Perception, 2011, volume 40, pages 887 – 904

doi:10.1068/p6917

Static and dynamic visual information about the size and passability of an aperture

Aaron J Fath, Brett R Fajen¶
Department of Cognitive Science, Carnegie Building 308, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, NY 12180, USA; e-mail: fajenb@rpi.edu
Received 24 January 2011, in revised form 26 August 2011

Abstract. The role of static eyeheight-scaled information in perceiving the passability of and guiding locomotion through apertures is well established. However, eyeheight-scaled information is not the only source of visual information about size and passability. In this study we tested the sufficiency of two other sources of information, both of which are available only to moving observers (ie are dynamic) and specify aperture size in intrinsic body-scaled units. The experiment was conducted in an immersive virtual environment that was monocularly viewed through a head-mounted display. Subjects walked through narrow openings between obstacles, rotating their shoulders as necessary, while head and shoulder position were tracked. The task was performed in three virtual environments that differed in terms of the availability of eyeheight-scaled information and the two dynamic sources of information. Analyses focused on the timing and amplitude of shoulder rotation as subjects walked through apertures, as well as walking speed and the number of collisions. Subjects successfully timed and appropriately scaled the amplitude of shoulder rotation to fit through apertures in all three conditions. These findings suggest that visual information other than eyeheight-scaled information can be used to guide locomotion through apertures.

1 Introduction

Many natural and human-made environments contain obstacles that form openings through which one may or may not be able to pass. When people navigate through such environments, it may be necessary to make gait and postural modifications to pass through narrow apertures or to follow an alternative route if passage is not possible. In such situations, the ability to perceive the size and passability of apertures in advance is necessary both for selecting safe and efficient routes and for guiding locomotion among obstacles.

When people are moving at a comfortable walking speed, they begin to make gait and postural adjustments when aperture width is less than 1.3 times shoulder width, suggesting that they leave themselves a small safety margin to account for lateral body sway during locomotion (Warren and Whang 1987). Such adjustments include slowing down to reduce body sway, rotating the shoulders, or turning completely sideways (Higuchi et al 2006).

It is sometimes undesirable to slow down or rotate the body, in which case the ability to perceive the passability of apertures in advance allows one to consider alternative routes. Indeed, passability is accurately perceived across a wide variety of conditions. Stationary observers perceive apertures as passable without shoulder rotation when the width of the aperture is 15% – 20% greater than the width of their shoulders, reflecting a safety buffer consistent with that which is observed when people actually walk through apertures (Warren and Whang 1987). Furthermore, passability is perceived in a way that accurately takes into account increases in lateral sway that accompany running (Wagman and Malek 2007) and increases in effective body width that result from carrying an object or walking side-by-side with another person (Chang et al 2009; Wagman and Malek 2007; Wagman and Taylor 2005). Wheelchair users can accurately account for their extended bodily dimensions, albeit only after extensive experience (Higuchi et al 2004, 2006; Savelsbergh et al 1998).

¶ Author to whom all correspondence should be addressed.

1.1 *Visual information about the size and passability of apertures*

In this study we consider three sources of visual information about the size and passability of apertures. All three sources of information are body-scaled in that they specify aperture size in intrinsic units related to dimensions of the observer's body (eg eyeheight) or movement (eg head-sway amplitude, stride length). [See Warren (2007) for an excellent review of body-scaled information.] Assuming that the perceptual system can calibrate itself to the relation between these measures and body width, body-scaled information allows for the perception of aperture size in units of body width. Thus, the advantage of body-scaled information is that it sidesteps the problem of having to know body size in extrinsic units, which would be necessary if aperture size was perceived in non-intrinsic units. For this reason, we did not consider sources of information other than those that specify size in intrinsic units.

The three sources of information identified below are described using complex combinations of optical angles and their temporal derivatives. Nonetheless, we do not want to imply that the visual system actually computes these variables. Our aim in this section is merely to provide mathematical descriptions of the sources of information that are available in the optic-flow field, which we believe to be an important first step toward a complete account that also includes the means by which such information is detected.

1.1.1 *Static, eyeheight-scaled information.* The most widely investigated source of visual information about the size of fixed-width apertures is static eyeheight-scaled information (Sedgwick 1980; Warren and Whang 1987) which specifies aperture width (G) in units of eyeheight (E):

$$\frac{G}{E} = \frac{2\tan(\alpha/2)}{\tan\gamma}\,, \qquad (1)$$

where α is the angle subtended by the inside edges of the obstacles and γ is the angle of declination of the base of the obstacles (figure 1). Because shoulder width (W) is a constant proportion of standing eyeheight, equation (1) also specifies aperture size in units of shoulder width. Strong evidence exists to support the hypothesis that people rely on eyeheight-scaled information to perceive passability, the most compelling of which is that subtle decreases in eyeheight make apertures appear more passable (Warren and Whang 1987; see also Wraga 1999). In addition, the accuracy with which passability is perceived does not significantly improve when apertures are viewed while walking rather than while stationary, attesting to the sufficiency of static information about passability.

Although static eyeheight-scaled information is sufficient to support the perception of passability, it is only reliable when the base of support is resting on the ground surface, which is not always the case. For example, such information could not be used

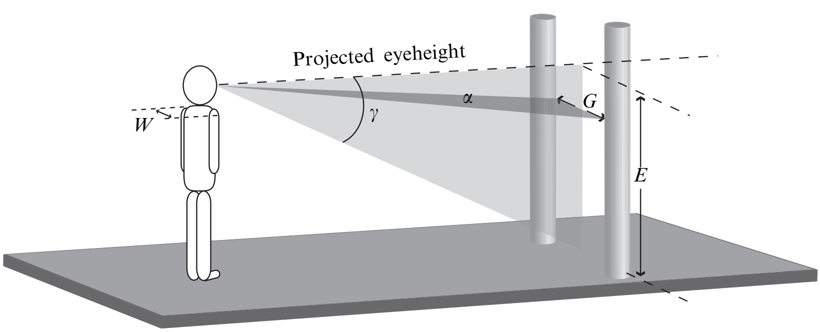


Figure 1. The geometry of eyeheight-scaling. G is the distance between the inside edges of the cylinders, E is eyeheight, W is shoulder width, α is the visual angle subtended by the inside edges of the cylinders, and γ is the angle of declination of the base of the cylinders.

to perceive the size of an aperture between a pair of overhanging tree branches. Next, we identify two sources of information that specify aperture size and passability regardless of whether the obstacles are resting on the ground plane. Unlike eyeheight-scaled information, both information sources are dynamic, such that they are only available when the observer is moving, which is often the case when people need to perceive the size or passability of apertures.

1.1.2 *Dynamic head-sway-scaled information.* The first form of dynamic information is illustrated in figure 2, which shows an observer moving toward an aperture between a pair of stationary obstacles. Note that the head follows a roughly sinusoidal path due to lateral sway that accompanies locomotion, and that the lateral position of the head at time t is designated by $x_h(t)$. Figure 2 shows the observer at the moment that the lateral position of the head is at peak amplitude, which we designate as $x_h(t_{x_h=A})$.

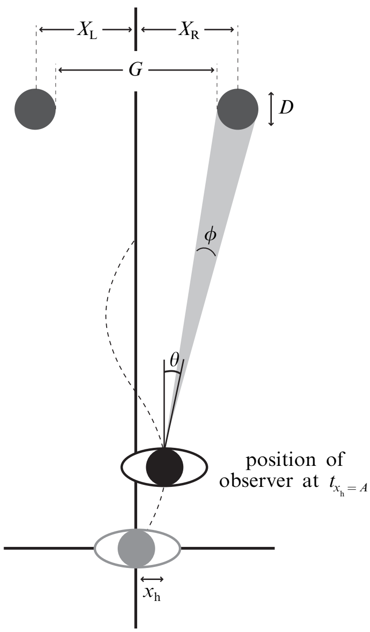


Figure 2. The geometry of head-sway-scaled information. X_R and X_L are the lateral positions of the centers of the right and left obstacles, respectively. G is the distance between the inside edges of the obstacles (ie the size of the aperture), D is the diameter of each obstacle, $x_h(t)$ is the lateral position of the head at time t, θ is the visual angle between the direction of locomotion and the center of the obstacle, and ϕ is the visual angle subtended by the edges of the obstacle.

The lateral distance between the position of the head at time $t_{x_h=A}$ and the position of the center of the right obstacle (X_R) is optically specified in units of obstacle diameter (D) by:

$$\frac{X_R - x_h(t_{x_h=A})}{D} \approx \frac{\dot{\theta}_R(t_{x_h=A})}{\dot{\phi}_R(t_{x_h=A})}\,, \tag{2}$$

where $\dot{\theta}_R$ is the azimuthal angle between the center of the right obstacle and the direction of locomotion, and $\dot{\phi}_R$ is the local visual angle of the right obstacle itself.

The size of the aperture is the distance between the inside edges of the right and left obstacles (ie $G = X_R - X_L - D$, where $X_R = -X_L$), which is optically specified in units of D by:

$$\frac{G}{D} \approx \frac{\dot{\theta}_R(t)}{\dot{\phi}_R(t)} - \frac{\dot{\theta}_L(t)}{\dot{\phi}_L(t)} - 1\,. \tag{3}$$

The amplitude of head sway (A) is equal to:

$$= -1/2\{[X_R - x_h(t_{x_h=A})] + [X_L - x_h(t_{x_h=A})]\}\,. \tag{4}$$

Dividing both sides of equation (4) by D, substituting equation (2) and rearranging terms yields:

$$D \approx -2A \Bigg/ \left[\frac{\dot{\theta}_{\mathrm{R}}\left(t_{x_{\mathrm{h}}=A}\right)}{\dot{\phi}_{\mathrm{R}}\left(t_{x_{\mathrm{h}}=A}\right)} + \frac{\dot{\theta}_{\mathrm{L}}\left(t_{x_{\mathrm{h}}=A}\right)}{\dot{\phi}_{\mathrm{L}}\left(t_{x_{\mathrm{h}}=A}\right)} \right] . \tag{5}$$

Substituting equation (5) into equation (3) and rearranging terms yields:

$$\frac{G}{A} \approx -1/2 \left[\frac{\dot{\theta}_{\mathrm{R}}(t)}{\dot{\phi}_{\mathrm{R}}(t)} - \frac{\dot{\theta}_{\mathrm{L}}(t)}{\dot{\phi}_{\mathrm{L}}(t)} - 1 \right] \Bigg/ \left[\frac{\dot{\theta}_{\mathrm{R}}\left(t_{x_{\mathrm{h}}=A}\right)}{\dot{\phi}_{\mathrm{R}}\left(t_{x_{\mathrm{h}}=A}\right)} + \frac{\dot{\theta}_{\mathrm{L}}\left(t_{x_{\mathrm{h}}=A}\right)}{\dot{\phi}_{\mathrm{L}}\left(t_{x_{\mathrm{h}}=A}\right)} \right] . \tag{6}$$

Equation (6) demonstrates that the size of the aperture is optically specified in units of head-sway amplitude. Assuming that the perceptual system can calibrate itself to the relation between head-sway amplitude and shoulder width, equation (6) means that aperture size can also be perceived in units of shoulder width.

Equation (2), which is the root of equation (6), was originally identified in the context of interceptive actions involving a stationary observer and a moving target, such as reaching out to catch a ball (Bootsma 1991; Peper et al 1994; Regan and Kaushal 1994). In that context, the ratio of the change in azimuthal angle to the change in angular size specifies the future lateral passing distance of the approaching object in units of ball size. Recent psychophysical and behavioral studies provide evidence that people can detect such information (Arzamarski et al 2007; Gray and Sieffert 2005; Jacobs and Michaels 2006), and use it to guide lateral hand movements during catching. Because equation (2) is the root of equation (6), such studies suggest that observers may also be able to detect the information captured by equation (6).

1.1.3 *Dynamic, stride-length-scaled information.* The second form of dynamic information is stride-length-scaled information (Lee 1980), which specifies aperture width in terms of the observer's current stride length (L):

$$\frac{G}{L} \approx \frac{\alpha\tau}{t_{\mathrm{s}}} , \tag{7}$$

where t_{s} is stride duration and τ is the ratio of α to the first temporal derivative of α (see figure 3). Assuming that the perceptual system can calibrate itself to the relation between stride length and shoulder width, equation (7) means that aperture size can also be perceived in units of shoulder width.

Although stride-length-scaled information was first identified in 1980, its contribution to the perception of passability has not been experimentally tested. This is likely due to the fact that the kinds of experimental manipulations that are needed to test for

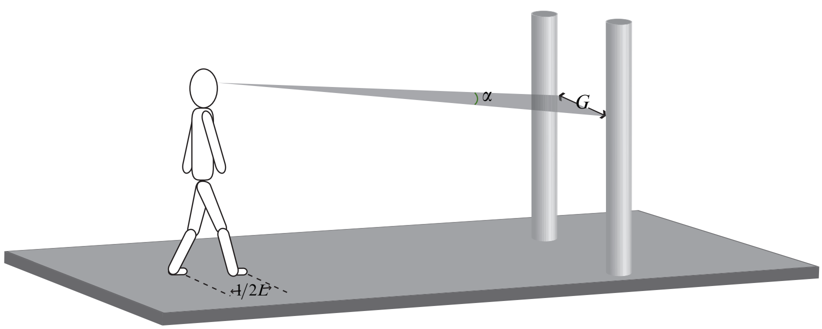


Figure 3. The geometry of stride-length-scaled information. G is the distance between the inside edges of the cylinders, L is stride length (2× step length), and a is the visual angle subtended by the inside edges of the cylinders.

the use of stride-length-scaled information are difficult if not impossible to implement in the real world. In this study, we used an immersive virtual environment to isolate the effects of stride-length-scaled information.

1.2 *Testing the contributions of eyeheight-scaled, head-sway-scaled, and stride-length-scaled information*

The aim of this experiment was to investigate the contributions of the aforementioned three sources of information to the perception of aperture size and the guidance of locomotion through narrow apertures. The experiment was conducted in a virtual environment that was viewed through a head-mounted display (HMD). Because the specific focus was on the sufficiency of each of the three sources of information, subjects viewed the virtual environment monocularly. Had subjects been allowed to use binocular vision, the contribution of binocular cues could not be ruled out. Therefore, monocular viewing was used to isolate the three candidate sources of information. A motion capture system was used to track the positions of subjects' left and right shoulders as they walked through apertures the size and position of which varied across trials. Subjects were instructed to rotate their shoulders as necessary to avoid colliding with the edges of the aperture, as illustrated in figure 4a.

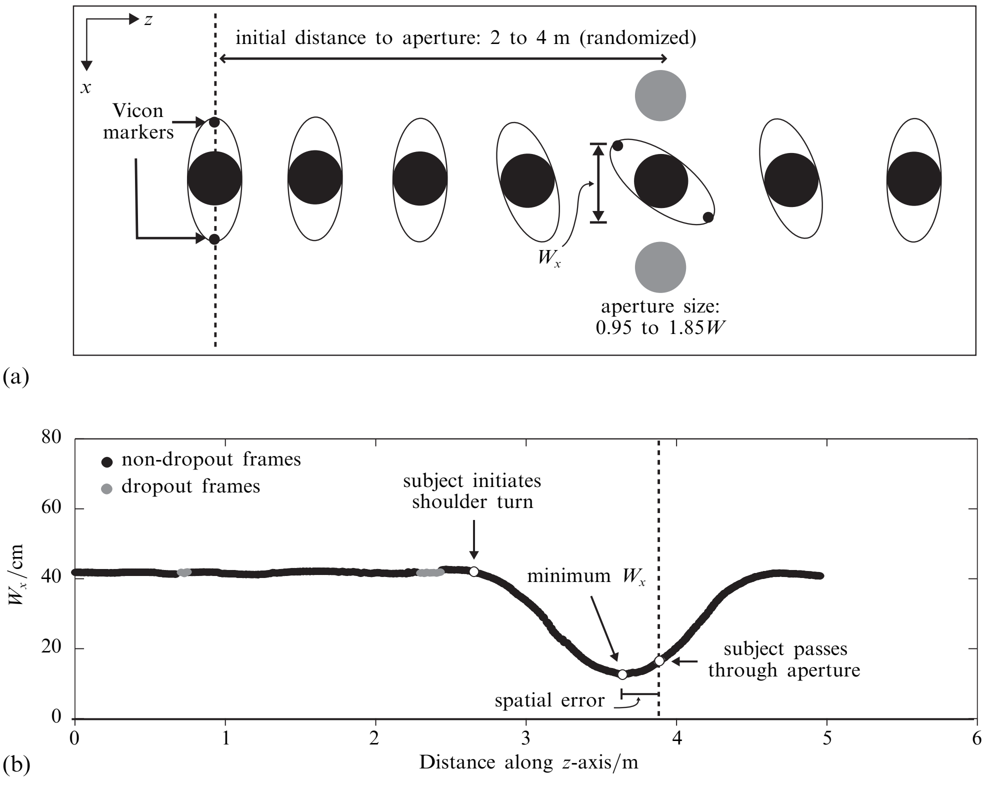


Figure 4. (a) Top – down schematic view of subject walking through an aperture, showing locations of shoulder markers, and definition of W_x. (b) Data from sample trial showing W_x as a function of distance along the *z*-axis. Black and gray segments represent non-dropout and dropout frames, respectively. Dashed vertical line indicates location of aperture.

The contributions of eyeheight-scaled, head-sway-scaled, and stride-length-scaled information were investigated by comparing performance across three conditions (see figure 5). In the post condition, the aperture was defined by a pair of cylindrical posts resting on a textured ground plane (see screenshot in top row of figure 5). Once subjects

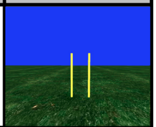
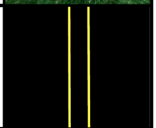
Condition	Information			Screenshots
	EH	HS	SL	
Post	✓	✓	✓	
Tall-post		✓	✓	
Wall			✓	

Figure 5. [In colour online at http://dx.doi.org/10.1068/p6917] Sources of information available in each condition used in the experiment. EH = eyeheight; HS = head-sway; SL = stride-length. Screenshots from each condition are shown in the right column.

began walking toward the aperture in the post condition, all three sources of information were available. In the tall-post condition, there was no ground plane and the vertical cylinders that defined the edges of the aperture spanned the entire visual field from top to bottom (see screenshot in middle row of figure 5). Once walking was initiated, both head-sway-scaled and stride-length-scaled information were available in the tall-post condition. However, because the ground plane was removed and the obstacles had no visible base, eyeheight-scaled information was not available. In the wall condition, there was no ground and the edges of the aperture were defined by a solid yellow frontal wall that spanned the entire visual field from top to bottom and from the edge of the aperture to the periphery (see screenshot in bottom row of figure 5). Stride-length-scaled information, which is based on the visual angle subtended by the aperture, was the only source of information among the three that were considered in this study that was available in the wall condition. Eyeheight-scaled information was not available for the same reason that it was not available in the tall-post condition (ie because the walls to the left and right of the aperture spanned the entire visual field from top to bottom). Head-sway-scaled information was not available because the walls also spanned the entire visual field to the left and right of the aperture, which made it impossible to detect the local optical expansion of the obstacles to the left and right of the aperture [ie $\dot{\phi}$ in equation (6) and figure 2].(1)

Performance in all three conditions was measured by analyzing the amplitude and timing of shoulder rotation, walking speed, and the number of collisions with the edges of the aperture. If people rely on eyeheight-scaled information and are unable to use head-sway-scaled or stride-length-scaled information, subjects should perform well in the post condition but not in the tall-post or wall conditions. In the two latter conditions, subjects may have difficulty appropriately scaling the amount of shoulder rotation to accommodate different aperture sizes. They may compensate for their inability to accurately perceive aperture size by always rotating their shoulders by the same amount—that which allows them to pass through the smallest aperture that they encounter. They may also have difficulty timing the rotation of their shoulders, they may decrease walking speed, and the number of collisions with the edges of the aperture may increase. If subjects are able to use head-sway-scaled information but not stride-length-scaled information,

(1) Absolute distance to an isolated object can be estimated by monocular observers making head movements in the frontal or sagittal planes (Panerai et al 2002; Peh et al 2002). Therefore, although the head-sway-scaled information identified in equation (6) was not available in the wall condition, we cannot rule out the possibility that head sway played some role in the perception of aperture size in the wall condition.

they should perform well in the post and tall-post conditions but not in the wall condition. Lastly, if subjects are able to use stride-length-scaled information, they should perform well in all three conditions.

2 Method

2.1 *Participants*

Twelve students (five female, seven male) participated in this study in exchange for extra credit in an undergraduate Psychology course. All participants had normal or corrected-to-normal vision, and no visual or motor impairments.

2.2 *Equipment and virtual environment*

The experiment was conducted in a 6.5 m × 9 m immersive virtual environment laboratory. Participants viewed the virtual environment through an nVis nVisor SX stereoscopic HMD. An eye patch was placed over their left eye to ensure monocular viewing. The virtual environment was created using Sense 8 World Tool Kit software running on a Dell Workstation 650 with a Wildcat 7110 dual-head graphics card. The HMD weighed 1000 g, had a resolution of 1280 pixels × 1024 pixels per eye, and a diagonal field-of-view of 60°. Participants' head position and orientation were tracked with an Intersense IS-900 motion tracking system. Data from the head tracker were used to update the position and orientation of the simulated viewpoint with an end-to-end latency of approximately 70 ms, which was measured by a method similar to that described in Di Luca (2010). The cables from the HMD and tracking system were bundled together and attached to a backpack worn by the experimenter, who walked alongside the participants as they moved to ensure that their movement was not restricted by the cables.

The positions of the left and right shoulders were tracked by a 14-camera Vicon MX motion-tracking system. The system tracked the positions of two lightweight, retro-reflective markers that were affixed to subjects' shoulders with double-sided tape. A third marker was affixed near the trapezius muscle on one side to disambiguate the left and right shoulder markers.

In the post conditions, the virtual environment consisted of a blue sky with a green grass-textured ground plane with 2 m tall × 10 cm wide solid yellow cylindrical posts. In the tall-post condition, there was no ground plane and the posts extended from the bottom to the top of the subject's visual field such that neither the base nor the cap was visible. In the wall condition, there was no ground plane and the virtual environment consisted of two textureless solid-yellow frontal walls that extended from the bottom to the top and to the periphery of the subject's visual field such that the only edges that were visible were the inside vertical edges.

Given the precision that is demanded by the task of walking through a narrow aperture, it was important to consider the possible effects of perceptual distortions introduced by the use of virtual environments. Significant underestimations of egocentric distance have been observed in subjects viewing virtual environments through HMDs (Knapp and Loomis 2004; Thompson et al 2004). Although misperceptions of exocentric distance tend to be smaller (Richardson and Waller 2005; Waller 1999), there still appears to be a bias to underestimate (Geuss et al 2010). Consistent with an underestimation of exocentric distance, pilot studies revealed that subjects often rotated their shoulders even when the aperture was wider than 1.3 × shoulder width, the critical value above which subjects in Warren and Whang (1987) did not rotate their shoulders. Therefore, we chose a range of aperture sizes that extended from 0.95 × shoulder width to 1.85 × shoulder width. Note that, although there was an overall tendency for subjects to rotate their shoulders more than necessary, there is no reason to believe that the bias to underestimate aperture size in virtual environments should affect subjects differently in the post, tall-post, and wall conditions. Therefore, this bias does not present any problems for evaluating the contributions of eyeheight-scaled, head-sway-scaled, and stride-length-scaled information, which was the main goal of the study. In addition, subjects completed a short (20 trial) warm-up block in which they performed the same task that was performed during the actual experiment. The warm-up block was completed in the post condition and was intended to allow subjects to familiarize themselves with the virtual environment, which has been found to reduce the effects of compression (Waller and Richardson 2008).

2.3 *Procedure*

At the beginning of each experimental session, the three retro-reflective markers were placed on the subject's left and right shoulders and trapezius muscle. The distance between the two markers was recorded, and entered into the computer as the estimate of shoulder width to which aperture sizes were scaled.

Subjects were read instructions describing the task that they were about to perform. The eye patch and HMD were then placed on the subject's head and he or she was told to walk to the home location to prepare for the first trial. The home location was a 0.4 m wide $\times$ 0.4 m long $\times$ 2 m tall box with translucent walls that changed color from red to yellow once the subject was inside the box. To ensure that subjects were properly oriented, they were instructed to face a vertical alignment marker, which appeared as a narrow red line in the distance. Once the subject was in the start box and properly aligned, the experimenter pressed a button on a handheld remote mouse. Upon pressing the mouse button, the home location and alignment marker disappeared and the obstacles or wall defining the aperture appeared at a random distance between 2 and 4 m from the start box. The size of the aperture was manipulated as an independent variable with seven levels: 0.95, 1.10, 1.25, 1.40, 1.55, 1.70, and 1.85 times the subject's shoulder width.

Subjects were instructed to approach and pass through the aperture, rotating their shoulders as necessary to avoid colliding with the edges of the aperture. If the subject successfully passed through the aperture without colliding with either side, a pleasant auditory tone played. If either shoulder collided with the edges of the aperture, a thud sound was played at the moment that the collision took place. After passing through the aperture, subjects continued to walk straight ahead. The experimenter pressed a button on the mouse and the home location reappeared in front of the subject. At this point, the subject entered the home location and turned 180° to face the alignment marker in preparation for the next trial.

2.4 *Design*

During the experiment, trials were completed in blocks with environment (post, tall-post, or wall) fixed within each block. Blocks 1, 3, and 5 were performed in the post environment and included 14 trials per block. Blocks 2 and 4 were performed in the tall-post and wall environments and included 42 trials per block. The post environment was always used in block 1 because the virtual environment was so impoverished in the other two conditions. Our concern was that being exposed to one of the two impoverished environments without first performing the task in the richer post environment would have been confusing. Had subjects performed block 1 in the tall-post or wall conditions, they may not have understood the task or adopted an unusual strategy that was specific to that environment. Although the post environment was always used in block 1, the 42 trials in the post condition were evenly distributed between blocks 1, 3, and 5 by presenting 14 trials per block to control for order effects. By the end of the experiment, each subject had completed six repetitions of each of the seven aperture widths for a total of 42 trials in each condition. The order of tall-post and wall environments was counterbalanced across subjects. There was a mandatory 3 to 5 min break between the 2nd and 3rd blocks. The entire experiment took approximately 60–75 min to complete, which included a 20-trial warm-up block in the post condition that was completed prior to the first trial of the experiment as well as at least one break.

2.5 *Data analysis*

Several dependent measures were based on the distance along the x-axis between the left and right shoulders, which we abbreviate W_x (see figure 4a). When the subject's torso was perfectly aligned with the z-axis, W_x was equal to the subject's shoulder width (W). When the subject rotated his or her shoulders to pass through the aperture, W_x decreased, reaching a minimum at the same time that shoulder rotation reached its maximum (see figure 4b).

For trials in which subjects rotated their shoulders, the initiation of shoulder rotation was identified by searching for the first frame in which the ratio of the change in W_x to the change in position along the z-axis exceeded 0.05 (see first white marker in figure 4b). The time-to-contact (TTC) at shoulder turn initiation was estimated by calculating the ratio of the distance to the aperture to the walking speed at the moment that shoulder rotation was initiated. Second, we calculated the distance between the subject and the aperture at the moment that W_x reached its minimum (ie at the moment of maximum shoulder rotation; see second white marker in figure 4b). If subjects are able to properly coordinate shoulder rotation with walking through the aperture, then W_x should reach its minimum at the moment that the subject passes through the aperture. Therefore, this provides a measure of spatial error. Third, we estimated W_x at aperture crossing by calculating the distance along the x-axis between the left and right shoulder markers at the moment that the subject passed through the aperture (see third white marker in figure 4b). Lastly, walking speed at aperture crossing was measured by calculating the mean speed of the subject from 0.5 m before the aperture to 0.5 m after the aperture.

Note that both TTC at shoulder turn initiation and spatial error can only be calculated on trials in which the subject actually rotated his or her shoulders. On trials in which subjects walked through the aperture without rotating their shoulders, these dependent measures could not be calculated. This occurred most often when aperture size was large. Therefore, the analysis of spatial error does not include data from the two largest aperture sizes. In addition, data from one subject were excluded from the analyses of TTC at shoulder turn initiation and spatial error because this subject did not turn her shoulders frequently enough to obtain a reliable estimate of spatial error.

On a small percentage ($M = 6.1\%$) of frames, the Vicon system failed to detect the position of the left or right shoulder markers, resulting in dropouts. When dropouts occurred at critical points during the trial (eg during shoulder turn initiation), certain dependent measures (eg TTC at shoulder turn initiation) could not be calculated in the usual way. Rather than excluding trials with dropouts, we developed an algorithm to estimate W_x using the data from the shoulder that did not drop out along with the data from the head tracker.

The method used to estimate W_x on dropout frames is illustrated in figure 6. Suppose that the Vicon system failed to record the position of the marker on the right shoulder. On such frames, we estimated the position of the axis of rotation of the shoulders ($\times$ in figure 6) using the position and orientation data recorded by the head tracker. This was achieved by moving from the position of the head in the direction opposite the orientation of the head by a distance d (see dotted line in figure 6). The value of d was varied on each trial to minimize the error between the estimated and actual values of W_x on frames in which dropouts did not occur. The next step was to use the position of the left shoulder and the position of the estimated axis of shoulder rotation to estimate the position of the right shoulder (see dashed line in figure 6).

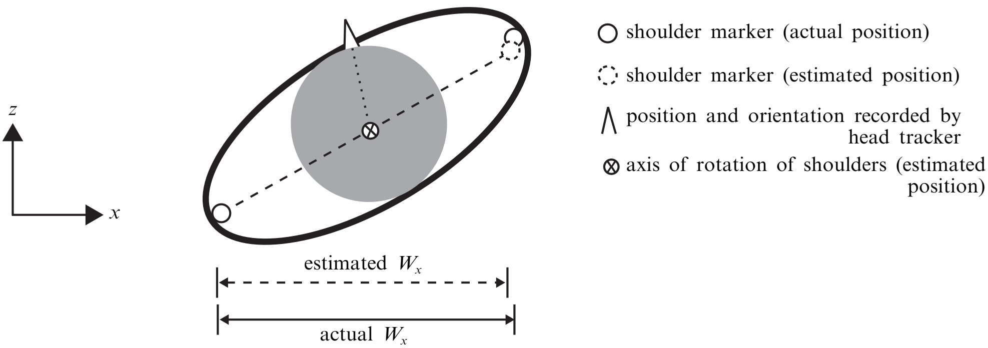


Figure 6. Method used to estimate W_x on dropout frames. See text for details.

The final step was to calculate the distance along the x-axis between the position of the left shoulder marker and the estimated position of the right shoulder.

The following procedure was used to measure the accuracy of the algorithm used to recover W_x on dropout frames. First, we identified for each trial the set of frames in which the Vicon system successfully recorded the position of both the left and right shoulder markers; that is, we identified frames in which dropouts did not occur. On each non-dropout frame, we used the data from the left shoulder to estimate the position of the right shoulder and W_x. We then calculated the difference (ie the error) between the estimated and actual values of W_x. Next, we repeated the procedure for the same frame using the data from the right shoulder to estimate the position of the left shoulder and W_x. Again, we calculated the difference between the estimated and actual values of W_x. The average of these two errors was considered the magnitude of error in estimation for that frame. This procedure was repeated for each non-dropout frame on each trial for all subjects. Averaged across frames, trials, and subjects, the overall mean absolute error was 0.9 cm, the overall mean constant error (ie estimated W_x minus actual W_x) was -0.1 cm, and the mean correlation (r) between estimated W_x and actual W_x was 0.93. Thus, for the small percentage of frames on which dropouts occurred, we were able to estimate W_x with excellent precision.

3 Results

Because the aim of this study was to investigate the perception of aperture size, we begin with the analysis of shoulder rotation and W_x at aperture crossing—the measures that best reflect perceived aperture size. We then present the analyses of shoulder turn initiation, spatial error, walking speed, and collision frequency. The data from these measures were averaged across repetitions within each condition prior to statistical analysis.

3.1 *Shoulder rotation at aperture crossing*

As expected, subjects generally rotated their shoulders as they passed through the aperture, and the angle of shoulder rotation increased as the size of the aperture decreased. Figure 7a shows shoulder rotation angle at aperture crossing as a function of aperture size in the post, tall-post, and wall conditions. A two-way (aperture size × environment) repeated-measures ANOVA revealed a significant main effect of aperture size ($F_{2.03, 20.25} = 30.15$, $p < 0.01$, partial $\eta^2 = 0.75$) and a significant aperture size × environment interaction ($F_{12, 120} = 6.05$, $p < 0.01$, partial $\eta^2 = 0.38$). The main effect of environment was not significant ($F_{2, 20} = 1.29$, $p = 0.30$, partial $\eta^2 = 0.11$). Follow-up analyses revealed that the simple main effect of aperture size was significant ($p < 0.05$) in all three conditions. Thus, the amount of shoulder rotation varied systematically with aperture size in all three environments, albeit to a lesser degree in the wall condition.

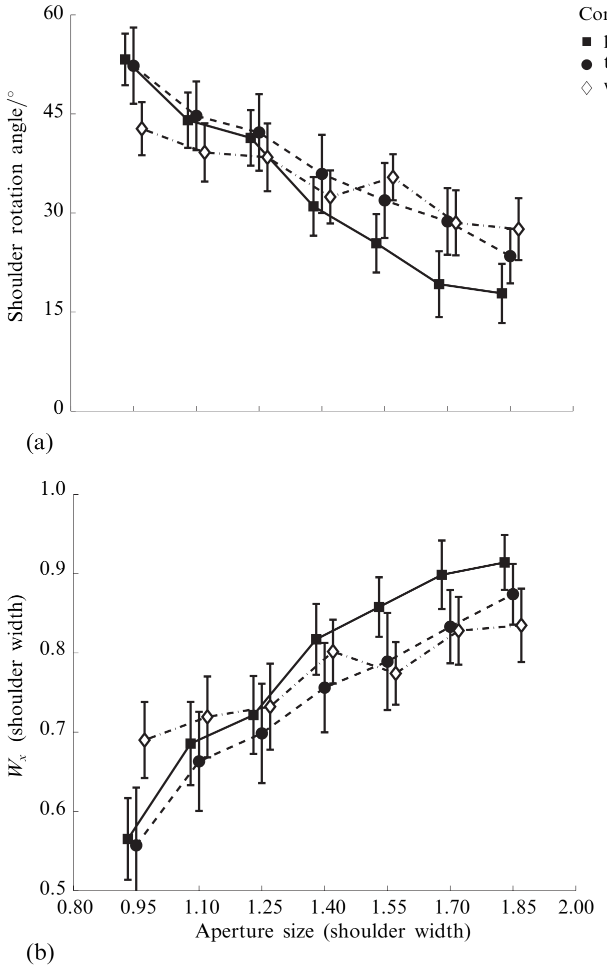


Figure 7. (a) Shoulder rotation angle (°) and (b) W_x (in units of shoulder width) at aperture crossing as a function of aperture size for the post, tall-post, and wall conditions.

It can be assumed that the amount by which subjects rotated their shoulders to reduce W_x while passing through the aperture reflects the perceived size of the aperture. As such, insight into the accuracy with which subjects were able to perceive aperture size across the three conditions can be gleaned by analyzing W_x at aperture crossing (see figure 7b). Consistent with the analysis of shoulder rotation, there was a significant main effect of aperture size ($F_{1.70,\,16.96} = 28.36$, $p < 0.01$, partial $\eta^2 = 0.74$) and a significant aperture size × environment interaction ($F_{12,\,120} = 6.24$, $p < 0.01$, partial $\eta^2 = 0.38$), but the main effect of environment was not significant ($F_{2,\,20} = 1.68$, $p = 0.21$, partial $\eta^2 = 0.14$). Follow-up analyses revealed that the simple main effect of aperture size was significant ($p < 0.05$) in all three conditions. Thus, W_x decreased as aperture size decreased in all three conditions, but less systematically in the wall condition.

3.2 *Initiation of shoulder rotation*

Subjects tended to initiate shoulder rotation between 0.5 and 0.7 s before reaching the aperture. Shoulder rotation began slightly earlier for smaller apertures, yielding a significant main effect of aperture size ($F_{4,\,36} = 2.90$, $p < 0.05$, partial $\eta^2 = 0.24$). This is most likely due to the fact that subjects rotated their shoulders more to walk through smaller apertures. Neither the main effect of environment ($F < 1$) nor the environment × aperture size interaction ($F_{8,\,72} = 1.54$, $p = 0.16$, partial $\eta^2 = 0.15$) were significant.

3.3 *Spatial error*

Recall that spatial error refers to the distance along the *z*-axis between the aperture and the subject at the moment that W_x reaches its minimum, as shown in figure 4b. We analyzed both absolute spatial error and signed spatial error. For absolute spatial error, neither of the main effects ($F < 1$ for environment; $F_{1.95, 17.54} = 1.53$, $p = 0.24$, partial $\eta^2 = 0.15$ for aperture size) nor the interaction were significant ($F_{8, 72} = 1.81$, $p = 0.09$, partial $\eta^2 = 0.17$). Maximum shoulder rotation tended to occur within 10 to 15 cm of the aperture, regardless of the size of the aperture or the environment.

Signed spatial error is shown in figure 8, with negative and positive values indicating that maximum shoulder rotation occurred before and after passing through the aperture, respectively. A two-way repeated-measures ANOVA revealed significant main effects of aperture size ($F_{4, 36} = 6.13$, $p < 0.01$, partial $\eta^2 = 0.41$) and environment ($F_{2, 18} = 4.87$, $p < 0.05$, partial $\eta^2 = 0.35$), and a marginally significant environment × aperture size interaction ($F_{8, 72} = 2.07$, $p = 0.05$, partial $\eta^2 = 0.19$). As illustrated in figure 8, W_x tended to reach its minimum slightly after passing through the aperture for small aperture sizes and slightly before reaching the aperture for large aperture sizes. This trend was greatest in the wall condition. The effect of aperture size could be due to the fact that subjects had to rotate their shoulders more to fit through smaller apertures. Larger shoulder rotations would be expected to take more time, which could account for the effect of aperture size. The effect of aperture size being greatest in the wall condition could be due to difficulty of perceiving aperture size in that condition, as suggested by figure 7. Nonetheless, mean signed spatial error varied between 10 cm before and 5 cm after the aperture, indicating that systematic errors were small in all three conditions.

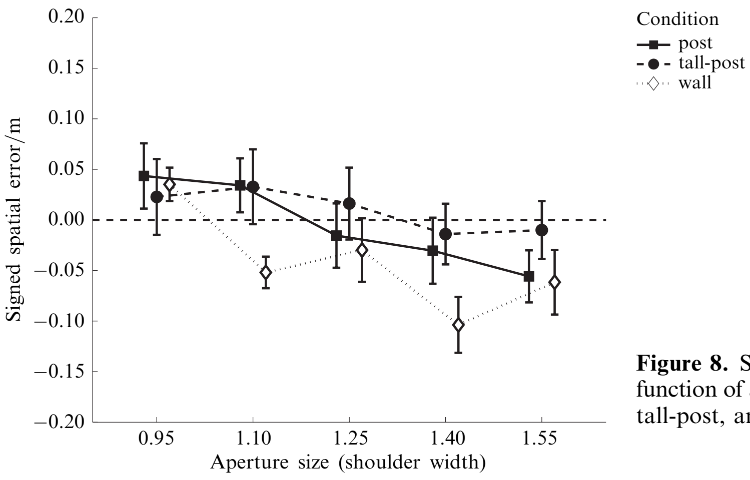


Figure 8. Signed spatial error as a function of aperture size for the post, tall-post, and wall conditions.

3.4 *Walking speed*

Not surprisingly, subjects walked slower when passing through narrow apertures compared to wider apertures, as confirmed by a significant main effect of aperture size ($F_{1.57, 15.67} = 10.41$, $p < 0.01$, partial $\eta^2 = 0.51$; see figure 9). Although mean walking speed was slightly higher in the post condition compared to the tall-post and wall conditions, the difference was not statistically significant ($F_{1.11, 11.08} = 2.92$, $p = 0.11$, partial $\eta^2 = 0.23$). The aperture size × environment interaction was not significant either ($F_{12, 120} = 1.64$, $p = 0.09$, partial $\eta^2 = 0.14$). Thus, the pattern of walking speed was similar in all three environments.

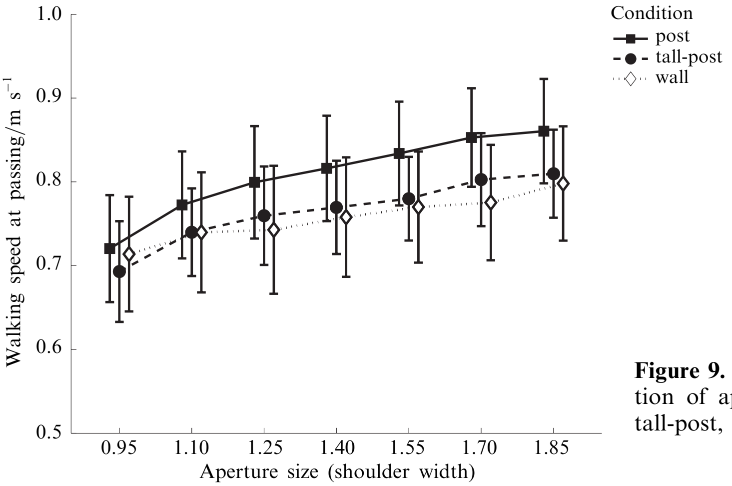


Figure 9. Walking speed as a function of aperture size for the post, tall-post, and wall conditions.

3.5 *Collisions*

Figure 10a shows the percentage of trials with collisions as a function of aperture size for all three conditions. Not surprisingly, the frequency of collisions increased as aperture size decreased ($F_{2.20,\,21.98} = 16.44$, $p < 0.01$, partial $\eta^2 = 0.62$). There was also a significant main effect of environment ($F_{2,20} = 4.07$, $p < 0.05$, partial $\eta^2 = 0.29$), reflecting the fact that collisions occurred more frequently in the tall-post ($M = 14.94\%$) and wall ($M = 19.48\%$) conditions compared to the post condition ($M = 6.93\%$). The increase in collision frequency in the tall-post and wall conditions may seem puzzling in light of the analyses reported above, all of which revealed similar performance across all three conditions. Consider, for example, the analysis of body width at aperture crossing. It can be assumed that the amount by which subjects rotated their shoulders to reduce body size while passing through the aperture reflects the perceived size of the aperture. If so, the fact that body width at aperture crossing decreased with aperture size in all three conditions suggests that subjects were able to differentiate apertures of different sizes in all three conditions. Therefore, the increase in collisions in the tall-post and wall conditions is not likely to be due to difficulty perceiving aperture size in the absence of eyeheight-scaled information. By the same logic, the fact that spatial error was similar in all three conditions suggests that the increase in collisions was not due to difficulty in timing shoulder rotation in the absence of eyeheight-scaled information.

So why were collisions more frequent in the tall-post and wall conditions? Recall that the virtual environments in these two conditions were extremely impoverished (see figure 5). Although all three environments contained sufficient information to perceive the size of the aperture, other aspects of guiding one's body through the aperture may have been more difficult due to the lack of visual structure (a textured ground plane) in the tall-post and wall conditions. For example, the only visible objects in the tall-post and wall environments were the two posts and two walls, respectively. As subjects approached and passed through the aperture, they may have had more difficulty centering themselves due to the lack of visual structure. In the absence of visual structure, the so-called flow-equalization strategy, which involves centering oneself by equalizing the rate of optic flow on the left and right sides (Duchon and Warren 2002), could not be used. Thus, having already shown that subjects were able to accurately perceive aperture size and rotate their shoulders by the appropriate amount in all three conditions, we now consider the possibility that the increase in collision frequency in the tall-post and wall conditions was due to difficulty centering oneself in the aperture.

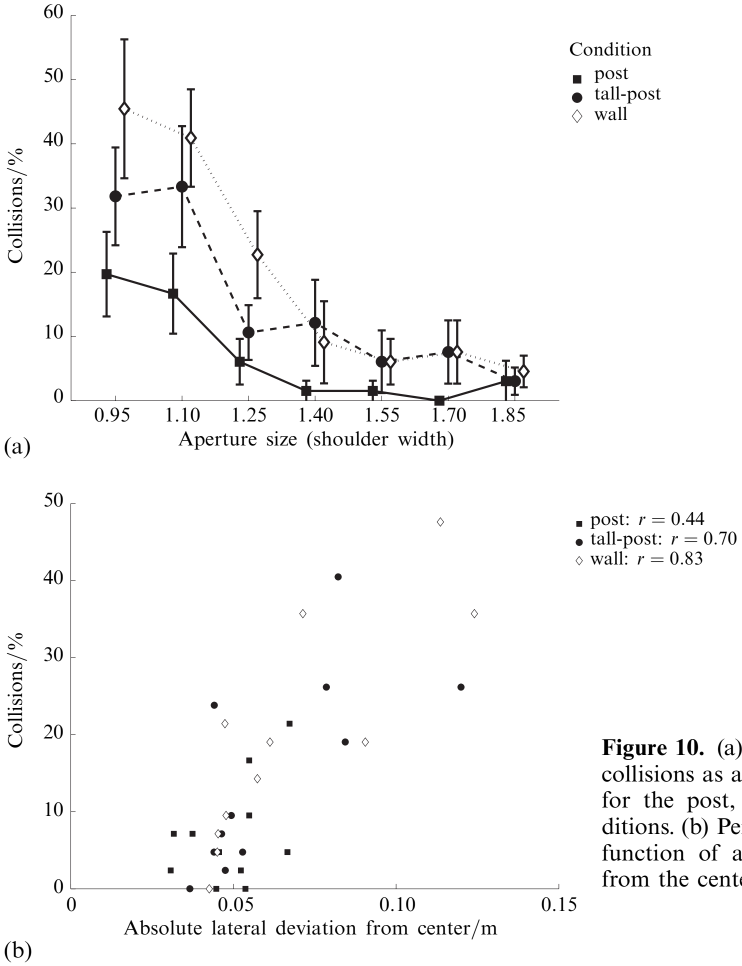


Figure 10. (a) Percentage of trials with collisions as a function of aperture size for the post, tall-post, and wall conditions. (b) Percentage of collisions as a function of absolute lateral deviation from the center of the aperture.

We tested this explanation by measuring the distance between the point midway between the two shoulder markers and the center of the aperture at the moment that subjects passed through the aperture. The larger the distance between these two points, the larger the error in centering. Indeed, there was a strong positive correlation between absolute deviation from the center of the aperture and the number of collisions in all three conditions (see figure 10b), confirming that collision frequency increased with deviation from the center of the aperture. Moreover, the mean absolute deviation from aperture center was significantly greater in the tall-post ($M = 6.2$ cm) and wall ($M = 6.8$ cm) conditions compared to the post condition ($M = 4.9$ cm) ($F_{2,20} = 3.56$, $p < 0.05$, partial $\eta^2 = 0.26$). Thus, the increase in the number of collisions in the tall-post and wall conditions appears to be primarily due to difficulty of centering oneself while passing through the aperture, which was most likely due to the absence of visual structure. In other words, the increase in collision frequency does not appear to reflect difficulty using body-sway-scaled or stride-length-scaled information to perceive aperture size.

One might also wonder why the frequency of collisions was as high as it was in the post condition. Although the virtual environment used in the post condition contained more visual structure than the other conditions, it was still impoverished compared to the real world. Moreover, the virtual environment was viewed monocularly in all three conditions through a head-mounted display with a field of view that is narrower than the unrestricted field of view. Thus, difficulty centering oneself while passing through the aperture may have been a factor even in the post condition, which would account for the overall high rate of collisions. Another factor that may have contributed to the high collision rate is the fact that collisions with virtual objects do not cause bodily harm as do real-world objects. Therefore, subjects may have been less careful about colliding with the virtual objects.

4 Discussion

The aim of this study was to investigate the sufficiency of three different sources of visual information about size for controlling walking through apertures: static eyeheight-scaled information, dynamic body-sway-scaled information, and dynamic stride-length-scaled information. Their contributions were tested by instructing subjects to walk through narrow apertures in three different virtual environments in which the availability of the three candidate sources of information was manipulated.

As expected, subjects rotated their shoulders as necessary to decrease body width as they passed through the aperture, and the amplitude of shoulder rotation varied systematically with the size of the aperture—subjects rotated their shoulders more on trials with small apertures. The degree to which shoulder rotation varied with aperture size was similar in the post and tall-post conditions, suggesting that subjects were able to perceive aperture size in the absence of eyeheight-scaled information. In the wall condition, the relation between shoulder rotation and aperture size was less systematic. Nonetheless, the amplitude of shoulder rotation was greater for small apertures than for large apertures even in the wall condition, suggesting that subjects were able to use stride-length-scaled information to perceive aperture size, albeit with less accuracy.

Analyses of spatial error and shoulder turn initiation revealed that subjects were able to coordinate shoulder rotation with passing through the aperture in all three conditions. Although mean walking speed while passing through the aperture was lower in the tall-post and wall conditions, the difference was not statistically significant. Taken together, the results suggest that dynamic head-sway-scaled and stride-length-scaled information are sufficient to perceive aperture size and coordinate shoulder rotation with passing through apertures. People can also perform the task using stride-length-scaled information alone, but with less accuracy.

4.1 *Contributions of static and dynamic information*

Although performance was similar in all three conditions, this does not mean that people do not rely on static, eyeheight-scaled information. To the contrary, judgments of passability made while moving are not different than judgments made while stationary (Warren and Whang 1987), suggesting that the availability of dynamic information does not improve the perception of passability. If eyeheight-scaled information is sufficient to perceive aperture size, then in what situations would dynamic information play a role? One answer is that dynamic information is used when eyeheight-scaled information is unavailable. As mentioned in section 1, eyeheight-scaled information is only available when the base of support is resting on the ground surface, which is not always the case. For example, when the aperture is formed by a pair of overhanging tree branches or a rock that juts out into an opening, dynamic head-sway-scaled information and stride-length-scaled information specify aperture size but eyeheight-scaled information does not. Thus, dynamic information about size may play a role when static eyeheight-scaled information is unavailable.

The contribution of static and dynamic information about size may also vary with distance to the aperture. When a person is several steps from an aperture, the base of support of the edges of the aperture is likely to be well within the field of view. As the person approaches the aperture, however, the angle of declination of the base of support (ie γ in figure 1) increases. If people rely on eyeheight-scaled information during the last couple of steps before passing through the aperture, one might expect gaze to shift downward toward the base of support. However, eye-movement studies suggest that people look at the aperture itself or the edges of the aperture, but not its base of support (Higuchi et al 2009).

Conversely, dynamic information may be more difficult to detect at larger distances because the rate of optical expansion of the aperture and the rate of optical motion of the edges of the aperture ($\dot{\theta}$ and $\dot{\phi}$ in figure 2) are slower at greater distances. As optical expansion rate increases (as distance to the aperture decreases), the contribution of dynamic information may increase. Thus, people may switch from static eyeheight-scaled information when they are several steps from the aperture to dynamic head-sway-scaled or stride-length-scaled information when they are closer. In other words, static and dynamic information about aperture size may serve as complementary sources of information that allow for robust size perception across a range of distances.

4.2 *Intrinsic units*

The claim that head-sway-scaled information specifies aperture size in intrinsic units of shoulder width rests on the assumption that the perceptual system can calibrate itself to the relation between the amplitude of lateral head sway and shoulder width. Similarly, the claim that stride-length-scaled information specifies aperture size in intrinsic units of shoulder width rests on the assumption that the perceptual system can calibrate itself to the relation between stride length and shoulder width. In this section, we consider the validity of these assumptions.

In the case of head-sway-scaled information, one possibility is to assume that the variation in lateral head sway amplitude during normal locomotion is negligible. If this assumption holds, then because shoulder width is also constant, the relation between lateral head sway amplitude and shoulder width is fixed and can therefore be reflected in the calibration of the perceptual system. Thus, just as eyeheight-scaled information specifies aperture size in units of shoulder width by virtue of the invariant relation between eyeheight and shoulder width, head-sway-scaled information specifies aperture size in units of shoulder width by virtue of the invariant relation between the amplitude of head sway and shoulder width. Similarly, to the degree that the variation in stride length during normal locomotion is negligible, the relation between stride length and shoulder width is also fixed, which is sufficient for stride-length-scaled information to specify aperture size in units of shoulder width.

Of course, both lateral head-sway amplitude and stride length vary under certain circumstances. However, changes in head-sway amplitude and stride length relative to normal are accompanied by changes in kinesthetic, vestibular, and visual information. Thus, head-sway-scaled and stride-length-scaled information could specify aperture size in intrinsic units even when head-sway amplitude and stride length vary.

4.3 *Conclusion*

Navigating through cluttered environments requires people to select routes and guide movements in a way that takes into account their body dimensions. The availability of body-scaled information provides a means by which the dimensions of the environment can be perceived in intrinsic units, allowing for the perception of action-relevant properties without having to rely on stored knowledge of body dimensions. Although the significance of body-scaled information has been appreciated for over 30 years, most empirical research on the perception of affordances has focused on eyeheight-scaled information. As anticipated by Lee (1980), however, eyeheight is not the only yardstick by which properties of the world are optically specified in intrinsic units. Dimensions of the actor's movements provide other possible yardsticks. The findings of the present study suggest that visual information scaled by head-sway amplitude and stride length may play an important role in perceiving action-relevant properties of the environment.

Acknowledgments. This research was supported by grants from the National Institutes of Health (1R01EY019317) and the National Science Foundation (BCS 0545141). We thank Sean Sullivan for programming the virtual environments, Julia van Cleve for assistance analyzing the data, and Jon Matthis for assistance measuring end-to-end latency.

References

Arzamarski R, Harrison S J, Hajnal A, Michaels C F, 2007 "Lateral ball interception: hand movements during linear ball trajectories" *Experimental Brain Research* **177** 312 – 323, doi:10.1007/s00221-006-0671-8

Bootsma R J, 1991 "Predictive information and the control of action: What you see is what you get" *International Journal of Sports Psychology* **22** 271 – 278

Chang C H, Wade M G, Stoffregen T A, 2009 "Perceiving affordances for aperture passage in an environment – person – person system" *Journal of Motor Behavior* **41** 495 – 500

Di Luca M, 2010 "New method to measure end-to-end delay of virtual reality" *Presence: Teleoperators and Virtual Environments* **19** 569 – 584

Duchon A P, Warren W H Jr, 2002 "A visual equalization strategy for locomotor control: of honeybees, robots, and humans" *Psychological Science* **13** 272 – 278

Geuss M N, Stefanucci J, Creem-Regehr S H, Thompson W B, 2010 "Can I pass? Using affordances to measure perceived size in virtual environments" paper presented at the 7th Annual ACM SIGGRAPH Symposium of Applied Perception in Graphics and Visualization, Long Beach, CA

Gray R, Sieffert R, 2005 "Different strategies for using motion-in-depth information in catching" *Journal of Experimental Psychology: Human Perception and Performance* **31** 1004 – 1022

Higuchi T, Cinelli M E, Greig M A, Patla A E, 2006 "Locomotion through apertures when wider space for locomotion is necessary: adaptation to artificially altered bodily states" *Experimental Brain Research* **175** 50 – 59

Higuchi T, Cinelli M E, Patla A E, 2009 "Gaze behavior during locomotion through apertures: the effect of locomotion forms" *Human Movement Science* **28** 760 – 771

Higuchi T, Takada H, Matsuura Y, Imanaka K, 2004 "Visual estimation of spatial requirements for locomotion in novice wheelchair users" *Journal of Experimental Psychology: Applied* **10** 55 – 66; doi:10.1037/1076-898x.10.1.55

Jacobs D M, Michaels C F, 2006 "Lateral interception I: operative optical variables, attunement, and calibration" *Journal of Experimental Psychology: Human Perception and Performance* **32** 443 – 458, doi:10.1037/0096-1523.32.3.443

Knapp J M, Loomis J M, 2004 "Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments" *Presence: Teleoperators and Virtual Environments* **13** 572 – 577

Lee D N, 1980 "Visuo-motor coordination in space-time", in *Tutorials in Motor Behavior* Eds G E Stelmach, J Requin (Amsterdam: North-Holland) pp 281 – 295

Panerai F, Cornilleau-Pérès V, Droulez J, 2002 "Contribution of extraretinal signals to the scaling of object distance during self-motion" *Perception and Psychophysics* **64** 717 – 731

Peh C-H, Panerai F, Droulez J, Cornilleau-Pérès V, Cheong L-F, 2002 "Absolute distance perception during in-depth head movement: calibrating optic flow with extra-retinal information" *Vision Research* **42** 1991 – 2003

Peper L, Bootsma R J, Mestre D R, Bakker F C, 1994 "Catching balls: how to get the hand to the right place at the right time" *Journal of Experimental Psychology: Human Perception and Performance* **20** 591 – 612

Regan D, Kaushal S, 1994 "Monocular discrimination of the direction of motion in depth" *Vision Research* **34** 163 – 177; doi:0042-6989(94)90329-8[pii]

Richardson A R, Waller D, 2005 "The effect of feedback training on distance estimation in virtual environments" *Applied Cognitive Psychology* **19** 1089 – 1108, doi:10.1002/acp.1140

Savelsbergh G J, Douwes Dekker L, Vermeer A, Hopkins B, 1998 "Locomoting through apertures of different width: a study of children with cerebral palsy" *Pediatric Rehabilitation* **2** 5 – 13

Sedgwick H A, 1980 "The geometry of spatial layout in pictorial representation", in *The Perception of Pictures* Ed. M Hagen (New York: Academic Press) pp 33 – 90

Thompson W B, Willemsen P, Gooch A A, Creem-Regehr S H, Loomis J M, Beall A C, 2004 "Does the quality of the computer graphics matter when judging distances in visually immersive environments?" *Presence: Teleoperators and Virtual Environments* **13** 560 – 571

Wagman J B, Malek E A, 2007 "Perception of whether an object can be carried through an aperture depends on anticipated speed" *Experimental Psychology* **54** 54 – 61, doi:10.1027/1618-3169.54.1.54

Wagman J B, Taylor K R, 2005 "Perceiving affordances for aperture crossing for the person-plus-object system" *Ecological Psychology* **17** 105 – 130

Waller D, 1999 "Factors affecting the perception of interobject distances in virtual environments" *Presence: Teleoperators and Virtual Environments* **8** 657 – 670

Waller D, Richardson A R, 2008 "Correcting distance estimates by interacting with immersive virtual environments: effects of task and available sensory information" *Journal of Experimental Psychology: Applied* **14** 61 – 72

Warren W H, 2007 "Action-scaled information for the visual control of locomotion", in *Closing the Gap: The Scientific Writings of David N. Lee* Eds G Pepping, M A Grealy (Mahwah, NJ: Lawrence Erlbaum Associates) pp 243 – 278

Warren W H, Whang S, 1987 "Visual guidance of walking through apertures: body-scaled information for affordances" *Journal of Experimental Psychology: Human Perception and Performance* **13** 371 – 383

Wraga M, 1999 "The role of eye height in perceiving affordances and object dimensions" *Perception & Psychophysics* **61** 490 – 507