PERGAMON

Vision Research 41 (2001) 3587–3596

Vision Research

www.elsevier.com/locate/visres

Oculomotor behavior and perceptual strategies in complex tasks

Jeff B. Pelz *, Roxanne Canosa

Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY 14623, USA

Received 18 December 2000; received in revised form 19 June 2001

Abstract

While we know a great deal about the dynamics and characteristics of eye movements in relatively simple tasks performed under reduced laboratory conditions, we know less about oculomotor behavior in complex, multi-step tasks. Complex tasks are not necessarily difficult. Part of the transition from 'hard' to 'easy' in completing complex tasks is the gradual reduction in conscious effort required to complete the sub-tasks. We are interested in learning whether high-level perceptual strategies can aid that transition. In the past, subjects performed relatively simple tasks or the eye movements themselves were the instructed task. But outside the laboratory vision is a tool, not the task. To study the oculomotor system in its native mode, we developed a wearable eyetracker that allows natural eye, head and whole-body movements. Using the over-learned, common task of hand-washing, we measured the global characteristics of fixation duration, saccade amplitude, and the spatial distribution of fixation positions. An important observation was the emergence of higher-order perceptual strategies in the complex task: while most fixations were related to the immediate action, a small number of fixations were made to objects relevant only to future actions. Based on a control task that differed only in the high-level goal, we conclude that the *look-ahead fixations* represent a task-dependent strategy, not a general behavior elicited by the salience or conspicuity of objects in the environment. We propose that the strategy of looking ahead to objects of future relevance supports the conscious percept of an environment seamless in time as well as in space.

Keywords: Complex tasks; Oculomotor behavior; Perceptual strategies; Eye movements

1. Introduction

Despite the seeming ease with which we perceive the world around us, vision is a set of complex processes that unfold over time, constantly gathering and refining information about the environment, and leading to the perception of a complete, stable environment. These processes occur at a level below conscious awareness, so their complexities do not yield to introspective report. Part of the complexity of the human visual system is due to the constraints imposed on it during its evolution. Finite neural resources and the competing evolutionary pressures for high acuity and a wide field-of-view led to the compromise that is the human retina. The illusion of a high-resolution, wide field-of-view scene is made possible by a highly anisotropic retina with high acuity in only a small central region (the fovea), surrounded by a low-resolution, large field-of-view periphery. This foveal/peripheral design allows a small region of regard to be imaged at high resolution, while at the same time providing a large peripheral field-of-view at significantly lower resolution.

Because only a small part of the field is imaged with high acuity at any given time, a mechanism is needed to reorient the eyes to sample the environment even in the simplest case of a stationary observer viewing a static scene. If the observer is in motion, and/or the scene is dynamic, mechanisms to stabilize the retinal image between the reorienting eye movements are also necessary. The oculomotor system provides a rich suite of eye movements that provide these capabilities.

In many experiments designed to understand the oculomotor system, simple tasks were performed by stationary observers viewing static scenes, and subjects interacted with the surrounding environment only by verbal response or button-press. While such experi-

* Corresponding author. Tel.: + 1-716-475-2783; fax: + 1-716-475-5988.

E-mail address: pelz@cis.rit.edu (J.B. Pelz).

0042-6989/01/$ - see front matter © 2001 Published by Elsevier Science Ltd.
PII: S0042-6989(01)00245-0

ments have contributed much to our understanding of eye movement mechanisms, control structures, and metrics, they have provided little insight into human behaviors in the real world where vision is a tool, not a task in its own right. While it is tempting to understand visual perception by breaking it down into component sub-tasks, it is impossible to break down tasks with high-level cognitive components into meaningful elements without losing the very nature of the complex tasks under study (see, Collewijn, Steinman, Erkelens, Pizlo, & van der Steen, 1992; Kowler et al., 1992).

A task can present a static or dynamic scene to a static or dynamic observer and the observer can interact with the environment. This allows the state of observer and scene to be used to categorize tasks. Some experiments using static observers and static scenes have investigated high-level influences on the oculomotor system. In one such task, Buswell (1935) studied the differences in viewing patterns between trained artists and untrained subjects, reporting that the two groups showed significantly different eye movement patterns viewing the same images. Yarbus (1967) looked at the difference in oculomotor behaviors based on the instructions to the subjects. He recorded the eye movements of subjects as they viewed the same painting, but with different instructions defining the subject's task. He showed that the pattern of eye movements was not only dependent on the scene, but was clearly dependent on the instructed task.

Experiments with static observers viewing dynamic scenes have shown that the oculomotor system can make use of experience and expectations about the environment. Kowler and McKee (1987) reported anticipatory smooth pursuit eye movements when subjects knew the direction but not the time of onset or velocity of target motion. Kowler (1989) showed that subjects' smooth pursuit eye movements were influenced by visual and aural cues relevant to previous trails. While the experiments revealed sophisticated abilities of the oculomotor system when completing complex tasks, the instructed tasks were still the eye movements themselves.

Other insights into oculomotor behaviors are available by studying dynamic observers interacting with static scenes. In a study of subjects performing a task requiring complex coordination of the eye, head, and hand, Epelboim et al. (1997) had subjects perform one of two tasks in identical environments. Subjects were seated in front of a table containing a set of raised pegs that were illuminated in a sequential pattern as their eye, head, and hand movements were recorded. In one condition, subjects were instructed to tap the pegs in the indicated order; in the other condition, they were instructed to look at the pegs in the same order. Despite the fact that in the *tap* task the pegs were fixated, the eye/head dynamics differed significantly between the two conditions. The *look only* instruction may be unnatural because observers rarely look at a target without some information-gathering or guiding goal. Nevertheless, the oculomotor behaviors differed based solely on the instruction despite the fact that the demands on the oculomotor system were identical.

Studies with natural tasks performed by mobile observers interacting with the environment have begun to reveal oculomotor behaviors used in the real world. In a study by Land, Mennie, and Rusted (1999), subjects' eye movements were recorded as they made a pot of tea. They found that nearly all fixations were directly task-related.

Information-gathering eye movements formed the bulk of the fixations, with eye movements made to targets for manipulation about 600 ms before contact with the object (e.g. reaching to turn the water faucet on). A number of fixations were clearly not task-related, as when the subject looked about as the teakettle was filled at the sink. Land et al. (1999) reported that only a small fraction ($\sim 5\%$) of the fixations were irrelevant to the task.

In all of these tasks, low-level metrics of eye movements provided externally visible indicators of performance and task demands because of their role in maintaining foveal vision for tasks requiring high acuity. In natural environments, eye movements are made not only to regions requiring foveal acuity, but also toward task-relevant targets even when peripheral acuity would suffice. Such 'attentional' eye movements, made without conscious intervention, can reveal attentional mechanisms and may provide a window into *perceptual strategies* employed by observers performing complex tasks. Perceptual strategies are defined here as actions that do more than support the immediate task; they can simplify and optimize performance.

There has been concern regarding the ability to infer high-level strategies from oculomotor behaviors. Noton and Stark (1971) described the *scanpaths* made by viewers, inferring the underlying cognitive goals of the viewers. But without knowledge of the intentions and pre-conceptions of the subject, (knowledge usually not available even to the subject in many complex tasks) the value of scanpaths in discovering cognitive strategies may be limited. While there is some similarity in scanpaths of observers viewing the same scene with the same instructions, there is also significant variability within and between subjects. Viviani (1990) cautioned against attempts to infer cognitive strategies from eye movement records, suggesting that experimental paradigms should constrain movements by limiting the complexity of the task. Viviani described single and double-step paradigms considered sufficiently constrained.

While acknowledging the challenge of inferring strategies from oculomotor behaviors, we posit that careful study of some complex tasks can indeed provide a window into preconscious strategies employed by observers. The challenge is to find complex tasks like those performed in daily life during which natural perceptual strategies might emerge, yet are sufficiently constrained to allow analysis. The constraints can take the form of task definition, subject instructions, and/or control experiments that dissociate strategy from performance. These constraints allow the study of complex behaviors to probe preconscious strategies using eye movements.

For example, Pelz (1995) and Ballard, Hayhoe, and Pelz (1995) monitored the eye movement patterns of subjects as they copied a colored pattern of building blocks. The oculomotor behaviors observed were task-specific, and their spatial and temporal patterns revealed a common strategy. Rather than memorizing a pattern and then replicating the pattern from memory, subjects made frequent eye movements, referring to the model pattern an average of more than 1.5 times per block. This perceptual strategy allowed subjects to optimize their performance, completing the task more quickly than when they were constrained from referring frequently to the model while copying. The pattern the observed behavior was not a conscious strategy employed by subjects to optimize or ease the task; subjects were unaware of the frequent model references, and were often surprised to see their eye movement records after completing the experiment. A control condition that increased the cost of frequent model fixations, and another that reduced the information necessary to copy each pattern supported the hypothesis that the frequent model fixations were part of a high-level perceptual strategy rather than low-level responses to the conspicuity of targets in the field.

In another example, Land and Lee (1994) reported higher-level strategies employed in real-world driving tasks. Land reported that as drivers approached and drove around a corner, they fixated the tangent to the curve at its radius. Fixating the tangent point is not necessary to perform the task; it is apparently evidence of a more sophisticated strategy. The information available at that point allows the driver to enhance task performance by simplifying the computational load that would otherwise be required to navigate the car smoothly around the curve. In the terminology adopted here, the eye movement patterns reported by Land go beyond oculomotor behavior and reveal a higher-level perceptual strategy.

The driving task used by Land and Lee (1994) and the block-copying task used by Ballard et al. (1995) both revealed perceptual strategies. While neither task was difficult, they were complex tasks that had rigid constraints requiring focused attention. The constraint in the block-copying task was imposed by the instruction to complete the task as quickly as possible without error. The driving task was constrained by the need to successfully negotiate the turn. We were interested in finding out whether optimizing perceptual strategies is evident in less demanding tasks with fewer constraints, and in whether high-level perceptual strategies emerge in other complex tasks.

Our goal in the experiments described here was to examine the eye movements of subjects as they perform complex, everyday tasks in natural environments to better understand the process, rather than the mechanics, of visual perception. While offering less control than experiments performed in reduced laboratory conditions, we were able to observe natural oculomotor behaviors and emergent perceptual strategies.

The experiment consisted of monitoring the eye movements of subjects as they walked to a washroom, washed and dried their hands, then walked back to the laboratory. In a control condition, the subjects received the same instructions, except that they were to fill a cup with water rather than wash their hands. The environment, path, and actions of the two tasks were identical up to the first contact with the water faucet. Consequently, differences in fixation patterns between the hand-washing and fill-cup tasks could be ascribed to a high-level perceptual strategy. Hand-washing, like many tasks performed in the course of daily life, is easy but complex. Its ease comes not from simplicity, but from familiarity. The overall task requires that a series of sub-tasks be performed in sequence; e.g. navigate a hallway, open a door, locate, reach for and manipulate the water faucets, sample and adjust the water temperature, locate, reach for, and grasp the soap, etc. Analysis of such tasks can occur at many levels; the highest is the instructed task, the lowest is arguably made up of individual eye movements.

2. Methods

2.1. Wearable eyetracker

In order to perform these experiments, we developed a self-contained, wearable eyetracker for monitoring complex tasks while interfering only minimally with natural eye, head, and whole-body movements. The eyetracker can be worn for an extended period of time (up to 2 h before battery recharging is necessary), does not restrict natural movements or behavior, and preserves peripheral vision.

The primary component of the eyetracker is the wearable headgear shown in Fig. 1. To the right of the subject's eye is a module containing an infrared illuminator, a miniature video eye camera, and a beam-splitter to align the camera to be coaxial with the illuminating beam. Retro-reflection at the retina back-illuminates the pupil, producing a bright-pupil image.

An external mirror folds the optical path toward the front of the goggles, where a hot mirror directs the IR illumination toward the eye and reflects the eye image back to the eye camera.

A second miniature camera is mounted on the goggles just above the right eye to capture the scene from the subject's perspective. This scene image creates the frame of reference for the line of gaze, which is indicated by a cross-hair superimposed over the scene image. The position of the scene camera eliminates horizontal parallax errors and minimizes vertical errors. The miniature eye and scene cameras are CMOS video cameras which offer slightly lower image quality than CCD cameras, but are smaller, lighter, and draw less power. The goggles also house a small semiconductor LASER that is used for calibration (see below).

The system uses a customized version of the Applied Science Laboratory (ASL) Model 501 controller. The line of gaze is computed in real time based on the vector difference between the center of the pupil and the first Purkinje image, making the output less sensitive to movement of the headgear with respect to the head. The controller is placed in the backpack shown in Fig. 1. The backpack also contains a digital picture-in-picture unit that can superimpose the eye image into one quadrant of the image, an LCD display to monitor the tracker signal, and a video recorder to capture the scene image. The video cameras and control unit operate at 60 Hz interlaced, recording a video field every 16.7 ms and a video frame every 33.3 ms. The eye image superimposed by the picture-in-picture unit theoretically allows fixations as short as two video fields (33 ms) to be resolved using a computer-controlled videotape player capable of displaying single video fields (16.7 ms). A fixation was defined as a sequence of at least two video fields during which the pupil is stationary (or moving slowly due to VOR or smooth pursuit eye movements), bounded by video fields in which the pupil image is blurred by saccadic eye movements. While the videotape was the primary record used for analysis in these experiments, the system can also capture the data stream including horizontal and vertical eye position, pupil diameter, and a time-stamp.

The eyetracker was calibrated by entering the location of nine calibration points with respect to the scene camera's view, then recording the horizontal and vertical position of the pupil centroid and first Purkinje image (P_1) as the subject looked at each point. Once calibrated, the differential pupil/P_1 calculation corrects for small movements of the eyetracker with respect to the subject's head, but the accuracy of the calibration is degraded by motion of the subject's head during the calibration process. A semiconductor LASER with a two-dimensional diffraction grating was attached to the headgear so the calibration pattern projected by the LASER was fixed with respect to the head and scene camera, and subtended a fixed visual angle regardless of the distance to the surface. This system eased the calibration procedure, making it possible to quickly calibrate the subject without the need to restrain the subject's head during calibration.

The wearable eyetracker is accurate to within 1° in the central 15° field, within 2° in the central 20° field, and can track within a 40° field. The precision of the system is approximately $\frac{1}{2}$°. After a subject was calibrated on the target, calibration was checked by monitoring the tracker signal as the subject fixated the calibration points and looked about the scene. The calibration was acceptable if the signal was within 1° on each of the calibration points, and track was maintained over a field of approximately 40°. If after recalibration the signal did not satisfy those criteria, the subject was dismissed.

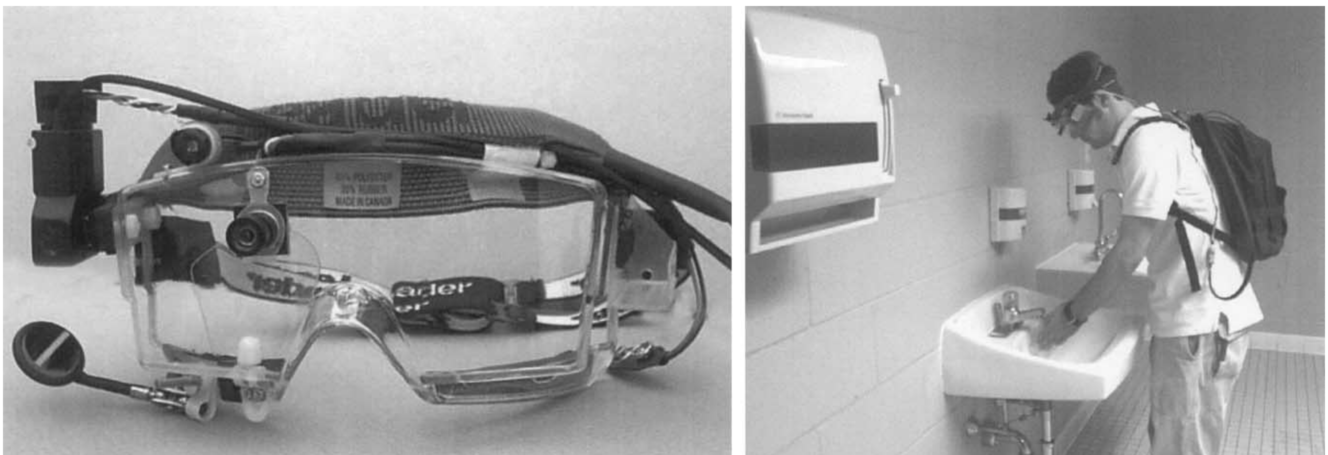

Fig. 1. Left: Headgear contains cameras to image the subject's eye and the scene from the subject's perspective. A semiconductor laser mounted above the eye projects a two-dimensional calibration pattern that remains stable with respect to the subject's head and the scene camera. Right: The controller, display unit, and recording systems are contained in a backpack worn by the subject.

2.2. The tasks

2.2.1. Experiment 1: hand-washing

The hand-washing task was one of several tasks completed by the subjects during an experimental session, but only the hand-washing task is considered here. The subjects were instructed to walk to the washroom, wash their hands, and return. The instructions were deliberately brief to elicit natural behavior. No other person was in the washroom at the time the subjects performed the task. The men's and women's washrooms were similarly appointed with sinks, soap dispensers, a mirror, a paper towel dispenser, and a waste bin.

The subjects walked to the washroom approximately 100 ft from the lab, but for the purposes of the present analysis, the beginning of the trial was defined as the first fixation after entering the washroom. The trial ended when the subject opened the door to leave the washroom.

2.2.2. Experiment 2: control fill-cup and hand-washing

In order to contrast performance under different high-level goals, a second group of subjects performed a control task before the hand-washing task. The control *fill-cup* task was similar to the hand-washing task except for the instructed task. Subjects were given a plastic cup and instructed to go to the washroom, fill the cup with water, and return to the lab. Until the initial contact with the water faucet, subjects had to walk the same path, enter the same environment, and perform the same actions. Once the subject contacted the water faucets, the tasks diverged. After completing the first task, the subjects performed two unrelated intervening tasks taking approximately 5 min. They were then instructed to walk to the washroom, wash their hands, and return to the lab.

Again, each trial began with the first fixation after entering the washroom and ended when the subject opened the door.

2.3. Eye movement data analysis

Because the subjects' body and head were in motion, we use the term *fixation* here to refer to a period during which a given point of regard is maintained over the fovea. This may occur when the eye and head are stationary; when the eyes rotate to compensate for lateral and rotational movements of the head and body; or when smooth pursuit eye movements stabilize the retinal image of a moving target, such as the hands. In this context oculomotor behaviors are divided into just two categories: fixations that stabilize the retinal image, and saccadic eye movements (Steinman, Kowler, & Collewijn, 1990).

The videotaped records were viewed in a Hi8 VTR (Sony EVO-9650) controlled by a lab computer. The VTR read the frame-accurate time-codes from frames selected during analysis and transmitted them to the computer, where they were converted to millisecond since the start of a trial. The duration of each fixation was determined by finding the elapsed time between the first frame in which the gaze was stable and the last frame before the gaze was shifted to a new point of regard. Because the precision of the eyetracker was limited to approximately $\frac{1}{2}°$, saccades were coded only for movements of greater than 1°. Note that a series of brief fixations separated by saccadic eye movements with amplitude less than 1° could be coded as a single longer fixation.

Saccade amplitude was gauged on the scene monitor by marking the beginning and end point of each eye movement on the scene camera image. The field of view of the scene camera was fixed at $75° \times 50°$, so saccadic amplitude was directly related to the extent of the cursor movement on the display monitor. Because this measurement was made with respect to the scene camera, it did not include the contributions due to head movements.

Fixation duration and saccadic amplitude were coded manually by viewing the videotape records of each trial. When two individuals coded the same videotape records for three subjects, there were no significant differences in the distribution of fixation durations ($P > 0.50$) or saccade amplitudes ($P > 0.70$). The identity of fixated objects did not vary as they were relatively large objects in an otherwise sparse environment, and there was no significant difference in coding look-ahead or look-back fixations ($P > 0.40$).

2.4. Subjects

2.4.1. Experiment 1: hand-washing

The hand-washing task was completed as part of a pilot experiment that included a number of tasks, including walking through spaces with a wide range of illumination and manipulating objects near the bottom of the field of view. Five unpaid subjects (three female, two male) were calibrated successfully and exhibited a good track in the extreme lower field. Those five subjects participated in Experiment 1, completing only the hand-washing task. Two of the five subjects (one female, one male) were inexperienced observers, naïve to the experiment. Two subjects (one female, one male) were experienced observers, naïve to the experiment, and one subject (female) was experienced, with knowledge of the experiment. All had normal or corrected to normal vision. They were fitted with the wearable tracker and calibrated as described above. If the accuracy and field met the criteria described, the subject received instructions and began a trial. Calibration was checked after each trial, and repeated if necessary.

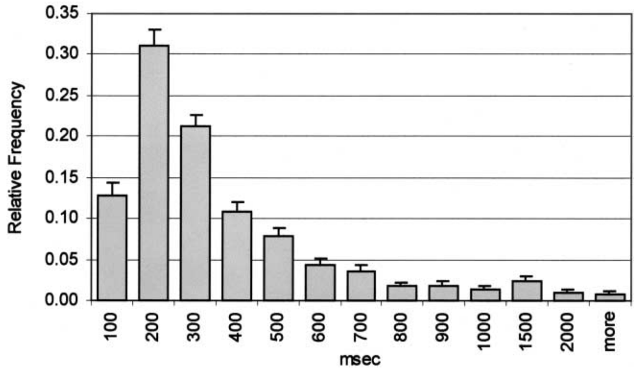

Fig. 2. Relative frequency histogram of fixation duration for 19 subjects in the hand-washing task ($N = 923$). Error bars represent one standard error of the mean between subjects.

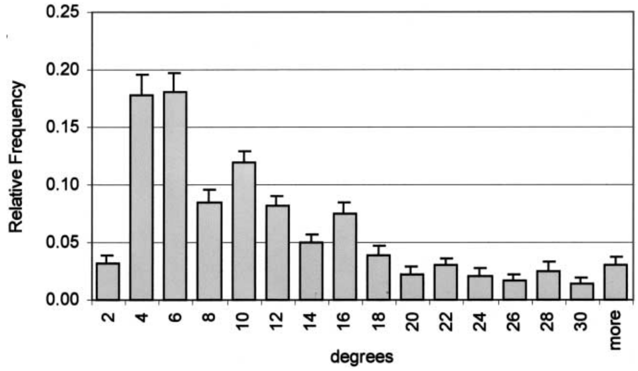

Fig. 3. Relative frequency of saccade amplitude for 19 subjects in hand-washing task ($N = 923$). Error bars represent one standard error of the mean between subjects.

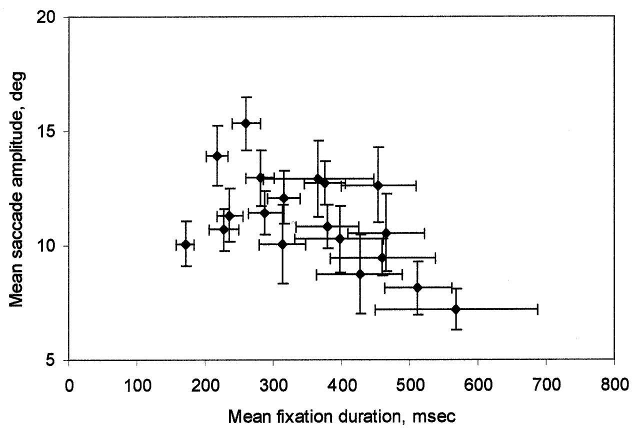

Fig. 4. Subjects' mean saccade amplitude was weakly correlated with mean fixation duration in the hand-washing task ($R^2 = 0.3$, error bars represent one standard error of the mean within subjects).

2.4.2. Experiment 2: control fill-cup and hand-washing

Fourteen undergraduate student subjects (seven female, seven male) were paid for their participation In Experiment 2, performing both hand-washing and control fill-cup tasks. All were inexperienced observers, naïve to the experiment, and had normal or corrected to normal vision. They were fitted with the wearable tracker and calibrated. If the accuracy and field met the criteria described above, the subject received verbal instructions and began a trial. Calibration was checked after each trial, and repeated if necessary.

3. Results

In Section 3.1 fixation duration and saccade amplitude statistics are considered for all subjects performing the hand-washing task in Experiment 1 and 2. The spatial distribution of fixations is described in Section 3.2. In Section 3.3 the higher-level perceptual strategies are examined, first for the five subjects in Experiment 1, then for the 14 subjects in Experiment 2.

3.1. Fixation duration and saccade amplitude

The data from five subjects participating in Experiment 1 and the 14 in Experiment 2 were pooled for analysis of fixation duration and saccade amplitudes. Fig. 2 shows the relative frequency of fixation durations pooled across all subjects. The error bars represent one standard error of the mean between subjects within each bin. The population mean was 327 ms (317). The median fixation duration was 233 ms; individuals' median values varied from 133 to 400 ms.

Fig. 3 shows the relative frequency of saccade amplitude pooled across the 19 subjects. The error bars represent one standard error of the mean between subjects within each bin. The population mean was 11°(9). The median saccade amplitude was 9°, with individuals' median amplitudes ranging from 6° to 14°.

Recall that the precision of the tracker precluded the coding of saccades smaller than 1°, so Fig. 2 may underrepresent short fixation durations and Fig. 3 may underrepresent small saccades.

Subjects' mean saccade amplitude was weakly correlated with mean fixation duration ($R^2 = 0.3$); subjects with longer average fixation durations tended to make smaller saccades. Fig. 4 shows saccade amplitude vs. fixation duration for all 19 subjects. Error bars indicate one standard error of the mean within subjects. Note that potential miscoding of brief fixations separated by short saccades could affect the correlation.

3.2. Spatial distribution of fixations

The majority of fixations were in the lower visual field due to the need to monitor and guide the hands, but significant head movements resulted in a peak eye-in-head fixation density only approximately 8° below the level. Fig. 5 shows the cumulative distribution of fixations within the visual field, centered on that peak. Seven percent of the fixations were within the central 5° (within a radius of 2.5°), 55% within the central 20°, 73% within the central 25°, and 95% within the central 45°.

3.3. Perceptual strategies: look-ahead fixations

In addition to the fixation duration, saccade amplitude, and fixation density statistics described above, the videotapes were coded in terms of the foveated object for each fixation.

3.3.1. Experiment 1: hand-washing

During analysis of the hand-washing task in Experiment 1, a common eye movement pattern became apparent; while nearly all fixations were on objects related to the immediate action, a small number of fixations were made to objects that would become relevant only in actions to be performed in the near future.

An example of a common pattern is illustrated in Fig. 6. The figure shows six video frames captured over a 3300 ms portion of a hand-washing trial. In Fig. 6 a, the subject fixates the water faucets while approaching the sink. Some 733 ms later, before reaching the sink, Fig. 6 b shows the subject fixating the soap dispenser above and to the right of the sink. Note that this fixation does not serve the immediate task (turning on the water faucets); rather it is a *look-ahead* to information that will be needed in the future. Look-ahead fixations were defined as fixations on objects not relevant to the immediate sub-task, but relevant for a future sub-task. Such fixations on the soap dispenser, paper towel dispenser, waste bin, and door handle were scored as look-ahead fixations in the hand-washing task. While fixations on the first three objects are irrelevant to the control fill-cup task, they were scored as look-ahead fixations to compare the frequency of such fixations to the hand-washing task.

In Fig. 6 c, 1400 ms after the initial fixation on the soap dispenser, the subject is still fixating the sink. Fig. 6 e shows a typical guiding fixation on the soap dispenser 633 ms before the reach toward the soap dispenser, and 2 s after the initial look-ahead fixation. In addition to the fixations on the soap dispenser while walking toward the sink before the initial reach to the water faucets, subjects often fixated the towel dispenser and waste bin during preceding sub-tasks. These eye movements occurred seconds before the reach toward the corresponding targets, and did not replace the guiding eye movements made ~ 500 ms before those reaches. The targeting eye movements, occurring approximately 500–1000 ms before a reach, have been reported in other natural tasks (Epelboim et al., 1997; Land & Furneaux, 1997; Land et al., 1999) and are typical of reaching tasks requiring visual guidance (Biguer, Jeannerod, & Prablanc, 1982).

All five subjects participating in Experiment 1 showed at least one look-ahead fixation during the trial, with an average of 3.6 look-ahead fixations per trial. For comparison, a measure of the number of times that any of the objects were fixated after use (a *look-back* fixation) was made. Two of the five subjects made such fixations, with a significantly lower average of 0.6 look-backs per trial ($P < 0.001$).

While a very significant difference, the comparison of look-ahead to look-back fixations may be misleading for two reasons. First, the objects were in the field for a shorter period after use than before, limiting the opportunity for subjects to fixate the objects again. Second, because the objects are fixated in the course of the hand-washing task, there may be a low-level inhibition of return (IOR) or a tendency not to revisit objects whose memory representations are fresh.

3.3.2. Experiment 2: control fill-cup and hand-washing

To address the issues related to the look-ahead vs. look-back comparison in Experiment 1, the control fill-cup task was designed to provide visual input similar to that in the hand-washing task.

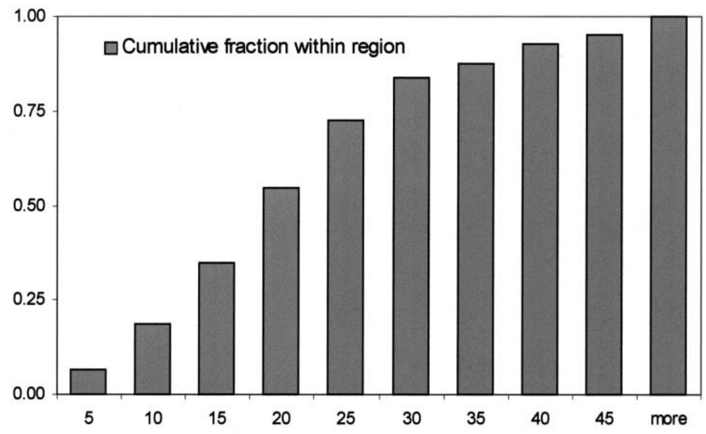

Fig. 5. Seven percent of the fixations were within the central 5° (within a radius of 2.5°), 55% within the central 20°, 73% within the central 25°, and 95% within the central 45°.

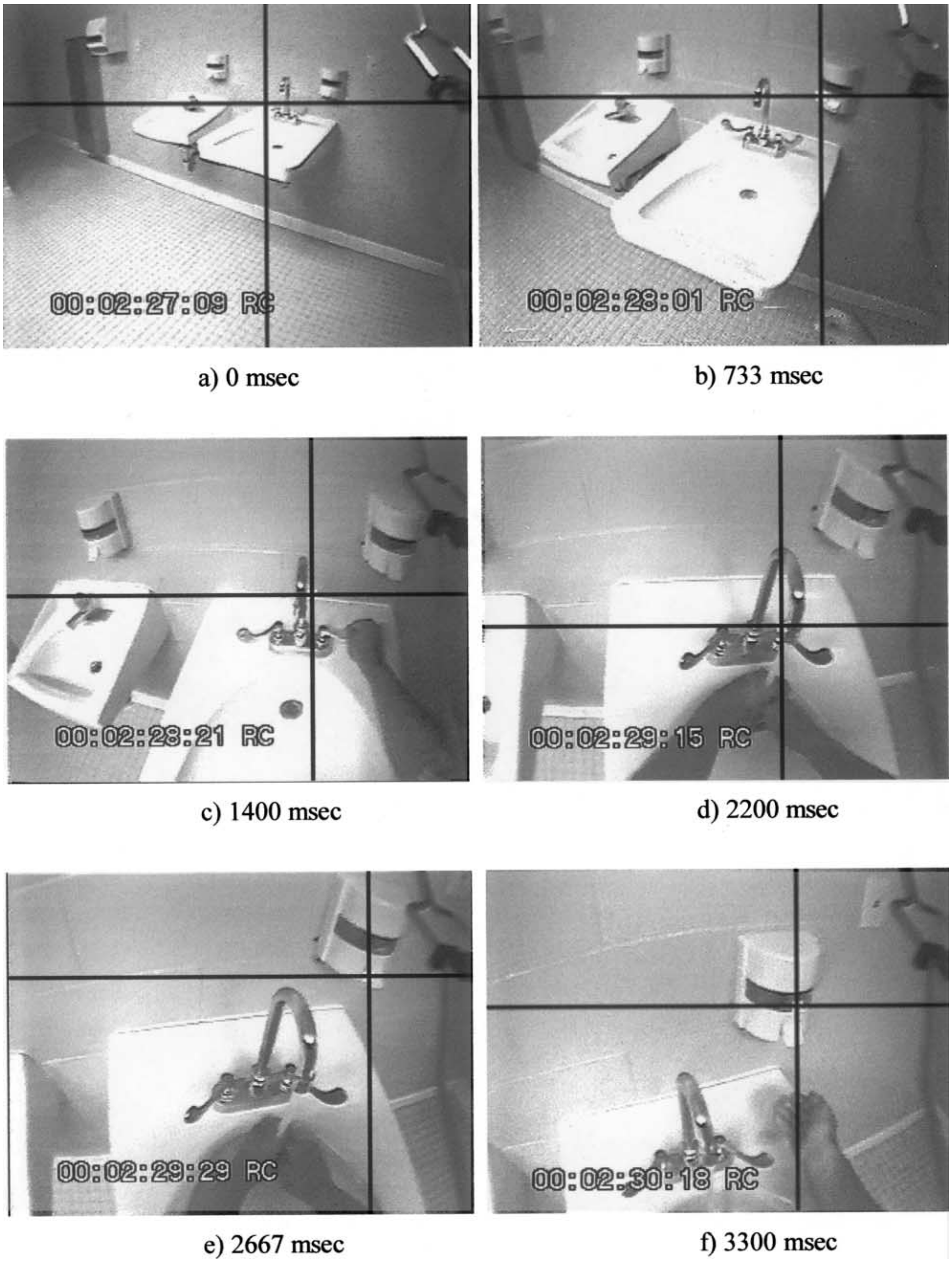

Fig. 6. Fixation records during hand-washing task illustrate the look-ahead pattern: (a) fixation on water faucet handles during approach to the sink; (b) fixation on soap dispenser during approach to sink, before initial contact with the faucet handle; (c) and (d) fixation returns to water faucet for manipulation; (e) fixation on soap dispenser preceding reach; (f) reach toward soap dispenser while maintaining fixation.

Because the visual fields and the objects' duration in the environment were similar for both tasks, and the control fill-cup task was performed first, any bias due to IOR or decaying memory representations would weaken evidence for a strategy of looking ahead toward objects of future interaction.

The control fill-cup task and the hand-washing task were identical from the start of the trial until the tasks diverged at the initial contact with the water faucet. Therefore, we assign particular significance to the pattern of fixations in that interval. Eight of the 14 subjects participating in the control fill-cup task in Experiment 2 fixated on the soap dispenser, towel dispenser, or waste bin, which had been scored as look-ahead fixations in the hand-washing task. The average number was 0.64 fixations per trial. When the same subjects in the control fill-cup task returned to the washroom approximately 5 min later with the instruction to wash their hands, 12 of the 14 subjects made look-ahead fixations to those objects. The average number of look-ahead fixations increased significantly from 0.64 to 2.00 per trial ($P < 0.005$). The number of look-backs in these subjects was 0.21 per trial, significantly fewer than the number of look-ahead fixations ($P < 0.001$).

4. Discussion

With the observation that the look-ahead eye movements are common in some tasks, it is useful to consider a number of hypotheses regarding their utility in a given task. It may be that the eye movement patterns may be a general behavior elicited by low-level cues such as the conspicuity or salience of objects in the environment. We reject this hypothesis based on the fixation patterns in the hand-washing task and the results of the control condition. As seen in the sequence of fixations in Fig. 6, the look-ahead fixations were on objects relevant to future tasks. While the adjacent sink and other soap dispensers were equal or greater in size, luminance, and contrast, subjects fixated the soap dispenser that would be used in the future. The same targets were rarely fixated after they had been used, indicating that it is not simply the conspicuity of the targets that attracts the fixations. The control condition offers further evidence; while the identical visual field was present in the hand-washing and control tasks, fixations to the soap dispenser fell dramatically in the control task.

Another possibility is that the look-ahead fixations are part of a visual search process. Search may have two meanings in this context. The eye movements could be part of a general 'sweep' of the surround to identify objects in the environment, or they may indicate a targeted search for a specific object. The first alternative is rejected based on the observed pattern of fixations. While subjects do look at some objects as they enter the washroom, they do not initiate a sweeping sequence of eye movements, fixating on a large number of objects. Further evidence is provided by the results of the control condition, as the pattern is tied to the high-level goal rather than the environment. An alternate hypothesis is that the fixations could be errors. If a subject has pre-programmed a sequence of actions, a part of the program may be executed out of order. Lashley (1951) described such *errors of anticipation* in the domains of typing and speech. In typing, it is common for a character from the following word to intrude into the current word. Spoonerisms can also be thought of in the same manner. Such errors of anticipation, however, would likely be accompanied by 'erroneous' arm and/or hand movements as well, a pattern we did not observe in any subjects. In addition, errors of anticipation interfere with normal actions, consuming attentional resources and degrading performance.

We propose that the look-ahead fixations provide a mechanism to 'stitch together' the stream of visual input resulting from the sequence of actions that make up daily life. Analogous to purported mechanisms supporting visual stability, look-ahead fixations may be part of a strategy that supports our subjective experience of an environment that is continuous in time as well as space. In the spatial domain, there is a large difference between the rapid sequence of retinal images due to eye movements and the subjective experience of a stable environment. The scene does not appear to jitter or jump as the retinal images certainly do, nor is there uncertainty about where the observer or scene is located in space. This phenomenon has drawn the attention of investigators for many years, and has led to several hypotheses about the mechanism(s) by which we achieve this visual stability. Knowledge of eye position via proprioception (Steinbach, 1987; Gauthier, Nommay, & Vercher, 1988; Matin, 1976), efference copy of oculomotor commands (Sperry, 1950; Stark & Bridgeman, 1983), and regularities in the scene (O'Regan, 1992; Irwin, 1992; Nakayama, 1990; Pelz & Hayhoe, 1995) have all been shown to play a role in the perceived stability of the visual scene.

Beyond the laboratory, we have another dimension to contend with; not only must vision provide spatial stability, it must support temporal seamlessness as well. As we move through and interact with the world, building a stable representation is not sufficient. The visual system must supply a steady, reliable stream of information to support our conscious experience of an environment. This temporal dimension has not arisen with experimental tasks in the past because task complexity and duration were purposely restricted.

The strategy may also ease the task-switching between the serial sub-tasks used to gather information from, and interact with, the environment. Like Land et al. (1999) report, we found that $\sim 95\%$ of fixations are dedicated to completing the immediate sub-task. A portion of the remaining 5% may be used to help bridge the task-switching that would otherwise be evident to subjects, as they are for complex tasks before proficiency is gained. Part of the transition from 'hard' to 'easy' in completing complex tasks is the gradual reduction in the conscious effort required to complete the discrete sub-tasks in sequence.

The oculomotor system evolved as part of the foveal compromise, supporting the illusion of a large, clear field by sampling the retinal image with the high-acuity fovea. A stationary observer in a static environment uses attention and eye movements to sample the spatial periphery, gathering and integrating information from successive views. Perception in the real world is more complicated, as the observer moves through and interacts with a dynamic environment. Sophisticated perceptual strategies, such as look-ahead fixations, may have evolved to ease the cognitive and attentive loads of perception in the real world. Eye movements serialize complex operations, optimizing performance by deploying foveal acuity as required by the task, as in the case of eye movements to objects ~ 500 ms before reaching. The serialization takes place at a level below consciousness; our perception is seamless in time as well as space.

Eye movements are used not only to serialize complex tasks, but also to merge the rapid sequence of retinal images into a stable percept.

The strategy we are reporting only emerged when we observed subjects performing complex natural tasks in natural environments. Further exploration in domains where vision is examined in its native role as a tool supporting high-level perception will surely identify other perceptual strategies that are fundamental to better understanding visual perception.

Acknowledgements

The authors wish to acknowledge Jason Babcock, who is largely responsible for the current design and construction of the wearable eyetracker, and Amy Silver, who performed much of the data analysis. The work was supported in part by an RIT College of Science Project Initiation Grant.

References

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*(1), 68–82.

Biguer, B., Jeannerod, M., & Prablanc, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. *Experimental Brain Research*, *46*, 301–304.

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

Collewijn, H., Steinman, R. M., Erkelens, C. J., Pizlo, Z., & van der Steen, J. (1992). Effect of freeing the head on eye movement characteristics during three-dimensional shifts of gaze and tracking. In A. Berthoz, W. Graf, & P. P. Vidal, *The head–neck sensory motor system*. Oxford University Press Chapter 64.

Epelboim, J., Steinman, R. M., Kowler, E., Pizlo, Z., Erkelens, C. J., & Collewijn, H. (1997). Gaze-shift dynamics in two kinds of sequential looking tasks. *Vision Research*, *37*, 2597–2607.

Gauthier, G., Nommay, P., & Vercher, J. (1988). Ocular muscle proprioception and visual localization in man. *Journal of Physiology*, *406*, 24.

Irwin, D.E. (1992). Visual memory within and across fixations. In K. Raynor. *Eye movements and visual cognition*; *scene perception and reading*. New York: Springer-Verlag.

Kowler, E. (1989). Cognitive expectations, not habits, control anticipatory smooth oculomotor pursuit. *Vision Research*, *29*, 1049–1057.

Kowler, E., & McKee, S. P. (1987). Sensitivity of smooth eye movement to small differences in target velocity. *Vision Research*, *27*, 993–1015.

Kowler, E., Pizlo, Z., Zhu, G., Erkelens, C. J., Steinman, R. M., & Collewijn, H. (1992). Coordination of head and eyes during the performance of natural (and unnatural) visual tasks. In A. Berthoz, W. Graf, & P. P. Vidal, *The head–neck sensory motor system*. Oxford University Press Chapter 65.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transaction of the Royal Society of London. Series B*, *352*, 1231–1239.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, *369*, 742–744.

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

Lashley, K. S. (1951). The problem of serial order in behavior. In W. A. Jeffress, *Cerebral mechanisms in behavior*: *the Hixon symposium* (pp. 112–136). New York: Wiley.

Matin, L. (1976). Saccades and extraretinal signal for visual direction. In R. A. Monty, & J. W. Senders, *Eye movements and psychological processes*. Hillsdale, NJ: Erlbaum.

Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore. *Vision: cooling and efficiency*. (pp. 411–422). Cambridge University Press.

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

O'Regan, J. K. (1992). Solving the 'real' mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychoiogy*, *46*, 461–488.

Pelz, J.B., 1995. Visual representations in a natural visuo-motor task. Doctoral Dissertation, University of Rochester, Rochester, NY.

Pelz, J. B., & Hayhoe, M. M. (1995). The role of exocentric reference frames in the perception of visual direction. *Vision Research*, *35*, 2267–2275.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*, 482–489.

Stark, L., & Bridgeman, B. (1983). Role of corollary discharge in space constancy. *Perception and Psychophysics*, *34*, 371–380.

Steinbach, M. J. (1987). Proprioceptive knowledge of eye position. *Vision Research*, *27*, 1737–1744.

Steinman, R. M., Kowler, E., & Collewijn, H. (1990). New directions in oculomotor research. *Vision Research*, *30*, 1845–1864.

Viviani, P. (1990). Eye movements in visual search: cognitive, perceptual, and motor control aspects. In E. Kowler, *Eye movements and their role in visual and cognitive processes*. *Reviews of oculomotor research V*4 (pp. 353–383). Elsevier.

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