

A Breadth-First Survey of Eye Tracking Applications

Andrew T. Duchowski

Department of Computer Science, Clemson University

Eye tracking applications are surveyed in a breadth-first manner, reporting on work from the following domains: Neuroscience, Psychology, Industrial Engineering and Human Factors, Marketing/Advertising, and Computer Science. Following a review of traditionally diagnostic uses, emphasis is placed on interactive applications, differentiating between selective and gaze-contingent approaches.

Introduction

Eye tracking research is entering its fourth era, distinguished by the emergence of interactive applications. Rayner (1998) summarizes the characteristics of eye tracking's first three eras: (ca. 1879–1920) defined by the discovery of many basic eye movement facts, including saccadic suppression, saccade latency, and the size of the perceptual span; (ca. 1930–1958) characterized by a more applied research focus, coinciding with the behaviorist movement in experimental psychology; (ca. 1970–1998) marked by improvements in eye movement recording systems facilitating increasingly accurate and easily obtained measurements. A wide variety of eye tracking applications now exist, which can broadly be dichotomized from a system analysis point of view as *diagnostic* or *interactive*.

In its diagnostic role, the eye tracker provides objective and quantitative evidence of the user's visual and (overt) attentional processes. In this capacity, eye movements are generally recorded to ascertain the user's attentional patterns over a given stimulus. Diagnostic applications are typically represented by the unobtrusive use of the eye tracking device. Furthermore, the stimulus being displayed does not usually need to change or react to the viewer's gaze. In this scenario, the eye tracker is used to record eye movements for post-trial, off-line assessment of the viewer's gaze during the experiment.

Bolstered by advancements in computational power, richness of graphical displays, and robustness of interactive styles, as an interface modality, the eye tracker serves as a powerful input device in a host of visually-mediated applications. An interactive system must respond to or interact with the user based on the observed eye movements. Such interactive systems may fall into two application subtypes: *selective* and *gaze-contingent*. Selective systems use the point of gaze as analogous to a pointing device such as the mouse, whereas gaze-contingent systems exploit knowledge of the user's gaze to facilitate the rapid rendering of complex dis-

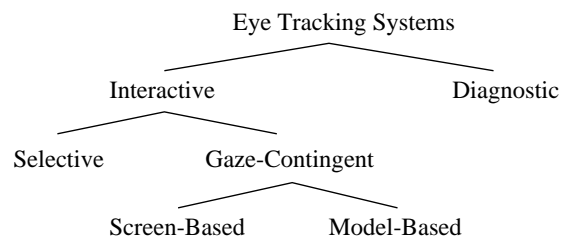


Figure 1. Hierarchy of eye tracking applications.

plays (e.g., graphical environments). The latter can be further delineated in terms of display processing, as shown in the hierarchy in Figure 1.

This paper reports on eye tracking methodologies spanning three disciplines: neuroscience, psychology, and computer science, with brief remarks about industrial engineering and marketing. A breadth first survey is given, proceeding from the diagnostic use of eye trackers to interactive applications. Material in the paper is condensed from Duchowski (2002), © 2002 Springer-Verlag, reprinted with permission.

Neuroscience

Neuroscientific research has identified numerous interconnected neural components of vision, starting at the retinal photoreceptors, and (more or less) ending in the cortical regions implicated in low-level vision. Retino-geniculate pathways of vision as well as deeper visual brain regions have also been identified. In the context of vision and eye movements, knowledge of the physiological organization of the optic tract as well as of cognitive and behavioral aspects of vision are both indispensable in obtaining a complete understanding of human vision (see for examples the papers by Robinson (1968) and Findlay and Walker (1999) which present frameworks for understanding human saccadic eye movement generation motivated by consideration of the convergence between low level studies of the saccadic system from a physiological standpoint and studies investigating saccadic eye movements in a behavioral context). Current neuroscientific trends related to the study of eye movements are briefly discussed, touching on exemplary work investigating attention and brain imaging.

This work was supported in part by a University Innovation grant (# 1-20-1906-51-4087), NASA Ames task (# NCC 2-1114), and NSF CAREER award IIS-9984278.

Attentional Neuroscience

The dissociation of the “spotlight of attention” (Posner, Snyder, & Davidson, 1980) from ocular fixation poses a problem for eye tracking researchers. It is possible to visually fixate one location while simultaneously diverting attention to another. For example, astronomers perform this fairly regularly when looking for faint stars or star clusters with the naked eye. When examining a scanpath over a visual stimulus, we can often say that specific regions were looked at, perhaps even fixated (following analysis of eye movements), however, we can not be fully confident that these specific regions were fully perceived. There is (currently) no simple way of telling what the brain is doing during a particular visual scan of the scene. Ideally, we would have to not only record the point of one’s gaze, but also of one’s brain activity.

Investigating neuronal activity related to fixational eye movements (saccades and intersaccadic (drift) intervals), Snodderly, Kagan, and Gur (2001) show that responses of a monkey’s V1 neurons to fixational eye movements are specific and diverse. Some cells are activated only by saccades, others discharge during drifts, and most show a mixture of these two influences. The patterns of activity reflect the interactions between the stimulus, the receptive-field activating region, the temporal response characteristics of the neuron, and the retinal positions and image motions imparted by eye movements. The diversity of the activity patterns suggests that during natural viewing of a stationary scene some cortical neurons carry information about saccadic occurrences and directions while other neurons code details of the retinal image. Snodderly et al. argue that, theoretically, saccade neurons could participate in saccadic suppression by inhibiting other neurons that carry stimulus information (reducing signal strength) or by adding noise to the signal, thereby raising thresholds and making stimulus events undetectable at the times of saccades.

The prefrontal (PF) cortex is largely thought to be central to the ability of shifting attention and choosing actions appropriate not only to the sensory information at hand but also according to the specific sensory, motor, and cognitive demands of the task (the behavioral context) in which it is encountered (Asaad, Rainer, & Miller, 2000). Asaad et al. report that most of the 305 neurons recorded in two monkeys displayed a task-dependent change in overall activity, particularly in the fixation interval preceding cue presentation. Results show that, for many PF neurons, activity was influenced by the task being performed. Asaad et al. suggest that formal demands of behavior are represented within PF activity and thus support the hypothesis that one PF function is the acquisition and implementation of task context and the “rules” used to guide behavior.

Eye Movements and Brain Imaging

Recently, eye movement recording and functional brain imaging have been used to track a subject’s fixation point while simultaneously recording cortical activation during attentional tasks, in order to identify functional brain structures implicated in attentional behavior. Presently, possibly due to

prohibitive cost, combined eye tracking and brain imaging equipment are not in widespread use, although such devices are beginning to appear.

Özyurt, DeSouza, West, Rutschmann, and Greenlee (2001) compared the neural correlates of visually guided saccades in the step and gap paradigms while recording saccadic eye movements during task performance. Results from Özyurt et al.’s study indicate significant task-related activity in the striate and extrastriate cortex, the frontal eye fields, the supplementary motor area, parietal cortex and angular gyrus, the frontal operculum, and the right prefrontal area 10. This type of research helps identify functional brain structures that participate in attentional mechanisms.

Psychology

Perhaps the first well-known use of eye trackers in the study of human (overt) visual attention were those conducted during reading experiments. Rayner (1992, 1998) provides an excellent survey of eye tracking applications in reading and other information processing tasks.

Reading

Rayner (1998) synthesizes over 100 years of research. While the reader is referred to Rayner’s article for the complete review, three interesting examples of eye movement characteristics during reading are summarized here. First, when reading English, eye fixations last about 200–250 ms and the mean saccade size is 7–9 letter spaces. Second, eye movements are influenced by textual and typographical variables, e.g., as text becomes conceptually more difficult, fixation duration increases and saccade length decreases. Factors such as the quality of print, line length, and letter spacing influence eye movements. Third, eye movements differ somewhat when reading silently from reading aloud: mean fixation durations are longer when reading aloud or while listening to a voice reading the same text than in silent reading. There is of course a good deal more that has been learned (e.g., Reichle, Pollatsek, Fisher, and Rayner (1998) summarize some basic data on reading and eye movements related to their subsequent effort on modeling of the reading process), here, the methodology behind such discoveries is what is of primary interest.

In addition to descriptive studies, where eye movements are simply recorded during reading of text, eye trackers have been used to modulate the stimulus display in real-time depending on where the reader is looking. Three experimental paradigms, the *moving window*, *boundary*, and *foveal mask*, have been developed to explore eye movements (see Reichle et al. (1998) for a concise review and examples of these techniques). Although first developed for reading studies, these paradigms have since been adapted to other contexts such as scene perception (see below).

In the moving window paradigm, or *gaze-contingent display change* paradigm, developed by McConkie and Rayner (1975), a window is sized to include a number of characters (e.g., 14) to the left and right of a fixated word. The assumption with this technique is that when the window is as large

as the region from which the reader can obtain information, there will be no difference between reading in that situation and when there is no window.

Conversely, a foveal mask, developed by Rayner and Bertera (1979) (see also Bertera and Rayner (2000)), is placed over a number of fixated characters (e.g., 7), creating an artificial foveal scotoma. Eye movement behavior in this situation is quite similar to the eye movement behavior of patients with real scotomas (Rayner, 1998).

In the boundary technique, developed by Rayner (1975) to investigate the use of peripheral information in reading, the stimulus changes as fixation crosses a predefined boundary. In this technique, a word or nonword letter string is initially presented in a target location. When the reader's gaze moves to that location, the initially presented stimulus is replaced by the target word. The fixation duration on the target allows the experimenter to make inferences about the type of information acquired from the target location when it was in parafoveal vision.

The motivation behind Rayner's (1975, 1998) use of eye movements for the study of reading was due in part to his having found tachistoscopic methods to be inadequate. Tachistoscopic (strobe-like) presentation of letters and words relies on the presentation of material for very brief exposures to exclude the possibility of an eye movement during the presentation. Prior to eye tracked reading studies, the tachistoscopic exposure was often thought of as being analogous to a single fixation during reading. Rayner argued that what subjects report from a tachistoscopic presentation cannot be taken as a complete specification of what they saw. The argument for eye movement recording over tachistoscopic displays carries over to scene perception and is discussed in greater detail in the next section.

Scene Perception

Although certain reading patterns are easily recognized (e.g., left-to-right, top-to-bottom for English readers, or right-to-left for Hebrew), no apparent strategies for scene viewing have been easily discerned. Contrary to reading, there appears to be no canonical scanpath for particular objects (Kennedy, 1992). There may be context differences at play. Kroll (1992) states that while there may be similarities between reading and scene viewing tasks, the tasks are very different.

According to Henderson and Hollingworth (1998), there are at least three important reasons to understand eye movements in scene viewing. First, eye movements are critical for the efficient and timely acquisition of visual information during complex visual-cognitive tasks, and the manner in which eye movements are controlled to service information acquisition is a critical question. Second, how we acquire, represent, and store information about the visual environment is a critical question in the study of perception and cognition. The study of eye movement patterns during scene viewing contributes to an understanding of how information in the visual environment is dynamically acquired and represented. Third, eye movement data provide an unobtrusive, online measure

of visual and cognitive information processing.

Rayner (1998) recounts the traditionally held belief that examining the fine details of eye movements during scene perception is a high-cost, low-yield endeavor. Experiments using tachistoscopic presentations and eye movement recordings have led to the conclusion that participants get the gist of a scene very early in the process of looking, sometimes even from a single brief exposure. Thus it has been advocated that the gist of the scene is abstracted on the first few fixations, and the remainder of the fixations on the scene are used to fill in details. Rayner reviews several findings which support the contention that important conclusions about temporal aspects of scene perception can be obtained from eye movement analysis.

Loftus (1981) presented results of a masked tachistoscopic study suggesting a model of picture encoding that incorporates the following propositions: (1) a normal fixation on a picture is designed to encode some feature of the picture; (2) the duration of a fixation is determined by the amount of time required to carry out the intended feature encoding; and (3) the more features are encoded from a picture, the better the recognition memory will be from the picture. A major finding of Loftus' experiments was that with exposure time held constant, recognition performance increased with increasing numbers of fixations. When eye fixations are simulated tachistoscopically and their durations experimentally controlled, all traces of this phenomenon disappear. Moreover, Loftus' experiments suggest that within a fixation, visual information processing ceases fairly early, that is, acquired information reaches asymptote soon after the start of a fixation. The problem identified with tachistoscopic exposures, which were meant to simulate saccades to new places, is that there is no guarantee that new places in the picture are fixated; it is entirely possible that subjects were simply holding their eyes steady throughout all tachistoscopic flashes. From an eye tracking experiment, Loftus draws the argument that given more places to look at in the picture, more information can be acquired from the picture. Additional (tachistoscopic) flashes are only useful insofar as they permit acquisition of information from additional portions of the picture.

Rayner and Pollatsek (1992) concede that much of the global information about the scene background or setting is extracted during the initial fixation. Some information about objects or details throughout the scene can be extracted far from fixation. However, if an object is important, it is usually fixated. Rayner and Pollatsek indicate that this foveal identification is aided significantly by the information extracted extrafoveally and conclude that it is necessary to study eye movements to achieve a full understanding of scene perception.

Henderson and Hollingworth (1998) suggest several metrics for evaluation of relative informativeness of scene regions. At a macro level analysis, the *total time* that a region is fixated in the course of scene viewing (the sum of the durations of all fixations in that region) is correlated with the number of fixations in that region. At a micro level analysis, several commonly used measures include *first fixation*

duration (the duration of the initial fixation in a region), *first pass gaze duration* (the sum of all fixations from first entry to first exit in a region), and *second pass gaze duration* (the sum of all fixations from second entry to second exit in a region). Generally, first pass gaze durations are longer for semantically informative (i.e., inconsistent) objects. Semantically informative objects also tend to draw longer second pass and total fixation durations. The influence of semantic informativeness on the duration of the very first fixation on an object is less clear. That is, scene context has an effect on eye movements: fixation duration on an object that does not belong in the scene is longer than fixation duration on an object that does belong (Rayner, 1998). However, it is not clear whether the longer fixations on objects in violation of the scene reflect longer times to identify those objects or longer times to integrate them into a global representation of the scene (it could also reflect amusement of the absurdity of the violating object in the given context). Recently, Greene and Rayner (2001) showed that familiarity with distractors around an unfamiliar target facilitates visual search. There were comparably long, but fewer, fixations when distractors were familiar, contradicting the theory that unfamiliar distractors need longer processing.

Henderson (1992) offers two criticisms of the eye movement paradigm. First, it is likely that global measures of fixation time, such as the total time spent on an object during the course of scene viewing, and the gaze duration on an object (the time of all initial fixations on an object prior to leaving that object for the first time) reflect post-identification processes. Thus, it is likely that gaze duration in scene processing reflects other processes beyond object identification. Henderson suggests that the preferred fixation measure is the true first fixation duration, or the duration of time from the initial landing of the eyes on an object until the eyes move to any other location, including another location on the object. Second, the basic premise of the eye movement paradigm is that the results will reflect normally occurring visual-cognitive processes because subjects can view scenes in a natural manner. However, unlike reading, where the overall task is arguably transparent, subjects must be given an orienting task when they view a scene. Unfortunately, viewing behavior and eye movement patterns change as a function of the viewing task given to the subject (Yarbus, 1967). One way to address the orienting task issue would be to give subjects a task that did not force the creation of a coherent memory representation for the scene, and look for similar scene context effects on fixation time across tasks.

Perception of Art. A particularly interesting subset of scene perception studies is the examination of gaze over a specific set of contextual images, namely art. The first systematic exploration of fixation positions in scenes was reported by Buswell (1935) (as cited by Henderson and Hollingworth (1998)). An important observation was that fixation positions were found to be highly regular and related to the information in the pictures (Henderson & Hollingworth, 1998). These data provided some of the earliest evidence that eye movement patterns during complex scene per-

ception are related to the information in the scene, and by extension, to perceptual and cognitive processing of the scene.

Another example of a study of eye movements directed at works of art, by Molnar (1981) (as reported by Solso (1999)), shows small differences in scanpaths between groups of subjects viewing the artwork for its semantic meaning or for its aesthetic appeal. Remarkably, however, both sets of scanpaths are very similar in terms of fixated image features. Fixations made by two groups of fine art students, given different sets of questions pertaining to the artwork, appear to coincide on important elements even though the order of fixations may differ.

A recent large-scale eye tracking study of art was conducted by Wooding (2002). In three months, eye movements were successfully collected from 5,638 subjects while they viewed digitized images of paintings from the National Gallery collection. Since a composite representation of eye movements from so many subjects posed a problem, Wooding devised a *fixation map* method of analysis, which might be descriptively termed a landscape or terrain map of fixations, and is in fact similar to the *landscape map* developed independently by Velichkovsky, Pomplun, and Rieser (1996). The value at any point on the map indicates the height or amount of a particular property at that point (e.g., the number of fixations).

Approaching art perception from a different starting point, eye movements have recently been considered in the *generation* of art by computers. Perceptual principles and eye tracking methodology (among other techniques) were employed by DeCarlo and Santella (2002) to process aesthetically pleasing images by computer. Additional exemplary uses of eye trackers in computer graphics are presented in subsequent sections.

Perception of Film. An interesting example of an eye tracking study over film, essentially a dynamic form of artistic media, is given by d'Ydewalle, Desmet, and Van Rensbergen (1998). d'Ydewalle et al. distinguish three levels of film editing errors in sequencing successive shots. First-order editing errors refer either to small displacements of the camera position or to small changes of the image size, disturbing the perception of apparent movement and leading to the impression of jumping. Second-order editing errors follow from a reversal of the camera position, leading to a change of the left-right position of the main actors (or objects) and a complete change of the background. With third-order editing errors, the linear sequence of actions in the narrative story is not obeyed. d'Ydewalle et al.'s experiment shows that there is an increased spatial distribution of eye movements 200–400 ms following both second- and third-order editing errors. Such an increase is not obtained after a first-order editing error, suggesting that increased distribution of eye movements after second- and third-order editing errors is due to postperceptual, cognitive effects.

Visual Search

How humans perceive the visual scene in natural or free tasks such as picture viewing can be modeled by the visual

search task. This task requires the report of the presence of a target in a display. In comparison to reading, there have not been nearly as many studies dealing with visual search (Rayner, 1998).

Because the nature of the search task influences eye movement behavior, any statement about visual search and eye movements needs to be qualified by the characteristics of the search task. Specifically, visual search tasks vary widely, and tasks in which eye movements have been monitored consist of at least the following: search (a) through text or textlike material, (b) with pictorial stimuli, (c) with complex arrays such as X-rays, and (d) with randomly arranged arrays of alphanumeric characters or objects. There is considerable variability in fixation time and saccade length as a function of the particular search task (Rayner, 1998). When eye movements are recorded during extended search, fixations tend to be longer than in reading.

As demonstrated by Yarbus (1967) and then by Noton and Stark (1971a, 1971b), eye tracked scanpaths strongly suggest the existence of a serial component to the picture viewing process. However, serial scanpaths do not adequately explain the brain's uncanny ability to integrate holistic representations of the visual scene from piecemeal (foveal) observations. That is, certain perceptual phenomena are left unaccounted for by scanpaths, including perception of illusory images such as the Kanizsa (1976) figure or the Necker (1832) cube. While scanpaths cast doubt on a purely Gestalt view of visual perception, it would seem that some sort of holistic mechanism is at work which is not revealed by eye movements alone. Models of visual search attempt to answer this dilemma by proposing a parallel component, which works in concert with the serial counterpart exhibited by eye movements.

The consensus view is that a parallel, pre-attentive stage detects four basic features: color, size, orientation, and presence and/or direction of motion (Doll, 1993; Wolfe, 1993). Todd and Kramer (1993) suggest that attention (presumably in the periphery) is captured by sudden onset stimuli, uniquely colored stimuli (to a lesser degree than sudden onset), and bright and unique stimuli. However, there is doubt in the literature whether human visual search can be described as an integration of independently processed features (Van Orden & DiVita, 1993). Visual search, even over natural stimuli, is at least partially deterministic, rather than completely random (Doll, Whorter, & Schmieder, 1993). Determinism stems from either or both of two kinds: the observer's strategy determines the search pattern (as in reading), and/or the direction of the next saccade is based on information gained through peripheral vision about surrounding stimuli. Features that are likely to be fixated include edges, corners, spatially high frequency components, but not plain surfaces.

Relatively few studies have addressed the relationship between eye movements and the search process (Findlay & Gilchrist, 1998). Findlay and Gilchrist argue that the tradition in search research to pay little attention to eye movements and instead to use the concept of covert visual attention movements (redirecting attention without moving the

eyes; an important component of the Feature Integration Theory of visual attention (Treisman & Gelade, 1980)) is misguided. Findlay and Gilchrist demonstrate that when viewers are free to move their eyes, no additional covert attentional scanning occurs. They show that unless instructions explicitly prevent eye movements, subjects in a search task show a natural propensity to move their eyes, even in situations where it would be more efficient not to do so. Findlay and Gilchrist suggest that the reason for this preference is that, in naturally occurring search situations, eye movements form the most effective way of sampling the visual field.

Findlay (1997) recorded eye movements during tasks of a simple feature search and a color shape feature conjunction search. Findlay's work provides an impressive confirmation that search for a pre-specified color target can be carried out in parallel. When two targets are presented simultaneously in neighboring positions, the first saccade is directed towards some "center of gravity" position. Results suggest that the control of the initial eye movement during both simple and conjunction searches is through a spatially parallel process.

Results from a conjunction search experiment (color and shape) are particularly relevant since this is a situation where serial scanning would be expected according to the classical search theory of Treisman and Gelade (1980). If a rapid serial scanning with covert attention could occur before the saccade is initiated, it is not clear why incorrect saccades would occur as frequently as observed in Findlay's (1997) experiment. Moreover, Findlay argues that the data place constraints on the speed of any hypothetical serial scanning process since it would be necessary for a number of locations to be scanned before the target is located, given the accuracy obtained. Alternative accounts of the visual search process have appeared which assign much more weight to the parallel processes and avoid the postulation of rapid serial scanning (e.g., see Wolfe (1994), Wolfe and Gancarz (1996), and Findlay and Walker (1999)).

Recently, Bertera and Rayner (2000) had viewers search through a randomly arranged array of alphanumeric characters (i.e., letters and digits) for the presence of a target letter. They used both the moving window technique and the moving mask technique. In the moving window study, search performance was equivalent for all windows at least as large as 5°. The moving mask had a deleterious effect on search time and accuracy, indicating the importance of foveal vision during search; the larger the mask, the longer the search time. The moving window paradigm for scene perception and visual search is discussed below as an instance of gaze-contingent display technology.

Natural Tasks

Studies of visual search are expanding to consider more complex stimuli such as natural scenery (Hughes, Nozawa, & Kitterle, 1996). However, viewing of pictures projected on a laboratory display still constitutes a somewhat artificial task. Recent advancements in wearable and virtual displays now allow collection of eye movements in more natural situations, usually involving the use of generally unconstrained

eye, head, and hand movements.

Important work in this area has been reported by Land, Mennie, and Rusted (1999) and Land and Hayhoe (2001). The aim of the first study was to determine the pattern of fixations during the performance of a well-learned task in a natural setting (making tea), and to classify the types of monitoring action that the eyes perform. Results of this study indicate that even automated routine activities require a surprising level of continuous monitoring. Foveal direction was always close to the object being manipulated, and very few fixations were irrelevant to the task. Land et al. conclude that although the actions of tea-making are 'automated' and proceed with little conscious involvement, the eyes closely monitor every step of the process.

Investigating a similar natural task, Land and Hayhoe (2001) examined the relations of eye and hand movements in extended food preparation tasks, tea-making and making peanut butter and jelly sandwiches. Gaze usually reached the next object in the sequence before any sign of manipulative action, indicating that eye movements are planned into the motor pattern and lead each action. The eyes usually fixated the same object throughout the action upon it, although they often moved on to the next object in the sequence before completion of the preceding action. General conclusions provided by Land and Hayhoe are that the eyes provide information on an 'as needed' basis, and the relevant eye movements usually precede the motor acts they mediate by a fraction of a second. Eye movements are thus in the vanguard of each action plan, and are not simply responses to circumstances. Land and Hayhoe conclude that their studies lend no support to the idea that the visual system builds up a detailed model of the surroundings and operates from that. Rather, most information is actively obtained from the scene as it is needed.

Ballard, Hayhoe, and Pelz (1995) investigated the use of short-term memory in the course of a natural eye-hand task. The investigation focused on the minimization of subjects' use of short-term memory by employing deictic primitives through serialization of the task with eye movements (e.g., using the eyes to "point to" objects in a scene in lieu of memorizing all of the objects' positions and other properties). Ballard et al. argue that a deictic strategy in a pick-and-place task employs a more efficient use of a frame of reference centered at the fixation point rather than a viewer-centered reference frame which might require memorization of objects in the world relative to coordinates centered at the viewer. By recording eye movements during the block pick-and-place task, Ballard et al. were able to show that subjects frequently directed gaze to the model pattern before arranging blocks in the workspace area, suggesting that information is acquired incrementally during the task and is not acquired *in toto* at the beginning of the task. That is, subjects appeared to use short-term memory frugally, acquiring information just prior to its use, and did not appear to memorize the entire model block configuration before making a copy of the block arrangement.

In a similar block-moving experiment, Smeets, Hayhoe, and Ballard (1996) were able to show that horizontal move-

ments of gaze, head, and hand followed a coordinated pattern. A shift of gaze was generally followed by a movement of the head, which preceded the movement of the hand. This relationship is to a large extent task-dependent. In goal-directed tasks in which future points of interest are highly predictable, Smeets et al. hypothesize that while gaze and head movements may decouple, the actual position of the hand is a likely candidate for the next gaze shift.

A recent example of such intentional visual attention was demonstrated by Pelz, Canosa, and Babcock (2000). Using a wearable eye tracker, Pelz et al. observed intentionally-based eye movements, which they termed "look-ahead" eye movements. During a simple hand-washing task, recorded eye movements showed that gaze moved to a location (soap dispenser) prior to the movement of the hands to the same location.

To further examine issues raised by observations of natural behavior, Hayhoe et al. (2002) have recently begun using complex virtual environments which can be manipulated by the experimenter at critical points during task performance. In a virtual environment where subjects copy toy models, Hayhoe et al. show that regularities in the spatial structure are used by subjects to control eye movement targeting. Other experiments in a virtual environment with haptic feedback show that even simple visual properties like size are not continuously available or processed automatically by the visual system, but are dynamically acquired and discarded according to the momentary task demands.

Auditory Language Processing

In this paradigm, eye movements are recorded as people listen to a story or follow instructions regarding an array they are looking at. Cooper (1974) found that when people are simultaneously presented with spoken language and a visual field containing elements semantically related to the informative items of speech, they tend to spontaneously direct their line of sight to the elements that are most closely related to the meaning of the language currently heard.

The eye movement paradigm has also been applied to auditory language processing by Allopenna, Magnuson, and Tanenhaus (1998). Participants followed spoken instructions to manipulate either real objects or pictures displayed on a computer screen. Eye movements to objects in the workspace were closely time-locked to referring expressions in the unfolding speech stream, providing a sensitive and nondisruptive measure of spoken language comprehension during continuous speech. Allopenna et al. argue that the sensitivity of the response measure coupled with a clear linking hypothesis between lexical activation and eye movements indicates that this methodology will be invaluable in exploring questions about the microstructure of lexical access during spoken word recognition.

Other Information Processing Tasks

Rayner (1998) provides a comprehensive review of eye movement work from multiple domains, including: (i) mathematics, numerical reading, and problem solving; (ii) dual

task situations; (iii) face perception; (iv) brain damage, and (v) dynamic situations such as driving, basketball foul shooting, golf putting, table tennis, baseball, gymnastics, walking in uneven terrain, mental rotation, and human-computer interaction. Some of these latter applications are covered in the following sections.

Industrial Engineering and Human Factors

Eye tracking is particularly important for evaluating present and future environments in which humans do and will work. Traditional measurement methods of human performance often include measures of reaction time and accuracy, e.g., how fast a person completes a task and how well this task is performed. To study the steps taken to perform the tasks requires analysis of the individual procedures performed. For this analysis, eye movements are particularly interesting since they present measures which can provide insights into the visual, cognitive, and attentional aspects of human performance. Here, three broad experimental domains are presented in which eye tracking can play an important analytical role: aviation, driving, and visual inspection.

Aviation

An example of a recent combined use of relatively new eye tracking technology in a sophisticated flight simulator was reported by Anders (2001). Eye and head movements of professional pilots were recorded under realistic flight conditions in an investigation of human-machine interaction behavior relevant to information selection and management as well as situation and mode awareness in a modern glass cockpit. Analysis of eye movements illustrates the importance of the Primary Flight Display (PFD) as the primary source of information during flight (the PFD is the familiar combined artificial horizon and altimeter display for combined altitude and attitude awareness). As a proof of concept, this study shows the potential of eye movements for judgment of pilots' performance and future training of novice pilots.

Eye movements have also been used to evaluate the usability of specific instruments such as newly developed electronic maps. In their study, Ottati, Hickox, and Richter (1999) compared eye movement patterns on different terrain features between experienced and novice pilots during a Visual Flight Rules (VFR) simulation. Based on previous studies of pilots' eye movements during Instrument Flight Rules (IFR) navigation, the authors expected experienced aviators to spend less time finding and fixating on their navigational landmarks, while novices were expected to have greater difficulty finding landmarks and extracting useful data from them, causing greater dwell times. As expected, a greater tendency in novice pilots was found to "fly out of the window" (i.e., devoting visual attention outside the cockpit) than in experienced pilots.

In a study of electronic maps for taxiing, Graeber and Andre (1999) suggest that training is necessary to assure proper usage of, and optimal visual attention interaction with Electronic Moving Maps (EMMs). The main objective of

Graeber and Andre's study was to understand how pilots visually interact with the EMM. Results indicate that atmospheric visibility significantly affected the amount of time pilots dwelled on the EMM, with dwell time significantly higher under high visibility conditions. That is, as visibility degrades, pilots spend more time eyes-out and less time dwelling on the EMM with no loss in taxi performance. A potential explanation for this surprising result is that pilots need to be eyes-out to maintain lateral and directional loop closure, scan for hazards, and maintain information gathering "out the window".

Driving

It is widely accepted that deficiencies in visual attention are responsible for a large proportion of road traffic accidents (Chapman & Underwood, 1998). Eye movement recording and analysis provide important techniques for understanding the nature of the driving task and are important for developing driver training strategies and accident countermeasures. Chapman and Underwood (1998) indicate that at their simplest, drivers' fixation patterns on straight roads can be described as concentrating on a point near to the focus of expansion (the point in the visual field in front of the driver where objects appear stationary) with occasional excursions to items of road furniture and road edge markers. It is clear that dangerous events generally evoke long fixation durations.

Dishart and Land (1998) discuss previous work showing that experienced drivers obtain visual information from two sections of their view of the road ahead, in order to maintain a correct position in lane whilst steering their vehicle around a curve. The more distant of these two sections is used to predict the road's future curvature. This section is optimally 0.75–1.00 s ahead of the driver and is used by a feedforward (anticipatory) mechanism which allows the driver to match the curvature of the road ahead. The other, nearer, section is about 0.5 s ahead of the driver and is used by a feedback (reactive) mechanism to 'fine tune' the driver's position in lane. Eye tracking experiments have shown that the feedback mechanism is present in most people regardless of their driving experience (although its accuracy is higher in those with experience), but that the feedforward mechanism is learned through experience of steering tasks (that can include riding a bicycle, computer driving games, etc.).

A fundamental problem in visual search in driving research is defining and controlling demand on the driver as an independent variable (Crundall, Underwood, & Chapman, 1998). Possible confounding factors are increases in visual demand, such as an increase in visual clutter or complexity, and increases in cognitive demand, such as an increase in the processing demands of a particular stimulus perhaps due to an increase in its relevance to a current context. Despite the lack of consistency in the manipulations of visual demand, it seems fairly well documented that general increases in task demands and visual complexity tend to reduce mean fixation duration and increase the sampling rate.

In a driving study examining the effects of clutter, lu-

minance, and aging, Ho, Scialfa, Caird, and Graw (2001) recorded subjects' visual search for traffic signs embedded in digitized images of driving scenes. On average, older adults were less accurate than younger ones. Accuracy for daytime scenes was independent of target presence, but errors for nighttime scenes were more common on target-present trials than on target-absent trials. In daytime scenes errors were generally more common in high clutter than in low clutter, while in nighttime scenes no consistent clutter effect was detected. Fixation number data, a measure that is strongly correlated with reaction time, indicated that older adults made more fixations than did younger adults, and more fixations were needed for high-clutter and target-absent scenes than for low-clutter and target-present scenes.

Experiments where gaze is monitored in a simulated driving environment demonstrate that visibility of task relevant information depends critically on active search initiated by the observer according to an internally generated schedule, which depends on learnt regularities in the environment (Hayhoe et al., 2002). Hayhoe et al.'s findings suggest fixation patterns and attentional control in normal vision is learnt.

Research on the effects of mental activity during driving suggests the convenience of raising drivers' awareness about the possible consequences of driving while their attention is focused on their own thoughts, unrelated to driving (Recarte & Nunes, 2000). Recarte and Nunes studied the consequences of performing verbal and spatial-imagery tasks on visual search when driving. Visual functional-field size decreased horizontally and vertically, particularly for spatial-imagery tasks. Compared with ordinary driving, fixations were longer during the spatial-imagery task. With regard to driving performance, glance frequency at mirrors and speedometer decreased during the spatial-imagery tasks. Performing mental tasks while driving caused an increased attentional workload on ordinary thought, as shown by pupillary dilation. With regard to the implications for driving, Recarte and Nunes suggest that the spatial reduction of the visual inspection window, including the reduction of the inspection of mirrors, could be interpreted as a predictor of decreased probability of detecting traffic events, particularly when performing mental spatial-imagery tasks.

Visual Inspection

Schoonard, Gould, and Miller (1973) state that "visual inspection pervades the lives of all people today. From poultry, meat, and fish inspection, to drug inspection, to medical X-ray inspection, to production line inspection, to photo interpretation, the consequences of inspection directly affect people's lives through their effects on the quality and performance of goods and services." Tracking eye movements during visual inspection may lead to predictive analyses, if certain re-curring patterns or statistics can be found in collected scanpaths. For example, an expert inspector's eye movements may clearly exhibit a systematic pattern. If so, then this pattern may be one that can be used to train novice inspectors. In their study of visual inspection of integrated

circuit chips, Schoonard et al. (1973) found that good inspectors are characterized by relatively high accuracy and relatively high speed, and make many brief eye fixations (as opposed to fewer longer ones) during the time they have to inspect.

In a survey of eye movements in industrial inspection, Megaw and Richardson (1979), identify the following relevant eye movement parameters: fixation times, number of fixations, spatial distribution of fixations, interfixation distances, direction of eye movements, and sequential indices (i.e., scanpaths). Megaw and Richardson review previous inspection studies where eye movements were recorded, including inspection of sheet metal, empty bottles, integrated circuits, and tapered roller bearings.

Besides being useful for gaging inspection performance, eye movements may play a part in training visual search strategy (Wang, Lin, & Drury, 1997). Visual search strategy training can be effective in adoption of a desirable systematic search strategy. To train visual search strategy, eye movements may be used as both a feedback mechanism and as confirmation of adoption of the new search strategy.

Duchowski, Medlin, Gramopadhye, Melloy, and Nair (2001) investigated the utility of eye movements for search training in a virtual environment simulating aircraft inspection. The user's gaze direction, as well as head position and orientation, were tracked to allow recording of the user's fixations within the environment. Analysis of eye movements leads to two observations. First, mean fixation times do not appear to change significantly following training. Second, the number of fixations appears to decrease following training. These results generally appear to agree with the expectation of reduced number of fixations with the adoption of an improved visual search strategy (e.g., due to learning or familiarization of the task) (Drury, Gramopadhye, & Sharit, 1997).

Recently, Reingold, Charness, Pomplun, and Stampe (2002) review the motivation behind using chess as an ideal task environment for the study of skilled performance. Reingold et al. study visual span as a function of chess skill (expert vs. intermediate vs. novice) and configuration type (chess configuration vs. random configuration) using a gaze-contingent window technique. The paper extends classic work demonstrating that after viewing structured, but not random, chess positions for five seconds, chess masters reproduced these positions much more accurately than lesser skilled players. These results provide strong evidence for a perceptual encoding advantage for experts attributable to chess experience, rather than to a general perceptual or memory superiority.

Marketing/Advertising

Eye tracking can aid in the assessment of ad effectiveness in such applications as copy testing in print, images, video, or graphics, and in disclosure research involving perception of fine print within print media and within television displays. Eye tracking can provide insight into how the consumer disperses visual attention over different forms of advertising.

Applied research organizations may routinely examine the eye movements of consumers as they look at advertisements, however, this work tends to be proprietary (Rayner, Rotello, Stewart, Keir, & Duffy, 2001).

Copy Testing

A particularly good example of analysis of eye movements over advertisements in the Yellow Pages™ is given by Lohse (1997). In this experiment, eye movement data is collected while consumers chose businesses from telephone directories. The study addresses (1) what particular features cause people to notice an ad, (2) whether people view ads in any particular order, and (3) how viewing time varies as a function of particular ad features. Eye movement analysis revealed that, consistent with previous findings, ad size, graphics, color, and copy all influence attention to advertisements.

Print Advertising

In a study of consumers' visual attention over print advertisements conducted by Rosbergen, Wedel, and Pieters (1990), an eye tracker was used to gain insight into attentive processes over repeated exposure to print advertisements. The authors explore the phenomenon of repeated advertising's "wearout", i.e., consumers' diminishing attentional devotion to ads with increased repetition. Analyses show that whereas duration decreases and attention onset accelerates during each additional exposure to the print ad, the attentional scanpath remains constant across advertising repetitions and across experimentally varied conditions.

In a recent study of eye movements over advertisements, Wedel and Pieters (2000) report that across two magazines, fixations to the pictorial and the brand systematically promote accurate brand memory, but text fixations do not. Brand surface has a particularly prominent effect. The more information is extracted from an ad during fixations, the shorter the latency of brand memory is (e.g., recollection is faster following longer fixations). A systematic recency effect was found: when subjects are exposed to an ad later, they tend to identify it better. The effect of the ad's location on the right or left of the page depends on the advertising context.

Considering text and pictorial information in advertisement, Rayner et al. (2001) performed a study where viewers looked at print advertisements as their eye movements were recorded. Rayner et al. found that viewers tended to spend more time looking at the text than the picture part of the ad, though they did spend more time looking at the type of ad they were instructed to pay attention to. Fixation durations and saccade lengths were both longer on the picture part of the ad than the text, but more fixations were made on the text regions. Viewers did not alternate fixations between the text and picture part of the ad, but they tended to read the large print, then the smaller print, and then they looked at the picture (although some viewers did an initial cursory scan of the picture).

Rayner et al. (2001) suggest that the presented data have some striking implications for applied research and adver-

tisement development and indicate that consumers may be paying much more attention to the text in ads than previously thought. Rayner et al. note, however, that the data also indicates that while an advertisement captures and holds participants' attention, this may be caused by the instructions given to those participants. Rayner et al. suggest that participants' goals must be considered in future research conducted by advertising agencies and researchers in the area.

Computer Science

Following the hierarchy of eye tracking systems given earlier, this section focuses on two types of interactive applications: selective and gaze-contingent. The former approach uses an eye tracker as an input device, similar in some ways to a mouse. The latter gaze-contingent application is a type of display system wherein the information presented to the viewer is generally manipulated to match human processing capability, often matching foveo-peripheral perception in real-time. It should be noted that a good deal of the work discussed in this section is based on conference proceedings.

Selective Systems

Interactive uses of eye trackers typically employ gaze as a pointing modality, e.g., using gaze in a similar manner to a mouse pointer. Prominent applications involve selection of interface items (menus, buttons, etc.), as well as selection of objects or areas in Virtual Reality (VR). A prototypical application of gaze as an interactive modality is eye typing, particularly for handicapped users. Other uses involve gaze as an indirect pointing modality, for example as a deictic reference in the context of collaborative systems, or as an indirect pointing aid in user interfaces. Diagnostic uses of eye trackers are being adopted for usability studies, i.e., testing the effectiveness of interfaces as evidenced by where users look on the display.

Eye-Based Interaction. One of the first eye-based interactive systems, introduced by Jacob (1990), demonstrated an intelligent gaze-based informational display. In Jacob's system a text window would scroll to show information on visually selected items. Jacob's paper was one of the first to use eye tracking technology interactively and is well known for its identification of an important problem in eye-based interactive systems: the Midas Touch problem. Essentially, if the eyes are used in a manner similar to a mouse, a difficulty arises in determining intended activation of foveated features (the eyes do not register button clicks!). To avoid the Midas Touch problem, Jacob discusses several possible solutions including blinks, finally promoting the use of dwell time to act as a selection mechanism.

Another early eye-tracked interactive system was presented by Starker and Bolt (1990). This system provided the user with gaze-controlled navigation in a three-dimensional graphics world. The graphical environment contained story world characters who responded in interesting ways to the user's gaze. Fixations activated character "behaviors", since

the system would maintain and increase the user's visual interest level of a fixated character. When fixated, characters would blush and/or provide a verbal narrative. Unlike Jacob's use of dwell time, in this system dwell time was used to zoom into the graphics world.

Recently, Tanriverdi and Jacob (2000) presented a new gaze-based interactive system, this time with gaze acting as a selective mechanism in VR. In this system Tanriverdi and Jacob compared eye-based interaction with hand-based interaction and found that performance with gaze selection was significantly faster than hand pointing, especially in environments where objects were placed far away from the user's location. In contrast, no performance difference was found in "close" environments. Furthermore, while pointing speed may increase with gaze selection, there appears to be a cognitive tradeoff for this gain in efficiency: subjects had more difficulty recalling locations they interacted with when using gaze-based selection than when using hand selection.

The archetypical gaze-based pointing application is eye typing. Eye typing is and has been a useful communication modality for the severely handicapped. Most eye typing systems are implemented by presenting the user with a virtual keyboard, either on a typical computer monitor, or in some cases projected onto a wall. Based on analysis of tracked gaze, the system decides which letter the user is looking at and decides (e.g., by dwell time) whether or not to type this letter. The system may provide feedback to the user by either visual or auditory means, or a combination of both. Majaranta and Raiha (2002) provide an excellent survey of additional selection and feedback techniques employed in eye typing systems.

Gaze-based communication systems such as those featuring eye typing offer certain (sometimes obvious) advantages but are also problematic. Gaze may provide an often faster pointing modality than a mouse or other pointing device, especially if the targets are sufficiently large. However, gaze location is not as precise as with a mouse since the fovea limits the accuracy of the measured point of regard. Another significant problem is accuracy of the eye tracker. Following initial calibration, eye tracker accuracy may exhibit significant drift, where the measured point of regard gradually falls off from the actual point of gaze. Together with the Midas Touch problem, drift remains a significant problem for gaze input.

Zhai, Morimoto, and Ihde (1999) take another approach to gaze-based interaction and test the use of gaze as a sort of predictive pointing aid rather than a direct effector of selection. This is a particularly interesting and significant departure from "eye pointing" since this strategy is based on the assertion that loading of the visual perception channel with a motor control task seems fundamentally at odds with users' natural mental model in which the eye searches for, and takes in information, while coordinating with the hand for manipulation of external objects. In their paper on Manual Gaze Input Cascaded (MAGIC) Pointing, Zhai et al. present an acceleration technique where a (two-dimensional) cursor is warped to the vicinity of a fixated target. The acceleration is either immediately triggered by eye movement (liberal

MAGIC pointing), or delayed until the mouse moves as well (conservative MAGIC pointing). Zhai et al. report that while the speed advantage is not obvious over manual (mouse) pointing, almost all users subjectively felt faster with either MAGIC pointing technique.

Usability. Besides the use of gaze for interactive means, diagnostic eye tracking is gaining acceptance within the Human-Computer Interaction (HCI) community as another means to test usability of an interface. It is believed that eye movements can significantly enhance the observation of users' strategies while using computer interfaces (Goldberg & Kotval, 1999). Among various experiments, eye movements have been used to evaluate the grouping of tool icons, compare gaze-based and mouse interaction techniques (Sibert & Jacob, 2000), evaluate the organization of click-down menus, and more recently to test the organization of web pages.

Byrne, Anderson, Douglass, and Matessa (1999) tested the arrangement of items during visual search of click-down menus contrasting two computational cognitive models designed to predict latency, accuracy, and ease of learning for a wide variety of HCI-related tasks: EPIC (Executive Process Interactive Control) (Kieras & Meyer, 1995) and ACT-R (Adaptive Control of Thought-Rational) (J. R. Anderson, personal communication, July 2002). The EPIC architecture provides a general framework for simulating a human interacting with their environment to accomplish a task (Hornof & Kieras, 1997). ACT-R is a framework for understanding human cognition whose basic claim is that cognitive skill is composed of production rules (Anderson, 1993). These models specifically predict, in different ways, the relationship between eye and mouse movement. A visual search experiment, conducted by Byrne et al., presented a target item prior to display of a click-down menu. Based on the number of fixations, both models were supported since in some cases visual search exhibited exclusively top-down search (ACT-R) and in others both top-down and random patterns were observed (EPIC). This particular study is informative for two reasons. First, it shows the importance of a good model which can be used for designing user interfaces. Lacking such a model can lead to a design based on inaccurate assumptions. Second, the paper points out that eye tracking is an effective usability modality and that a new model of visual search over menus is needed.

In a usability study of web pages, Goldberg, Stimson, Lewnstein, Scott, and Wichansky (2002) derive specific recommendations for a prototype web interface tool and discuss gaze-based evaluation of web pages in which the system permits free navigation across multiple pages. This is a significant advancement for usability studies of web browsers since prior to this study recording of gaze over multiple web pages has been difficult due to the synchronization of gaze with windows that scroll or hide from view.

Collaborative Systems. Gaze can also be utilized to aid multiparty communication in collaborative systems. In Vertegaal's (1999) GAZE Groupware system, an eye tracker

is used to convey gaze direction in a multiparty teleconferencing and document sharing system, providing a solution to two problems in multiparty mediated communication and collaboration: knowing who is talking to whom, and who is talking about what. The system displays 2D images of remotely located participants in a virtual world. These images rotate to depict gaze direction alleviating the problem of turn-taking in multiparty communication systems. Furthermore, a gaze-directed “lightspot” is shown over a shared document indicating the users’ fixated regions and thereby providing a deictic (“look at this”) reference.

Gaze-Contingent Displays

In general, eye-based interactive applications can be thought of as selective, since gaze is used to select or point to some aspect of the display, whether it is two-dimensional (e.g., desktop), collaborative, or immersive (such as a virtual environment). Mixing both directly interactive and indirectly “passive” usage styles of gaze are gaze-contingent displays. Here, gaze is used not so much as a pointing device, but rather as a passive indicator of gaze. Given the user’s point of regard, a system can tailor the display so that the most informative details of the display are generated at the point of gaze but are degraded in some way in the periphery. The purpose of these displays is usually to minimize bandwidth requirements, as in video telephony applications, or in graphical applications where complex data sets cannot be fully displayed in real-time, by matching the display’s spatial frequency to functional visual acuity across the visual field.

When evaluating Gaze-Contingent Displays (GCDs), it is often necessary to distinguish between two main types of influences of fixation-dependent changes: those affecting perception and those affecting performance. As a general rule, perception is more sensitive than performance. That is, it may be possible to degrade a display to a quite noticeable effect without necessarily degrading performance. In either case, one of the main difficulties that must be addressed is the latency of the system. Without predictive capabilities, most gaze-contingent displays will lag behind the user’s gaze somewhat, usually by a constant amount of time proportional to both measurement of gaze (which may take up to the time it takes to display a single frame of video, typically 16 ms for a 60 Hz video-based tracker), and the subsequent time it takes to refresh the gaze-contingent display (this may take another 33 ms for a system with update rate of 30 frames per second).

Two main types of gaze-contingent applications are discussed: screen-based and model-based. The former deals with image (pixel) manipulation, while the latter is concerned with the manipulation of graphical objects or models prior to rendering.

Screen-Based Displays. Loschky and McConkie (2000) conducted an experiment on a gaze-contingent display investigating spatial, resolutional, and temporal parameters affecting perception and performance. Two key issues addressed are the timing of GCDs and the detectability of the peripherally degraded component of the GCD. In all experiments,

monochromatic photographic scenes were used as stimuli with a circular, high-resolution window surrounded by a degraded peripheral region. In one facet of the experiment, it was found that for an image change to go undetected, it must be started within 5 ms after the end of the eye movement. Detection likelihood rose quickly beyond that point. In another facet of the study concerning detection of peripheral degradation, results showed that the least peripheral degradation went undetected even at the smallest window size (2°), where the opposite was true with the highest level of degradation—it was quite detectable at even the largest window size (5°). It was found that the generation of an imperceptible GCD was quite difficult in comparison to the generation of a GCD which does not deteriorate performance. While greater delays (e.g., 15 ms) and greater degradation produce detectable visual artifacts, they appear to have minimal impact on performance of visual tasks when there is a 4.1° high-resolution area centered at the point of gaze.

Parkhurst, Culurciello, and Niebur (2000) investigated behavioral effects of a two-region gaze-contingent display. A central high-resolution region, varying from 1 to 15 degrees, was presented at the instantaneous center of gaze during a visual search task. Parkhurst et al. found that reaction time and accuracy co-varied as a function of the central region size and note this as a clear indicator of a strategic speed/accuracy tradeoff where participants favor speed in some conditions and accuracy in others. For small central region sizes, slow reaction times are accompanied by high accuracy. Conversely, for large central regions sizes, fast reaction times are accompanied by low accuracy. A secondary finding indicated that fixation duration varies as a function of central region size. For small central region sizes, participants tend to spend more time examining each fixation than under normal viewing conditions. For large central regions, fixation durations tend to be closer to normal. In agreement with reaction time and accuracy, fixation duration is approximately normal (comparable to that seen for uniform resolution displays) with a central region size of 5° .

By distinguishing the effects of gaze-contingent windowing and peripheral filtering, Reingold and Loschky (2002) (in this issue) have recently shown that peripheral target detection with GCDs may be negatively impacted by not only the loss of perceptual detail due to filtering, but also by the mere presence of the gaze-contingent window (the *window effect*). An important consideration in Reingold and Loschky’s study is the saliency of the gaze-contingent window boundary: whether the gaze-contingent window is sharply defined or smoothly gradated appears to have no effect on peripheral target acquisition.

For screen-based VR rendering the work of Watson et al. (1997) is particularly relevant. Watson et al. studied the effects of Level Of Detail (LOD) peripheral degradation on visual search performance, evaluating both spatial and chrominance detail degradation effects in Head Mounted Displays (HMDs). To sustain acceptable frame rates, two polygons were texture mapped in real-time to generate a high resolution inset within a low resolution display field. Watson et al. suggest that visual spatial and chrominance complexity can

be reduced by almost half without degrading performance.

In an approach similar to Watson et al.'s, Reddy (1998) used a view-dependent LOD technique to evaluate both perceptual effects and system performance gains. Reddy reported a perceptually modulated LOD system which affords a factor 4.5 improvement in frame rate.

For excellent reviews of Gaze-Contingent Displays, see Reingold, Loschky, McConkie, and Stampe (2002) as well as Parkhurst and Niebur (2002).

Model-Based Graphical Displays. As an alternative to the screen-based peripheral degradation approach, model-based methods aim at reducing resolution by directly manipulating the model geometry prior to rendering. The technique of simplifying the resolution of geometric objects as they recede from the viewer, as originally proposed by Clarke (1976), is now standard practice, particularly in real-time applications such as VR (Vince, 1995). Clarke's original criteria of using the projected area covered by the object for descending the object's LOD hierarchy is still widely used today. However, as Clarke suggested, the LOD management typically employed by these polygonal simplification schemes relies on pre-computed fine-to-coarse hierarchies of an object. This leads to uniform, or *isotropic*, object resolution degradation.

A gaze-contingent model-based adaptive rendering scheme was proposed by Ohshima, Yamamoto, and Tamura (1996), where three visual characteristics were considered: central/peripheral vision, kinetic vision, and fusional vision. The LOD algorithm generated isotropically degraded objects at different visual angles. Although the use of a binocular eye tracker was proposed, the system as discussed used only head tracking as a substitute for gaze tracking.

Isotropic object degradation is not always desirable, especially when viewing large objects at close distances. In this case, traditional LOD schemes will display an LOD mesh at its full resolution even though the mesh may cover the entire field of view. Since acute resolvability of human vision is limited to the central 5°, object resolution need not be uniform. This is the central tenet of gaze-contingent systems.

Numerous multiresolution mesh modeling techniques suitable for gaze-contingent viewing have recently been developed (Zorin & Schröder, 2000). Due to the advancements of multiresolution modeling techniques and to the increased affordability of eye trackers, it is now becoming feasible to extend the LOD approach to gaze-contingent displays, where models are rendered *nonisotropically*.

An early example of a nonisotropic model-based gaze-contingent system, where gaze direction is directly applied to the rendering algorithm, was presented by Levoy and Whitaker (1990), where a spatially adaptive near real-time ray tracer for volume data displayed an eye-slaved ROI by modulating both the number of rays cast per unit area on the image plane and the number of samples drawn per unit length along each ray as a function of local retinal acuity. The ray-traced image was sampled by a nonisotropic convolution filter to generate a 12° foveal ROI within a 20° mid-resolution transitional region. Based on preliminary estimates, Levoy and Whitaker suggested a reduction in image generation time

by a factor of up to 5.

For environments containing significant topological detail, such as virtual terrains, rendering with multiple levels of detail, where the level is based on user position and gaze direction, is essential to provide an acceptable combination of surface detail and frame rate (Danforth, Duchowski, Geist, & McAliley, 2000). Danforth et al. (2000) used an eye tracker as an indicator of gaze in a gaze-contingent multiresolution terrain navigation environment. A surface, represented as a quadrilateral mesh, was divided into fixed-size (number of vertices) sub-blocks, allowing rendering for variable LOD on a per-sub-block basis. Resolution level was chosen per sub-block, based on viewer distance and direction of gaze. To exaggerate the gaze-contingent effect, fractal mountains disappear when not in view.

More recent work on gaze-contingent LOD modeling has been carried out by Luebke and Erikson (1997). Luebke and Erikson present a view-dependent LOD technique suitable for gaze-contingent rendering. While simplification of individual geometric objects is discussed in their work, it appears the strategy is ultimately directed toward solving the interactive "walkthrough" problem (Funkhouser & Séquin, 1993). In this application, the view-dependent LOD technique seems more suitable to the (possibly) gaze-contingent rendering of an entire scene or environment. Recently, Luebke, Hallen, Newfield, and Watson (2000) have developed a gaze-directed LOD technique to facilitate the gaze-contingent display of geometric objects. To test their rendering approach Luebke et al. employed a table-mounted monocular eye tracker to measure the viewer's real-time location of gaze over a desktop display.

A new object-based LOD method has been developed by Murphy and Duchowski (2001). The technique is similar to Luebke and Erikson's and to Ohshima et al.'s, where objects are modeled for gaze-contingent viewing. Unlike the approach of Ohshima et al., resolution degradation is applied nonisotropically. The spatial degradation function for LOD selection differs significantly from the area-based criteria originally proposed by Clarke. Instead of evaluating the screen coverage of the projected object, the degradation function is based on the evaluation of visual angle in world coordinates. A three-dimensional spatial degradation function was obtained from human subject experiments in an attempt to imperceptibly display spatially degraded geometric objects. System performance measurements indicate an approximate overall 10-fold average frame rate improvement during gaze-contingent viewing.

Another interesting approach to gaze-contingent modeling for real-time graphics rendering was taken by O'Sullivan and Dingliana (2001) and O'Sullivan, Dingliana, and Howlett (2002). Instead of degrading the spatial resolution of peripherally located geometric objects, O'Sullivan and Dingliana (2001) consider a degradable collision handling mechanism to effectively limit object collision resolution outside a foveal ROI. When the viewer is looking directly at a collision, it is given higher (computational) priority than collisions occurring in the periphery. Object collisions given higher priority are allocated more processing time so that

the contact model and resulting response is more believable. O'Sullivan and Dingliana tested viewers' sensitivity to collision resolution (size of gap between colliding objects) at eccentricity and noted a significant fall-off in detection accuracy at about 4° visual angle. Based on their previous psychophysical findings, O'Sullivan et al. (2002) developed a gaze-contingent collision handling system. Two variants of the gaze-contingent system were compared, each containing a high-priority ROI wherein collisions were processed at higher resolution than outside the ROI. In the tracked case, the high-priority ROI was synchronized to the viewer's tracked gaze position, while in the random case, the ROI position was determined randomly every 5 frames. O'Sullivan et al. report an overall improvement in the perception of the tracked simulation.

Conclusion

As the present review demonstrates, eye trackers have traditionally shown themselves to be valuable in *diagnostic* studies of reading and other information processing tasks. The diagnostic use of an eye tracker, as exemplified by the research reviewed here, can be considered the eye tracker's mainstay application at the present time and probably in its near future. As also documented in this review, because eye trackers are able to provide a quantitative measure of real-time overt attention, they are valuable components of *interactive* systems. Several key problems and solutions have been identified in interactive eye tracking systems (e.g., Midas Touch problem). For disabled users, an eye tracking interface may be an indispensable form of communication (e.g., eye typing). In a more general interactive setting, however, there is some debate whether it makes sense to overload a perceptual organ (the eye) by a motor task (e.g., mouse-like pointing). As an auxiliary interface modality, the eye tracker may serve better as an indirect indicator of the user's future selective intent (i.e., serving as a mouse pointing accelerator rather than the mouse pointer itself). This type of indirect use of eye movement may also be exploited in gaze-contingent scenarios where gaze is not used as an interaction modality *per se* (i.e., no particular action is performed), rather, gaze is used to alter the scene in some manner to either manage computational or bandwidth resources or to test human visual perception or cognitive skill. In the latter sense, while the beginning of the fourth eye tracking era may coincide with an increased number of interactive applications driven by increasingly sophisticated imagery (e.g., real-time video or Virtual Reality), gaze-contingent applications are simply extensions of gaze-contingent paradigms of the third era. However, due to the richness and flexibility of graphical environments, it is likely that novel interactive uses of eye trackers within increasingly complex contextual situations will allow investigation of a broader class of applications than seen in the past.

Acknowledgments

I am grateful for helpful suggestions concerning the manuscript made by Keith Rayner and a second anonymous

referee.

References

- Alloppenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, 38(4), 419-439.
- Anders, G. (2001). Pilot's Attention Allocation During Approach and Landing—Eye- and Head-Tracking Research in an A330 Full Flight Simulator. In *International Symposium on Aviation Psychology (ISAP)*. Columbus, OH. (URL: <http://www.geerdanders.de/literatur/2001_ohio.html>. Last accessed 07/12/02.)
- Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-Specific Neural Activity in the Primate Prefrontal Cortex. *Neurophysiology*, 84, 451-459.
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory Representations in Natural Tasks. *Journal of Cognitive Neuroscience*, 7(1), 66-80.
- Bertera, J. H., & Rayner, K. (2000). Eye Movements and the Span of the Effective Stimulus in Visual Search. *Perception & Psychophysics*, 62(3), 576-585.
- Buswell, G. T. (1935). *How People Look At Pictures*. Chicago, IL: University of Chicago Press.
- Byrne, M. D., Anderson, J. R., Douglass, S., & Matessa, M. (1999). Eye Tracking the Visual Search of Click-Down Menus. In *Human Factors in Computing Systems: CHI '99 Conference Proceedings* (p. 402-409). ACM Press.
- Chapman, P. R., & Underwood, G. (1998). Visual Search of Dynamic Scenes: Event Types and the Role of Experience in Viewing Driving Situations. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (p. 369-394). Oxford, England: Elsevier.
- Clarke, J. H. (1976). Hierarchical Geometric Models for Visible Surface Algorithms. *Communications of the ACM*, 19(10), 547-554.
- Cooper, R. M. (1974). The Control of Eye Fixation by the Meaning of Spoken Language: A New Methodology for the Real-Time Investigation of Speech Perception, Memory, and Language Processing. *Cognitive Psychology*, 6, 84-107.
- Crundall, D. E., Underwood, G., & Chapman, P. R. (1998). How Much Do Novice Drivers See? The Effects of Demand on Visual Search Strategies in Novice and Experienced Drivers. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (p. 395-418). Oxford, England: Elsevier.
- Danforth, R., Duchowski, A., Geist, R., & McAliley, E. (2000). A Platform for Gaze-Contingent Virtual Environments. In *Smart Graphics (Papers from the 2000 AAAI Spring Symposium, Technical Report SS-00-04)* (p. 66-70). Menlo Park, CA.
- DeCarlo, D., & Santella, A. (2002). Stylization and Abstraction of Photographs. *Transaction on Graphics*, 21(3), 769-776.
- Dishart, D. C., & Land, M. F. (1998). The Development of the Eye Movement Strategies of Learner Drivers. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (p. 419-430). Oxford, England: Elsevier.
- Doll, T. J. (1993). Preattentive Processing in Visual Search. In *Proceedings of the Human Factors and Ergonomics Society, 37th Annual Meeting* (p. 1291-1249). Santa Monica, CA.

- Doll, T. J., Whorter, S. W., & Schmieder, D. E. (1993). Simulation of Human Visual Search in Cluttered Backgrounds. In *Proceedings of the Human Factors and Ergonomics Society, 37th Annual Meeting* (p. 1310-1314). Santa Monica, CA.
- Drury, C. G., Gramopadhye, A. K., & Sharit, J. (1997). Feedback Strategies for Visual Inspection in Airframe Structural Inspection. *International Journal of Industrial Ergonomics, 19*, 333-344.
- Duchowski, A., Medlin, E., Gramopadhye, A., Melloy, B., & Nair, S. (2001). Binocular Eye Tracking in VR for Visual Inspection Training. In *Virtual Reality Software & Technology (VRST)*. Banff, AB, Canada.
- Duchowski, A. T. (2002). *Eye Tracking Methodology: Theory & Practice*. London, UK: Springer-Verlag (London), Inc.
- d'Ydewalle, G., Desmet, G., & Van Rensbergen, J. (1998). Film Perception: The Processing of Film Cuts. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (p. 357-368). Oxford, England: Elsevier.
- Findlay, J. M. (1997). Saccade Target Selection During Visual Search. *Vision Research, 37*(5), 617-631.
- Findlay, J. M., & Gilchrist, I. D. (1998). Eye Guidance and Visual Search. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (p. 295-312). Oxford, England: Elsevier.
- Findlay, J. M., & Walker, R. (1999). A Model of Saccade Generation Based on Parallel Processing and Competitive Inhibition. *Behavioral and Brain Sciences, 22*(4), 661-721.
- Funkhouser, T. A., & Séquin, C. H. (1993). Adaptive Display Algorithm for Interactive Frame Rates During Visualization of Complex Virtual Environments. In *Computer Graphics (SIGGRAPH '93)*. New York, NY.
- Goldberg, J., Stimson, M., Lewnstein, M., Scott, N., & Wichansky, A. (2002). Eye Tracking in Web Search Tasks: Design Implications. In *Eye Tracking Research & Applications (ETRA) Symposium*. New Orleans, LA.
- Goldberg, J. H., & Kotval, X. P. (1999). Computer Interface Evaluation Using Eye Movements: Methods and Constructs. *International Journal of Industrial Ergonomics, 24*, 631-645.
- Graeber, D. A., & Andre, A. D. (1999). Assessing Visual Attention of Pilots While Using Electronic Moving Maps for Taxiing. In *International Symposium on Aviation Psychology (ISAP)* (p. 791-796). Columbus, OH.
- Greene, H. H., & Rayner, K. (2001). Eye Movements and Familiarity Effects in Visual Search. *Vision Research, 41*(27), 3763-3773.
- Hayhoe, M. M., Ballard, D. H., Triesch, J., Shinoda, H., Rogriguez, P. A., & Sullivan, B. (2002). Vision in Natural and Virtual Environments (Invited Paper). In *Eye Tracking Research & Applications (ETRA) Symposium*. New Orleans, LA.
- Henderson, J. M. (1992). Object Identification in Context: The Visual Processing of Natural Scenes. *Canadian Journal of Psychology, 46*(3), 319-341.
- Henderson, J. M., & Hollingworth, A. (1998). Eye Movements During Scene Viewing: An Overview. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (p. 269-294). Oxford, England: Elsevier.
- Ho, G., Scialfa, C. T., Caird, J. K., & Graw, T. (2001). Visual Search for Traffic Signs: The Effects of Clutter, Luminance, and Aging. *Human Factors, 43*(2), 194-207.
- Hornof, A. J., & Kieras, D. E. (1997). Cognitive Modeling Reveals Menu Search is Both Random and Systematic. In *Human Factors in Computing Systems: CHI '97 Conference Proceedings* (p. 107-144). Atlanta, GA.
- Hughes, H. C., Nozawa, G., & Kitterle, F. (1996). Global Precedence, Spatial Frequency Channels, and the Statistics of Natural Images. *Journal of Cognitive Neuroscience, 8*(3), 197-230.
- Jacob, R. J. (1990). What You Look at is What You Get: Eye Movement-Based Interaction Techniques. In *Human Factors in Computing Systems: CHI '90 Conference Proceedings* (p. 11-18). ACM Press.
- Kanizsa, G. (1976). Subjective Contours. *Scientific American, 234*(4), 48-52,138.
- Kennedy, A. (1992). The Spatial Coding Hypothesis. In K. Rayner (Ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading* (p. 379-396). New York, NY: Springer-Verlag. (Springer Series in Neuropsychology)
- Kieras, D., & Meyer, D. E. (1995). *An Overview of the EPIC Architecture for Cognition and Performance with Application to Human-Computer Interaction* (EPIC Tech. Rep. No. 5 No. TR-95/ONR-EPIC-5). Ann Arbor, University of Michigan, Electrical Engineering and Computer Science Department.
- Kroll, J. F. (1992). Making a Scene: The Debate about Context Effects for Scenes and Sentences. In K. Rayner (Ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading*. Springer-Verlag. (Springer Series in Neuropsychology)
- 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), 1307-1432.
- Land, M. F., & Hayhoe, M. (2001). In What Ways Do Eye Movements Contribute to Everyday Activities. *Vision Research, 41*(25-26), 3559-3565. ((Special Issue on *Eye Movements and Vision in the Natural World*, with most contributions to the volume originally presented at the 'Eye Movements and Vision in the Natural World' symposium held at the Royal Netherlands Academy of Sciences, Amsterdam, September 2000))
- Levoy, M., & Whitaker, R. (1990). Gaze-Directed Volume Rendering. In *Computer Graphics (SIGGRAPH '90)* (p. 217-223). New York, NY.
- Loftus, G. R. (1981). Tachistoscopic Simulations of Eye Fixations on Pictures. *Journal of Experimental Psychology: Human Learning and Memory, 7*(5), 369-376.
- Lohse, G. L. (1997). Consumer Eye Movement Patterns on Yellow Pages Advertising. *Journal of Advertising, 26*(1), 61-73.
- Loschky, L. C., & McConkie, G. W. (2000). User Performance With Gaze Contingent Multiresolutional Displays. In *Eye Tracking Research & Applications Symposium* (p. 97-103). Palm Beach Gardens, FL.
- Luebke, D., & Erikson, C. (1997). View-Dependent Simplification of Arbitrary Polygonal Environments. In *Computer Graphics (SIGGRAPH '97)*. New York, NY.
- Luebke, D., Hallen, B., Newfield, D., & Watson, B. (2000). *Perceptually Driven Simplification Using Gaze-Directed Rendering* (Tech. Rep. No. CS-2000-04). University of Virginia.
- Majaranta, P., & Raiha, K.-J. (2002). Twenty Years of Eye Typing: Systems and Design Issues. In *Eye Tracking Research & Applications (ETRA) Symposium*. New Orleans, LA.
- McConkie, G. W., & Rayner, K. (1975). The Span of the Effective Stimulus During a Fixation in Reading. *Perception & Psychophysics, 17*, 578-586.
- Megaw, E. D., & Richardson, J. (1979). Eye Movements and Industrial Inspection. *Applied Ergonomics, 10*(3), 145-154.

- Molnar, F. (1981). About the Role of Visual Exploration in Aesthetics. In H. Day (Ed.), *Advances in Intrinsic Motivation and Aesthetics*. New York, NY: Plenum Press.
- Murphy, H., & Duchowski, A. T. (2001). Gaze-Contingent Level Of Detail. In *Eurographics (short presentations)*. Manchester, UK.
- Necker, L. A. (1832). Observations on Some Remarkable Optical Phaenomena Seen in Switzerland, and on an Optical Phaenomenon Which Occurs on Viewing a Figure or a Crystal or Geometrical Solid. *Philosophical Magazine and Journal of Science*, 1, 329-337. (Third Series.)
- Noton, D., & Stark, L. (1971a). Eye Movements and Visual Perception. *Scientific American*, 224, 34-43.
- Noton, D., & Stark, L. (1971b). Scanpaths in Saccadic Eye Movements While Viewing and Recognizing Patterns. *Vision Research*, 11, 929-942.
- Ohshima, T., Yamamoto, H., & Tamura, H. (1996). Gaze-Directed Adaptive Rendering for Interacting with Virtual Space. In *Proceedings of VRAIS'96* (p. 103-110). New York, NY.
- O'Sullivan, C., & Dingliana, J. (2001). Collisions and Perception. *ACM Transactions on Graphics*, 20(3).
- O'Sullivan, C., Dingliana, J., & Howlett, S. (2002). Gaze-Contingent Algorithms for Interactive Graphics. In J. Hyöna, R. Radach, & H. Deubel (Eds.), *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movement Research*. Oxford, England: Elsevier Science.
- Ottati, W. L., Hickox, J. C., & Richter, J. (1999). Eye Scan Patterns of Experienced and Novice Pilots During Visual Flight Rules (VFR) Navigation. In *Proceedings of the Human Factors and Ergonomics Society, 43rd Annual Meeting*. Minneapolis, MN.
- Özyurt, J., DeSouza, P., West, P., Rutschmann, R., & Greenlee, M. W. (2001). Comparison of Cortical Activity and Oculomotor Performance in the Gap and Step Paradigms. In *European Conference on Visual Perception (ECVP)*. Kusadasi, Turkey.
- Parkhurst, D., Culurciello, E., & Niebur, E. (2000). Evaluating Variable Resolution Displays with Visual Search: Task Performance and Eye Movements. In *Eye Tracking Research & Applications Symposium* (p. 105-109). Palm Beach Gardens, FL.
- Parkhurst, D. J., & Niebur, E. (2002). Variable Resolution Displays: A Theoretical, Practical, and Behavioral Evaluation. *Human Factors*. (In Press.)
- Pelz, J. B., Canosa, R., & Babcock, J. (2000). Extended Tasks Elicit Complex Eye Movement Patterns. In *Eye Tracking Research & Applications (ETRA) Symposium* (p. 37-43). Palm Beach Gardens, FL.
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the Detection of Signals. *Experimental Psychology: General*, 109(2), 160-174.
- Rayner, K. (1975). The Perceptual Span and Peripheral Cues in Reading. *Cognitive Psychology*, 7, 65-81.
- Rayner, K. (Ed.). (1992). *Eye Movements and Visual Cognition: Scene Perception and Reading*. New York, NY: Springer-Verlag. (Springer Series in Neuropsychology)
- Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124(3), 372-422.
- Rayner, K., & Bertera, J. H. (1979). Reading Without a Fovea. *Science*, 206, 468-469.
- Rayner, K., & Pollatsek, A. (1992). Eye Movements and Scene Perception. *Canadian Journal of Psychology*, 46(3), 342-376.
- Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating Text and Pictorial Information: Eye Movements When Looking at Print Advertisements. *Journal of Experimental Psychology: Applied*, 7(3), 219-226.
- Recarte, M. A., & Nunes, L. M. (2000). Effects of Verbal and Spatial-Imagery Tasks on Eye Fixations While Driving. *Journal of Experimental Psychology: Applied*, 6(1), 31-43.
- Reddy, M. (1998). Specification and Evaluation of Level of Detail Selection Criteria. *Virtual Reality: Research, Development and Application*, 3(2), 132-143.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a Model of Eye Movement Control in Reading. *Psychological Review*, 105(1), 125-157.
- Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2002). Visual Span in Expert Chess Players: Evidence from Eye Movements. *Psychological Science*. (In Press.)
- Reingold, E. M., & Loschky, L. C. (2002). Saliency of Peripheral Targets in Gaze-Contingent Multi-Resolutional Displays. *Behavior Research Methods, Instruments, & Computers*. (In Press.)
- Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2002). Gaze-Contingent Multi-Resolutional Displays: An Integrative Review. *Human Factors*. (In Press.)
- Robinson, D. A. (1968). The Oculomotor Control System: A Review. *Proceedings of the IEEE*, 56(6), 1032-1049.
- Rosbergen, E., Wedel, M., & Pieters, R. (1990). *Analyzing Visual Attention to Repeated Print Advertising Using Scanpath Theory* (Tech. Rep.). University Library Groningen, SOM Research School. (Technical Report 97B32.)
- Schoonard, J. W., Gould, J. D., & Miller, L. A. (1973). Studies of Visual Inspection. *Ergonomics*, 16(4), 365-379.
- Sibert, L. E., & Jacob, R. J. (2000). Evaluation of Eye Gaze Interaction. In *Human Factors in Computing Systems: CHI 2000 Conference Proceedings*. ACM Press.
- Smeets, J. B. J., Hayhoe, H. M., & Ballard, D. H. (1996). Goal-Directed Arm Movements Change Eye-Head Coordination. *Experimental Brain Research*, 109, 434-440.
- Snodderly, D. M., Kagan, I., & Gur, M. (2001). Selective Activation of Visual Cortex Neurons by Fixational Eye Movements: Implications for Neural Coding. *Visual Neuroscience*, 18, 259-277.
- Solso, R. L. (1999). *Cognition and the Visual Arts* (3rd ed.). Cambridge, MA: The MIT Press.
- Starker, I., & Bolt, R. A. (1990). A Gaze-Responsive Self-Disclosing Display. In *Human Factors in Computing Systems: CHI '90 Conference Proceedings* (p. 3-9). ACM Press.
- Tanriverdi, V., & Jacob, R. J. K. (2000). Interacting with Eye Movements in Virtual Environments. In *Human Factors in Computing Systems: CHI 2000 Conference Proceedings* (p. 265-272). ACM Press.
- Todd, S., & Kramer, A. F. (1993). Attentional Guidance in Visual Attention. In *Proceedings of the Human Factors and Ergonomics Society, 37th Annual Meeting* (p. 1378-1382). Santa Monica, CA.
- Treisman, A., & Gelade, G. (1980). A Feature Integration Theory of Attention. *Cognitive Psychology*, 12, 97-136.
- Van Orden, K. F., & DiVita, J. (1993). Highlighting with Flicker. In *Proceedings of the Human Factors and Ergonomics Society, 37th Annual Meeting* (p. 1300-1304). Santa Monica, CA.

- Velichkovsky, B., Pomplun, M., & Rieser, J. (1996). Attention and Communication: Eye-Movement-Based Research Paradigms. In W. H. Zangemeister, H. S. Stiehl, & C. Freksa (Eds.), *Visual Attention and Cognition* (p. 125-154). Amsterdam, Netherlands: Elsevier Science.
- Vertegaal, R. (1999). The GAZE Groupware System: Mediating Joint Attention in Mutiparty Communication and Collaboration. In *Human Factors in Computing Systems: CHI '99 Conference Proceedings* (p. 294-301). ACM Press.
- Vince, J. A. (1995). *Virtual Reality Systems*. Reading, MA: Addison-Wesley.
- Wang, M.-J. J., Lin, S.-C., & Drury, C. G. (1997). Training for strategy in visual search. *Industrial Ergonomics, 20*, 101-108.
- Watson, B., Walker, N., Hodges, L. F., & Worden, A. (1997). Managing Level of Detail through Peripheral Degradation: Effects on Search Performance with a Head-Mounted Display. *ACM Transactions on Computer-Human Interaction, 4*(4), 323-346.
- Wedel, M., & Pieters, R. (2000). Eye Fixations on Advertisements and Memory for Brands: A Model and Findings. *Marketing Science, 19*(4), 297-312.
- Wolfe, J. M. (1993). Guided Search 2.0: The Upgrade. In *Proceedings of the Human Factors and Ergonomics Society, 37th Annual Meeting* (p. 1295-1299). Santa Monica, CA.
- Wolfe, J. M. (1994). Visual Search in Continuous, Naturalistic Stimuli. *Vision Research, 34*(9), 1187-1195.
- Wolfe, J. M., & Gancarz, G. (1996). GUIDED SEARCH 3.0: A Model of Visual Search Catches Up With Jay Enoch 40 Years Later. In V. Lakshminarayanan (Ed.), *Basic and Clinical Applications of Vision Science* (p. 189-192). Dordrecht, Netherlands: Kluwer Academic.
- Wooding, D. (2002). Fixation Maps: Quantifying Eye-Movement Traces. In *Eye Tracking Research & Applications (ETRA) Symposium*. New Orleans, LA.
- Yarbus, A. L. (1967). *Eye Movements and Vision*. New York, NY: Plenum Press.
- Zhai, S., Morimoto, C., & Ihde, S. (1999). Manual and Gaze Input Cascaded (MAGIC) Pointing. In *Human Factors in Computing Systems: CHI '99 Conference Proceedings* (p. 246-353). ACM Press.
- Zorin, D., & Schröder, P. (2000). *Course 23: Subdivision for Modeling and Animation*. New York, NY. (URL: <<http://www.mrl.nyu.edu/dzorin/sig00course/>> (last accessed 12/30/00))