Saccadic localization in the presence of cues to three-dimensional shape

Dhanraj Vishwanath

Eileen Kowler

Saccades directed to simple two-dimensional (2D) target shapes under instructions to look at the target as a whole land near the center of gravity (COG) of the shape with a high degree of precision (He & Kowler, 1991; Kowler & Blaser, 1995; McGowan, Kowler, Sharma, & Chubb, 1998; Melcher & Kowler, 1999; Vishwanath, Kowler, & Feldman, 2000). This pattern of performance has been attributed to the averaging of visual signals across the shape. Natural objects, however, are three-dimensional (3D), and the shape of the object can differ dramatically from its 2D retinal projection. This study examined saccadic localization of computer-generated perspective images of 3D shapes. Targets were made to appear two-dimensional (2D) or 3D by manipulating shading, context, and contour cues. Average saccadic landing positions (SD ~ 10% eccentricity) fell at either the 2D or 3D COG, and occasionally in between, depending on the nature of the 3D cues and the subject. The results show that saccades directed to objects are not compelled to land at the 2D COG, but can be sensitive to other visual cues, such as cues to 3D structure. One way to account for these results, without abandoning the averaging mechanism that has accounted well for performance with simple 2D shapes, is for saccadic landing position to be computed based on averaging across a weighted representation of the shape in which portions projected to be located at a greater distance receive more weight.

Keywords: eye movements, saccades, localization, 3D shape and object perception, depth perception

Introduction

Saccadic eye movements are used to bring the line of sight to selected objects in visual scenes. Programming of saccades to look around scenes requires two processing stages: selection of an object to look at, followed by the computation of a saccadic landing position within the selected object. The first stage, selection, is accomplished by the allocation of attention to the object or region (Kowler & Blaser, 1995; Hoffman & Subramaniam, 1995; Deubel & Schneider, 1996). The second stage, computation of the landing position, is based on pooling information across the attended region. A significant problem for understanding saccadic control is to determine the nature of the visual representation over which pooling operates.

When a saccade is made to a spatially extended two-dimensional (2D) target under instructions to “look at the target as a whole” (instructions designed to capture what happens when people look from object to object during natural viewing), the line of sight lands, on average, near the center of gravity (COG) with precision comparable to that obtained with single target points (He & Kowler, 1991; Kowler & Blaser, 1995; McGowan et al., 1998; Melcher & Kowler, 1999; Vishwanath et al., 2000; Vishwanath & Kowler, 2003). It is, of course, possible to aim a saccade at a variety of selected locations within target objects (He & Kowler, 1991), but when saccades are made to the target as a whole, the consistency of landing positions (SDs 7-10% of eccentricity) and the proximity of landing positions to the COG (average errors < 10% eccentricity) suggest that landing positions are computed by averaging across all stimulated locations within the selected target. Such performance has been obtained with structured targets, such as simple forms (Kowler & Blaser, 1995; Melcher & Kowler, 1999), and with unstructured targets, such as configurations of random dots (McGowan et al., 1998), and is obtained even when the COG lies outside the boundaries of the form (Vishwanath & Kowler, 2003). The COG is a better predictor of landing position than alternatives based on features, such as local landmarks, bisectors, or the symmetric axis (McGowan et al., 1998; Melcher & Kowler, 1999; Vishwanath & Kowler, 2003). Earlier studies (e.g., Findlay, 1982; Ottes, Van Gisbergen, & Eggermont, 1985; Coëffé & O’Regan, 1987) implicated COG tendencies to explain the saccadic errors that result when saccades are made to targets in the presence of distractors (for a critique, see He & Kowler, 1989; Vishwanath & Kowler, 2003).

Averaging has also been invoked to explain the perceptual alignment of spatially extended targets, such as random dot patterns, Gabor patches, or Gaussian blobs. These perceptual studies showed that for the most part, the reference position on which the alignment was based coincided with the centroid of the luminance or contrast distribution of the target (Westheimer & McKee, 1977; Whitaker & Walker, 1988; Morgan, Hole, & Glennerster, 1990; Levi & Tripathy, 1995; Whitaker, McGraw, Pacey, & Barrett, 2003).
Averaging models of saccadic localization generally supported an averaging model, but with some modification. For example, when saccades are directed to shapes with different internal distributions of luminance or texture, landing positions are near the COG of the shape, independent of the internal characteristics (Melcher & Kowler, 1999; for a related result, see Findlay, Brogan, & Wenban-Smith, 1993). This suggests that saccadic landing position is computed from a representation of shape, rather than from more primitive distributions of retinal luminance or contrast. The representation of shape used by saccades, however, may differ from that available for perception. Vishwanath et al. (2000) found that saccades directed to triangles with one or more vertices occluded landed reliably at the COG of the visible fragment, despite efforts on the part of subjects to direct saccades to the perceptually completed shape, thus suggesting that there may be distinctions between perceived shape and the representations guiding saccades.

Landing position may also depend on the task. Vishwanath and Kowler (2003) found that saccades made to an L-shape, a target whose COG lies outside the bounding contour of the shape, were biased toward the intersection of the limbs of the L when subjects either made a single saccade to the target L, or when they choose what they felt was the most natural fixation position within the L (for more about chosen fixation positions within shapes, see Kaufman & Richards, 1969; Murphy, Haddad, & Steinman, 1974; Steinman, 1965). In a third task that required a sequence of saccades to be made to a set of targets including the L, mean saccadic landing positions coincided with the COG of the L with no biases toward the intersection of the limbs or any other features. The task of scanning targets in sequence, the case where landing positions were closest to the COG, may be more representative of how saccades are used naturally than either single saccades or choosing a single fixation position. (The present experiments will also test landing positions of saccades made as part of sequences.)

One of the outstanding differences between objects encountered in natural viewing and the targets used in the localization tasks described above is that natural objects are 3D. Depending on the vantage point from which an object is viewed, the retinal location of the COG of the 3D object and the COG computed by unweighted averaging of the 2D retinal image projection can be quite different. If the determination of the saccadic endpoint is based exclusively on the shape of the image on the retina, then saccades should land at the COG of the 2D retinal projection, regardless of the 3D structure conveyed by the appropriate visual cues. On the other hand, if computation of the saccadic landing position takes into account 3D information, then other landing positions, such as the 3D COG, which coincide more closely with the center of the perceived 3D object, may be obtained. One situation in which there can be large differences between the COG of the 3D object and the COG of the 2D retinal image occurs when the object is viewed from vantage points that produce large perspective scaling in the retinal image (see Figure 1). Under such viewing conditions, the projected 3D COG is displaced from COG of the 2D image toward the part of the image representing more distal portions of the object.

Figure 1. a and c. Examples of displays containing the 3D target and reference objects with two different luminance distributions. b and d. 2D target and reference objects. The 2D shapes in all cases (a-d) are identical. Red and black crosses are the 3D COG and 2D COG, respectively (see text). Dashed rectangle indicates target region within which saccadic landing positions were identified as having been directed to the target (see “Methods”). Neither the crosses nor the dashed rectangle were present in actual displays. c and d show representative individual saccadic landing positions (subject AM).
In the present study, sensitivity of saccadic localization to perceived 3D structure was tested using computer-generated perspective images of simulated 3D objects, such as the one shown in Figure 1. These images contain cues, such as perspective, shading, and shadows, that create a vivid illusion of a 3D object oriented in depth. One advantage of testing stimuli such as these is that the illusion of the 3D object oriented in depth can be readily abolished by modifications to the shading, texture, and surrounding visible frameworks. This allows a comparison of saccadic localization in the presence of 3D cues with saccadic localization of the same shapes without such cues. Finding sensitivity of saccades to 3D shaped cues, (specifically, a shift in mean saccadic landing position toward the COG of the inferred 3D shape without loss of saccadic precision) would provide evidence that the saccades are not guided by an averaging process that operates solely on an unweighted 2D retinal representation, but is instead using representations that may incorporate other visual cues, such as those signaling the shape of the 3D object.

**Experiment 1**

In this experiment, saccades were made to 2D perspective images of a capsule-shaped 3D object. Visual cues were manipulated such that the same image shape appeared either as a 3D capsule-shaped object, or as a flat, 2D, pears-shaped object (Figure 1). The stimulus shapes were chosen so that there would be a substantial difference between the COGs of the 2D image shape and the 3D capsule. Prior research on saccadic localization, reviewed in the “Introduction,” showed that saccadic landing positions in simple 2D shapes coincide well with the COG. Sensitivity of saccades to the cues to 3D shape would thus be indicated by departures from the 2D COG in the 3D, but not in the flat, 2D version of the stimulus.

**Methods**

**Subjects**

Two naive subjects (BS and AM) were tested, both of whom had some prior experience as eye movement subjects. Both had normal vision, with no spectacle correction.

**Eye movement recording**

Two-dimensional movements of the right eye were recorded by a Generation IV SRI Double Purkinje Image Tracker (Crane & Steele, 1978). The subject’s left eye was covered and the head was stabilized on a dental bite-bar.

The voltage output of the Tracker was fed online through a low pass 50-Hz filter to a 12-bit analog to digital converter (ADC). The ADC, controlled by a PC, sampled eye position every 10 ms. The digitized voltages were stored for later analysis. The PC controlled the timing of the stimulus display via a serial link to the SGI computer. Voltage from a photocell that recorded stimulus onset and offset directly from the display monitor was fed into a channel of the ADC and recorded along with the eye position samples to ensure accurate temporal synchronization between stimulus display and eye movement recording.

Tracker noise level was measured with an artificial eye after the tracker had been adjusted to have the same first and fourth image reflections as the average subject’s eye. Filtering and sampling rate were the same as those used in the experiment. Noise level, expressed as a SD of position samples, was 0.4’ for horizontal and 0.7’ for vertical position.

Recordings were made with the tracker’s automatically movable optical stage (auto-stage) and focus-servo disabled. These procedures are necessary with Generation IV Trackers because motion of either the auto-stage or the focus-servo introduces larger artifactual deviations of Tracker output. The focus-servo was used, as needed, only during inter-trial intervals to maintain subject alignment. This can be done without introducing artifacts into the recordings or changing the eye position/voltage analog calibration. The auto-stage was permanently disabled because its operation, even during inter-trial intervals, changed the eye position/voltage analog calibration.

**Stimuli**

Stimuli were displayed on an SGI GDM 17-E21 17” color monitor (1.9 pixels/minarc, refresh rate 65 Hz) controlled by an SGI Iris O2 (viewing distance, 119 cm) and were viewed monocularly. The critical saccadic target shapes were 2D perspective images of a simulated 3D capsule-shaped object generated with the ARRIS CAD software (Sigma Design). Visual cues were manipulated such that the objects appeared either 3D or 2D.

For the 3D version of the stimulus (Figure 1a and 1c), the perspective images were rendered with shading and shadow consistent with a principal light source and ambient lighting using the ARRIS program (diffuse illumination model). The COG of the 3D-capsule shape was calculated assuming a solid of uniform density (centroid), and the red crosses in Figure 1 indicate its location on the image.

The 2D versions of the stimuli were created by editing the base 3D-Bitmap images in Adobe Photoshop (utilizing the drop-shadow feature) such that they appeared as flat 2D shapes parallel to the display surface (Figure 1c). The COG of the 2D shape was calculated as the centroid of a plane surface of uniform density, and is indicated by the black crosses in Figure 1.

The visual display contained the critical target shape (either the 3D or 2D version) along with two 45°-diameter reference targets shown against a uniform gray background (11.3 x 9 deg) for the 2D version or the simulated interior of a gray room for the 3D version. The reference targets were either spheres (presented along with the 3D shape) or disks (presented along with the 2D shapes). One reference target appeared above the critical shape, the other either to its left or right. In Figure 1, the lower sphere or disk appears to the right of the critical shape.
**Characteristics of the critical shapes**

The critical image shapes ranged in size from 200' to 245' as measured along their major axis. Centers-of-gravity of 2D and 3D versions differed by at least 30'.

The following variables were also manipulated: Orientation. Three different orientations of the shapes were tested (Figures 2 and 3). For each orientation, the 2D and 3D versions of the target had identical object boundary contours. Lighting direction. Two different positions of the primary light sources were tested to produce different luminance patterns in the 2D projection of the stimulus. The region with the highest luminance was located either in the lower right (Figure 1a) or upper left (Figure 1c) portion of the 2D projection. Location. The critical target shape was located either to the left or to the right of the center of the display. The relative position of the reference spheres/disks with respect to the critical target, and the relative position of the entire configuration with respect to the display boundaries, was varied to produce four different spatial configurations. The critical target shapes were presented at a horizontal retinal eccentricity of either 240' or 265' from the lower reference target and a vertical eccentricity of 240' from the upper reference target. Additionally, the upper reference target was located in two possible horizontal positions (separation of 50') and the lower reference target in two possible vertical positions (separation of 50'). Data will be collapsed across the four different configurations for each orientation and location (left or right) of the target.

**Disk targets**

Saccadic localization was also tested for a 200'-diameter circular disk target. This was done to assess the idiosyncratic saccadic undershoots and overshoots that occur when saccades are directed to either single point or spatially extended targets. In prior work we showed that these idiosyncratic undershoots and overshoots did not depend on the target shape (Melcher & Kowler, 1999; Vishwanath & Kowler, 2003). To isolate effects of the shape on saccades, measured saccadic endpoints for the critical targets will be corrected for the undershoots and overshoots determined from trials with the disk targets, as was done previously (Melcher & Kowler, 1999; Vishwanath & Kowler, 2003).

**Luminance and color**

For the 3D-shaded displays, the target and reference shapes appeared yellow (CIE \(x = 0.4, y = 0.5\)) and the walls of the “room” were gray (CIE \(x = 0.25, y = 0.35\)). Luminance of target varied from 3 to 13 cd/m\(^2\), and the room varied from 0.1 to 10.0 cd/m\(^2\). For the 2D displays, target luminance was 5.6 cd/m\(^2\) and background was 1.5 cd/m\(^2\).

**Procedure**

One reference target (either the sphere for the 3D condition or the disk for the 2D condition) was displayed before the start of each trial to act as an initial fixation stimul-
Results

Saccades directed to the flat 2D shape landed near the COG as expected (Figure 2). Departures from the COG (mean vector error) were 8° for AM and 5° for BS (less than 5% of average saccade size). No consistent direction for the error vectors was found for the 2D target shapes. Variability of saccades was small, in keeping with prior reports. Average SDs were 20′ (horizontal) and 24′ (vertical) for AM, and 21′ on both meridians for BS (8-10% of saccade size).

Results were different for the 3D shape (Figure 3). AM’s saccades were consistently shifted toward the 3D COG, with average landing position closer to the 3D COG for two of the three orientations tested (vector shifts of mean saccadic landing positions from the 2D COG were 21′ and 24′ for two of the three orientations tested). For the remaining orientation, she landed between the 2D and 3D centers of gravity (vector shift 17′). BS landed even fur-


determined empirically by examining a large sample of analog records of eye position. Saccades as small as the micro saccades that may be observed during maintained fixation (Steinman, Haddad, Skavenski & Wyman, 1973) could be reliably detected by the algorithm.

When instructions to use a single saccade were successfully followed, the trial should contain exactly four saccades, two of which should land at the target shape. Some trials contained exactly the four saccades required. Others contained one or more secondary, corrective saccades made after reaching the target region with a larger primary saccade from one of the disks. Results are based on landing positions of all saccades (including the small corrective saccades) that landed within a 3.6° x 4.4° acceptance region centered on the target and containing the entire target shape. (Results were essentially the same when corrective saccades were eliminated.) The rare saccades that landed outside the acceptance region (such as the few stray saccades seen in Figure 1) were eliminated because the errors were so large that they did not seem to be genuine attempts to reach the targets.

Mean departures in landing positions from the COG, determined from trials in which a sphere or disk was tested in place of the critical target shape in the same experimental sessions, were subtracted from saccade offset positions obtained with the critical target shape to isolate the effects of the shape on landing position (see “Disk Targets” above) (Melcher & Kowler, 1999; Vishwanath et al., 2000; Vishwanath & Kowler, 2003). The departures were determined separately, and corrections were applied separately, for the left- and right-hand target positions (see “Stimuli” above). Corrections for AM were horizontal 9′ and vertical 4′ for leftward target positions, and horizontal 4′ and vertical 0′ for rightward target positions. For BS, corrections were horizontal 17′ and vertical -16′ for leftward target positions, and horizontal -8′ and vertical -24′ for rightward target positions. Negative corrections shifted mean saccadic landing positions either down or to the left.

Experiment 2

Experiment 1 showed that saccadic landing positions could be affected by cues to 3D shape, as evidenced by the observed shifts in mean landing position toward the 3D COG. Experiment 2 takes a different approach to the same
issue. Instead of comparing landing positions in 2D and 3D versions of the same shape, landing positions were compared for two 3D shapes that had identical 2D images, but different 3D interpretations. The different 3D interpretations were conveyed by outline cues drawn within the shape itself as well as the background, rather than by shading and shadow.

**Methods**

Methods were the same as Experiment 1 except for the following aspects:

**Stimuli and display**

The critical target shapes were perspective images of a 3D-curved capsule (Figure 4). Shading and contextual cues were removed, leaving only a silhouette of the target object and reference spheres against a black background (Figure 6 shows the original capsule with shading and contextual cues included). Two different shapes were constructed, each with the same contour, and thus the same 2D images, but a different 3D interpretation. The different 3D interpretations were conveyed by the addition of a small line drawn inside the boundary of the target shape. They will be denoted as “neck away” (Figure 4a) and “neck toward” (Figure 4b). Additional lines outside the boundary indicated the 3D context.

The neck away shape was created from a simulated curved capsule shape created on the ARRIS CAD software; thus, its 3D COG could be computed and is indicated by the red cross in Figure 4a. The neck toward shape (Figure 4b) was created by removing the short line segment inside the neck away shape, which contributed to the 3D appearance, and substituting a hand-drawn short segment with a different curvature selected to create the impression of a different orientation in depth. The 3D COG of this neck toward shape thus cannot be computed exactly without establishing the inferred tilt of the object toward the viewpoint. Nevertheless, an approximation of the location of the 3D COG of the neck toward shape was made based on an assessment of differences in the location of the 2D and projected 3D COG for images of truncated cones (of comparable dimensions) at different tilts. This assessment showed that differences between 2D and 3D centers of gravity for the neck toward shape would be on the order of a few minutes of arc. For illustration purposes, an estimated location of the projected 3D COG is indicated by the dashed cross in Figure 4b.

Two different orientations of each of the two kinds of targets shapes were tested, as shown in upper and lower panels of Figure 5. The 2D images of the shapes for each orientation subtended 230’ and 260’, as measured between the most distal points on the bounding contour of the shapes. The reference disks, presented along with the critical target shapes, subtended 50’ in diameter.

Six different spatial configurations of the critical target and reference disks were tested. Specifically, the critical
target shape was presented either to left or to the right of the center of the display. In addition, for each direction (left or right), three different relative spatial positions of the critical target shape and reference disks with respect to the display boundaries were tested. Horizontal retinal eccentricity of the critical target was 240' or 265' from the lower reference target and vertical eccentricity was 250' from the upper reference target. Vertical position of the lower reference and horizontal position of the upper reference was varied by 50'. As a result, there was no consistent relationship between the position of critical target shape with respect to reference disks and with respect to the display boundaries. The target and reference disks were yellow (CIE x = 0.4, y = 0.5) and the background was black. Luminance of the target and disks was 9.8 cd/m².

Control trials were also included in which saccadic localization was tested for a 200'-diameter circular disk target in place of the critical shape to assess idiosyncratic saccadic undershoots and overshoots (see “Experiment 1”). Corrections for AM were horizontal 12' and vertical -14' for leftward target positions, and horizontal 1’ and vertical -16’ for rightward target positions. For BS, corrections were horizontal 20’and vertical -16’ for leftward target positions, and horizontal -5’ and vertical -23' for rightward target positions. Negative corrections indicate that saccadic landing positions were shifted either down or to the left.

Results

Figure 5 shows the mean saccadic landing positions for the two target orientations tested, corrected for the idiosyncratic undershoots and overshoots measured in the control trials. The individual red dots are mean landing positions for the six different spatial configurations tested for each orientation, and the open red circle is the mean landing position averaging over all the configurations. BS’s mean landing positions were close to the 3D COG for both versions of the stimulus, consistent with her performance in Experiment 1. AM, on the other hand, did not show any reliable differences and landed near the 2D COG for both interpretations.

Individual differences in perceptual localization performance in the presence of several available visual cues have been observed in the past (Akutsu et al., 1999; Akutsu & Levi, 1998; Koenderink, van Doorn, Kappers, & Todd, 2001). The differences between BS and AM’s performance prompted a third experiment in which a larger group of subjects were tested on a different set of 3D shapes.

Experiment 3

Experiment 3 explores the sensitivity of saccades to 3D cues with the neck away shape tested in Experiment 2, but with a larger set of 3D cues available and with more subjects.

Methods

Methods were the same as Experiment 2 except for the following aspects.

Subjects

AM and BS were tested once again, along with three additional naïve subjects (NR, JR, and AG) who had not participated in prior experiments on saccadic localization.

Stimulus and display

Stimuli were 2D perspective images of the simulated curved, capsule shape tested in Experiment 2 and presented with two reference spheres and with shading and contextual cues (Figure 6). The perspective images were rendered with shading and shadow consistent with a principal light
source and ambient lighting using the ARRIS program (diffuse illumination model). Two different viewing orientations of the same curved capsule shape were generated (Figures 6a and 6b). The size of the shapes and the distances between the shapes and the reference spheres were the same as in Experiment 2, as were the six different spatial configurations of the display tested for each orientation. Two different positions of the primary light sources were tested to produce different luminance patterns in the 2D projection of the stimulus. The region with the highest luminance was located either in the lower right (Figure 6b and 6d) or upper left (Figure 6a and 6c) portion of the 2D projection.

Target and reference shapes were yellow (CIE $x = 0.4$, $y = 0.5$) and the walls of the “room” were gray (CIE $x = 0.25$, $y = 0.35$). Luminance of target varied from 3 to 13 cd/m$^2$, and the room varied from 0.1 to 10 cd/m$^2$.

Landing positions were corrected for departures from the COG obtained in trials in the same experimental sessions in which a sphere/disk was tested in place of the critical shape, as in Experiments 1 and 2. These departures were all less than 20’ with two exceptions. JR and NR each landed about 40’ above the COG of the sphere, requiring a downward correction in landing positions obtained with the critical shapes.

**Results**

Figure 7 shows mean saccadic landing positions for all five subjects for the two different orientations. The two circles within each panel show mean landing positions obtained for each of the two luminance distributions, averaged over the six spatial configurations ($N = 70-200$).

The results for the five subjects fell into two categories. The saccades of two of the subjects (BS and JR) landed near the 3D COG. The saccades of the other three (AM, RA, and NR) were close to the 2D COG, although AM showed reliable shifts toward the 3D COG. The mean vector displacements from the 2D COG for the five subjects, averaged over the two target orientations, were BS 57’, JR 50’, AM 18’, RA 10’, and NR 9’. For four of the five subjects (BS, JR, AM, and RA), the mean vector displacements were toward the 3D COG. Average SDs of horizontal saccades were 7-10% of eccentricity, and SDs of vertical saccades were 9-12% of eccentricity. SDs did not depend on the proximity of the mean landing position to either the 2 or 3D COG.

Figure 6. a and c. Perspective images of the two curved capsule-shaped targets, Experiment 3. b and d. The same but with a different luminance distribution. Red crosses indicate 3D COG and black crosses 2D COG.

Figure 7. Mean saccadic landing position for five subjects for each target shape tested in Experiment 3. White and yellow circles are mean saccadic landing positions ($N = 70-200$) for each of the two luminance distributions. The luminance distribution shown in the figure is the same as Figure 6a and 6c. SEs are smaller than the open circle and are not shown.
Figure 8 shows all individual saccadic landing positions for all stimulus presentations of the two shapes. The red dashed lines are the bivariate normal ellipses, showing the region containing 68% of the observations. Different patterns of landing were observed in that for some of the subjects, saccades landed in an approximately circular region surrounding the mean landing position, whereas for others saccades were distributed in a region surrounding medial axis of the shape. The difference in these patterns was not correlated with whether the mean landing position coincided more closely with the 2 or the 3D COG. Figure 8 also shows that saccades rarely landed at any of the available local features, such as the discontinuities in the contour near the perceptually far end of the shape.
Individual differences in perceived 3D structure: A demonstration

A vivid percept of 3D structure from a 2D image depends on the congruence of available 3D cues, such as perspective, shading, motion parallax, disparity, viewing geometry, as well as intrinsic properties of image shape such as symmetry (Buckley & Frisby, 1993; Koenderink et al., 2001; Todd, Koenderink, van Doorn, & Kappers, 1996; Leyton, 1992). Computer generated 2D images of 3D objects present on a flat computer screen present inherent cue conflicts in the visual stimulation. In such situations, individual differences in the reported 3D shape are often observed and have been attributed to variations in the assignment of depth/orientation values to specific locations of the shape (e.g., Buckley & Frisby, 1993; Koenderink et al., 2001). A brief perceptual experiment was performed to find out whether displays containing cues similar to those used in Experiments 1-3 would produce individual variations on a perceptual task.

Subjects were asked to compare the size of the left and right spatial interval defined by the three dots superimposed on the image of a cylinder (Figure 9a). The task was to ignore the background and make a judgment only on the 2D separation. Though the separation between the dots is equal, Figure 9a shows that when the dots are superimposed on the 3D figure, the right-hand interval can appear larger than the left-hand interval, even in 2D space, presumably due to a “capture” effect of the inferred depth of the cylinder. The size of this illusion was measured for the five subjects tested in Experiment 2 (two-alternative forced-choice procedure; 200-ms exposure). All showed an illusion (Figure 9b), but individual differences were clear, with three of the five subjects (BS, JR, and NR) showing a greater illusion than the other two. These observations show that individual differences were not restricted to saccades. They may have originated in different ways of coding the 3D cues, which may depend on the nature of the available cues, the individual, and the task.

Discussion

Much of the literature on saccadic eye movements is based on studies in which saccades are directed to targets so small that the desired endpoint of the saccade is unambiguous. With larger objects characteristic of natural scenes, the endpoint of the saccade has to be determined from visual information distributed across the target. Prior work has shown that when saccades are directed to larger targets, the line of sight lands near the target’s COG, with a relatively high level of precision. This finding suggests that saccadic landing position is determined by pooling (averaging) visual signals over the selected object (He & Kowler, 1991; Kowler & Blaser, 1995; Melcher & Kowler, 1999; Vishwanath et al., 2000; Vishwanath & Kowler, 2003). Earlier work had also implicated averaging as a way of interpreting the influence of nontarget stimuli on the landing position of the saccade (e.g., Ottes et al., 1985; Coëffé & O’Regan, 1987; for discussion, see He & Kowler, 1989). The main virtue of averaging is that it ensures accurate and precise
saccadic localization of objects without requiring a contribution of either the attentive or cognitive systems in any way other than the delineation of the region over which averaging is to occur (He & Kowler, 1991). Thus, an efficient averaging mechanism, which brings the line of sight to precise locations in selected objects without imposing undue burdens on attentive or cognitive systems, is a key component of effective saccades in natural environments.

Prior studies of saccadic localization in general, and averaging in particular, have been restricted to fairly simple target configurations; thus, a major unresolved issue about saccadic localization is the nature of the visual representation on which the averaging is based. The present study shows that in the presence of cues that signal a 3D object, such as perspective, shading, context, and shape, saccades can land closer to the projected COG of the 3D object than the COG of the 2D retinal shape. The consistent departures from the 2D COG indicate that, at the very least, unweighted averaging across the 2D image shape is not adequate to account for saccadic localization of objects.

Shifts in saccadic landing position away from the 2D COG and toward the projected 3D COG of the shape were prominent in some subjects and weaker or absent in others. Individual differences in shape 3D perception or perceptual localization in the presence of multiple cues have been observed before (Akutsu & Levi, 1998; Akutsu et al., 1999; Koenderink et al., 2001; Todd & Norman, 2003). We found that even in individuals who landed close to the 2D COG, there were normally small displacements of landing position toward the 3D COG (AM and RA in Experiment 3). By contrast, shifts away from the 2D COG in shapes lacking 3D cues showed no consistent direction (Figure 2; see also Vishwanath & Kowler, 2003).

What might account for the systematic displacements of saccadic landing positions toward the 3D COG?

It is not likely that the 3D COG was determined from a fully reconstructed, view-invariant, 3D-object representation. Finding the 3D COG in such a representation would require a complete recovery of 3D-object shape before the saccade was programmed, while the object was imaged extrafoveally. Moreover, even with such recovery, it is not obvious how the 3D COG would be determined. A view-invariant 3D representation of object shape would probably involve higher cortical areas, such as V4 and the lateral occipital complex (LOC) (Janssen, Vogel & Orban, 2000; Logothetis & Pauls, 1995; Pasupathy & Conner, 2002; Kourtzi et al., 2003), which have been implicated in object recognition and may be ill suited for representing location via retinotopic population responses. Thus, a localization model that requires a view-invariant representation of the 3D object, either in place of or along with the 2D COG, is not plausible because visual coding of 3D shape as it is currently understood does not appear to be well suited to extraction of this reference position.

It is also unlikely that saccades were deliberately aimed to the 3D COG. In our experiments, as in prior work, subjects were instructed to look at the object as a whole, an instruction that resembles how saccades are used naturally. Under these instructions, saccades have been found to land close to the 2D COG, with modest scatter, even with targets in which the COG could not be readily determined from local cues, or the COG was located outside the boundaries of the shape, or there were large variations in luminance (He & Kowler, 1991; Kowler & Blaser, 1995; McGowan et al., 1998; Melcher & Kowler, 1999; Vishwanath & Kowler, 2003). The present results also showed the same level of scatter characteristic of the prior work, regardless of whether the mean landing position was closer to the 2D or the 3D COG. Deliberate selection of a particular landing position, on the other hand, should have led to increased variability because of the expected difficulty of maintaining consistent strategies across trials and experimental sessions (for this argument applied to perceptual localization, see Morgan et al., 1990).

Perhaps the strongest reason to doubt that deliberate selection of a particular landing position drew saccades to the 3D COG is that there were no local cues indicating its position. The 3D COG (horizontal and vertical coordinates) did not coincide with any prominent landmark or feature (Figures 1, 6, and 7); thus, it is not obvious how it would be possible to select and maintain attention to the 3D COG during saccadic preparation. In fact, the results showed little influence of local features. Saccades rarely landed at regions of highest luminance, regions of local concavities or convexities, or the larger and more prominent region located at the “perceptually near” portion of the shape (Figure 8). Thus, the pattern of results obtained with the 3D shape suggests that landing position was based on pooling of information across the shape, rather than selection of a distinct landing position.

An explanation for the displacement of saccadic landing positions toward the 3D COG that seems more likely than either of the alternatives considered above is weighted averaging. While the 3D COG is not readily available from local cues, one way of obtaining landing positions close to the 3D COG is by weighted averaging over the 2D image, where more weight is assigned to the portions projected to be further in depth. Figure 10 compares the COGs obtained by application of five different weighting functions. In each panel, lighter colors indicate higher weights and the blue crosses show the weighted average position. For comparison, the 2D COGs (black crosses) and 3D COGs (red crosses) reproduced from Figure 6 are also shown. Figure 10a shows that a weighted average position near the 3D COG can be obtained by applying a monotonically increasing weighting function in which weight increases along the bisector of shape (see inset). A similar outcome can be obtained by increasing weight abruptly, as shown in Figure 10b. These two weighting models can be compared to alternatives in which weight was assigned based on proximity to a prominent local feature, namely, the contour disconti-
nuity on the left side of the shape. In Figure 10c, the boundary of the local region receiving additional weight is marked by the dashed blue line, and COGs were computed as a function of the weight assigned to this local region. Figure 10c shows that as the relative weight assigned to this local region increased, the COG moved along the dashed

line from the unweighted COG (black cross) to the weighted COG (blue cross) that was obtained by averaging exclusively within the marked local region. Figures 10d and 10e are the same, except the local region receiving higher weight is larger. A comparison of Figures 10c-10e shows that the weighted average approaches the 3D COG only when the highly weighted region is so large that it encompasses most of the “far” (upper) portion of the shape (Figure 10e), thus resembling the weighting in Figure 10a and 10b. Although it is possible for different and more complex weighting functions to be explored, the interesting aspect of the simple models illustrated in Figure 10 is that averaging across relatively large regions of the shape, while assigning more weight to the “far” portions (Figures 10a, 10b, and 10e), produces a result closer to the 3D COG than assigning additional weight to small regions surrounding local features.

Assigning more weight to the portions of a shape at greater projected distances is consistent with processes the visual system might use as part of a transformation of the retinal representation of a shape into a representation of the 3D object (Schwartz, 1980, 1999). Dobbins Jeo, Fiser, and Allman (1998) showed that neurons in V1 are sensitive to cues to depth, firing at higher rates to more distant targets. With such “distance scaling,” a population-averaging response would reflect distortions in the retinal image due to perspective projection, and thus bring the saccadic landing positions closer to the 3D rather than the 2D center of the object. A key property of a distance-scaled averaging model is that finding different landing positions (e.g., 2D COG with some displays or individuals, 3D COG with others) does not imply different rules for determining the saccadic endpoint. The process used to determine landing position within the selected object—population averaging—remains the same. The only difference from one situation to the next is the weighting function imposed on the region over which averaging takes place.

**Conclusion**

Previous treatments of the saccadic localization of simple target shapes suggested that the landing position is computed by averaging locations across the selected target. The present study shows that with more complex target shapes, containing visual cues implying an object in depth, individuals show two patterns of performance, with some landing near the 2D COG and others near the 3D COG. Saccadic landing positions near the 3D COG can be obtained by modifying the averaging model to allow spatial weighting, including spatial weighting according to implied depth. Spatial weighting is a useful option, allowing flexible control of saccadic endpoints without requiring effortful selection of a precise landing position each time a saccade is launched. In the presence of 3D cues, bringing the line of sight to the 3D COG may be critical for hand-eye coordination in object manipulation.
Acknowledgments

This work was supported by the Air Force Office of Scientific Research, Grant AF 4962002-140112, Life Sciences Directorate. We thank Professors Doug De Carlo, Jacob Feldman, Michael Leyton, Thomas Papathomas, and Manish Singh for their comments and suggestions.

Commercial relationships: None.
Corresponding author: Dhanraj Vishwanath.
Email: dhanraj@john.berkeley.edu.
Address: Vision Science Program, UC Berkeley, School of Optometry, Berkeley, CA, USA.

Footnotes

1 The term 2D center of gravity (COG) used in this work implies the centroid of the target shape taken as a plane region of uniform density. Others (e.g., Whitaker et al., 1996) use the term “centroid of the luminance distribution” for the average of nonuniform luminance variation across the target shape.

2 The 3D COG of the capsule shape was determined assuming a solid of uniform density (centroid). The pixel location of the 3D COG in the perspective image (i.e., its projection onto the picture plane) was determined by marking the COG on a wire-frame perspective image of the simulated object. Note that alternative definitions of the 3D COG, such as defining the 3D COG over only the visible 3D surface area or for a hollow object, may generate a different 3D COG position but also shifted away in depth from the 2D COG.

References


