1** (left) PUPIL from Pupil Labs is an accessible, affordable, and extensible open source platform for mobile eye-tracking, gaze-based interaction, and egocentric vision research. (right) MEME from J!NS is an integrated eyewear computer for measuring eye movements using Electrooculography (EOG) and head movements using integrated inertial sensors.

camera additionally records the participant's field of view. Because the eye camera is mounted on the head and moves with it, the position of the IR source relative to the eye is nearly fixed. This allows participants to move around freely, i.e., without any constraints on head or body movements. The output of the eye tracker is a video of the participant's field of view with the 2D gaze point overlaid in real-time in the form of a cursor or cross-hair. The two most important applications of mobile eye-tracking, so far, are in the analysis of visual attention and behavior in human vision research, and for gaze-based human-computer interaction in computer science.

### *Video-Based Eye-Tracking Using Visible Light Cameras*

While PCR-based eye trackers can provide high tracking accuracy, they do require special-purpose equipment (IR light sources and special cameras) that is not commonly available. This requirement triggered research into methods that only require off-the-shelf cameras in combination with computer vision algorithms. These video-based methods can generally be categorized as model-based or learning-based (Hansen & Ji, 2010). Model-based methods use a geometric model of the human eye as a basis for estimating the direction of gaze. The contour of the pupil and iris, which is a circle in three dimensions, takes the form of an ellipse when projected on a 2D image plane. An ellipse fitted algorithmically to the pupil and/or iris can, in turn, be used to reconstruct the original sphere in three dimensions, which allows the orientation of the eyeball to be calculated and the gaze direction to be predicted (Chen & Ji, 2008; Valenti, Sebe, & Gevers, 2012). Although model-based video-oculography methods have recently been applied in more practical scenarios (Cristina & Camilleri, 2016; Funes Mora & Odobez, 2014; Wood & Bulling, 2014), their gaze estimation accuracy is still low, since they depend on accurate eye feature detection for which high-resolution images and homogeneous bright illumination are required. This has largely prevented these methods from being widely used in real-world settings or on commodity devices.

In contrast, learning-based gaze estimation methods do not rely on explicit detection of eye features but directly map the pixel information contained in images obtained from the user to 3D gaze directions using machine learning algorithms. Because they do not rely on explicit eye feature detection, learning-based methods can handle low-resolution images and longer distances from the object of gaze. While early methods assumed a fixed head pose, more recent methods allow for free 3D head movement in front of the camera (Gao, Harari, Tenenbaum, & Ullman, 2014). An open research challenge in learning-based gaze estimation is to train gaze estimators that make minimal assumptions regarding the user, environment, or camera.

The need to collect person-specific training data represents a fundamental limitation for both model-based and learning-based gaze estimation methods. To reduce the burden on the user, several previous works used events that can be observed when the user interacts with computing systems, such as mouse clicks or key presses, as a proxy for users' on-screen gaze position. Alternatively, visual saliency maps (Sugano & Bulling, 2015) or pre-recorded human gaze patterns on defined visual stimuli (Alnajar, Gevers, Valenti, & Ghebreab, 2013) can be used as training data to learn the gaze estimation function. However, the need to acquire user input fundamentally limits the extent to which these approaches can be applied to interactive settings. Thus, another line of work aims to train gaze estimators that generalize to arbitrary users without requiring explicit user input (Funes Mora & Odobez, 2013).

Despite significant advances in such person-independent gaze estimation, all of these previous works only considered gaze estimation tasks in which training and test data are assumed to come from the same respondents. Zhang, Sugano, Fritz, and Bulling (2015, 2017, 2019) were first to study the practically most relevant but also significantly more challenging task of unconstrained gaze estimation via cross-dataset evaluation. They introduced a method based on a multimodal deep convolutional neural network that outperformed the state-of-the-art methods by a large margin. These latest methods were also shown to have significant potential for other tasks, such as estimation of audience attention (Sugano et al., 2016) or detection of eye contact in everyday settings (Zhang, Sugano, & Bulling, 2017). Later works demonstrated that large-scale methods for unconstrained gaze estimation can benefit from advances in computer graphics techniques for eye region modelling. These models can be used to synthesize large amounts of highly realistic and perfectly annotated eye region images, thereby significantly reducing both data collection and annotation efforts (Wood et al., 2015). The latest model is fully morphable and can synthesize more than 40 eye images per second on commodity hardware (Wood, Baltrušaitis, Morency, Robinson, & Bulling, 2016).

The cameras used for video-based eye-tracking typically capture video images at a frame rate of between 30 and 120 Hz. Therefore, pervasive eye trackers currently provide lower spatio-temporal precision than PCR-based systems, and face additional challenges with respect to accurate detection of fast eye movements, so-called saccades (Rayner, 1998). Studies employing pervasive eye-tracking therefore mostly report dwell times on manually defined areas of interest. Burton, Albert, and Flynn (2014) compared video-based with PCR-based eye-tracking and found that the former is less accurate especially for smaller regions of interest (around 1% of the screen or less), even more when they are located in the periphery of the computer screen. Video-based eye-tracking may underestimate dwell time by as much as 50% for these smaller areas of interest. For larger areas of interest that comprise of 5% of the screen or more, dwell times may be underestimated by about 25%. Video-based technology, however, may realize accuracies comparable to IR eye-tracking when interest focuses on hit rates, that is, percentages of participants who fixated at least once on a larger area of interest. The advantages of video-based eye-tracking are the very low cost, the possibility of eye-tracking in natural settings (at home, at work, or any other location where respondents are in front of a desktop or laptop computer), and across dispersed geographic locations. The lower spatial and temporal precision may partially be offset by using much larger samples of participants.

### *Information That Pervasive Eye-Tracking Systems Provide*

As mentioned before, eye-tracking provides a plethora of information about the user and has, consequently, been used for a long time as both a measurement technique and input modality. Arguably, the two most important applications of eye-tracking, so far, are in the analysis of visual attention and behavior in human vision research, and for gaze-based human-computer interaction in computer science (see Figure 2.2).

Analysis of visual attention in human vision research has traditionally focused on analyzing the deployment of gaze to different areas of interest (AOIs) on a defined stimulus, e.g., a natural image, visual pattern, or website, displayed to the user on a computer monitor (see Chapters 4–5). The AOIs can be either content-based (face, text, image, object, etc.) or space-based (grid, image pixels). The analysis either involves statistically testing for differences in eye movement characteristics, or aggregating and visualizing fixations in fixation density maps (see Figure 2.3) or other graphical displays (Holmqvist et al., 2011). All of these analyses are readily provided by commercial software shipped together with the eye trackers or free software downloadable from the web. If the analysis is done for a large number of users, robust measures of potential differences in visual attention towards a given stimulus for two or more user groups can be obtained. The eye movement characteristics commonly used for statistical analysis are the average fixation duration, the total gaze duration, time until the first fixation, or the total number of fixations on a set of given AOIs (Holmqvist et al., 2011). See Pieters and Wedel (2004) for an example of this approach.

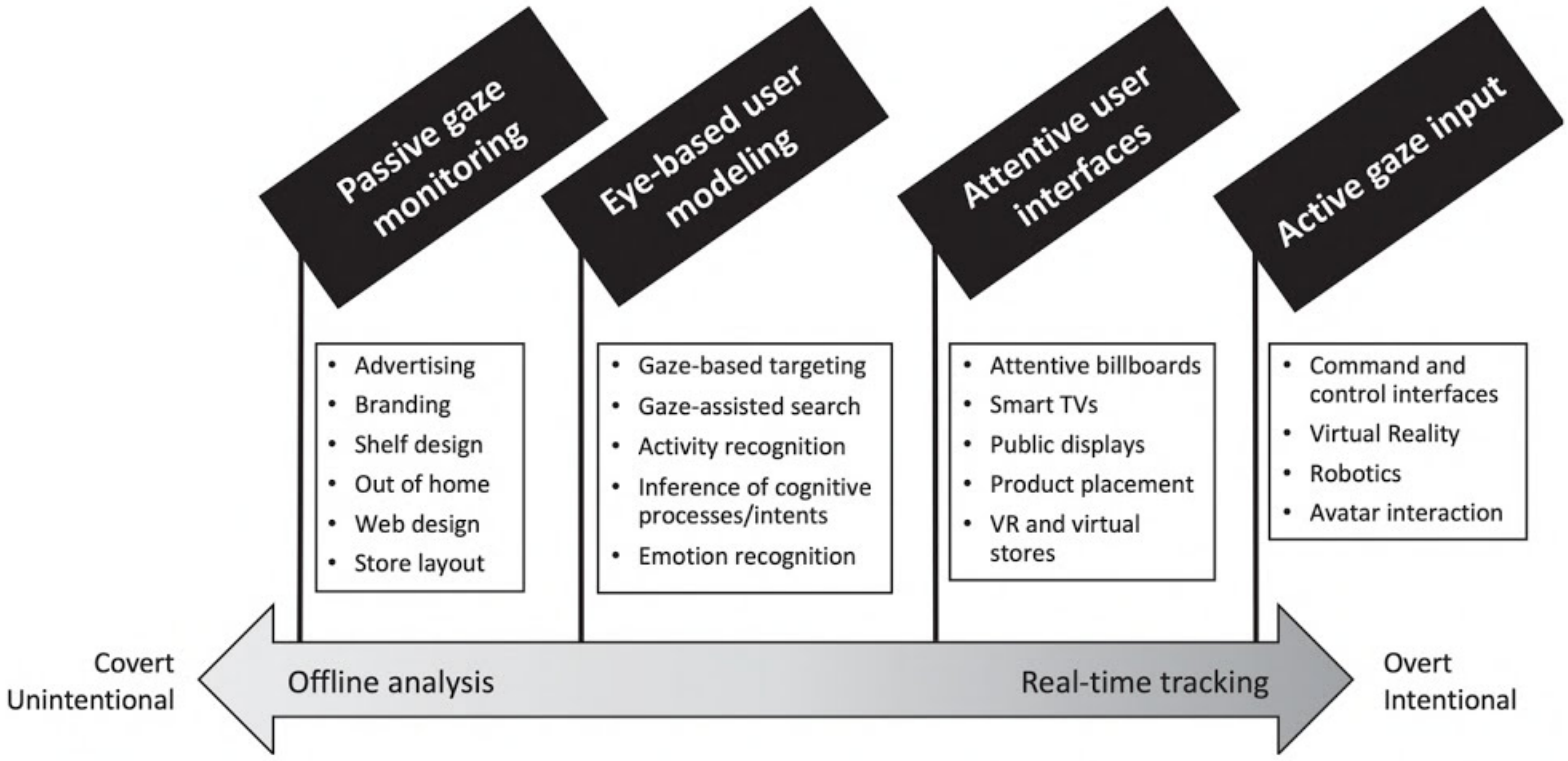


**FIGURE 2.2** Spectrum of possible eye-tracking applications and use of gaze information in consumer behavior analysis and decision making (inspired by Majaranta & Bulling, 2014; Author generated).

Statistical testing provides a principled and well-established way of identifying differences in attentive behavior between AOIs. One drawback of this approach, however, is that temporal information is not considered. Scan-path analysis was devised to address this limitation and thus provides a complementary form of attention analysis (Noton & Stark, 1971; Pieters, Rosbergen, & Wedel, 1999). Scan-paths are sequences of multiple fixations on a stimulus for a given amount of time, typically no more than a few seconds. To calculate scan-paths, fixation sequences are first encoded, averaged per user if desired, and then programmatically or statistically compared to the scan-paths of other users or experimental conditions. An application to decision-making tasks, for example, was provided by Day (2010). The programmatic comparison involves calculating how similar different scan-paths are, for example using normalized scan-path similarity (Le Meur & Baccino, 2013). In addition to retaining temporal information of fixation sequences, scan-paths can also be visualized easily by overlaying the fixations and their connections with lines onto the stimulus (see Figure 2.3). The key drawbacks of this type of scan-path analysis are that visualization does not scale well to a large number of scan-paths and users and scan-path visualizations become cluttered very quickly. An alternative approach is a statistical approach to scan-path analysis, which may involve either Markov (Pieters, Rosbergen, & Wedel, 1999), or Hidden Markov Models (Liechty, Pieters, & Wedel, 2003) to describe first-order transitions between AOIs. Finally, if neither difference between individual eye movement characteristics nor temporal information of fixation sequences are desired, fixation density maps provide yet another means to summarize the eye movement data (Holmqvist et al., 2011). Further, statistical models that predict probability distributions of fixations from low-level image characteristics can be used to estimate saliency maps that display regions in the stimulus that are salient to each viewer (van der Lans, Pieters, & Wedel, 2008).

## *Gaze-Based Interaction*

In human-computer interaction, gaze has a long history as a means for hands-free interaction with computing systems (Majaranta & Bulling, 2014; Sibert & Jacob, 2000). Gaze has, for example,

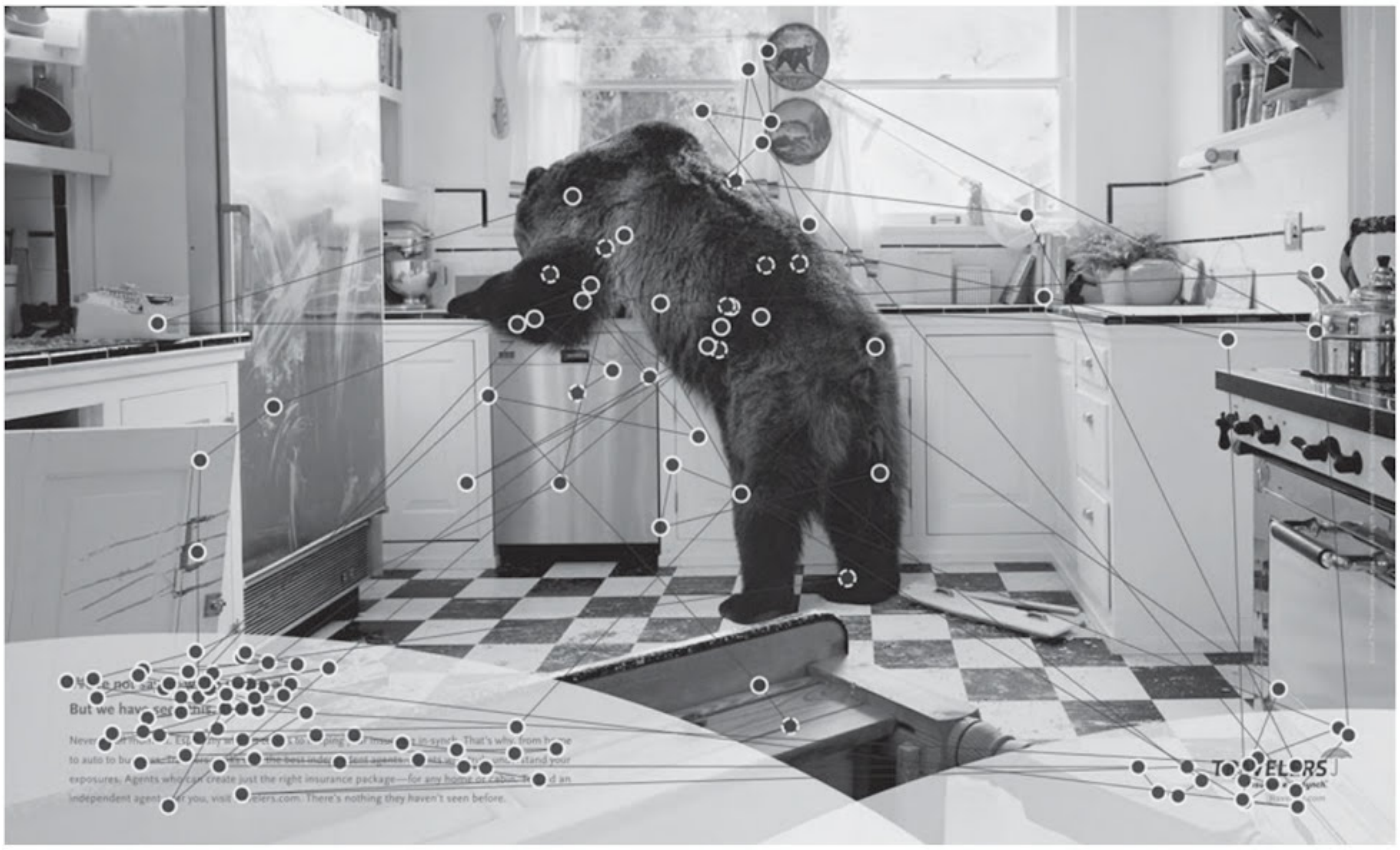

**FIGURE 2.3** Eye movements of multiple participants on advertisement for Traveler (reprinted with permission of MIT Press; Fawcett, Jonathan M., Evan F. Risko, and Alan Kingstone, (eds.) *The Handbook of Attention*, Figure 25.1, p. 571, © 2015 Massachusetts Institute of Technology, published by the MIT Press.).

been used for fast, accurate, and natural interaction with ambient and body-worn devices and displays (Esteves, Velloso, Bulling, & Gellersen, 2015; Huang, Li, Ngai, & Leong, 2017; Vaitukaitis & Bulling, 2012; Wood & Bulling, 2014) for a variety of tasks including, but not limited to, pointing (Zhai, Morimoto, & Ihde, 1999), object selection and transfer (Sibert & Jacob, 2000; Stellmach & Dachselt, 2012; Turner, Alexander, Bulling, Schmidt, & Gellersen, 2013; Vidal, Bulling, & Gellersen, 2013; Zhang, Bulling, & Gellersen, 2013; Zhang, Müller, Chong, Bulling, & Gellersen, 2014), or text entry (Majaranta & Räihä, 2002). Prior work also investigated means to combine gaze input with other modalities, such as touch (Simeone, Bulling, Alexander, & Gellersen, 2016; Stellmach & Dachselt, 2012; Turner, Alexander, Bulling, & Gellersen, 2015), mouse and keyboard input (Kumar, Paepcke, & Winograd, 2007) or mid-air gestures (Velloso, Turner, Alexander, Bulling, & Gellersen, 2015). Due to the aforementioned prior limitations of eye-tracking, much of this work was done in desktop settings or, in general, settings in which users moved relatively little in front of the display. Complementing these active uses of gaze for interaction is a relatively large body of work on attentive user interfaces (Bulling, 2016; Vertegaal, 2003; Xu, Sugano, & Bulling, 2016), i.e., interfaces that monitor user attention passively and adapt to users' current attentional capacity and state in different ways. These attentive user interfaces are increasingly explored in everyday settings, in particular on public displays (Alt, Bulling, Mecke, & Buschek, 2016; Khamis, Alt, & Bulling, 2016; Khamis, Bulling, & Alt, 2015; Sugano et al., 2016; Walter, Bulling, Lindlbauer, Schüssler, & Müller, 2015).

The key underlying measure of the vast majority of all of these works is the 2D point of gaze on the display itself or, in cases in which 3D information can be related to specific objects of interests, 3D gaze direction measured using both mobile and remote eye trackers. More recently, a new line of work has started to explore machine learning approaches on top of the "raw" gaze data, i.e., methods to encode gaze into rich higher-level representations that can subsequently be linked to user behavior and, thus, be used for implicit human-computer interaction (Majaranta & Bulling, 2014). A growing body of work has demonstrated that, for example, spatio-temporal information on gaze can be used to automatically predict users' everyday activities (Bulling, Ward, & Gellersen, 2012; Bulling et al., 2011; Kunze, Bulling, Utsumi, Yuki, & Kise, 2013), also in an unsupervised fashion during full-day mobile gaze recordings (Bulling et al., 2013; Steil & Bulling, 2015), cognitive processes and states (Bulling & Roggen, 2011; Bulling & Zander, 2014; Tessendorf et al., 2011), intentions and goals (Bednarik et al., 2012), social interactions (Pfeiffer, Vogeley, & Schilbach, 2013), or even aspects of users' personality and decision-making processes (Hoppe et al., 2015).

All of these analyses rely on the 2D point of gaze on a stimulus or the 3D gaze direction as input. In addition, eye trackers may also provide other information about the user. For example, the videos of the users' faces that video-based eye trackers obtain can serve as input to emotion recognition (Cohen, Sebe, Garg, Chen, & Huang, 2003). Although not directly eye-tracking, this is a rich source of auxiliary data that may provide important information on users' underlying emotional states (see Chapter 23).

In addition, eye trackers also provide measurements of the pupil diameter (see Chapter 3), which may depend on the cognitive load and/or arousal of the respondent. The pupil tends to dilate when users are aroused or deploy more cognitive resources to process the information (Bradley, Miccoli, Escrig, & Lang, 2008). Pupil size also depends on other factors, however, including lightning conditions, because of which pupil diameter is not an unambiguous indicator of cognitive load or arousal (Loewenfeld, 1993).

Further, an image of the scene that the person is looking at can be reflected on the cornea and may thus be obtained from the corneal image recorded using mobile eye trackers (Nakazawa & Nitschke, 2012). The cornea reflects not only the incoming light, but also the entire surrounding scene over a wide field of view. The corneal reflection itself allows for the analysis of the entire field of view. It enables the reconstruction and analysis of the scene and the 3D environment of the viewer (Nishino & Nayar, 2006), and eliminates the need for a separate camera to capture the users' field of view, thus further enabling miniaturization of pervasive eye-tracking.

Other measures that are obtained as a corollary of eye-tracking are micro-saccades, blinks and vergence movements. First, micro-saccades are very small, involuntary movements of the eyes (less than 1° of visual angle) that have been shown to be associated with attentional load, onset of new or oddball visual stimuli, and the preparation of motor response (Engbert, 2006; Pastukhov & Braun, 2010; Rolfs, Kliegl, & Engbert, 2008). Despite their potential, as of yet, micro-saccades are difficult to record in everyday settings and require highly accurate and high-speed PCR-based eye trackers. Second, eye blinks are recorded as an interruption of the corneal reflection in PCR-based eye trackers and can be detected with video-based trackers (Grauman, Betke, Gips, & Bradski, 2001). An increase in blink rate, i.e., the number of blinks for a particular time duration, is associated with higher levels of arousal (Bradley, Codispoti, Cuthbert, & Lang, 2001), while a decrease can be observed during attentional focus, high cognitive load (Stern, Walrath, & Goldstein, 1984), or increased drowsiness (Caffier, Erdmann, & Ullsperger, 2003). Third, vergence eye movements are movements where both eyes turn inward or outward, in order to keep an object that moves towards or from us in focus. Vergence eye movements can thus be used to determine the distance of a visually attended object (Choi, Jung, Ban, Niitsuma, & Lee, 2006) or, if controlled voluntarily, as a means for user input (Kirst & Bulling, 2016).

All of the auxiliary measures discussed earlier provide additional information on users' underlying cognitive states and can be used together with users' gaze to model underlying cognitive states and/or traits. It may additionally enable the optimization of visual design and user interfaces, filling the void in areas such as optimization of sponsored search, movie and video clips, banner ads, product reviews, text and image search, and product comparison layouts (Pieters, Wedel, & Zhang, 2007). Bayesian models can be used to represent cognition from first principles and make predictions on how multiple unobserved attentional processes may have affected the recorded eye movements. This enables inferences on multiple underlying cognitive processes from eye movement data, and has been shown to result in accurate forecasts of downstream behavior such as memory (Wedel & Pieters, 2000), search (van der Lans et al., 2008), consideration (Chandon, Hutchinson, Bradlow, & Young, 2009), choice (Stüttgen et al., 2012), and even sales (Zhang et al., 2009) from eye movements.

## Pervasive Eye-Tracking Applications in Consumer Behavior Analysis

Most recent mobile eye trackers that rely on video cameras can be implemented as lightweight and fully embedded mobile systems and therefore have become suitable for recordings in everyday settings (see Figure 2.1 for an example). Such systems now also allow, for the first time, to record gaze over long periods of time, e.g., over a full day of a person's life (Bulling et al., 2013; Steil & Bulling, 2015). The low cost of this new generation of eye-tracking systems, easy calibration, and unobtrusive measurement in natural exposure conditions are beginning to contribute to the growth of applications in practice, and theory development and testing in academic research.

Consequently, several recent works have started to explore the use of mobile eye-tracking to analyze consumer behavior in these and other natural everyday settings.

Classic work using PCR-based mobile eye-tracking (Land & Hayhoe, 2001; Land, Mennie, & Rusted, 1999; Smeets, Hayhoe, & Ballard, 1996) showed that routine goal-directed activities require continuous monitoring with the eyes and have revealed a tight linkage between eye movements and the motor actions that are performed. Eye movements during these tasks are thus mostly directed top-down towards task-relevant objects. The pioneering work of Yarbus had already shown this early on for static contexts (Yarbus, 1967). The eyes usually reach an object before any hand action towards the object, and move on to the next object before the preceding action is completed. A shift of the eyes is often followed by a movement of the head, which is followed by the movement of the hand. Thus, research with mobile eye trackers has shown that eye movements are a fundamental component of the motor pattern and are leading indicators of goal-directed motion.

Using mobile eye-tracking, research has investigated human performance in real-world tasks such as driving (Shinoda, Hayhoe, & Shrivastava, 2001), making tea (Land, Mennie, & Rusted, 1999), walking (Jovancevic-Misic & Hayhoe, 2009), and playing sports (Vickers, 2006, Vickers & Adolphe, 1997). As for consumer decision making, mobile eye trackers have been used to assess the effectiveness of in-store merchandising (Hendrickson & Ailawadi, 2014). Research with several hundreds of shoppers in multiple stores demonstrated that shoppers 1) look in a narrow window below and above eye level and as a consequence especially signage placed on the ceiling in stores is hardly noticed; 2) look at signage for about a second on average and process 3–5 words; 3) look at signage only when it is immediately relevant for and in close proximity to the shopping goal; and 4) process information on signage in the store in a left–right or top–bottom direction. Other research, using content analysis of data produced by mobile eye trackers (Harwood & Jones, 2014), has confirmed that 75–85% of fixations that shoppers make in a store fall on products, while signage receives a much lower number of fixations (0.5%). This research also revealed that the (vertical) line of sight and visual salience (brightness and color contrast) are two main factors affecting store navigation. A study on digital out-of-home advertising in public transport revealed that over

60% of participants looked at the digital screens placed in a tram during a 30-minute tram ride, fixating on it 16% of the time the screen was in their field of view (Höller, Schrammel, Tscheligi, & Paletta, 2009). While they are merely scratching the surface of applications of mobile eye-tracking, these studies illustrate how research in retail and out-of-home settings benefits from mobile pervasive eye-tracking and yields insights that would be difficult to obtain otherwise. Another area of research that can benefit from the application of mobile eye-tracking is research into multitasking and multi-screen behaviors. One study has shown that viewers preferentially attend to computer screens as compared to TV screens during media multitasking, which is manifested in longer gaze on the computer screen (Brasel & Gips, 2011). But, gaze times on both screens are limited to a few seconds only, and people switch between screens around four times per minute. Again, these insights would be very difficult to obtain without the use of pervasive eye-tracking.

Ever since the work of Yarbus (1967) it is evident that eye movements are dependent on the tasks and goals of consumers. It is increasingly recognized that decision making is embedded in perception-action cycles, and that attention plays an active role in constructing decisions (Orquin & Loose, 2013). However, most often decision making has been studied in isolation from the perception-action cycle in which it naturally occurs. A major step forward was made when recent research investigated eye movements during decision making (e.g., Glaholt & Reingold, 2011; Krajbich, Armel, & Rangel, 2010; Pieters & Warlop, 1999; Shi, Wedel, & Pieters, 2013; Shimojo, Simion, Shimojo, & Scheier, 2003), because it revealed the role of attention in decision making. This stream of research was initiated by the work of Russo and colleagues (Russo & Leclerc, 1994; Russo & Rosen, 1975) who demonstrated not only that eye movements were used to acquire information, but also how that information was used. This has been recently formalized in statistical models that describe eye movements and decisions jointly (Stüttgen et al., 2012; Yang, Toubia, & de Jong, 2015). However, in that research eye movements are still collected in a lab setting rather than a real-life decision context. Pervasive eye-tracking systems will be needed to study decision making embedded in the perception-action cycle in natural contexts. A first attempt was made by Gidlöf, Wallin, Dewhurst, and Holmqvist (2013), who extended the work by Russo and Leclerc (1994) to real-world settings, and revealed a deeper processing of the decision alternatives in the evaluation stage, as compared to lab settings.

## Emerging Applications

In a few years from now, eye-tracking will likely be an integrated part of our lives, via camera-based gaze estimation incorporated in laptops and desktop computers, billboards, kiosks, smart-TVs, tablets, smartphones, and so on. The incorporation of gaze recording in everyday digital devices will help make our daily lives simpler, safer, more efficient, and more enjoyable. It may also enable the optimization of visual design and user interfaces, filling the void in areas such as optimization of sponsored search, movie and video clips, banner ads, product reviews, text and image search, and product comparison layouts (Pieters et al., 2007). The rapid development of recording technology has already begun to see innovative application and is likely to create many more opportunities (see Figure 2.2). For example, eye-tracking systems that record what users look at on their digital screens may provide hands-free access to information via gaze control, thereby facilitating interaction with electronic devices. Today, users can already deploy their gaze to activate apps, scroll through web pages, and make a selection among options by fixating on one of them, thus improving their user experience. In combination with automated image analysis, pervasive eye-tracking may allow for automatic alerts if important information is overlooked, and may use visual cues to support or even interactively direct visual search (Sattar, Müller, Fritz, & Bulling, 2015; Wedel, Yan, Siegel, & Li, 2016). As more gaze data are collected and being used for marketing purposes, privacy and security become critical issues. Privacy laws have not kept pace with data collection and processing technologies. Several governments have been enacting stricter privacy laws. But, because their gaze, emotion, and other process tracing data is considered sensitive by most consumers, respecting customers' privacy is good business practice.

In reading, selective blurring of text may improve reading speed and focus. Explanations may pop-up when eye movements indicate comprehension is slow. Much of this was already implemented in Text 2.0, which is a framework for developing web-based eye-tracking applications to facilitate interactive reading (Biedert, Buscher, Schwarz, Hees, & Dengel, 2010). Also, words and sentences that a viewer looks at longer may be included in document summaries, which can be optimized to reflect a reader's personal interests and used to develop recommendation systems that recommend new articles, texts, or reviews.

*Pervasive eye-tracking* will also render computer games more immersive. Waterloo Labs allows players the use of eye movements to control EyeMario (http://waterloolabs.blogspot.com/), and Formula Face (http://games.redbull.com/int/en/game/formula-face) allows gamers to use blinks, smiles, and head movements to control their game. As a player moves her head, expresses emotions, or blinks, the head and face of the avatar moves synchronously by mirroring these movements and emotions. Thus, eye, head, and facial movements are recorded and analyzed in real time to render the avatar's movements and expressions more realistic and more in tune with the user's moment-to-moment feelings, producing games that are more immersive and appealing. Optitrack's (www.naturalpoint.com/optitrack/) face capture system already accomplishes much of that.

While eye-tracking has traditionally been used to measure people's visual attention, it is increasingly becoming clear that when used along with additional measures that can be extracted from the facial images, including pupil dilation (see Chapter 3), blinks, micro-saccades, emotion expressions, scene reflection, head movements, and body movements, a much richer picture of the behavioral and cognitive state of the viewer can be obtained, including attention, location and environment, emotions, goals and intentions, activities, and social interactions. Inferring such states presents opportunities of improving visual user interfaces in a large variety of everyday contexts (Bulling, 2016; Bulling & Zander, 2014).

New virtual and augmented reality applications may also benefit greatly from routine eye movement recording. Increasingly, companies are exploiting these technologies. Augmented reality applications allow consumers to see information that is important to them, but not present in the real-world context, by overlaying digital information on top of real-world settings. Virtual reality (VR) and augmented reality (AR) allow consumers to use their digital device to try on new clothes, explore a home, hotel or museum, visit or fly over new cities and countries, or test drive a car. VR and AR are already used in an increasing number of applications in advertising and selling. For example, for virtual test driving of cars, the car manufacturer Volvo has developed smartphone applications (www.volvocars.com/intl). There are also an increasing number of applications in hospitality and travel, through which prospective customers can virtually explore hotels and hotel rooms, museums, entertainment options, and tourist destinations via 360-degree views on their smartphones. The hotel chain Mariott has developed in-room VR travel applications (http://marriott-hotels.marriott.com/). Gaming is on the forefront of virtual and augmented reality applications, with Pokemon Go as a prime example (www.pokemongo.com/), but with current developments going already beyond that. The Chinese retailer Yihaodian has developed mobile phone applications that enable its customers to browse and shop in virtual stores at any location using their smartphones (https://en.wikipedia.org/wiki/Yihaodian). Eye-tracking has already been integrated with VR technology (Pfeiffer, 2008), where it enables rendering of the virtual context based on the real-time field of view and depth of field blur. It uses the viewer's gaze to analyze what he looks at inside the virtual space with the purposes of making the aim of actions more accurate, making the virtual environment more immersive, and the interactions with actors in it more life-like (Hillaire, Lécuyer, Cozot, & Casiez, 2008).

## Open Challenges

Mobile eye trackers are ideal for tracking eye movements during everyday activities in natural settings, where head, hand, and body movements need to be unconstrained. One challenge with the application of mobile eye trackers for research purposes is the analysis of data from multiple participants, because each participant has an idiosyncratic field of view at each point in time during the recording. The data resulting from pervasive eye-tracking essentially consists of movie clips of the visual field of the user, on which the gaze point is indicated. The degree of heterogeneity of the data can be extensive, each respondent having his/her own field of view at each point in time during the study. Computer vision methods are needed to process the individual data streams and aggregate them to enable the application of statistical methods that facilitate generalizable conclusions. Progress has been made in off-the-shelf software to identify and track AOIs across multiple videos, which makes quantitative analyses feasible especially when researchers have well-defined ideas about the objects and regions that are of interest in the analysis.

Another challenge is that mapping gaze to the 3D environment requires visual markers that have to be placed and detected in real time in the environment. Alternatively, sophisticated computer vision algorithms are required to detect and track objects such as displays or, in general, areas of interest in the egocentric video (Lander, Gehring, Krüger, Boring, & Bulling, 2015). Further, fully invisible integration of the eye tracker into ordinary glasses is not yet feasible due to the rather large imaging sensors currently used. The design of current mobile eye trackers can lead to low social acceptance and was shown to result in unnatural behavior of both the wearers and people they interact with (Nasiopoulos, Risko, Foulsham, & Kingstone, 2015).

## Conclusion

Over the past decade, eye movement research has increasingly relied on the integration of techniques and theories from visual computing, attention research, and in-lab eye-tracking—fields that have been relatively disparate before. Visual computing has developed powerful tools that enable the extraction of basic visual features from images, segment and describe images, and recognize forms, shapes, faces, and large numbers of object classes. Attention research offers theories that explain eye movements from underlying cognitive processes while laboratory eye-tracking experiments offer methods of recording, analyzing, and interpreting eye, face, and head movements. Combined, these fields have provided unprecedented insights into people's processing of, evaluation of, and behavior towards controlled visual stimuli.

Recent advances in mobile eye-tracking as well as stationary eye-tracking using video cameras readily integrated into handheld devices and ambient displays pave the way for a new generation of pervasive eye-tracking systems that allow researchers and practitioners to understand and analyze gaze information in real-world settings. As such, pervasive eye-tracking has not only significant potential to validate and complement existing theories and findings in the previous research areas but also to uncover entirely new behavioral, cognitive, and attention phenomena and enable new applications impossible before. The ramifications of this imminent paradigm shift are transformative, in particular for applications in consumer behavior analysis and decision making in offline, online, mobile, and VR settings. With pervasive eye-tracking and analysis of gaze behavior and facial expressions becoming a commodity, gaze will provide a unique source of information on the "inner workings" of consumers. Moreover, these measures will increasingly provide input that will be used to shape our digital environment.

## Recommended Reading List

- Land (2006): Early research on eye movements in everyday life
- Rayner (1998): A classic review of eye-tracking research in reading
- Wedel and Pieters (2008): A review of eye-tracking and visual attention research in marketing
- Majaranta and Bulling (2014): An introduction to eye-tracking and eye-based human-computer interaction

## References


Adams Jr., R. B., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science*, *14*(6), 644–647.

Alnajar, F., Gevers, T., Valenti, R., & Ghebreab, S. (2013). Calibration-free gaze estimation using human gaze patterns. In *Proceedings IEEE International Conference on Computer Vision* (pp. 137–144).

Alt, F., Bulling, A., Mecke, L., & Buschek, D. (2016). Attention, please! Comparing Features for Measuring Audience Attention Towards Pervasive Displays. In *Proc. ACM SIGCHI Conference on Designing Interactive Systems (DIS)* (pp. 823–828).

Bednarik, R., Vrzakova, H., & Hradis, M. (2012). What do you want to do next: A novel approach for intent prediction in gaze-based interaction. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 83–90). ACM.

Biedert, R., Buscher, G., Schwarz, S., Hees, J., & Dengel, A. (2010). Text 2.0. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 4003–4008). ACM.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607.

Brasel, S. A., & Gips, J. (2011). Media multitasking behavior: Concurrent television and computer usage. *Cyberpsychology, Behavior, and Social Networking*, *14*(9), 527–534.

Bulling, A. (2016). Pervasive attentive user interfaces. *IEEE Computer*, *49*(1), 94–98.

Bulling, A., & Gellersen, H. (2010). Toward mobile eye-based human-computer interaction. *IEEE Pervasive Computing*, *9*(4), 8–12.

Bulling, A., & Roggen, D. (2011). Recognition of visual memory recall processes using eye movement analysis. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (pp. 455–464). ACM.

Bulling, A., Ward, J. A., & Gellersen, H. (2012). Multimodal recognition of reading activity in transit using body-worn sensors. *ACM Transactions on Applied Perception*, *9*(1), 2:1–2:21.

Bulling, A., Ward, J. A., Gellersen, H., & Tröster, G. (2011). Eye movement analysis for activity recognition using electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(4), 741–753.

Bulling, A., Weichel, C., & Gellersen, H. (2013). Eyecontext: Recognition of high-level contextual cues from human visual behaviour. In *Proceedings ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)* (pp. 305–308).

Bulling, A., & Zander, T. O. (2014). Cognition-aware computing. *IEEE Pervasive Computing*, *13*(3), 80–83.

Burton, L., Albert, W., & Flynn, M. (2014). A comparison of the performance of webcam vs. infrared eye tracking technology. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 58, pp. 1437–1441). Los Angeles, CA: SAGE Publications.

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, *89*(3–4), 319–325.

Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing*, *73*(6), 1–17.

Chen, J., & Ji, Q. (2008). 3D gaze estimation with a single camera without IR illumination. In *International Conference on Pattern Recognition (ICPR)* (pp. 1–4).

Choi, S. B., Jung, B. S., Ban, S. W., Niitsuma, H., & Lee, M. (2006). Biologically motivated vergence control system using human-like selective attention model. *Neurocomputing*, *69*(4), 537–558.

Cohen, I., Sebe, N., Garg, A., Chen, L. S., & Huang, T. S. (2003). Facial expression recognition from video sequences: Temporal and static modeling. *Computer Vision and Image Understanding*, *91*(1), 160–187.

Cristina, S., & Camilleri, K. P. (2016). Model-based head pose-free gaze estimation for assistive communication. *Computer Vision and Image Understanding*, *149*, 157–170.

Day, R. F. (2010). Examining the validity of the Needleman–Wunsch algorithm in identifying decision strategy with eye-movement data. *Decision Support Systems*, *49*(4), 396–403.

Dodge, R. (1900). Visual perception during eye movement. *Psychological Review*, 7(5), 454.

Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and Practice* (Vol. 373). London, UK: Springer.

Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. *Progress in Brain Research*, *154*, 177–192.

Esteves, A., Velloso, E., Bulling, A., & Gellersen, H. (2015). Orbits: Enabling gaze interaction in smart watches using moving targets. In *Proceedings ACM Symposium on User Interface Software and Technology (UIST)* (pp. 457–466).

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*(17), 1920–1931.

Funes Mora, K. A., & Odobez, J. M. (2013). Person independent 3D gaze estimation from remote RGB-D cameras. In *Proceedings IEEE Int. Conf. on Image Processing* (pp. 2787–2791).

Funes Mora, K. A., & Odobez, J. M. (2014). Geometric generative gaze estimation (G3E) for remote RGB-D cameras. In *Proceedings IEEE Conf. Computer Vision and Pattern Recognition* (pp. 1773–1780).

Gao, T., Harari, D., Tenenbaum, J., & Ullman, S. (2014). *When Computer Vision Gazes at Cognition*. Cambridge, MA: Center for Brains, Minds, & Machines.

Gidlöf, K., Wallin, A., Dewhurst, R., & Holmqvist, K. (2013). Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment. *Journal of Eye Movement Research*, *6*(1), 1–14.

Glaholt, M. G., & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 125–146.

Grauman, K., Betke, M., Gips, J., & Bradski, G. R. (2001). Communication via eye blinks-detection and duration analysis in real time. In *Proceedings 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)* (Vol. 1, pp. 1010–1017). IEEE.

Hansen, D. W., & Ji, Q. (2010). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 478–500.

Harwood, T., & Jones, M. (2014). Mobile eye-tracking in retail research. In M. Horsley, M. Eliot, B. A. Knight, & R. Reilly (Eds.), *Current Trends in Eye Tracking Research* (pp. 183–199). London, UK: Springer.

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

Hendrickson, K., & Ailawadi, K. L. (2014). Six lessons for in-store marketing from six years of mobile eye-tracking research. In D. Grewal, A. L. Roggeveen, & J. Nordfält (Eds.), *Shopper Marketing and the Role of In-Store Marketing* (pp. 57–74). Bingley, UK: Emerald Group Publishing Limited.

Hillaire, S., Lécuyer, A., Cozot, R., & Casiez, G. (2008). Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. In *Virtual Reality Conference, 2008. VR'08. IEEE* (pp. 47–50). IEEE.

Höller, N., Schrammel, J., Tscheligi, M., & Paletta, L. (2009). The perception of information and advertisement screens mounted in public transportation vehicles-results from a mobile eye-tracking study. In *Proceedings of the 2nd Workshop on Pervasive Advertising* (Vol. 39).

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, UK: Oxford University Press.

Hoppe, S., Loetscher, T., Morey, S., & Bulling, A. (2015). Recognition of curiosity using eye movement analysis. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers* (pp. 185–188). ACM.

Huang, M. X., Li, J., Ngai, G., & Leong, H. V. (2017). Screenglint: Practical, in-situ gaze estimation on smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 2546–2557). ACM.

Huey, E. B. (1898). Preliminary experiments in the physiology and psychology of Reading. *American Journal of Psychology*, *9*(4), 575–586.

Jovancevic-Misic, J., & Hayhoe, M. (2009). Adaptive gaze control in natural environments. *Journal of Neuroscience*, *29*(19), 6234–6238.

Khamis, M., Alt, F., & Bulling, A. (2016). Challenges and design space of gaze-enabled public displays. In *Adjunct Proceedings ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)* (pp. 1736–1745).

Khamis, M., Bulling, A., & Alt, F. (2015). Tackling challenges of interactive public displays using gaze. In *Adjunct Proceedings ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)* (pp. 763–766).

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive ethology: A new approach for studying human cognition. *British Journal of Psychology*, *99*(3), 317–340.

Kirst, D., & Bulling, A. (2016). On the verge: Voluntary convergences for accurate and precise timing of gaze input. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human factors in Computing Systems* (pp. 1519–1525). ACM.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298.

Kumar, M., Paepcke, A., & Winograd, T. (2007). EyePoint: practical pointing and selection using gaze and keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 421–430). ACM.

Kunze, K., Bulling, A., Utsumi, Y., Yuki, S., & Kise, K. (2013). I know what you are reading: Recognition of document types using mobile eye tracking. In *Proceedings IEEE International Symposium on Wearable Computers (ISWC)* (pp. 113–116). ACM.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, *25*(3), 296–324.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25), 3559–3565.

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, *28*(11), 1311–1328.

Lander, C., Gehring, S., Krüger, A., Boring, S., & Bulling, A. (2015). Gazeprojector: Accurate gaze estimation and seamless gaze interaction across multiple displays. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (pp. 395–404). ACM.

Le Meur, O., & Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behavior Research Methods*, *45*(1), 251–266.

Liechty, J., Pieters, R., & Wedel, M. (2003). Global and local covert visual attention: Evidence from a Bayesian hidden Markov model. *Psychometrika*, *68*(4), 519–541.

Loewenfeld, I. E. (1993). *The Pupil: Anatomy, Physiology, and Clinical Applications*. Detroit, MI: Wayne State University Press.

Majaranta, P., & Bulling, A. (2014). Eye tracking and eye-based human–computer interaction. In *Advances in Physiological Computing* (pp. 39–65). London, UK: Springer.

Majaranta, P., & Räihä, K.-J. (2002). Twenty years of eye typing: Systems and design issues. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications* (pp. 15–22). ACM.

Marius't Hart, B., Vockeroth, J., Schumann, F., Bartl, K., Schneider, E., Koenig, P., & Einhäuser, W. (2009). Gaze allocation in natural stimuli: Comparing free exploration to head-fixed viewing conditions. *Visual Cognition*, *17*(6–7), 1132–1158.

Nakazawa, A., & Nitschke, C. (2012). Point of gaze estimation through corneal surface reflection in an active illumination environment. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), *Computer Vision – ECCV 2012* (pp. 159–172). Berlin/Heidelberg: Springer.

Nasiopoulos, E., Risko, E. F., Foulsham, T., & Kingstone, A. (2015). Wearable computing: Will it make people prosocial? *British Journal of Psychology*, *106*(2), 209–216.

Nishino, K., & Nayar, S. K. (2006). Corneal imaging system: Environment from eyes. *International Journal of Computer Vision*, *70*(1), 23–40.

Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American*, *224*(6), 34–43.

Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206.

Orschansky, J. (1899). Eine Methode die Augenbewegungen direct zu untersuchen. *Zentralblatt für Physiologie*, *12*, 785–790.

Pastukhov, A., & Braun, J. (2010). Rare but precious: Microsaccades are highly informative about attentional allocation. *Vision Research*, *50*(12), 1173–1184.

Pfeiffer, T. (2008). Towards gaze interaction in immersive virtual reality: Evaluation of a monocular eye tracking set-up. In M. Schumann, & T. Kuhlen (Eds.), *Virtuelle und Erweiterte Realität – Fünfter Workshop der GI-Fachgruppe VR/AR* (pp. 81–92). Aachen: Shaker Verlag.

Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience & Biobehavioral Reviews*, *37*(10), 2516–2528.

Pieters, R., Rosbergen, E., & Wedel, M. (1999). Visual attention to repeated print advertising: A test of scanpath theory. *Journal of Marketing Research*, *36*(4), 424–438.

Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of Research in Marketing*, *16*(1), 1–16.

Pieters, R., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, *68*(2), 36–50.

Pieters, R., & Wedel, M. (2007). Goal control of attention to advertising: The Yarbus implication. *Journal of Consumer Research*, *34*(2), 224–233.

Pieters, R., Wedel, M., & Zhang, J. (2007). Optimal feature advertising design under competitive clutter. *Management Science*, *53*(11), 1815–1828.

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

Rolfs, M., Kliegl, R., & Engbert, R. (2008). Toward a model of microsaccade generation: The case of microsaccadic inhibition. *Journal of Vision*, *8*(11), 5.1–5.23.

Russo, J. E., & Leclerc, F. (1994). An eye-fixation analysis of choice processes for consumer nondurables. *Journal of Consumer Research*, *21*(2), 274–290.

Russo, J. E., & Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition*, *3*(3), 267–276.

Sattar, H., Müller, S., Fritz, M., & Bulling, A. (2015). Prediction of search targets from fixations in open-world settings. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 981–990).

Shi, S. W., Wedel, M., & Pieters, F. (2013). Information acquisition during online decision making: A model-based exploration using eye-tracking data. *Management Science*, *59*(5), 1009–1026.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322.

Shinoda, H., Hayhoe, M. M., & Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research*, *41*(25–26), 3535–3545.

Sibert, L. E., & Jacob, R. J. (2000). Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 281–288). ACM.

Simeone, A., Bulling, A., Alexander, J., & Gellersen, H. (2016). Three-point interaction: Combining bi-manual direct touch with gaze. In *Proceedings of the International Conference on Advanced Visual Interfaces (AVI)* (pp. 168–175).

Smeets, J. B., Hayhoe, M. M., & Ballard, D. H. (1996). Goal-directed arm movements change eye-head coordination. *Experimental Brain Research*, *109*(3), 434–440.

Steil, J., & Bulling, A. (2015). Discovery of everyday human activities from long-term visual behaviour using topic models. In *Proceedings ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)* (pp. 75–85).

Stellmach, S., & Dachselt, R. (2012). Look & touch: Gaze-supported target acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2981–2990). ACM.

Stern, J. A., Walrath, L. C., & Goldstein, R. (1984). The endogenous eyeblink. *Psychophysiology*, *21*(1), 22–33.

Stüttgen, P., Boatwright, P., & Monroe, R. T. (2012). A satisficing choice model. *Marketing Science*, *31*(6), 878–899.

Sugano, Y., & Bulling, A. (2015). Self-calibrating head-mounted eye trackers using egocentric visual saliency. In *Proceedings ACM Symposium on User Interface Software and Technology (UIST)* (pp. 363–372).

Sugano, Y., Zhang, X., & Bulling, A. (2016). Aggregaze: Collective estimation of audience attention on public displays. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST)* (pp. 821–831). ACM.

Tessendorf, B., Bulling, A., Roggen, D., Stiefmeier, T., Feilner, M., Derleth, P., & Tröster, G. (2011). Recognition of hearing needs from body and eye movements to improve hearing instruments. In *Proceedings of the International Conference on Pervasive Computing* (pp. 314–331).

Tonsen, M., Steil, J., Sugano, Y., & Bulling, A. (2017). Invisibleeye: Mobile eye tracking using multiple low-resolution cameras and learning-based gaze estimation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, *1*(3), 106.

Turner, J., Alexander, J., Bulling, A., & Gellersen, H. (2015). Gaze+RST: Integrating gaze and multitouch for remote rotate-scale-translate tasks. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 4179–4188). ACM.

Turner, J., Alexander, J., Bulling, A., Schmidt, D., & Gellersen, H. (2013). Eye pull, eye push: Moving objects between large screens and personal devices with gaze and touch. In *14th International Conference on Human-Computer Interaction (INTERACT)* (pp. 170–186). Springer.

Vaitukaitis, V., & Bulling, A. (2012). Eye gesture recognition on portable devices. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (pp. 711–714). ACM.

Valenti, R., Sebe, N., & Gevers, T. (2012). Combining head pose and eye location information for gaze estimation. *IEEE Transactions on Image Processing*, *21*(2), 802–815.

van der Lans, R., Pieters, R., & Wedel, M. (2008). Eye-movement analysis of search effectiveness. *Journal of the American Statistical Association*, *103*(482), 452–461.

Velloso, E., Turner, J., Alexander, J., Bulling, A., & Gellersen, H. (2015). An empirical investigation of gaze selection in mid-air gestural 3D manipulation. In *Human-Computer Interaction* (pp. 315–330). Springer.

Vertegaal, R. (2003). Attentive user interfaces. *Communications of the ACM*, *46*(3), 30–33.

Vertegaal, R., Slagter, R., van der Veer, G., & Nijholt, A. (2001). Eye gaze patterns in conversations: There is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 301–308). ACM.

Vickers, J. N. (2006). Gaze of Olympic speedskaters skating at full speed on a regulation oval: Perception-action coupling in a dynamic performance environment. *Cognitive Processing*, 7(1), 102–105.

Vickers, J. N., & Adolphe, R. M. (1997). Gaze behaviour during a ball tracking and aiming skill. *International Journal of Sports Vision*, *4*(1), 18–27.

Vidal, M., Bulling, A., & Gellersen, H. (2013). Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 439–448). ACM.

Wade, N. J. (2010). Pioneers of eye movement research. *i-Perception*, *1*(2), 33–68.

Walter, R., Bulling, A., Lindlbauer, D., Schüssler, M., & Müller, H. J. (2015). Analyzing visual attention during whole body interaction with public displays. In *Proc. ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)* (pp. 1263–1267).

Wedel, M., & Pieters, R. (2000). Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Science*, *19*(4), 297–312.

Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. In N. Malhotra (Ed.), *Review of Marketing Research* (Vol. 4, pp. 123–147). New York, NY: M. E. Sharpe Inc.

Wedel, M., Yan, J., Siegel, E. L., & Li, H. A. (2016). Nodule detection with eye movements. *Journal of Behavioral Decision Making*, *29*(2–3), 254–270.

Wood, E., Baltrušaitis, T., Morency, L. P., Robinson, P., & Bulling, A. (2016). A 3D morphable eye region model for gaze estimation. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 297–313). Springer.

Wood, E., Baltrušaitis, T., Zhang, X., Sugano, Y., Robinson, P., & Bulling, A. (2015). Rendering of eyes for eye-shape registration and gaze estimation. In *Proc. IEEE International Conference on Computer Vision (ICCV)* (pp. 3756–3764).

Wood, E., & Bulling, A. (2014). Eyetab: Model-based gaze estimation on unmodified tablet computers. In *Proceedings of the International Symposium on Eye Tracking Research and Applications (ETRA)* (pp. 207–210). ACM.

Xu, P., Sugano, Y., & Bulling, A. (2016). Spatio-temporal modeling and prediction of visual attention in graphical user interfaces. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3299–3310). ACM.

Yang, L., Toubia, O., & de Jong, M. G. (2015). A bounded rationality model of information search and choice in preference measurement. *Journal of Marketing Research*, *52*(2), 166–183.

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye Movements and Vision* (pp. 171–211). New York, NY: Plenum Press.

Zhai, S., Morimoto, C., & Ihde, S. (1999, May). Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 246–253). ACM.

Zhang, J., Wedel, M., & Pieters, R. (2009). Sales effects of attention to feature advertisements: A Bayesian mediation analysis. *Journal of Marketing Research*, *46*(5), 669–681.

Zhang, X., Sugano, Y., & Bulling, A. (2017). Everyday eye contact detection using unsupervised gaze target discovery. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST)* (pp. 193–203). ACM.

Zhang, X., Sugano, Y., Fritz, M., & Bulling, A. (2015). Appearance-based gaze estimation in the wild. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4511–4520).

Zhang, X., Sugano, Y., Fritz, M., & Bulling, A. (2019). MPIIGaze: Real-World Dataset and Deep Appearance-Based Gaze Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (TPAMI), 41(1), pp. 162–175, 2019.

Zhang, X., Sugano, Y., Fritz, M., & Bulling, A. (2017). It's written all over your face: Full-face appearance-based gaze estimation. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).*

Zhang, Y., Bulling, A., & Gellersen, H. (2013). Sideways: A gaze interface for spontaneous interaction with situated displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)* (pp. 851–860). ACM.

Zhang, Y., Müller, H. J., Chong, M. K., Bulling, A., & Gellersen, H. (2014). GazeHorizon: Enabling passers-by to interact with public displays by gaze. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)* (pp. 559–563). ACM.