

Visual Distance Estimation in Static Compared to Moving Virtual Scenes

Harald Frenz and Markus Lappe

Westfälische Wilhelms-Universität Münster, Germany

Visual motion is used to control direction and speed of self-motion and time-to-contact with an obstacle. In earlier work, we found that human subjects can discriminate between the distances of different visually simulated self-motions in a virtual scene. Distance indication in terms of an exocentric interval adjustment task, however, revealed linear correlation between perceived and indicated distances but with a profound distance underestimation. One possible explanation for this underestimation is the perception of visual space in virtual environments. Humans perceive visual space in natural scenes as curved, and distances are increasingly underestimated with increasing distance from the observer. Such spatial compression may also exist in our virtual environment. We therefore surveyed perceived visual space in a static virtual scene. We asked observers to compare two horizontal depth intervals, similar to experiments performed in natural space. Subjects had to indicate the size of one depth interval relative to a second interval. Our observers perceived visual space in the virtual environment as compressed, similar to the perception found in natural scenes. However, the nonlinear depth function we found can not explain the observed distance underestimation of visual simulated self-motions in the same environment.

Keywords: virtual environment, space perception, depth cues

El movimiento visual se emplea en el control de la dirección y la velocidad de la auto-locomoción y, también, para conocer el tiempo de contacto con un obstáculo. En trabajos anteriores encontramos que los observadores humanos pueden discriminar entre las distancias de diferentes auto-locomociones simuladas visualmente en una escena virtual. La indicación de la distancia mediante una tarea de ajuste de intervalo exocéntrico, sin embargo, reveló una correlación lineal entre las distancias percibidas y las indicadas, pero con una gran subestimación de la distancia. Una posible explicación de esta subestimación se basa en las características de la percepción visual del espacio en ambientes virtuales. En las escenas naturales los humanos percibimos el espacio visual como curvado, y las distancias se subestiman con el incremento de la separación respecto al observador. Esta compresión espacial también puede existir en nuestro ambiente virtual. Por ello, se decidió evaluar el espacio visual percibido en una escena estática virtual. Pedimos a los observadores que comparasen dos intervalos de profundidad horizontal, similares a experimentos llevados a cabo en el espacio natural. Los sujetos debían indicar el tamaño de un intervalo de profundidad con respecto a un segundo intervalo. Nuestros observadores percibían el espacio visual en el ambiente virtual como comprimido, similar a la percepción encontrada en escenas naturales. Sin embargo, la función no lineal de profundidad que encontramos no puede explicar la subestimación observada de la distancia de las auto-locomociones visuales simuladas en el mismo ambiente.

Palabras clave: ambiente virtual, percepción del espacio, claves de profundidad

We perceive self-induced optic flow fields in our daily life as we move through the environment. The information in optic flow fields becomes especially important in the use of motor-driven vehicles. In these situations, precise information about the movement coming from the proprioceptive and vestibular system is missing, and people have to rely mainly on the information provided by the visual system. The question arises whether optic flow fields can provide the information necessary to successfully navigate. Previous research has shown that human subjects can use optic flow to detect the heading of self-motion (see Warren, 1998; and Lappe, Bremmer, & van den Berg, 1999, for extensive discussion) and the time-to-contact with environmental objects (Hecht & Savelsbergh, 2004; Lee, 1980). The upright stance (Bronstein & Buckwell, 1997; Lee, 1980) and the speed of walking (Prokop, Schubert, & Berger, 1997) are also controlled by optic flow information. Thus, optic flow can indeed provide important information for successful navigation through the environment and even affects normal navigational behavior when other kinds of information (e.g., proprioceptive information) are unavailable.

Time-to-contact can be derived unambiguously from optic flow alone because it combines speed and distance of an approaching object. Speed itself cannot be derived from optic flow because the speed of visual motion depends on the distance of the moving object. For that reason, it is also impossible to recover the distance traveled from optic flow alone. A scale factor is necessary to calibrate optic flow speed to the environment. However, if combined with information about the depth structure of the environment, optic flow can be used to gauge travel distance of a self-motion. Bremmer and Lappe (1999) showed that human subjects can reproduce and discriminate the travel distances of two sequentially simulated self-motions only on the basis of optic flow information. To solve such a discrimination task without absolute scales in the scene, observers have to assume that the depth structure of the virtual scene remained constant between the two motion simulations. Given this assumption, the participants could have either used the image velocity of environmental objects on the retina (2-D hypothesis) or the perceived velocity of the self-motion (3-D hypothesis) to estimate the traveled distance. To differentiate between these two hypotheses Frenz, Bremmer, and Lappe (2003) conducted experiments in which the optic flow was altered by varying the simulated ego velocity and the perspective on the virtual scene. If the observers did not notice the perspective change between the two motion sequences, they made predictable errors: They attribute the whole change in the optic flow field to a change of ego velocity. If they noticed the change of the environmental structure, they could divide the amount of change in the flow field into a part based

on this depth structure change and another part based on ego velocity change. This result supports the 3-D hypothesis. An immediate question arising from this result is whether human subjects can use this 3-D percept to build up an internal representation of the traveled distance in some arbitrary unit and, if this were the case, whether they can indicate this perceived distance in a static environment. This question is directed towards the calibration required to link accumulated visual speed to environmental depth. To answer this question, in earlier work (Frenz & Lappe, 2005), we visually simulated ego-motions over virtual ground planes on a projection screen (between 0.58 to 3.46 eye heights) and asked observers to indicate the perceived travel distance in terms of a virtual interval on the ground plane. Two virtual horizontal lines built this depth interval on the ground plane. We presented one of these lines at a constant distance of 1.54 eye heights to the observer's position. The observers had to move the second line in depth until the size of the interval matched the perceived travel distance of the motion sequence. We found linear correlation between simulated and indicated distances. Hence, observers could obtain a precise distance measure from optic flow. This linear correlation is consistent with other studies that investigated visual motion based on distance estimation while walking (Kearns, Warren, Duchon, & Tarr, 2002; Loomis et al., 1993; Witmer & Kline, 1998), riding a bike (Sun, Campos, & Chan, 2004), steering on a mobile robot (Berthoz, Israel, Georges-Francois, Grasso, & Tsuzuku, 1995), or navigating in a virtual environment (Peruch, May, & Wartenberg, 1997; Riecke, van Veen, & Bühlhoff, 2002; Witmer & Kline, 1998). However, the indicated distances in our experiments systematically undershot the simulated distances by about 30 percent. This undershoot even existed when observers first accurately reproduced the self-motion in an active control scenario and then indicated the distance of the seen (and reproduced) movement. We investigated several possible reasons why observers undershot the simulated distances. First, we varied the methods of distance indication. We used either active walking without visual information, verbal report in units of multiple eye-heights, or in terms of a second self-controlled visual motion prior to interval adjustment (Frenz & Lappe, 2005). Second, we provided additional static depth information by using stereoscopic scene presentation and an immersive environment (computer animated virtual environment, CAVE) (Lappe, Frenz, Bührmann, & Kolesnik, 2005). The results always revealed a strong correlation between simulated and indicated distances but the same undershoot of the indicated distance (Figure 1). Another possible explanation for the undershoot of distance judgment is the compressed percept of the visual space in virtual environments. To test this, we decided to survey the perceived visual space in a static virtual environment.

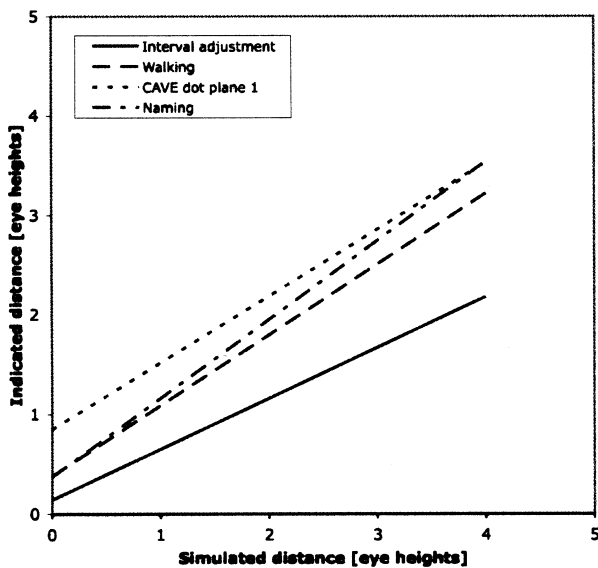


Figure 1. Relationship between indicated and simulated depth intervals in travel distance judgment from optic flow. The solid line corresponds to distance indication in terms of interval adjustment derived from optic flow. The dashed line shows the results when subjects indicated perceived distances in terms of active walking without vision. Naming is the verbal distance indication in terms of multiple eye heights. These three experiments were performed on a textured ground plane (see Frenz & Lappe, 2005). The dotted line corresponds to data obtained when subjects indicated the perceived travel distance in a Computer Animated Environment (CAVE) with stereoscopic stimulus presentation on a dot plane (see Lappe et al., 2005).

Physical space is Euclidean. Visual space in natural scenes, however, is not. Luneburg (1947) described visually perceived space as a Riemannian space of constant curvature. He based his theory on so-called “alley” experiments in which two rows of light points (e.g., street lights) in a frameless environment had to be arranged by the observer either equidistantly between the two sides of the street (“distance alley”) or equidistantly on either side of the street (“parallel alley”). In both tasks, rows that appeared parallel in perception were curved in physical space. Other researchers found perceived visual space to be curved even under full cue conditions (Koenderink, van Doorn, Kappers, & Todd, 2002; Norman, Todd, Perotti, & Tittle, 1996; Wagner, 1985). The exact curvature differed somewhat from Luneburg’s model, however, and varied depending on the task and the lighting conditions (Beusmans, 1998; Cuijpers, Kappers, & Koenderink, 2000, 2002; Foley, 1980; Foley, Ribeiro-Filho, & Da Silva, 2004; Indow, 1991; Wagner, 1985). These experiments were conducted in natural scenes, either outdoors or in lab rooms. But, as our optic flow experiments as well as many more experiments on human navigation are performed in virtual environments (e.g., heading estimation; Bertin, Israël, & Lappe, 2000), distance

estimation (Bremmer & Lappe, 1999; Frenz et al., 2003; Frenz & Lappe, 2005; Redlick, Jenkin, & Harris, 2001; Riecke et al., 2002; Sun, Campos, & Chan, 2004; Sun, Campos, Young, Chan, & Ellard, 2004), or triangle completion (Kearns et al., 2002; Peruch et al., 1997; Riecke et al., 2002), it is important to know how visual space is perceived in virtual scenes. Experiments on distance estimation and triangle completion in particular have demonstrated that observers systematically misperceive distances of simulated self-motion.

There are different ways to survey visual space (Loomis & Knapp, 2003). One can either use egocentric or exocentric distance judgment. In our previous studies of distance estimation from optic flow (Frenz & Lappe, 2005; Lappe et al., 2005), we initially attempted to use an egocentric distance judgment task. However, we found that observers used the lower part of the projection screen as reference for their position in the virtual scene, thus effectively employing an exocentric judgment. This strategy leads to misperception of the simulated distance. To circumvent this misperception, we decided to use an exocentric distance judgment task directly. While not necessarily identical, egocentric and exocentric distance judgments correlate with each other (Foley et al., 2004; Wu, Ooi, & He, 2004) so that we expect that measuring exocentric judgments is also somewhat informative for perceived space in egocentric terms. One procedure to measure exocentric distance is the bisection task. Observers have to divide a reference distance into two equal-sized test distances by changing the position of a marker along the reference distance. The disadvantage of this type of distance judgment is that changing the position of the marker changes the sizes of the two test distances. Another way to survey visual space is to use a reference depth interval and instruct the participants to indicate the perceived size in terms of a second depth interval. This test interval is adjacent to the reference interval, and it can be either closer or further away from the observer’s position than the reference interval. The advantage of this type of distance measure is that only one distance is adjusted (in contrast to the bisection task). A second advantage is that we were directly able to compare the results with the data obtained from distance judgments of visually simulated self-motions (Frenz & Lappe, 2005), as we also used an exocentric interval matching task in those experiments.

Our experimental approach for exocentric distance judgments follows the procedure introduced by Beusmans (1998). Besides allowing a comparison with the distance judgments in our optic flow studies, this approach also allows a direct comparison with distance judgments in real world scenes. In Beusmans’ experiments, observers had to match the 3-D size of two collinear depth intervals on an open field in a real world experiment. Three orange markers formed the two intervals. The nearest marker to the observer’s position (T1) and the second marker (T2) were rectangles (7 cm width, 10 cm high). The furthest marker (T3) was a

triangle (10 cm high, base 15 cm). The authors positioned the markers on the ground plane along a straight line in front of the observer. The size of the interval between T2 and T3 and the distance to the observer was set by the experimenter. The position of marker T1 and, therefore, the size of the near interval T1-T2 should be adjusted by the observers with a system of pulleys. Beusmans reported that interval sizes were increasingly underestimated with increasing interval size and with increasing distance to the observer.

We reproduced this experiment in a virtual scene using a textured ground plane and three virtual horizontal lines forming the depth intervals. Our participants had to match either the nearest or the furthest interval on the ground plane in size by adjusting either the furthest or the nearest of the three lines. With these experiments, we wanted to estimate the perceived depth structure of the virtual scene we used in the previous flow-based distance estimates and compare the perception of visual space in natural and virtual scenes. Although the task in the study of Beusmans (1998) and our experiments were identical, the experiments differed with respect to the depth information provided in the scenes. In natural scenes, people normally view the scene binocularly and get stereoscopic depth information in terms of binocular disparity. Additionally, in natural scenes small head or body movements provide depth information in terms of motion parallax. The virtual scenes we used lacked these depth cues as we presented them non-stereoscopically and without scene rendering according to the subject's head movement and orientation. We did this to allow comparison with our prior results, which were mostly collected with non-stereoscopic displays and fixed head position. Moreover, the distance estimation undershot in the optic flow experiments did not change with stereoscopic presentation of the stimuli or with displays that were yoked to head movements in a CAVE virtual environment (Lappe et al., 2005). Depth cues that were included in our virtual scenes comprised the density gradient of objects towards the horizon and the visual size of objects changed according to the distance to the observer. Because our experimental procedure is based on exocentric distance judgments completely within a virtual scene, judgments are ideally independent from the absolute scale of the environment. Moreover, absolute scale is not available in our set-up: The virtual scenes we used were ambiguous with respect to distance and length as we presented no scale or absolute depth cues. Therefore, distance judgment in meters did not seem to be a reasonable measure (see Loomis & Philibeck, 1999; and Loomis, Philibeck, & Zahorik, 2002, for discussions of this issue). Unambiguous information about distances in our virtual scenes is the observer's eye height above the ground plane. With a simulated gaze parallel to the ground plane, the distance of one eye height is always determined by a visual angle of 45° below the simulated gaze direction. Therefore, distance estimation in terms of eye heights seems to be a more reasonable measure.

Method

Participants

Four members of the department (age 24-30, 3 male 1 female), including the first author, voluntarily participated in these experiments. All had normal or corrected to normal vision. We gave no feedback about the participant's performance in the experiment.

Apparatus

We used the same experimental set-up as described in Frenz et al. (2003). We created all stimuli in real time on a Silicon Graphics Indigo2 workstation using custom-made OpenGL programs with the Iris Performer library. The spatial resolution of the stimuli was 1280 × 1024 pixel and the frame rate 36 Hz. The observer sat 60 cm in front of a back-projection screen (Dataframe, type CINEPLEX, 120 × 120 cm) on which the stimuli were projected using a CRT video projector (Electrohome ECP 4100). The resulting field of view was 90 × 90 degree. The observer's eye height was adjusted to be level with the virtual horizon. The participants viewed the stimuli binocularly. Neither head nor eye movements were recorded, but we instructed the participants to avoid head movements, as the stimuli were not yoked to head movements. The darkened experimental room was only illuminated by the stimulus (3.1 CD/m²).

Procedure

At the beginning of each trial, we presented a static virtual ground plane. We used a texture pattern (Iris Performer type "gravel") of 3.08 × 3.08 eye height size and mapped it onto a large (153.85 × 153.85 eye heights) virtual plane. Three horizontal white lines (15.38 eye heights wide to both sides of the observer, 2 pixels thick) were placed at ground level onto the plane. We placed the three horizontal lines at different distances from the observer's virtual position. Thus, the horizontal lines formed two depth intervals on the virtual ground plane. The screen-shots in Figure 2 illustrate the scene. To survey visual space, we positioned the two intervals on the ground plane at various depths and varied their sizes. The exact values of the distance between the observer and the second line and the various interval sizes are given in the parameters section. In each experimental trial, one of the two ground plane intervals was fixed, the other was adjustable by the participant. The observer's task was to match the size of the adjustable depth interval on the ground plane to the fixed one. Therefore, the participants could move one of the horizontal lines, either the nearest line (Experiment 1) or the furthest line (Experiment 2) towards or away from the observer's virtual position on the ground plane by moving a computer mouse. Resolution of the mouse movements corresponded to 0.01

eye heights as the smallest possible adjustment. The participants indicated their decision with a mouse button-press after they adjusted the interval on the ground. Before the next trial, the observer's virtual position was randomly shifted to either side to avoid recognition of same texture elements.

Parameters

In the first experiment, the near interval had to be matched in size to the far interval. We varied the size of the fixed interval and the distance to the observer (distance to the middle line). The distance to the fixed interval was 1.54,

1.92, 2.31, or 2.69 eye heights. The size of the interval ranged from 0.18 to 1.94 eye heights (see Table 1). We presented four interval sizes with each distance to the middle line, resulting in 16 conditions. Each condition was measured ten times. We presented all trials in a pseudo-randomized order.

In the second experiment, the size of the far interval had to be matched to the size of the near interval. The distance of the interval from the observer (distance to middle line) was again 1.54, 1.92, 2.31, or 1.69 eye heights. The sizes of the near interval in eye heights were the same as before (listed in Table 1). Thus, the "correct" positions of the first line of Experiment 1 were presented with the second line as the reference ground interval in Experiment 2.

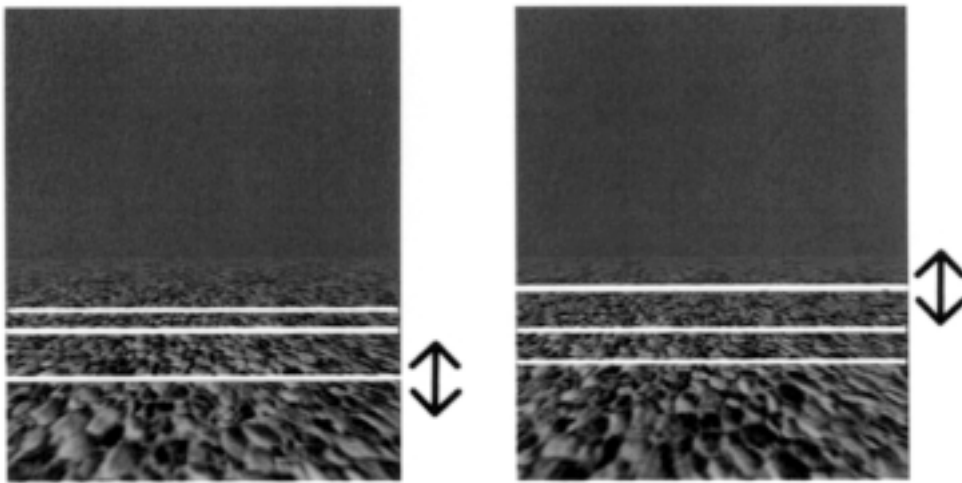


Figure 2. Screenshots of the visual scenes. The left image shows a scene from Experiment 1 (far interval is fixed in size), the right from Experiment 2 (near interval is fixed in size). The arrows next to the screenshots indicate the adjustable lines in the experiments.

Table 1

Distances of the Observer's Virtual Position on the Ground Plane to the Middle Line and Sizes of the Fixed Ground Intervals in Degree Viewing Angle (deg) and Eye Heights (EH)

Distance to middle line (EH)	Size of fixed interval (deg)	Distance to 3 rd line (EH)	Size of fixed interval (EH)
1.54	2	1.72	0.18
1.54	3	1.82	0.28
1.54	4	1.94	0.4
1.54	5	2.07	0.53
1.92	2	2.2	0.28
1.92	3	2.37	0.45
1.92	4	2.56	0.64
1.92	5	2.78	0.86
2.31	2	2.71	0.4
2.31	3	2.96	0.65
2.31	4	3.26	0.95
2.31	5	3.63	1.32
2.69	2	3.24	0.55
2.69	3	3.60	0.91
2.69	4	4.06	1.37
2.69	5	4.63	1.94

Results

To analyze the data, we calculated the distance ratio between the indicated interval size and the size of the fixed interval for each subject and condition. For perfect performance, the distance ratio should be 1, independent of the simulated interval size and distance to the observer.

Figure 3 shows the single-subject results of Experiment 1. Figure 4 shows the single-subject results of Experiment 2. Each data point in these figures represents the median distance ratio over ten trials of that condition. Error bars show the 30th and 70th percentile. The indices are plotted as a function of the size of the fixed ground interval. Note that the scaling of the y-axis in Figure 4 for subject *ms* is different from that of the other subjects. Symbols of data points indicate distances

between the observer's position and the position of the second horizontal line on the virtual ground plane (circle: 1.54 eye heights, square: 1.92 eye heights, triangle up: 2.31 eye heights, and triangle down: 2.69 eye heights).

In Experiment 1, distance ratios for all subjects decreased with increasing size of the fixed interval. Thus, subjects indicated smaller near intervals than the simulated far intervals and, therefore, underestimated the size of the far interval. Secondly, distance ratios for identical interval sizes (same symbols in Figure 3) were larger if the fixed intervals were further from the observer. These two effects are clear for two subjects (*hf* and *ms*). Subjects *kg* and *ps* showed the same tendency to increasingly underestimate the size of the reference interval with increasing size of the interval. However, these two subjects had distance ratios larger 1 for small interval sizes.

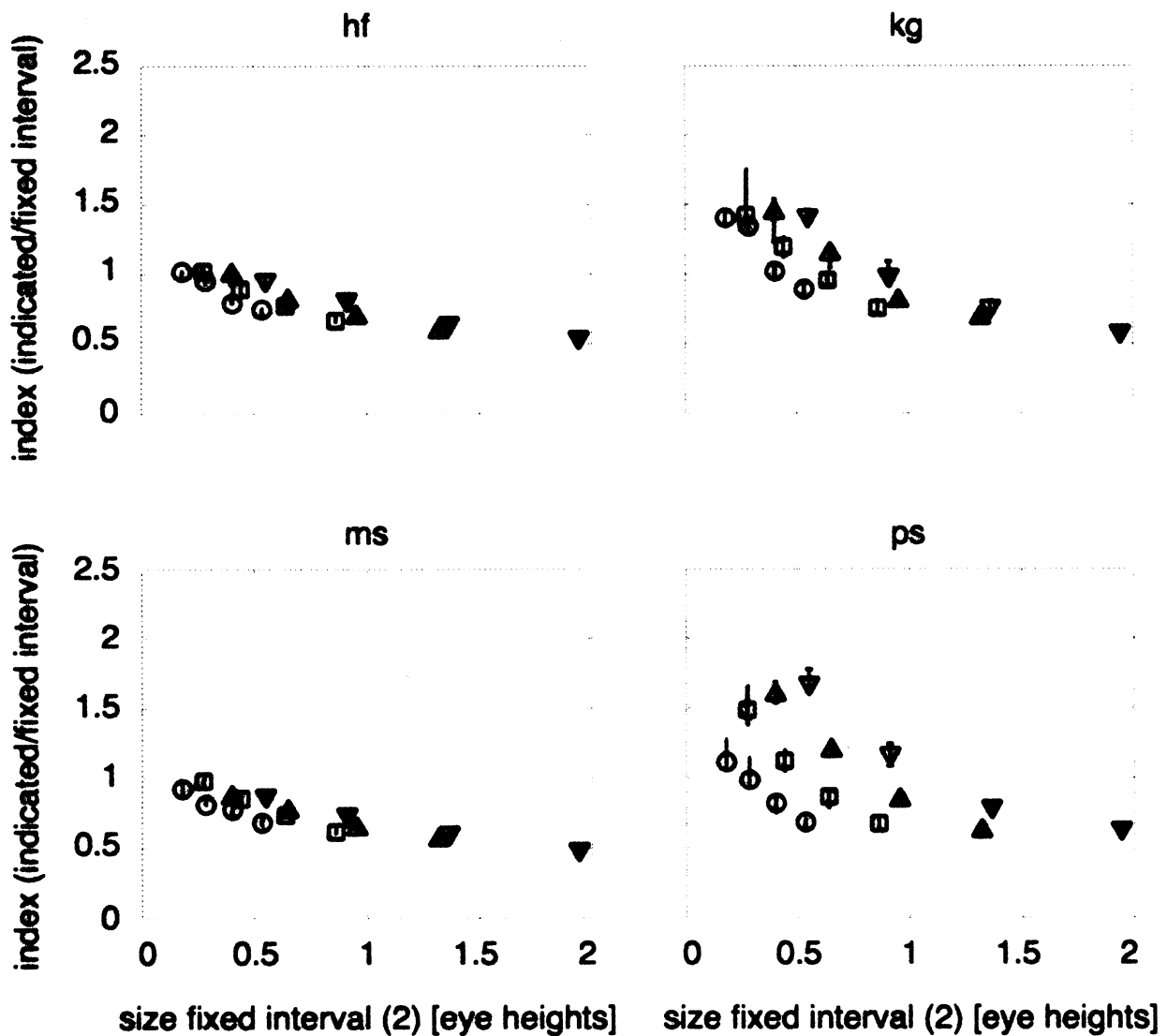


Figure 3. Ratios of indicated and fixed interval sizes (median) for each subject as a function of the fixed interval size. In this experiment, the far interval was fixed in size, and the subjects had to adjust the near interval. Symbols represent the distances to the middle line (circle: 1.54 eye heights, square: 1.92 eye heights, triangle up: 2.31 eye heights, and triangle down: 2.69 eye heights). Error bars are the 30th and the 70th percentile.

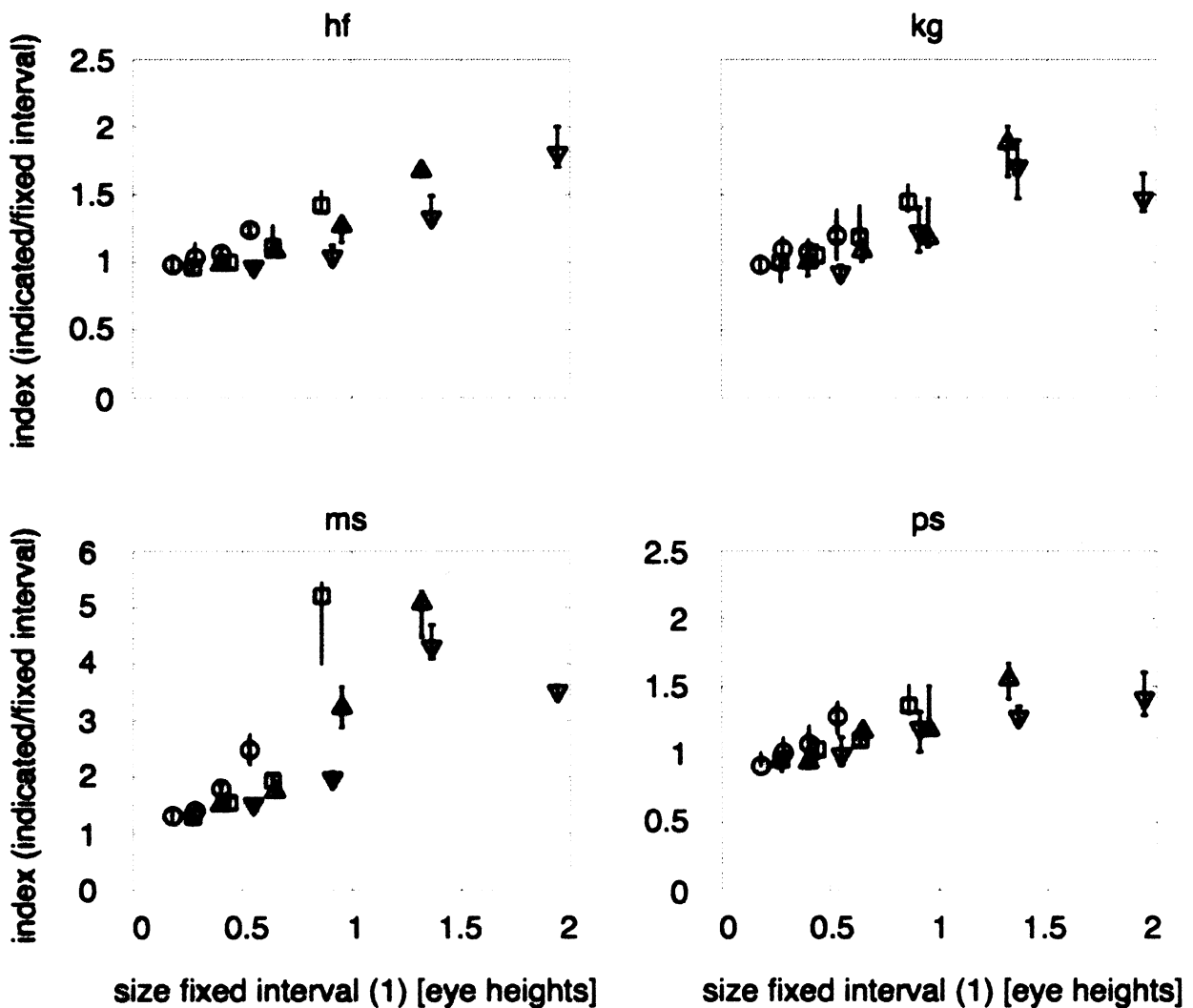


Figure 4. Ratios of indicated and fixed interval sizes (median) for each subject as a function of the fixed interval size. In this experiment, the near interval was fixed in size, and the far interval had to be adjusted. Symbols represent the distances to the middle line (circle: 1.54 eye heights, square: 1.92 eye heights, triangle up: 2.31 eye heights, and triangle down: 2.69 eye heights). Error bars are the 30th and the 70th percentile.

Distance ratios larger 1 indicate that the adjustable near interval was larger (in plane co-ordinates) than the reference interval. This offset vanishes with increasing size of the reference interval. For distances of 2.69 eye heights to the middle line and interval sizes of 1.94 eye heights, the amount of distance underestimation was the same for all subjects (around 50 %).

For distant intervals on the ground plane, the two indicator lines are closer together in screen co-ordinates than for near intervals. Therefore, the underestimation of the near interval size on the ground may originate from a bias to match the interval sizes in screen coordinates instead of ground plane coordinates. To test whether our subjects really matched interval sizes in screen coordinates, we compared the data (10 iterations \times 16 conditions \times 4 subjects) with the expected data of equalized interval sizes in screen coordinates. A Mann-Whitney-*U*-test revealed that the data

were significantly different ($p < .01$) from the expected data of equalizing the ground intervals in screen co-ordinates. Thus, our subjects must have judged the size of the virtual intervals on the ground plane but might have had a tendency towards uniformity in screen coordinates.

The upper panel of Figure 5 shows the pooled data of all subjects for Experiment 1. The pooled data gave the same pattern as the single-subject data: decreasing indices with increasing interval sizes and better estimation of same interval sizes for intervals in greater distance than for nearer intervals. An analysis of variance (two-way-ANOVA) revealed that the size of the fixed interval ($p < .05$) and the virtual distance of the fixed interval from the observer ($p < .05$) had a significant influence on the perceived size of the fixed virtual ground interval. Same was true for the interaction of both parameters ($p < .05$).

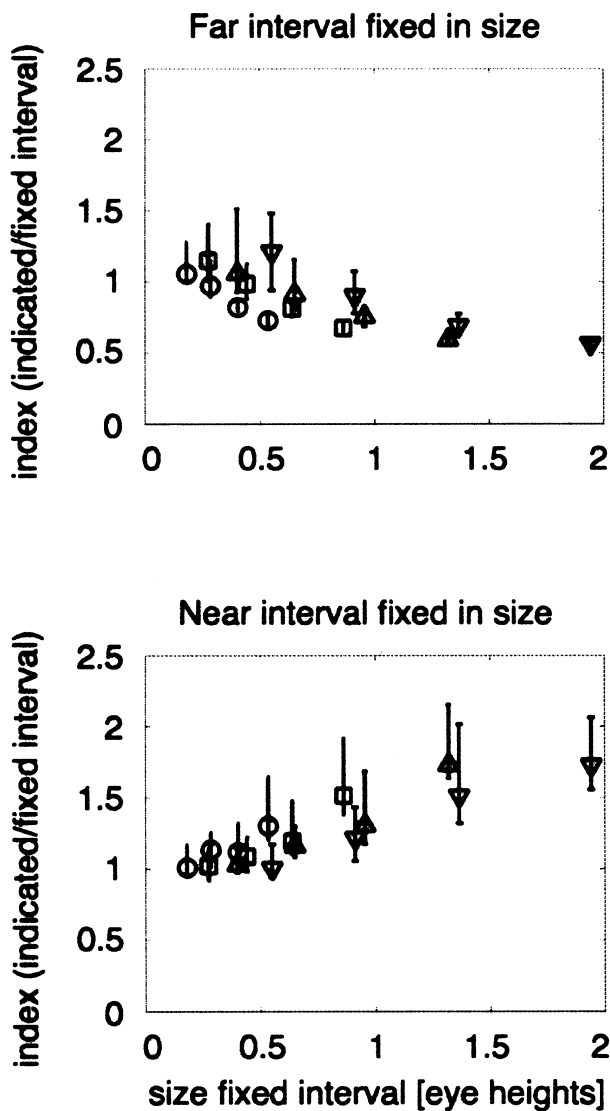


Figure 5. Distance ratios of the indicated and the fixed interval sizes (medians) are plotted as a function of the size of the fixed interval. Symbols represent the distances to the middle line (circle: 1.54 eye heights, square: 1.92 eye heights, triangle up: 2.31 eye heights, and triangle down: 2.69 eye heights). Error bars indicate the 30th and 70th percentile.

Figure 4 shows the single subject results from Experiment 2. In this experiment, only subject *ps* showed a small offset (distance ratio smaller than 1 for small interval sizes). Subject *ms* indicated farther intervals with about 5 times the size of the reference interval in two conditions. This is a strong bias towards equalizing the interval sizes in screen coordinates. For all subjects, the indices increased with increasing size of the fixed near interval. Thus, subjects adjusted larger far intervals than the simulated size of the near interval and, therefore, overestimated the size of the near interval. A two-way-ANOVA showed that the size of the fixed interval, the distance between the observer's virtual

position and the middle line, and the interaction of both had significant influences on the interval judgment (always $p < .05$). The variation of the distance to the middle line showed the same influence on interval size estimation as in Experiment 1. With increasing distance to the second interval, the size estimation for same interval sizes became more accurate. Again, we calculated a Mann-Whitney-*U*-test to test whether the subjects equalized the interval sizes in screen coordinates. The test showed that the data were significantly different from hypothetical data of equal interval sizes in screen coordinates ($p < .01$).

Figure 5 (lower panel) shows the pooled data of all subjects obtained with Experiment 2. For the pooled data the index also increased with increasing interval size and decreasing distance to the far interval. The influence of the fixed interval size, the distance to the far interval, and the interaction of both were significant (two-way-ANOVA, $p < .05$). Experiment 2 also showed that the interval sizes were increasingly underestimated with increasing position in depth in the virtual environment.

In summary, in both experiments, interval sizes were increasingly underestimated with increasing size of the interval and with increasing distance of the interval from the observer. Thus, subjects perceived the visual space compressed with increasing distance.

Our experimental procedure measures relative exocentric distance. In order to relate these measurements to the flow-based distance estimation in (Frenz & Lappe, 2005), one would want to quantify the amount of space-interval compression as a function of the egocentric distance of the interval from the observer. There is indication that egocentric distance perception can be derived from piecing together subsequent relative exocentric distance estimates. Wu et al. (2004) and He, Wu, Ooi, Yarbrough, and Wu (2004) showed that the visual system uses patches of exocentric distances to form a global representation of a ground plane and, therefore, also egocentric distances. To quantify the compression rate as a function of distance, we therefore attempted to plot subjective distances as a function of the physical distances and fit the data with depth functions described in literature. As subjects were required to position the adjustable line on the ground plane so that the variable interval had the same size as the fixed interval, we reasoned that the position of the adjusted line is subjectively at the correct position on the ground plane after the adjustment. Therefore, it corresponds to the correct position of accurate interval adjustment. In Experiment 1, the subjects adjusted the near interval (moved the first horizontal line) in 16 test conditions. In Experiment 2, the subjects adjusted the far interval (the third line moved). Altogether, in 32 conditions, either the first or the third line was adjusted. We used the physical locations of the adjustable line in the exocentric matching task and plotted these locations as egocentric distances to the observer. The egocentric distance between the adjustable line and the observer's virtual position is the

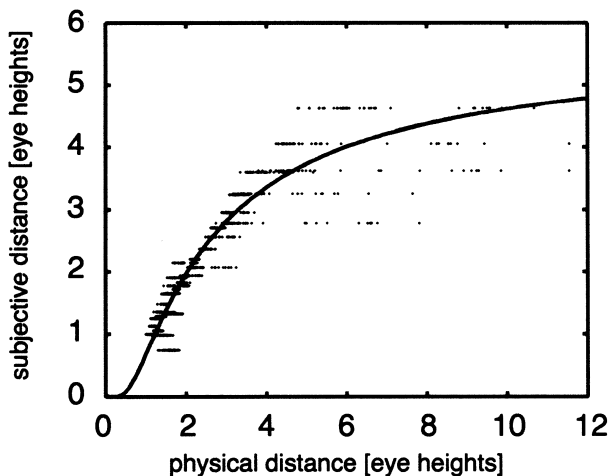


Figure 6. Subjective distance in virtual world coordinates as a function of the true distance. Black dots indicate the raw data of four subjects from both experiments in all conditions. The solid line is the fitted Luneburg function $f(r) = ae^{-\delta r}$ (see text for description).

physical distance in Figure 6. Different researchers have proposed different depth functions to describe this type of data (see for example Hecht, van Doorn, & Koenderink, 1999). The Luneburg function (Luneburg, 1947) is of the form $f(r) = ae^{-\delta r}$. The Gilinsky function (Gilinsky, 1951) is of the form $f(r) = ar/a+r$, where r is the distance to the observer and a the subject-dependent distance of straight horizon. A third possibility is a power function. For our data, the fit with the Luneburg function produced the least residuals and therefore described the data best. The fit revealed an a of 5.53 and a δ of 14.85. Figure 6 plotted the estimated subjective distances for every subject as a function of the assumed physical distances (black dots).

Discussion

We investigated the way that human subjects perceive spatial distances in virtual environments. First, we asked whether the perception of depth intervals can explain the observed underestimation of visually simulated self-motions that we found in earlier work (Frenz & Lappe, 2005). Second, we wanted to know whether the perception of spatial distances in virtual environments is comparable to that found in natural scenes.

We followed the experimental approach of Beusmans (1998). Observers had to match the size of a given depth interval by adjusting the size of a second depth interval, nearer (Experiment 1) or further (Experiment 2) to the observer's position. We used three horizontal lines instead of single markers. Two lines were fixed in distance to the observer. The observers could move a third line along the line of sight. Our results, like the results of Beusmans, show

that the perceived depth extent of ground plane intervals becomes smaller as the interval is located further away from the observer. This suggests that perceived distance is compressed. Loomis, Da Silva, Fujita, and Fukushima (1992) used a similar task to survey visually perceived space in natural scenes. They instructed their observers to indicate the perceived size of a frontal interval (1, 1.5, or 2 m or 0.63, 0.94, and 1.25 eye heights in size) in terms of a depth interval. The egocentric distances between the frontal interval and the observer's position ranged from 4 to 12 m (2.5 to 7.5 eye heights). Observers adjusted larger depth intervals than the reference frontal intervals depending on the egocentric distances and interval sizes. The error in size perception varied between 25% underestimation for egocentric distances of 4 m (2.5 eye heights) and interval sizes of 2 m (1.25 eye heights) to nearly 50% underestimation for 1-m intervals in a distance of 12 m (7.5 eye heights). This is also in line with the findings of Wagner (1985) who reported size underestimation of depth intervals about 50% compared to frontal intervals.

Norman et al. (1996) compared the perception of distances in natural scenes and computer displays. In virtual scenes, they presented two line segments sequentially. Participants had to indicate which of the line segments was longer. The position of the line segments in depth was varied using stereoscopic display. In addition, the lines were virtually rotated around a visually marked axis to provide further length and distance information. In the natural scene condition, the length of intervals formed by two red LEDs mounted on the top of a table had to be judged. The LEDs either formed intervals in the frontoparallel plane, intervals in depth, or in an oblique orientation in depth. The observers had to indicate the perceived length of the interval by adjusting the length of a line in the frontoparallel plane on a computer display. Although the authors only tested in near visual space less than 2 m (1.25 eye heights), they found increasing compression of the perceived size in depth with increasing viewing distance. This was similar in natural and computer-generated scenes. With the largest viewing distances, Norman et al. found distance underestimation up to 40%. This is comparable to our results.

We used such an exocentric distance indication task to make the results easy to compare with our previous studies (Frenz & Lappe, 2005; Lappe et al., 2005). In these studies, we asked observers to indicate the perceived travel distances of an ego-motion sequence in a static virtual scene in terms of an exocentric interval. We first simulated self-motions and afterwards asked the participants to indicate the perceived reference distance with an interval on the ground. This interval always started 1.54 eye heights in front of the observer. We found strong linearity between simulated distances and indicated distances, but distances were consistently underestimated.

On the basis of the results of the present study, we believe that this underestimation is unlikely to be the result

of general compression of perceived space for two reasons. First, as Figure 6 indicates, the perceived interval size is neither linearly correlated with the sizes of the reference intervals nor with the distance of the interval from the observer. Rather, the indicated distances are best described by a Luneburg function of the form $f(r) = ae^{-\delta/r}$. In experiments investigating distance perception from optic flow in virtual environments, the simulated and indicated distances were linearly correlated (Frenz & Lappe, 2005; Redlick et al., 2001; Riecke et al., 2002; Sun, Campos, & Chan, 2004). Therefore, the compression of visual space cannot explain the underestimation of travel distances in virtual environments. Second, within the distance range that we explored in the flow-based experiments (0.5-4 eye heights), static distance judgments are rather accurate (see Figure 6 and compare also Foley et al., 2004; Loomis & Knapp, 2003). In contrast, in the flow-based experiments, distances were substantially underestimated.

Our implementation of the exocentric distance indication task of Beusmans (1998) also allows us to compare results from natural and virtual scenes. We found errors in distance judgment up to 46% in Experiment 1 and 42% in Experiment 2 for an interval size of 1.94 eye heights in a distance of 2.69 eye heights to the observer. In the experiments of Beusmans, observers stood in an open field and had to match the size of a fixed ground interval in a distance up to 20 m (approximately 12.5 eye heights, assumed eye height 1.6 m) with a second ground interval, nearer to the observer's position. The arrangement of the two ground intervals in Beusmans' experiments was the same as in the present experiments except for the markers, which formed the intervals (orange rectangles—marker T1 and T2—or triangle—T3—in Beusmans' experiments). Beusmans also found increasing interval underestimation with increasing interval size and distance in depth: the error for interval sizes of 12 m (7.5 eye heights) in a distance of 20 m (12.5 eye heights) to the observer's position was about 35%. Thus, the compression of spatial distances is more pronounced in virtual than in real scenes. A stronger compression in virtual than in real scenes has been reported by a number of studies using different methods (Knapp & Loomis, 2004; Loomis & Knapp, 2003; Thompson et al., 2004). An explanation for the discrepancies in the amount of underestimation between Beusmans and our results may be the different depth cues in the scene. Beusmans allowed small head or body movements, which provide depth cues in terms of motion parallax. Motion parallax was not available for our virtual scenes. Moreover, the participants viewed the natural scenes binocularly and, therefore, had access to depth information from binocular disparity and vergence. Stereoscopic depth information was not available in our experiments. The type of error (distance underestimation depending on the egocentric distance to the observer and the size of the distance which has to be judged), however, was identical in natural and virtual scenes.

References

- Berthoz, A., Israel, I., Georges-Francois, P., Grasso, R., & Tszuku, T. (1995). Spatial memory of body linear displacement: What is being stored? *Science*, *269*, 95-98.
- Bertin, R.J., Israël, I., & Lappe, M. (2000). Perception of two-dimensional, simulated ego-motion from optic flow. *Vision Research*, *40*, 2951-2971.
- Beusmans, J. M. (1998). Optic flow and the metric of the visual ground plane. *Vision Research*, *38*, 1153-1170.
- Bremmer, F., & Lappe, M. (1999). The use of optic flow for distance discrimination and reproduction during visually simulated self motion. *Experimental Brain Research*, *127*, 33-42.
- Bronstein, A., & Buckwell, D. (1997). Automatic control of postural sway by visual motion parallax. *Experimental Brain Research*, *113*, 243-248.
- Cuijpers, R.H., Kappers, A.M.L., & Koenderink, J.J. (2000). Large systematic deviations in visual parallelism. *Perception*, *29*, 1467-1482.
- Cuijpers, R.H., Kappers, A.M.L., & Koenderink, J.J. (2002). Visual perception of collinearity. *Perception and Psychophysics*, *64*, 392-404.
- Foley, J.M. (1980). Binocular distance perception. *Psychological Review*, *87*, 411-434.
- Foley, J.M., Ribeiro-Filho, N.P., & Da Silva, J.A. (2004). Visual perception of extent and the geometry of visual space. *Vision Research*, *44*, 147-156.
- Frenz, H., Bremmer, F., & Lappe, M. (2003). Discrimination of travel distances from 'situated' optic flow. *Vision Research*, *43*, 2173-2183.
- Frenz, H., & Lappe, M. (2005). Absolute travel distances from optic flow. *Vision Research*, *45*, 1679-1692.
- Gilinsky, A.S. (1951). Perceived size and distance in visual space. *Psychological Review*, *58*, 460-482.
- He, Z.J., Wu, B., Ooi, T.L., Yarbrough, G., & Wu, J. (2004). Judging egocentric distance on the ground: Occlusion and surface integration. *Perception*, *33*, 789-806.
- Hecht, H., & Savelsbergh, G.J.P. (2004). *Advances in psychology 135 - Time-to-contact*. Amsterdam, Boston, Heidelberg: Elsevier.
- Hecht, H., van Doorn, A., & Koenderink, J.J. (1999). Compression of visual space in natural scenes and their photographic counterparts. *Perception and Psychophysics*, *61*, 1269-1286.
- Indow, T. (1991). A critical review of Luneburg's model with regard to global structure of visual space. *Psychological Review*, *98*, 430-53.
- Kearns, M.J., Warren, W.H., Duchon, A.P., & Tarr, M.J. (2002). Path integration from optic flow and body senses in a homing task. *Perception*, *31*, 349-374.
- 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*, *13*, 572-577.
- Koenderink, J.J., van Doorn, A., Kappers, A.M., & Todd, J.T. (2002). Pappus in optical space. *Perception and Psychophysics*, *64*, 380-391.

- Lappe, M., Bremmer, F., & van den Berg, A.V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, 3, 329-336.
- Lappe, M., Frenz, H., Bührmann, T., & Kolesnik, M. (2005). Virtual odometry from visual flow. *Proceedings of SPIE*, 5666, 493-502.
- Lee, D.N. (1980). The optic flow field: The foundation of vision. *Philosophical Transactions of the Royal Society of London B*, 290, 169-179.
- Loomis, J.M., Da Silva, J.A., Fujita, N., & Fukusima, S.S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception Performance*, 18, 906-21.
- Loomis, J.M., Klatzky, R.L., Golledge, R.G., Cicinelli, J.G., Pellegrino, J.W., & Fry, P.A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General*, 122, 73-91.
- Loomis, J.M., & Knapp, J.M. (2003). Visual perception of egocentric distance in real and virtual environments. In L.J. Hettinger & M.W. Haas (Eds.), *Virtual and adaptive environments* (pp. 21-46). Mahwah NJ: Erlbaum.
- Loomis, J.M., & Philibeck, J.W. (1999). Is the anisotropy of perceived 3-D shape invariant across scale? *Perception and Psychophysics*, 61, 397-402.
- Loomis, J.M., Philibeck, J.W., & Zahorik, P. (2002). Dissociation between location and shape in visual space. *Journal of Experimental Psychology: Human Perception Performance*, 28, 1202-1212.
- Luneburg, R.K. (1947). *Mathematical analysis of binocular vision*. Princeton, NJ: Princeton University Press.
- Norman, J.F., Todd, J.T., Perotti, V.J., & Tittle, J.S. (1996). The visual perception of three-dimensional length. *Journal of Experimental Psychology: Human Perception Performance*, 22, 173-186.
- Peruch, P., May, M., & Wartenberg, F. (1997). Homing in virtual environments: Effects of field of view and path layout. *Perception*, 26, 301-311.
- Prokop, T., Schubert, M., & Berger, W. (1997). Visual influence on human locomotion: Modulation to changes in optic flow. *Experimental Brain Research*, 114, 63-70.
- Redlick, F.P., Jenkin, M., & Harris, L.R. (2001). Humans can use optic flow to estimate distance of travel. *Vision Research*, 41, 213-219.
- Riecke, B.E., van Veen, H.A.H.C., & Bühlhoff, H.H. (2002). Visual homing is possible without landmarks: A path integration study in virtual reality. *Presence*, 11, 443-473.
- Sun, H.-J., Campos, J.L., & Chan, G.S. (2004). Multisensory integration in the estimation of relative path length. *Experimental Brain Research*, 154, 246-254.
- Sun, H.-J., Campos, J.L., Young, M., Chan, G.S., & Ellard, C.G. (2004). The contributions of static visual cues, nonvisual cues, and optic flow in distance estimation. *Perception*, 33, 49-65.
- 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*, 13, 560-571.
- Wagner, M. (1985). The metric of visual space. *Perception and Psychophysics*, 38, 483-495.
- Warren, W.H. (1998). Visually controlled locomotion: 40 years later. *Ecological Psychology*, 10, 177-219.
- Witmer, B.G., & Kline, P.B. (1998). Judging perceived and traversed distance in virtual environments. *Presence*, 7, 144-167.
- Wu, B., Ooi, T.L., & He, Z.J. (2004). Perceiving distance accurately by a directional process of integrating ground information. *Nature*, 428, 73-77.

Received December 15, 2006
 Review received February 20, 2006
 Accepted July 10, 2006