Estimation of Travel Distance from Visual Motion in Virtual Environments

HARALD FRENZ and MARKUS LAPPE
Westfälische-Wilhelms University, and Ruhr-University
MARINA KOLESNIK
Fraunhofer Institute for Media Communication IMK
and
THOMAS BÜHRMANN
Ruhr-University

Distance estimation of visually simulated self-motion is difficult, because one has to know or make assumptions about scene layout to judge ego speed. Discrimination of the travel distances of two sequentially simulated self-motions in the same scene can be performed quite accurately (Bremmer and Lappe 1999; Frenz et al., 2003). However, the indication of the perceived distance of a single movement in terms of a spatial interval results in a depth scaling error: Intervals are correlated with the true travel distance, but underestimate travel distance by about 25% (Frenz and Lappe, 2005). Here we investigated whether the inclusion of further depth cues (disparity/motion parallax/figural cues) in the virtual environment allows more veridical interval adjustment. Experiments were conducted on a large single projection screen and in a fully immersive computer-animated virtual environment (CAVE). Forward movements in simple virtual environments were simulated with distances between 1.5 and 13 m with varying speeds. Subjects indicated the perceived distance of each movement in terms of a depth interval on the virtual ground plane. We found good correlation between simulated and indicated distances, indicative of an internal representation of the perceived distance. The slopes of the fitted regression lines revealed an underestimation of distance by about 25% under all conditions. We conclude that estimation of travel distance from optic flow is subject to scaling when compared to static intervals in the environment, irrespective of additional depth cues.

Categories and Subject Descriptors: J.4 [**Psychology**]: Social and Behavioral Sciences; I.4.8 [**Image Processing and Computer Vision**]: Computing Methodologies—*Scene analysis*

General Terms: Human factors, Experimentation

Additional Key Words and Phrases: Optic flow, distance estimation, stereo, depth, virtual reality

ACM Reference Format:

Frenz, H., Lappe, M., Kolesnik, M., and Bührmann, T. 2007. Estimation of travel distance from visual motion in virtual environments. ACM Trans. Appl. Percpt. 4, 1, Article 3 (January 2007), 18 pages. DOI = 10.1145//1227134./1227137 http://doi.acm.org/10.1145./1227134./1227137

Authors' addresses: Harald Frenz and Markus Lappe, Psychological Institute II, Westfälische-Wilhelms University, Münster, Germany and Department of Zoology and Neurobiology, Ruhr-University, Bochum, Germany; email: mlappe@psy.uni-muenster.de. Marina Kolesnik, Fraunhofer Institute for Media Communications IMK, St. Augustin, Germany. Thomas Bührmann, Psychological Institute II, Westfälische-Wilhelms University, Münster, Germany.

1. INTRODUCTION

A moving observer receives a wealth of sensory information about his self-motion. Besides proprioceptive and vestibular signals, visual signals provide a powerful information source to control one's own movement. This visual information can be either positional (e.g., fixed landmarks, which can be used as reference objects in the scene) or based on self-generated motion signals (optic flow [Gibson 1950]) during the movement. Past research has shown that the optic flow can be useful for different aspects of motion control such as the time-to-contact [Lee 1980], the estimation of heading [Warren et al. 1988; Warren and Hannon 1990], the control of upright stance [Lee 1980; Bronstein and Buckwell 1997], and the control of walking speed [Prokop et al. 1997].

We are interested in the use of the optic flow for the estimation of travel distance. In order to judge travel distance on the basis of optic flow, the observer has to estimate the velocity and the duration of the self-motion. Judging the absolute speed of self-motion from the optic flow speed alone is not possible, because the depth distribution of objects in the scene must also be known. Thus, optic flow fields are ambiguous with respect to absolute travel distance without known scale of the scene, as identical optical flow fields can be achieved by covarying the observer speed and the distances to objects in the scene. Yet, human subjects can discriminate the travel distance of two sequential self-motion simulations in the same scene with the assumptions that the environment is identical in the two simulations [Bremmer and Lappe 1999; Frenz et al. 2003]. When subjects are asked to estimate the time of arrival at a previously defined position from the motion field [Redlick et al. 2001], or use information in motion displays for home finding such as in triangle completion tasks [Peruch et al. 1997; Witmer and Kline 1998; Kearns et al. 2002; Riecke et al. 2002; Sun et al. 2004], performance is, however, often inaccurate.

Since optic flow alone is ambiguous with respect to scale, observers in a travel distance task need information or have to make assumptions about scene layout. In previous travel distance-discrimination experiments, we asked observers to judge which one of two sequentially simulated self-motions covered a greater distance. We found accurate distance discrimination when the virtual scene did not change between the two motion sequences [Bremmer and Lappe 1999; Frenz et al. 2003]. In this case, subjects could assume that the depth structure of the scene was identical between the two motions. When we changed the scene layout unnoticed by the subject (altered viewing range, viewing angle in the scene or eye height above the ground plane), predictable errors occurred as subjects attributed the whole change in the flow field to a changed ego velocity and, therefore, perceived larger travel distances. When the subjects noticed the altered depth structure of the scene, they could extract this influence on the flow field and calculate the simulated ego velocity [Frenz et al. 2003]. We concluded that subjects use a three-dimensional (3D) percept of their self-motion to calculate the travel distance, rather than pure two-dimension (2D) image motion.

This 3D percept might serve as the basis for a true distance estimate in which subjects would convert the percept of their self-motion to a spatial interval within the scene. This conversion would also be independent from absolute scene scale as long as it is confined within the space of the scene. In subsequent experiments, we therefore investigated whether observers can use the 3D percept of the self-motion simulation to build up an internal representation of the traveled distance [Frenz and Lappe 2005] and report this in terms of a distance within the scene. We presented a visually simulated self-motion over a virtual ground plane and instructed subjects to later indicate the perceived travel distance in a static view of the scene. In one experiment, the subjects indicated the perceived travel distance with a virtual interval on the ground plane. We observed a strong linear correlation between simulated and indicated distances, which means that subjects were very precise in their report.

This linear correlation is consistent with other studies that investigated visual motion-based distance estimation during walking [Loomis et al. 1993; Witmer and Kline 1998; Kearns et al. 2002], riding a bike

[Sun et al. 2004], steering on a mobile robot [Berthoz et al. 1995], or navigating in a virtual environment [Peruch et al. 1997; Witmer and Kline 1998; Riecke et al. 2002]. However, the reported distances in our experiments undershot the simulated distances, on average, by about 25% [Frenz and Lappe 2005]. Similar undershoots have been observed by Sun et al. [2004] and Witmer and Kline [1998]. For some stimulus parameters, overestimations of travel distance also occurred [Redlick et al. 2001; Frenz and Lappe 2005]. These findings suggest that optic flow can be used to obtain an estimate of travel distance, but this estimate is subject to a scale factor when compared to static intervals in the environment.

We tested a number of possible reasons for why the distances were so strongly underestimated. We measured our observers ability to report distances in static scenes. Often distances in static scenes are perceived as compressed [Foley 1980; Wagner 1985; Indow 1991; Beusmans 1998; Cuijpers et al. 2000; 2002; Foley et al. 2004]. To investigate whether such compression of static distances could be the reason for the observed undershot in travel distance report, we presented two virtual depth intervals with different sizes and distances to the observer in a static view of the scene. The subjects had to indicate the size of one fixed interval by adjusting the other interval. The results showed a nonlinear correlation between given and adjusted size in depth. Within the range of the distances used in our travel judgment experiments, however, static distance judgments were rather accurate. Therefore, we concluded that compression effects in the perception of visual space, in general, are not the reason for underestimation of travel distance.

A second possible source for the observed error could be the particular manner of distance report. We, therefore, tested several different reporting procedures [Frenz and Lappe 2005]. First, since previous experiments showed accurate distance indication when subjects had to reproduce the distance of a simulated self-motion with a second, self controlled visual motion [Bremmer and Lappe 1999], we asked our observers to first reproduce the seen motion display with a force transducer to control the velocity of the movement simulation. Subjects then had to report the travel distance with the interval adjustment procedure. We found accurate distance reproduction, but, at the same time, an underestimation of reported distance in the static scene interval adjustment. The underestimation also occurred when subjects had to report the travel distance verbally in terms of eye heights and when they were instructed to reproduce the perceived travel distance by blindfolded walking. In the latter case, we first presented a motion sequence while the subjects stood on a wooden catwalk wearing a small high-resolution head-mount display (HMD) (Sony Glastron PLM-S700E). After the motion sequence, the HMD turned black and subjects had to walk the same distance in total darkness. As in the previous experiment we observed linear correlation between simulated and indicated distances, but distances were again underestimated by about 30%.

A third possible reason why subjects may have underestimated the travel distances of visually simulated self-motion might lie in insufficient depth information about scene layout. Recall that knowledge of scene layout is needed to disambiguate the scale problem of optic flow. In our experimental conditions, knowledge of absolute depth layout was not needed, because visual motion stimulation and report happened within the same virtual environment. Thus, the task is independent of absolute scene layout, because the mere assumption of scene identity between motion simulation and report suffices. However, it might still be that the visual system relies on depth cues in the scene directly for the build-up of the travel distance representation or for the control of interval adjustment in the reporting phase. Therefore, we wanted to investigate the influence of additional depth cues on travel distance estimation from visual motion in virtual scenes. In the following, we report results from four experiments. In experiment 1, we add vertical poles to the ground plane to increase motion parallax information and pictorial depth cues (c.f. [Kearns et al. 2002; Li and Warren 2004]). In experiment 2, we increase the viewing range of the scene to better specify the location of the horizon [c.f. Messing and Durgin 2005].

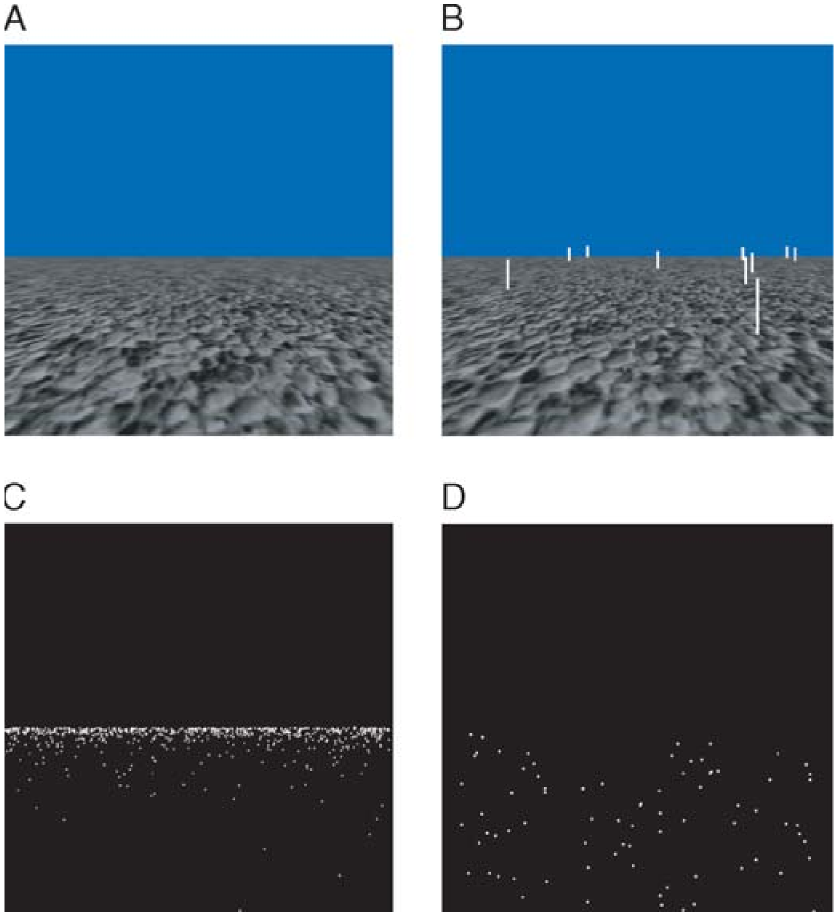

Fig. 1. Screenshots of the virtual ground planes. (A) Textured ground plane; (B) textured ground plane with poles; (C) dot plane 1; (D) dot plane 2.

In experiment 3, we enhance the scene presentation with binocular disparity. In experiment 4, we use a fully immersive computer-animated virtual environment (CAVE) with observer viewpoint tracking, stereoscopic presentation, and realistic motion parallax from the observers head movements. In all experiments, we find similar underestimation of travel distances. We conclude that the inaccuracy of travel distance estimation from optic flow is not because of a lack of depth information in the scene.

2. GENERAL METHODS

2.1 Virtual Scenes

We simulated movement on top of virtual ground planes with varying depth information. Screenshots of the different environments are shown in Figure 1.

2.1.1 *Textured Ground Plane.* We used a 8 m × 8 m texture pattern (Iris Performer type "gravel") and mapped it on a 400 m × 400 m virtual ground plane. Blue sky (RGB code: 0.1, 0.3, 0.7) covered the parts of the scene above the ground. The depth range of the viewing frustum for this and the other virtual scenes was restricted to 30 m in experiments 1, 3, and 4, and extended to 100 m in experiment 2.

Before each trial, we randomly varied the starting position of the movement sequence up to 10 m to either side of the ground plane origin (midpoint of the plane) to avoid recognition of texture elements in successive trials. The textured ground plane provided ample static depth cues, contained in gradients of texture density, and size of texture elements toward the horizon. It also provided dynamic depth cues in the motion simulation, most notably motion parallax and the change of size of texture elements as they approached the observer. In addition, the trajectory of ground plane elements gave information about the depth structure and simulated travel distance. The vertical refresh rate of the rendering was 36 Hz. The mean luminance was 3.1 cd/m^2.

2.1.2 *Textured Ground Plane with Poles.* This scene consisted of the textured ground plane, as described above, but with ten additional vertical poles placed on the ground. At the beginning of a trial, we calculated for each of the ten poles a random position in the scene, ranging up to 29 m to either side of the starting position of the observer and up to 29 m in front of the observer's virtual position. The height of the poles was 1.5 m and the thickness was 2 pixels, independent of the distance to the observer. This stimulus provides all the depth cues of the ground plane stimulus plus increased motion parallax and pictorial depth cues from the poles.

2.1.3 *Dot Plane 1.* Dot plane 1 consisted of 3300 white light points, which we positioned on a grating every 6 m to both sides of the observer's virtual position within 30 m. In depth, we positioned the light points every 2 m within 52 m. To achieve a balanced random distribution, we afterward shifted the position of the light points randomly up to 5 m to either side. The lifetime of each point was limited. With a probability of 10% in a single frame, each light point could vanish and reappear on a random position within the scene. The vertical refresh rate of the rendering was 72 Hz. Thus, the mean lifetime of the dots was 139 ms. On average, our subjects saw 970 light points on the screen in each frame. Size and luminance were constant during movement simulation. Dot plane 1 provided dynamic depth cues by motion parallax and static depth information in terms of texture density toward the horizon. Mean luminance was 2.0 cd/m^2.

2.1.4 *Dot Plane 2.* Dot plane 2 lacked all static depth information. We randomly distributed 150 light points on the lower half of the screen on a black background. Again, dot size and luminance remained constant regardless of the distance to the observer. During movement simulation the light points moved as if they lay on a ground plane, i.e., they obeyed the pattern of motion parallax with dots more distant to the observer visually moving slower than dots nearer to the observer's virtual position. The limitation of the dot's lifetime was the same as for dot plane 1. Because this virtual scene provided no static depth information, motion simulation continued during distance indication. This resulted in the impression of movement over a static ground plane and allowed the subject to indicate the perceived travel distance in terms of an interval on the ground plane. The virtual indicator lines (see below) were superimposed on the image of the moving dots and thus appeared to move over the plane along with the subject, always remaining in a static configuration with respect to the subject. The reference line remained fixed on the screen with the dots moving behind it. The movable line could be moved by the subject to adjust interval size. This motion simulation during distance indication had no effect on distance judgment (see [Frenz and Lappe 2005]). The mean luminance was 0.6 cd/m^2.

2.1.5 *The Virtual Indicator Lines.* The indicator lines spread over 40 m to both sides of the observer's virtual position. The thickness of the indicator lines was 2 pixels and remained constant regardless of the distance to the observer's virtual position. The subject controlled the position of the adjustable line by moving a computer mouse. We used the vertical coordinates of the invisible mouse pointer position on the screen (ranging from 0 to 1024 pixels) and calculated the corresponding virtual position on ground in the simulated environment (ranging from 0 to 30 m). Therefore, changing the position of the mouse pointer by one pixel altered the position of the line by 2.93 cm on the virtual ground plane (3000 cm/ 1024 pixels). The physical distance between the two lines in the virtual scene was calculated after the subject indicated the decision with a button press.

2.2 Procedure

Each trial started with a visual simulation of a linear self-motion with one of four different ego velocities and four possible durations (velocities: 1, 1.5, 2.5, or 3 m/s, durations: 1.5, 2, 2.5, or 3 s). Therefore, the simulated travel distances varied between 1.5 and 9 m. Four travel distances (3, 3.75, 4.5, and 7.5 m) were simulated with different combinations of the translation velocities and simulation duration to allow a comparison of travel distance estimates for identical distances with variable speed and duration. The velocity profile was rectangular so that the self-motion simulation started and ended instantaneously. The static virtual scene was presented for 300 ms before the motion simulation started in order to ensure that the subjects perceive the whole simulation. The simulated gaze direction was always in the direction of the self-motion. We presented the resulting 16 conditions in a pseudorandomized order with 10 repetitions each. After the self-motion simulation, two horizontal white lines appeared on the virtual ground plane. One line (reference line) was fixed on the ground plane in a distance of 4 m to the subject's virtual position. The second line appeared 3 m in front of the subject's position in the scene. The subject could adjust the position of this line in the scene. The task in all experiments was to indicate the travel distance of the visually simulated self-motion in terms of a virtual ground interval with the adjustable line more distant than the reference line. In pilot studies of earlier work [Frenz and Lappe 2005], we also tested an egocentric distance indication task, in which the distance between the subjects virtual position in the scene and one adjustable indicator line was adjusted. This procedure gave a constant error in distance indication because the subjects used the lower part of the projection screen as reference point. To eliminate this constant error, we decided to use the exocentric matching task with two indicator lines.

2.3 Data Analysis

We plotted the indicated travel distances as a function of the simulated distances and fitted linear regressions to the data points. We used three parameters to analyze the data. The first is the correlation coefficient ρ between the indicated and simulated distances. High correlation coefficients indicate that subjects can precisely use the movement sequence to build up an internal representation of the traveled distance. The second is the slope of the fitted regression, indicating the accuracy of the distance matching. Without any error in distance estimation and indication, the slopes of the fitted linear regressions would be 1 with an offset of 0. Slopes smaller than 1 indicate undershoot; slopes greater than 1 indicate overshoot of the travel distance of the self-motion simulation. We omitted the intercepts of the regression lines from further analysis as they represent constant errors and our main concern is the relationship between simulated and indicated distances. The third parameter is the difference in distance indication between identical virtual distances that were simulated with different combination of translation velocities and simulation duration. If the subjects based their judgment on the travel distance, identical simulated travel distances should be indicated with ground intervals of the same size independent of speed and duration of the simulation.

3. EXPERIMENT 1: ADDITIONAL MOTION PARALLAX AND PICTORIAL DEPTH CUES

In the first experiment, we tested whether additional motion parallax and pictorial depth cues can improve travel distance estimation from flow field simulations. Subjects had to indicate travel distances of a self-motion simulation over a textured virtual ground plane with poles fixed on the ground. We compare the results with data obtained in a previous study using the same textured ground planes without poles [Frenz and Lappe 2005].

3.1 Participants

Five subjects (1 female and 4 male, 24–30 years old) participated in this experiment; one author was also included. Three of these subjects also participated in previous experiments without poles on the textured ground plane [Frenz and Lappe 2005]. All subjects had normal or corrected to normal vision.

3.2 Procedure

The subjects sat 60 cm in front of a large back projection screen (120 × 120 cm, Dataframe, type CINEPLEX) with their eyes level with the virtual horizon. The resulting field of view was $90 \times 90^\circ$. The spatial resolution was 1280 × 1024 pixels. The stimulus was generated on a Silicon Graphics Indigo2 workstation and presented using a CRT video projector (Electrohome ECP 4100). Subjects viewed the stimulus binocularly and were instructed to avoid head movements to keep the distance to the screen constant. Rendering and presentation in this experiment was monoscopic. The time course of the experiment is described in Section 2.2.

3.3 Results

Figure 2 shows the sizes of the adjusted ground interval as a function of the simulated travel distance for the single subjects and the pooled data of all subjects. The fitted linear regressions are the solid lines in Figure 2; the dashed lines are hypothetical data of accurate distance indication. The correlation coefficient ρ varies between 0.66 and 0.88 among the subjects. For the pooled data of all subjects the correlation coefficient was 0.73. Therefore, the subjects possessed a precise internal representation of the simulated distance. Four distances were simulated with different combinations of ego velocity and duration (3, 3.75, 4.5, and 7.5 m; see Section 2.2). The subjects indicated same travel distances with same interval sizes irrespective of the combination of ego speed and simulation duration (compare black circles with corresponding grey circles in Figure 2). Therefore, the distance judgments are truly based on travel distance and not a particular combination of travel duration and self-motion speed. The fitted linear regressions described the data very accurately ($r^2 > 0.88$, $p < 0.05$ for all fitted regressions). Subject HF showed accurate distance matching (slope 0.97), whereas the other four subjects undershot the distances (slopes between 0.5 and 0.79). Fitted to all data of all subjects the slope is 0.74 (±0.08). Thus, subjects, on average, undershot the simulated distances of the movement sequence by 26%. This undershoot is comparable to the undershoot found without poles in the scene in the earlier study (slope of 0.79 using the same virtual environment without poles [Frenz and Lappe 2005]). Three of the subjects of experiment 1 also took part in that earlier study (HF, JL, and KG). For these subjects we also individually compared the slopes found here with the slopes found earlier in the same virtual environment without poles. For subject JL, the slope with poles was inside the confidence interval calculated in the condition without poles. The error in distance indication, therefore, did not change between the pole and without pole condition. Subject HF showed a steeper slope in the pole condition, i.e., a smaller error in distance estimation in the pole condition. Subject KG revealed a shallower slope, i.e., a larger error, in the pole condition. Therefore, the motion parallax and pictorial depth cues added by the poles did not result in an uniform increase of accuracy of travel distance reported on the single subject level.

4. EXPERIMENT 2: ADDITIONAL DEPTH OF THE FIELD OF VIEW

Here, we investigated the influence of increasing the maximum visible distance in the scene. We increased the depth range of the viewing frustum on the textured ground plane from 30 to 100 m. This

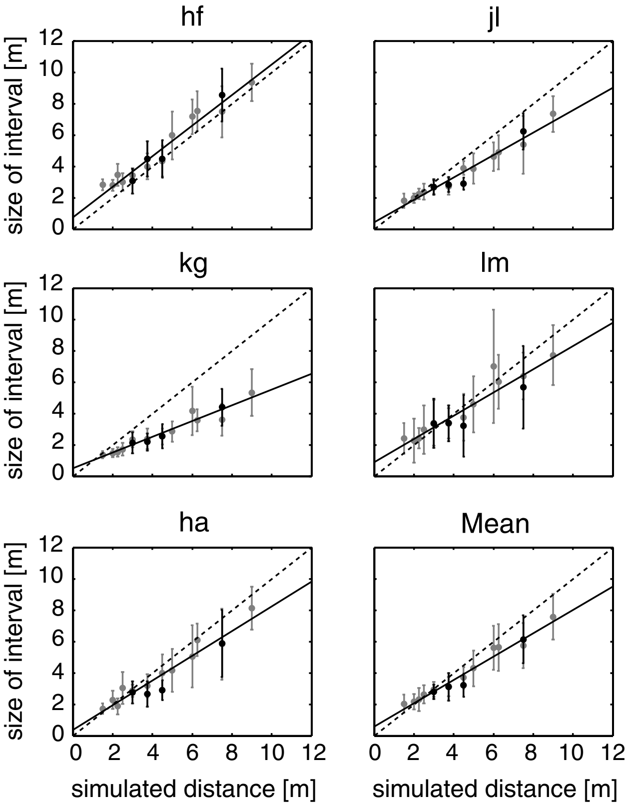

Fig. 2. Results of experiment 1. The indicated distances are plotted as a function of the simulated travel distances. The solid line marks the fitted linear regressions to the data. The dashed line gives the prediction for accurate distance matching. Black data points indicate travel distances simulated with lower ego velocity, but longer simulation duration than the corresponding grey data points. In the single subject subplots data points show the mean of all trials for each condition. In the subplot "Mean," each data plot is the mean over all subjects and corresponding trials. The error bars show the standard deviation.

gives more distant depth information and a better estimate of the height of the horizon. The angle of declination below the horizon is an important cue to distance perception in virtual scenes [Messing and Durgin 2005]. The time course of the experiments is the same as described in Section 2.2.

4.1 Participants

Six subjects (1 female, 5 male), including one author, participated in the experiment. Four of these subjects also participated in experiment 1.

4.2 Results

Figure 3 shows the single subjects results and the pooled data of all subjects. The fitted linear regression is illustrated by the solid lines; the dashed lines correspond to hypothetical data of accurate distance matching. All regressions are an appropriate description of the data ($r^2 > 0.75$, $p < 0.05$). The calculated correlation coefficient ρ varied between 0.55 and 0.80 among the subjects. For the pooled data of all subjects, we calculated a correlation coefficient of 0.66. This shows that the subjects can use their representation of the perceived travel distance to adjust an interval on the ground plane of equal size. In addition, subjects indicated same travel distances with same sized intervals on the ground plane

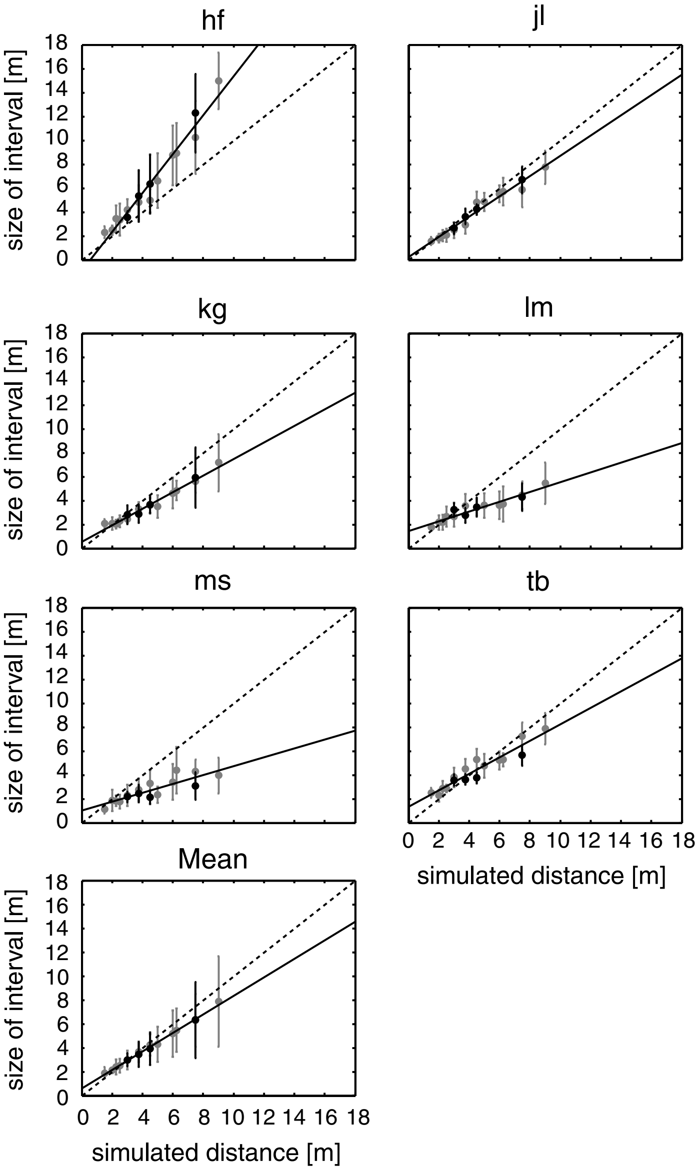

Fig. 3. Results of experiment 2. The indicated distances are plotted as a function of the simulated travel distances. Same convention as in Figure 2.

(black versus grey circles in Figure 3). As in previous experiments with the shorter viewing range up to 30 m, the reported distances undershot the simulated distances. The slopes of the fitted linear regressions varied between 0.37 and 1.67. One subject overshot and 5 subjects undershot the simulated travel distances. For the three subjects that also took part in the earlier study with a shorter viewing range (HF, JL, and KG) [Frenz and Lappe 2005], individual comparisons showed that for JL and KG were within the 95% confidence interval of their slopes in the earlier study, while for HF the error in distance estimation was larger in the present experiment with additional depth information than in the earlier experiment. The slope of the fitted regression of the pooled data was 0.77 (±0.03). This means, that the error in distance perception is 23% and, therefore, comparable to the error of 21% found with shorter viewing distances (slope of 0.79 [Frenz and Lappe 2005]).

5. EXPERIMENT 3: STEREOSCOPIC PRESENTATION

In this experiment, dot planes 1 and 2 were presented stereoscopically with the appropriate binocular disparity. The added depth information from the stereoscopic presentation of the stimuli may improve the perception of the self-motion. Palmisano [2002] reported that stereoscopic information about a simulated self-motion increased the perceived translation velocity. If, in our earlier experiments, the subjects misperceived the simulated translation velocity of the self-motion, they may have underestimated the traversed distance because of the underestimation of velocity. Moreover, van den Berg and Brenner [1994a; 1994b] demonstrated that the ability of human subjects to estimate the heading direction of a visually simulated self-motion became more tolerant to noise in the optic flow field when the authors added binocular disparity to the scene. Performance in noise-free conditions did not improve, however. Rushton et al. [1999] also found no benefit of disparity on steering. Grigo and Lappe [1998] described that human subjects used stereoscopic depth information as an additional source of information for the interpretation of the simulated flow field.

5.1 Methods

For the stereoscopic presentation, two camera viewpoints onto the scene were created with a virtual interocular distance of 6.4 cm. The resulting images from the two viewpoints were alternatingly presented on the projection screen each with a frame rate of 60 Hz. The disparities of the dots in the display ranged from -0.12 to 5.7°. Subjects wore liquid crystal shutter glasses (StereoGraphics; model CrystalEyes), which were synchronized with the stimulus presentation rate. The glasses for each eye were thus opened and shut with a rate of 60 Hz. This ensured that each eye got only images of its correct viewpoint on the scene. The shutter glasses reduced the subjects' field of view to $60^\circ \times 70^\circ$.

5.2 Participants

Six subjects (1 female, 5 male) participated in this experiment, including the first author. All subjects were already tested in experiment 2, four had also participated in experiment 1. All subjects reported an increased perceived depth of the scene in static viewing because of the stereoscopic presentation.

5.3 Results

Figures 4 and 5 illustrate the results. Each circle shows the mean over subjects; the error bars are the standard deviations. The dashed lines denote hypothetical data of accurate distance matching, while the solid lines are the fitted linear regressions. All regressions were an appropriate description of the data ($r^2 > 0.74$, $p < 0.05$).

Simulated and indicated travel distances were highly correlated. For dot plane 1, correlation coefficients ranged between 0.64 and 0.93 among subjects; for dot plane 2 correlation coefficients ranged between 0.55 and 0.89. In both virtual environments, when identical simulation distances were presented with different speeds, subjects mostly indicated the same intervals on the ground for all speeds used (compare black and grey symbols in Figures 4 and 5). The slopes of the regressions to the data of each subject for dot plane 1 varied between 0.23 and 1.1. One subject overshot the distances, one subject showed accurate performance, and four subjects undershot the simulated travel distances. The data of subject KG showed no improvement in distance estimation compared to the study without stereoscopic stimulus presentation [Frenz and Lappe 2005]. Her slope was within the 95% confidence interval of

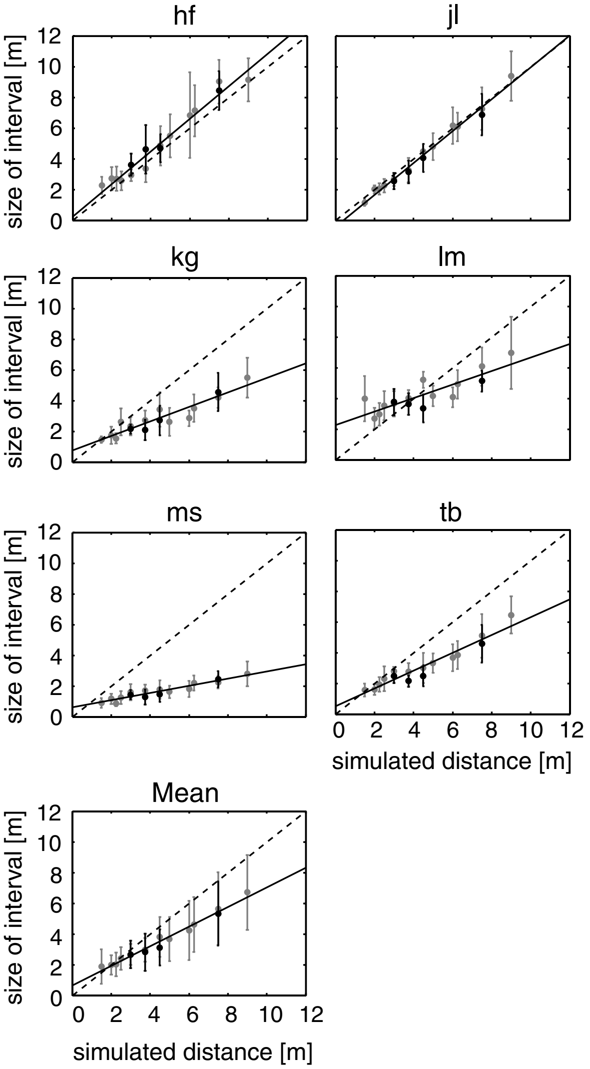

Fig. 4. Results of experiment 3 with motion simulation on dot plane 1. Same convention as in Figure 2.

her slope in the earlier study. Subjects HF and JL revealed smaller errors in distance estimation than previously. The pooled data of all subjects had a slope of 0.64. Therefore, on average, the subjects undershot the simulated travel distance by 36%. This underestimation is very similar to the underestimation observed using dot plane 1 without stereoscopic depth information (33% underestimation [Frenz and Lappe 2005]). With dot plane 2, slopes ranged between 0.34 and 1.1. One subject overshot the simulated distances, two subjects showed accurate distance matching, and three subjects undershot the distances of the simulated self-motion. Here, there was no improvement in distance estimation for subjects HF and JL over the nonstereoscopic presentation of the same environment in Frenz and Lappe [2005]. The results were within the 95% confidence interval from that study. For subject KG, the error in distance estimation increased with stereoscopic presentation of the virtual scene (comparison of slopes and 95% confidence intervals). The fitted linear regression to the data of all subjects had a slope of 0.79. Thus, the

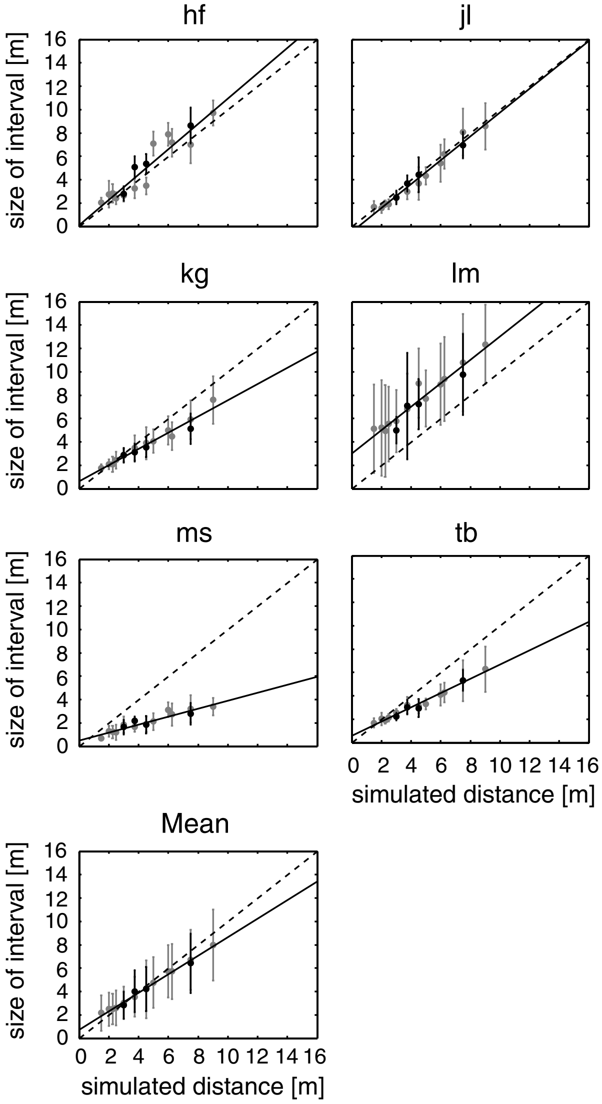

Fig. 5. Results of experiment 3 with motion simulation on dot plane 2. See Figure 2 and text for description.

subjects, on average, underestimated the travel distances by 21%. Without stereoscopic depth information, we found an undershot of the travel distance of 24% [Frenz and Lappe 2005]. We, therefore, conclude that binocular disparity did not increase the accuracy of travel distance estimates from optic flow.

6. EXPERIMENT 4: FULLY IMMERSIVE VIRTUAL ENVIRONMENT

6.1 Methods

6.1.1 *Stimuli.* This experiment was conducted in a CAVE, which is a 3 × 3 × 3 m large room with stimulus back-projection onto the front and the two side walls, as well as projection onto the floor via a mirror at the ceiling. Thus, the visual stimulation area extended more than 180° horizontally and included the floor space in front of the observer. Projections were realized with four three-color Electrohome projectors. Stimuli were rendered in real time by Silicon Graphics Hardware. Spatial resolution was 1280 by 1024 pixels for each projection wall. Frame rate was 120 Hz. Subjects wore shutter glasses (CrystalEyes, description see above) to generate the stereoscopic view on the virtual environment with 60 Hz per eye. Position and orientation of the shutter glasses were tracked and the view of the virtual environment was rendered accordingly. This ensured that the participant always had a correct view of the environment. The direction of motion was always constant with respect to the environment. During the experiment, the participant stood in the middle of the CAVE. Dot planes 1 and 2 were used. The self-motions simulated travel distances between 2.18 and 13.05 m with velocities of 1.45, 2.18, 3.63, and 4.35 m/s, and durations of 1.5, 2, 2.5, or 3 s. We simulated four travel distances (4.35, 5.44, 6.53, and 10.88 m) with two different combination of ego velocity and simulation duration.

6.1.2 *Interval Adjustment.* For the adjustment of the virtual ground interval in the reporting period, the subject used a Cubic Mouse. The Cubic Mouse consisted of a box with three perpendicular rods passing through the cube. Each rod represented one axis of the CAVE's coordinate system. Because the subjects had to adjust the interval on the ground only in depth, only one of the rods was used. The subjects controlled the virtual position of the adjustable line by pulling and pushing the rod. Before each indication of the traversed distance, the subjects had to pull the rod as far back as possible, whereupon the indicator lines appeared in the virtual environment. When the subject had adjusted the interval such that it appeared to match the travel distance of the reference self-motion, he pressed a button on the Cubic Mouse and the next trial started.

6.1.3 *Participants.* Six subjects (1 female, 5 male) participated in this experiment. All these subject already participated in experiment 3.

6.2 Results

The results are illustrated in Figures 6 and 7. For the single subject data, data points represent the means over ten trials, in the pooled data plot data points represent the collapsed data of all subjects. The solid lines show the fitted linear regressions; the dashed lines show hypothetical data of accurate distance matching. Correlation coefficient between simulated and indicated distance varied between 0.71 and 0.83 for dot plane 1 and between 0.67 and 0.91 for dot plane 2. Correlation coefficients for the pooled data were $\rho = 0.73$ for dot plane 1 and $\rho = 0.76$ for dot plane 2. Same travel distances were indicated with ground intervals of the same size regardless of the velocity and duration of the simulation (black and grey symbols in Figures 6 and 7). The slopes of the regression lines varied between 0.48 and 1 with dot plane 1 (see Figure 6). One subject gave accurate travel distances estimates, whereas four subjects underestimated the distances. For two of the subjects in this experiment, a comparison of their individual data with earlier nonstereoscopic presentation [Frenz and Lappe 2005] was possible. Subject KG showed no improvement in distance estimation compared to the nonstereoscopic scene, as her slope was within the 95% confidence interval of her results in the earlier study. For subject HF, the error increased beyond the 95% confidence interval from the nonstereoscopic condition. On average, the subjects undershot the simulated distances by 33% (slope of 0.67 for the fitted regression to the data of all subjects). With dot plane 2, all subjects underestimated the simulated travel distance (see Figure 7). The slopes of the fitted linear regressions varied between 0.46 and 0.85. The slope for subject HF was not different from that found in nonstereoscopic presentation. Subjects JL and KG showed larger errors in the stereoscopic environment than in the nonstereoscopic case. Distance underestimation was 36% on average (slope of regression 0.64). In comparison to the results obtained without stereoscopic stimulus presentation in earlier work [Frenz and Lappe 2005], the results show a similar error in distance underestimation (33% distance underestimation using dot plane 1 and 24% using dot plane 2).

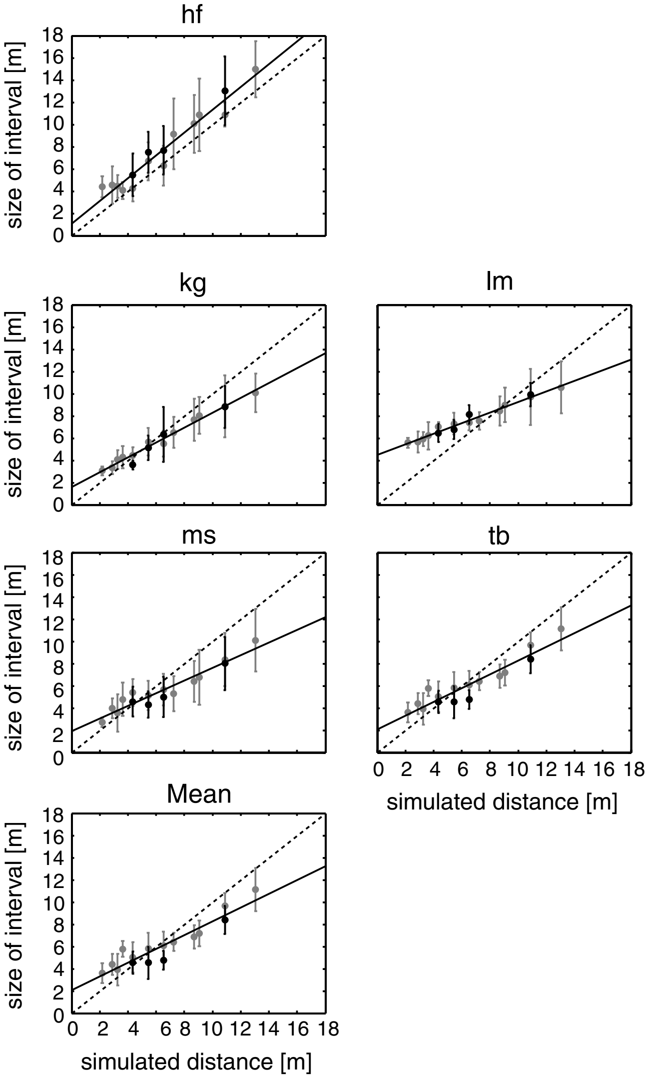

Fig. 6. Results of experiment 4 with motion simulation on dot plane 1. Same convention as in Figure 2.

7. DISCUSSION

We investigated how the visual motion induced by one's own movement through a structured space can be used to gauge travel distance. Estimating absolute travel distance from optic flow is problematic, because optic flow speeds covary with the dimensions of the environment and are, thus, subject to an environment-specific scale factor. Discrimination of the distances of two simulated self-motions of different speed and duration is possible from optic flow [Bremmer and Lappe 1999; Frenz et al. 2003]. When a distance estimate obtained from optic flow has to be transformed into a spatial interval in the visual environment, however, the estimate often undershoots the true distance [Frenz and Lappe 2005]. With the experiments described here, we investigated whether this underestimation can be

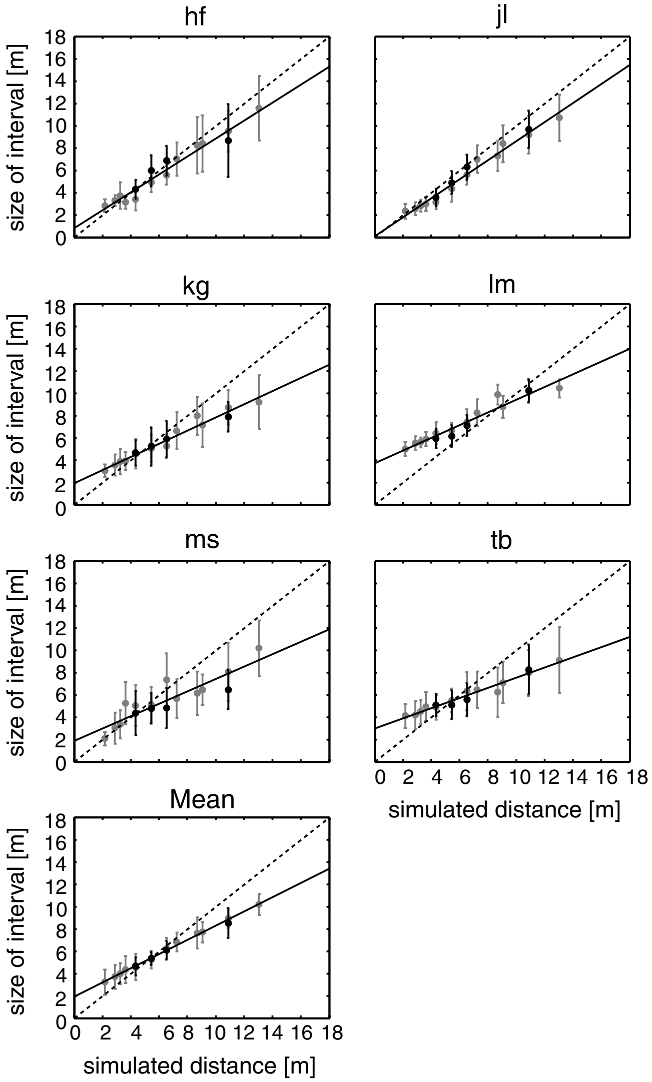

Fig. 7. Results of experiment 4 with motion simulation on dot plane 2. Same convention as Figure 2.

remedied with more and better information about the 3D layout of the scene. Experiments 1 and 2 tested whether pictorial depth cues and additional motion parallax increase the accuracy of flow-based travel distance estimation. This was not the case. Previous work [Frenz et al. 2003; Frenz and Lappe 2005] had shown that motion parallax in itself is necessary and sufficient to allow accurate distance discrimination. Motion parallax is present in the flow stimulus of the ground plane. Increasing the amount of motion parallax by adding the pole experiment 1 did not increase accuracy of the distance estimate. Additional pictorial depth cues like the size of the poles or the position of the horizon also did not improve accuracy. This is consistent with the earlier finding that pictorial depth cues are not needed for accurate discrimination. Yet, if the underestimation resulted from a failure to correctly perceive the scales in the scene, the added depth cues should have increased estimation accuracy. Thus, the undershoot is likely not resulting from errors in the perception of general scene layout. This conclusion is further supported by the results of experiments 3 and 4, in which the scene was presented in stereoscopic vision. The undershoot remained in these experiments even though scene layout was specified by binocular 3D information. A second reason why binocular disparity could have improved the accuracy of distance estimation from optic flow was that the apparent speed of a self-motion simulation is increased when the motion is simulated with stereoscopic information [Palmisano 2002]. If the stereoscopic presentation had improved the perception of the translation velocity in our experiments, we would have expected that the subjects indicated larger travel distances. This should have led to steeper slopes of the fitted linear regressions. This was not the case. Thus, stereoscopic presentation does not improve the perception of travel distance from optic flow.

Taken together, the results of this and earlier studies suggest that the undershoot is not because of errors in perceiving the static layout of the virtual scene. Nor does it seem a result of an inability to integrate visual motion from the optic flow into a distance percept, because distance discrimination is accurate and distance estimation is highly correlated to the true travel distance. Therefore, we suggest that the error arises in the transformation of the optic flow-based distance measure into a static scene interval. If this process is subject to a scale factor, then our results can be explained if this scale factor is commonly lower than one.

Support for this idea may be gained from observing the slopes from individual subjects. Comparison of Figures 2–7 reveals that the tendency to under- or overshoot the true travel distance is preserved across experiments for individual subjects. For instance, subject JL gives very accurate estimates in all experiments. Subjects HF slightly overshoots in most experiments. Subjects KG and LM show strong undershoots in all experiments. Subject MS severely and consistently underestimates distances in all experiments. This observation suggests that while, on average, subjects undershoot the true distance, the amount of undershoot is specific and consistent for each subject, and some subjects also consistently overshoot for in small amounts. This could be explained if the scale factor for conversion of the flow-based distance measure into a static scene interval is specific for each subject.

In this view, the scale factor would amount to a form of calibration between the integrated self-motion form the optic flow and the apparent distance in a particular scene. In the real world, this calibration would likely be obtained from active interaction with the environment, such as walking and head movements. Thus, it should depend on the presence of other sensory signals for self-motion. Indeed, the addition of vestibular sensory stimulation or proprioceptive signals during real self-motion influences the accuracy of travel distance judgments [Harris et al. 2000; Sun et al. 2004]. Moreover, the scale factor should be plastic and subject to adaptation. Indeed, it has been shown that the perceptual-motor calibration of human locomotion in the real world can be altered when the visual flow associated with self-motion is mismatched relative to biomechanical walking speed [Pelah and Barlow 1996; Durgin and Pelah 1999; Durgin et al. 2005; Thompson et al. 2005]. Therefore, an explanation for the underestimation of travel distance from optic flow may be that the subjects were not adapted to the self-induced optic flow for the virtual environments we used. Subjects did not move through the virtual environment to experience how their self-motion changed the virtual scene. Perhaps an adaptation to the optic flow field in this virtual environment according to the movement of the observer could improve the perception of how far the observer actually traveled with the self-motion simulation. The subjects would be able to calibrate the presented optic flow field to their actual movement. Such experiments of adapting the subjects to the flow fields, together with the results of the present work, could help clarify whether the subjects can calibrate the flow fields with their motion.

ACKNOWLEDGMENTS

ML is supported by the German Science Foundation, the German Federal Ministry of Education and Research BioFuture Prize and the EC Project Drivsco.

REFERENCES

BERTHOZ, A., ISRAEL, I., GEORGES–FRANCOIS, P., GRASSO, R., AND TSUZUKU, T. 1995. Spatial memory of body linear displacement: What is being stored? *Science 269*, 95–98.

BEUSMANS, J. M. 1998. Optic flow and the metric of the visual ground plane. *Vision Res. 38*, 1153–1170.

BREMMER, F. AND LAPPE, M. 1999. The use of optic flow for distance discrimination and reproduction during visually simulated self motion. *Exp. Brain Res. 127*, 33–42.

BRONSTEIN, A. AND BUCKWELL, D. 1997. Automatic control of postural sway by visual motion parallax. *Exp. Brain Res. 113*, 243–248.

CUIJPERS, R. H., KAPPERS, A. M. L., AND KOENDERINK, J. J. 2000. Large systematic deviations in visual parallelism. *Perception 29*, 1467–1482.

CUIJPERS, R. H., KAPPERS, A. M. L., AND KOENDERINK, J. J. 2002. Visual perception of collinearity. *Percep. Psychophys. 64*, 392–404.

DURGIN, F. H. AND PELAH, A. 1999. Visuomotor adaption without vision? *Exp. Brain Res. 127*, 12–18.

DURGIN, F. H., PELAH, A., FOX, L. F., LEWIS, J., KANE, R., AND WALLEY, K. A. 2005. Self-motion perception during locomotor recalibration: More than meets the eye. *J. Exp. Psychol. Hum. Percept. Perform. 31*, 398–419.

FOLEY, J. M. 1980. Binocular distance perception. *Psychol. Rev. 87*, 411–434.

FOLEY, J. M., RIBEIRO-FILHO, N. P., AND DA SILVA, J. A. 2004. Visual perception of extent and the geometry of visual space. *Vision Res. 44*, 147–156.

FRENZ, H. AND LAPPE, M. 2005. Absolute travel distances from optic flow. *Vision Res. 45*, 1679–1692.

FRENZ, H., BREMMER, F., AND LAPPE, M. 2003. Discrimination of travel distances from 'situated' optic flow. *Vision Res. 43*, 2173–2183.

GIBSON, J. J. 1950. *The Perception of the Visual World*. Houghton Mifflin, Boston, MA.

GRIGO, A. AND LAPPE, M. 1998. Interaction of stereo vision and optic flow processing revealed by an illusory stimulus. *Vision Res. 38*, 281–290.

HARRIS, L. R., JENKIN, M., AND ZIKOVITZ, D. C. 2000. Visual and non-visual cues in the perception of linear self motion. *Exp. Brain Res. 135*, 12–21.

INDOW, T. 1991. A critical review of Luneburg's model with regard to global structure of visual space. *Psychol Rev. 98*, 430–453.

KEARNS, M. J., WARREN, W. H., DUCHON, A. P., AND TARR, M. J. 2002. Path integration from optic flow and body senses in a homing task. *Perception 31*, 349–374.

LEE, D. N. 1980. The optic flow field: The foundation of vision. *Phil. Trans. R. Soc. Lond. B 290*, 169–179.

LI, L. AND WARREN, W. H. 2004. Path perception during rotation: influence of instructions, depth range, and dot density. *Vision Res. 44*, 1879 – 1889.

LOOMIS, J. M., KLATZKY, R. L., GOLLEDGE, R. G., CICINELLI, J. G., PELLEGRINO, J. W., AND FRY, P. A. 1993. Nonvisual navigation by blind and sighted: Assessment of path integration ability. *J. Exp. Psych. Gen. 122*, 73–91.

MESSING, R. AND DURGIN, F. 2005. Distance perception and the visual horizon in head-mounted displays. *ACM Transactions on Applied Perception 2*, 234–250.

PALMISANO, S. 2002. Consistent stereoscopic information increases the perceived speed of vection in depth. *Vision Res. 31*, 463–480.

PELAH, A. AND BARLOW, H. B. 1996. Visual illusion from running. *Nature (London) 381*, 283.

PERUCH, P., MAY, M., AND WARTENBERG, F. 1997. Homing in virtual environments: Effects of field of view and path layout. *Perception 26*, 301–311.

PROKOP, T., SCHUBERT, M., AND BERGER, W. 1997. Visual influence on human locomotion. modulation to changes in optic flow. *Exp. Brain Res. 114*, 63–70.

REDLICK, F. P., JENKIN, M., AND HARRIS, L. R. 2001. Humans can use optic flow to estimate distance of travel. *Vision Res. 41*, 213–219.

RIECKE, B. E., VAN VEEN, H. A. H. C., AND BÜLTHOFF, H. H. 2002. Visual homing is possible without landmarks: A path integration study in virtual reality. *Presence 11*, 443–473.

RUSHTON, S. K., HARRIS, J. M., AND WANN, J. P. 1999. Steering, optic flow, and the respective importance of depth and retinal motion distribution. *Perception 28*, 255–266.

SUN, H.-J., CAMPOS, J. L., AND CHAN, G. S. 2004. Multisensory integration in the estimation of relative path length. *Exp. Brain Res. 154*, 246–254.

THOMPSON, W. B., CREEM-REGEHR, S. H., MOHLER, B. J., AND WILLEMSEN, P. 2005. Investigations on the interactions between vision and locomotion using a treadmill virtual environment. *Proceedings of SPIE 5666*. 481–492.

VAN DEN BERG, A. V. AND BRENNER, E. 1994a. Humans combine the optic flow with static depth cues for robust perception of heading. *Vision Res. 34*, 2153–2167.

VAN DEN BERG, A. V. AND BRENNER, E. 1994b. Why two eyes are better than one for judgements of heading. *Nature (London) 371*, 700–702.

WAGNER, M. 1985. The metric of visual space. *Percept Psychophys. 38*, 483–495.

WARREN, W. H. AND HANNON, D. J. 1990. Eye movements and optical flow. *J. Opt. Soc. Am. A 7*, 160–169.

WARREN, W. H., MORRIS, M. W., AND KALISH, M. 1988. Perception of translational heading from optical flow. *J. Exp. Psychol. Hum. Percept. Perform. 14*, 646–660.

WITMER, B. G. AND KLINE, P. B. 1998. Judging perceived and traversed distance in virtual environments. *Presence 7*, 144–167.

Received December 2005; revised September 2006; accepted November 2006