Virtual Odometry From Visual Flow

Markus Lappe[a,b], Harald Frenz[a,b], Thomas Bührmann[b], Marina Kolesnik[c]

[a]Psychol. Inst. II, Westf.-Wilhelms University, Münster, Germany;
[b]Dept. Zoology & Neurobiology, Ruhr-University Bochum, Germany;
[c]Fraunhofer Inst. Media Communication IMK, St. Augustin, Germany

ABSTRACT

We investigate how visual motion registered during one's own movement through a structured world can be used to gauge travel distance. Estimating absolute travel distance from the visual flow induced in the optic array of a moving observer is problematic because optic flow speeds co-vary 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 reliably possible from optic flow, however, if the visual environment is the same for both motions, because the scale factors cancel in this case.[1,2] Here, we ask whether a distance estimate obtained from optic flow can be transformed into a spatial interval in the same visual environment. Subjects viewed a simulated self-motion sequence on a large (90 by 90 deg) projection screen or in a computer animated virtual environment (CAVE) with completely immersive, stereographic, head-yoked projection, that extended 180deg horizontally and included the floor space in front of the observer. The sequence depicted self-motion over a ground plane covered with random dots. Simulated distances ranged from 1.5 to 13 meters with variable speed and duration of the movement. After the movement stopped, the screen depicted a stationary view of the scene and two horizontal lines appeared on the ground in front of the observer. The subject had to adjust one of these lines such that the spatial interval between the lines matched the distance traveled during the movement simulation. Adjusted interval size was linearly related to simulated travel distance, suggesting that observers could obtain a measure of distance from the optic flow. The slope of the regression was 0.7. Thus, subjects underestimated distance by 30%. This result was similar for stereoscopic and monoscopic conditions. We conclude that optic flow can be used to derive an estimate of travel distance, but this estimate is subject to scaling when compared to static intervals in the environment, irrespective of steroscopic depth cues.

Keywords: optic flow, distance estimation, stereo, depth, virtual reality

1. INTRODUCTION

Because optic flow speeds induced by self-motion depend on the dimensions of the environment, distance estimation from optic flow is possible only when speeds can be calibrated from other sources of depth information, such as binocular disparity or known geometrical properties of the environment. For instance in a ground-plane environment with fixed eye height, discrimination of the distances of simulated self-motions of different speed and duration is reliably possible from optic flow[1,2] because the height above the ground can be used as a scale factor. In a study by Bremmer and Lappe[1] human subjects had to indicate, which one of two sequentially visually simulated self-motion covered a larger distance. Subjects correctly performed this discrimination under the assumption that the environment did not change between the two motions. Their results could be derived from either of two main strategies. The first hypothesis assumes that observers integrate and compare the image motions of the environment on the retina (2-D hypothesis). The second hypothesis involves a percept of the ego-motion (3-D hypothesis): the observer first estimates the simulated translation ego-velocity (V) from the amount of optic flow and then integrates V over time. To differentiate between these possibilities, Frenz et al.[2] conducted experiments in which the optic flow field was altered by varying the translation velocity and the perspective on the environment. If the environment changed between the two ego-motion simulations without the subjects noticing the change, the subjects made predictable errors in distance discrimination: they attributed the whole change in the optic flow field to a change of the translation velocity, assuming that the scene remained

---

Further author information: (Send correspondence to M.L.)
M.L.: E-mail: mlappe@psy.uni-muenster.de, Telephone: 49 251 83 34141

Human Vision and Electronic Imaging X, edited by Bernice E. Rogowitz,
Thrasyvoulos N. Pappas, Scott J. Daly, Proc. of SPIE-IS&T Electronic Imaging,
SPIE Vol. 5666 ©2005 SPIE and IS&T 0277-786X/05/$15

constant. If the subjects noticed the altered environmental structure, they could extract the amount of change in the flow field based on the change of the environment from the amount based on changes of translation velocity. These results support the 3-D hypothesis.

Given that observers can integrate translation velocity to discriminate distances, one must ask whether they can also use their 3-D percept of the self-motion to build up a distance measure in some arbitrary unit derived from visual motion. As we used virtual ground planes to simulate the environment, it is possible to make absolute judgements in terms of eye heights. To build up a correct distance measure subjects have to calibrate the optic flow on the basis of the environmental depth information and integrate the velocity of the ego-motion over time. The question then is whether subjects can indicate the perceived travel distances in a stationary environment. In this case, the judged distance has to be transferred to the environment in terms of a virtual distance.

Frenz and Lappe[3] instructed subjects to indicate the perceived distances in terms of a virtual ground interval. Observers viewed a stimulus depicting self-motion over a random dot ground plane with variable speed and duration. Simulated distances ranged from 2.2m to 13m. Afterwards the screen depicted a stationary view of the scene and two horizontal lines appeared on the ground in front of the observer. The subject had to adjust the interval between these lines with a pointing device to match the distance traveled during the movement simulation. Frenz and Lappe[3] found that interval size increased linearly with simulated travel distance. Hence, observers could obtain a coherent distance measure from optic flow.

This linear correlation is consistent with other studies that investigated visual motion based distance estimation during walking,[4–6] riding a bike,[7] steering on a mobile robot,[8] or navigating in a virtual environment[5, 9, 10] . However, the adjusted interval size in the experiments of Frenz and Lappe systematically undershot the true travel distance by about 30 percent. Similar undershoots have been observed by Sun et al.[7] and Witmer and Kline.[5] For some stimulus parameters, overestimations of travel distance also occurred.[3, 11] These findings suggest that optic flow can be used to derive an estimate of travel distance, but this estimate is subject to scaling when compared to static intervals in the environment.

The observed error of underestimation of the traversed distance was not due to a compressed metric of visual space in general (Frenz and Lappe, submitted), or the way of presentation of the reference distance (statically or in terms of a motion simulation), or the method by which the subjects indicated the perceived travel distance (in terms of a second self-controlled visual motion, in multiple virtual eye-heights, or by actively walking the same distance without visual information).[3] A remaining possible explanation for the occurrence of this constant error is a mis-perception of the visually simulated self-motion, i.e. a mis-perception of the optic flow field itself. A mis-perception of the visually simulated self-motion is not contrary to the previous findings: accurate distance estimation occurred only in experiments in which the subjects had to compare the travelled distances of two visually simulated self-motions (see[1] or[2]). A mis-perception of flow fields could occur in both motion simulation, and thus cancel out in the comparison.

The estimate of travel distance is derived from the translational ego-velocity of the observer. The perception of the translation velocity is closely linked to the structure of the environment, i.e. to the provided depth cues. Motion parallax (the relative optical velocity in an optic flow field depending on the distance to the moving observer) as the only available depth information is sufficient to discriminate the travelled distances of two visually simulated self-motions.[2] But, motion parallax is a dynamic depth cue and therefore only available when the observer is moving (either simulated or real motion) through the environment. For the perception of the depth structure, static information about the virtual environment might also be helpful. On the basis of static depth information, the optic flow field could possibly be calibrated with less error with respect to the travelled distances. Therefore, if subjects misperceived the flow field and hence the simulated travel distance, additional cues about the structure of the environment might improve the percept.

Another way to emphasise the depth structure of the environment is the use of binocular disparity. Disparity can be simulated by presenting each eye of the observer a different image, i.e., he or she receives a stereoscopic view of the virtual scene. Visualization of the virtual scene comprises two views that simulate the positions of the observers eyes. Each eye of the observer then gets the image of the corresponding view. The stereoscopic presentation of the stimuli with the additionally provided depth information can possibly improve the perception of the flow field. In a study by Palmisano,[12] stereoscopic information about a simulated self-motion increased the perceived translation velocity. Thus, in our earlier experiments the subjects could have misperceived the simulated translation velocity of the self-motion and therefore underestimated the traversed distance. Additionally, van den Berg and Brenner[13, 14] 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. Grigo and Lappe[15] also described that human subjects used stereoscopic depth information as an additional source of information for the interpretation of the simulated flow field.

With the following experiments we investigated whether stereoscopic presentation in an immersive environment reduces the error of distance estimation when the traversed distances were indicated in terms of a virtual ground interval. We performed experiments both with an experimental setup we used earlier[2] and in a Computer Animated Virtual Environment (CAVE). The CAVE consisted of a large room with stimulus projection on the front wall, the two sidewalls, and the floor. Thus, with the use of the stereoscopic presentation the stimuli appeared in their full scale. This accordingly enhanced the sensation of depth within the stimuli. With the use of this two experimental setups (single screen vs. CAVE) we could investigate whether the complete immersion into the virtual scene simulated in the CAVE improved the perception of the self-motion simulation and allowed better travel distance estimation.

2. METHODS

2.1. General Procedure

We visually simulate ego-motion in different artificial virtual environments with varying depth information and asked human observers to indicate the perceived travel distances in terms of a virtual interval on the ground. Each trial started with the visual simulation of a self-motion. We tested 4 different translation velocities (1, 1.5, 2.5 or 3 m/s in the single screen experiments, 1.45, 2,18, 3.63 and 4.35 m/s in the CAVE) and four different durations of the motion simulations (1.5, 2, 2.5 or 3 s). The simulated travel distances varied therefore between 1.5 and 9 m in the single screen experiments and between 2.18 and 13.05 m in the CAVE experiments. We presented the 16 conditions in a pseudorandomised order with ten repetitions each. Four travel distances (3, 3.75, 4.5 and 7.5 m in the single screen experiments and 4.35, 5.44, 6.53 and 10.88 m in the CAVE experiments) were simulated with different combinations of the translation velocities and simulation duration. After the self-motion simulation, two horizontal white lines appeared on the virtual ground plane. One line (reference) was always presented 4m in front of the observer's virtual position. The second line appeared 3m in front of the observer and was adjustable by moving a computer mouse. Both lines were positioned on the virtual ground level. The subject's task was to indicate the travelled distance of the visually simulated reference motion in terms of a virtual ground interval. Six subjects participated in the experiments performed with the single screen. Five of these also participated in the experiments in the CAVE.

2.2. Virtual Environments

Both environments consisted of a ground plane covered with random dots. The first (dot plane 1) comprised 3300 white light points, a subset of which were shown on the screen at any time. These light points were positioned on a grating every 2m within 52m in front of the observer and every 6m to either side within a distance of 30m. Thereafter, the position of each light point was shifted randomly up to 5m to either side and forward/backwards to achieve a balanced random distribution. Lifetime of the light points was controlled. With a probability of 10 % each point would vanish and reappear randomly in the scene in each frame. With a frame rate of 72 Hz the mean lifetime of each dot was 139 ms. Therefore, on average 970 light points were visible on the screen. The limitation of the dots lifetime ensures that the subjects could not get information about the travel distance from trajectories of the light points. Size and luminance of the light points remained constant during movement simulation, eliminating size change as a distance cue. Dynamic depth cues were provided by motion parallax. In the static scene, the gradient of texture density towards the horizon still served as depth cue. Frame rate was 72 Hz. Mean luminance was 2.0 CD/m^2.

Dot plane 2 also simulated motion over a ground plane. It consisted of 150 white light points on a black background. The points were evenly distributed on the lower part on the screen. During movement simulation the dots moved as if they lay on a ground plane, i.e., they obeyed the pattern of motion parallax. The velocity gradient during motion was the only cue to the depth and 3D structure of the scene. Without motion simulation

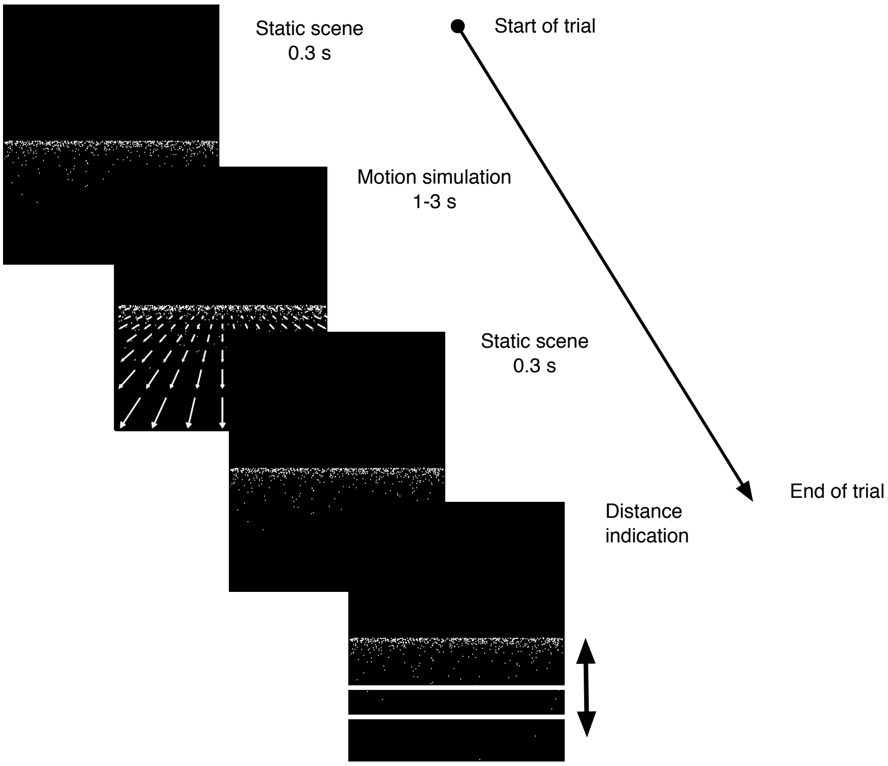

Figure 1. Temporal sequence of events in a single trial. First, the environment (a random dot ground plane) is presented statically for 300 ms. Then, forward motion is simulated for a given combination of speed and duration. The white arrows symbolise the optic flow experienced by the observer in this phase. Afterwards the environment is presented again statically for 300ms. Then two horizontal indicator lines appear in the environment. The subject has to adjust the second indicator line such that the distance on the ground between the two lines matches the distance covered by the motion simulation

the dot pattern provided no information about the distance between observer and light points or about the structure of the environment. Furthermore, the lifetime of the light points was limited in the same way as described for dot plane 1. The mean luminance was 0.6 CD/m^2.

2.3. Stimulus presentation on the single screen

The stimuli were created in real-time on a Silicon Graphics Indigo2 workstation and presented on a 120 x 120 cm back-projection screen (Dataframe, type CINEPLEX) using a CRT video projector (Electrohome ECP 4100). Resolution was 1280 x 1024 pixel. Frame rate of the rendered images was 72 Hz. The room was darkened and only illuminated by the stimuli. The subjects were positioned 0.6 m in front of the screen on a chair. The chair was adjusted in height so that the subject's physical eye height was 1.6m above the ground. The field of view was 90 x 90 deg. In the single screen monoscopic condition the participants viewed the scene binocularly but without stereoscopic depth information. For the stereoscopic presentation two camera viewpoints onto the scene were created with a virtual inter-ocular distance of 6.4 cm. The resulting images from the two viewpoints were alternatingly presented on the screen each with a frame rate of 60 Hz. Subjects wore liquid crystal shutter glasses (StereoGraphics; model CrystalEyes), which were synchronised with the stimulus presentation rate. The glasses for each eye were thus opened and shut with a rate of 60 Hz. The shutter glasses reduced the subjects' field of view to 60° x 70°.

2.4. Stimulus presentation in the CAVE

The CAVE is a room of 3 x 3 x 3 m size constructed of projection surfaces. The stimuli were back-projected onto the wall in front of the subjects and the two sidewalls. The projection onto the floor was done from the ceiling. Projections were realized with four three-color Electrohome projectors. Stimuli were rendered in real time by Silicon Graphics Hardware. Spatial resolution per screen and frame rate were the same as in the experiments performed with the single screen. Subjects wore shutter glasses (CrystalEyes, description see above) to generate the stereoscopic view on the virtual environment. The system tracked the position and orientation of the shutter glasses and rendered the simulated gaze on the virtual environment according to the movements of the subject's head. During the experiment, the participants stood in the middle of the CAVE. For the adjustment of the virtual ground interval we used a Cubic Mouse, which the subjects held in their hands. The Cubic Mouse consists of twelve degrees of freedom that can be manipulated, leaving the three-dimensional coordinate system in the hands of the user and allowing the simple and effective manipulation of spatial models. We used only one dimension of the device because only changes along the axis in depth of the virtual environment were used to vary the position of the adjustable virtual line. The subjects controlled the virtual position of the adjustable line by pulling and pushing one of the rods on the cubic mouse. Before each indication of the traversed distance, the subjects had to pull the rod as far as possible. Without the pulling of the rod, the horizontal lines did not appear in the virtual environment. When the participants thought they matched the travel distance of the reference self-motion, they pressed a button on the Cubic Mouse.

3. RESULTS

In the following, we will first describe the results obtained with the experiments performed with the single screen and afterwards the results obtained with the experiments performed in the CAVE. At the end of this section, we will compare the results obtained with the two setups. We used three parameters to analyse the data. The first was the correlation coefficient ρ between the indicated and simulated travel distances. The correlation coefficient served as an indicator for the subject's ability to use an abstract distance gauge for the indication of the perceived travel distance. The second parameter was the slope of the linear regressions fitted to the data points. 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 underestimation of the travel distance of the self-motion simulation. Slopes greater 1 indicate overestimation of the traversed distance. Note that the intercepts were omitted from the analysis, because they reflect only a constant subject depending error. The third parameter was the difference in distance indication between same virtual distances, simulated with different combination of translation velocities and simulation duration. If the subjects based their judgement on the travel distance, same simulated travel distances should be indicated with ground intervals of the same size.

3.1. Results of single screen experiments

Figure 2B illustrates the results obtained in the single screen stereoscopic experiments. For comparison, Figure 2A shows the results obtained in the single screen monoscopic experiments obtained earlier with the same methods.[3] The upper panels give the results obtained with dot plane 1. The lower panels give the results obtained with dot plane 2. Each panel plots the indicated distances as a function of the simulated distances. Each circle shows the mean over subjects. The error bars indicate standard deviations. Red circles show the mean indicated distances of a simulated distance travelled with higher translation velocities and shorter simulation durations than the blue circles. The dashed lines show hypothetical data of accurate distance estimation without errors. The solid lines correspond to linear regressions fitted to the actual data. Slopes, offsets, correlation coefficients, and significance levels of the fit are indicated in the figures. The parameters of the linear regressions fitted to the data of each individual subject are given in table 1. All regressions were an appropriate description of the data ($r^2 > 0.74$, $p < 0.05$).

Correlation coefficients between the indicated and simulated distances of the self-motion with dot plane 1 varied between 0.64 and 0.93 and with dot plane 2 between 0.55 and 0.89 among subjects. This shows that the subjects possess an abstract distance gauge as described in earlier work.[3] Four travel distances were simulated each with two different translation velocities and simulation durations. If the subjects based their distance estimation on an integral of the perceived translation velocity over time, one would expect that same traversed

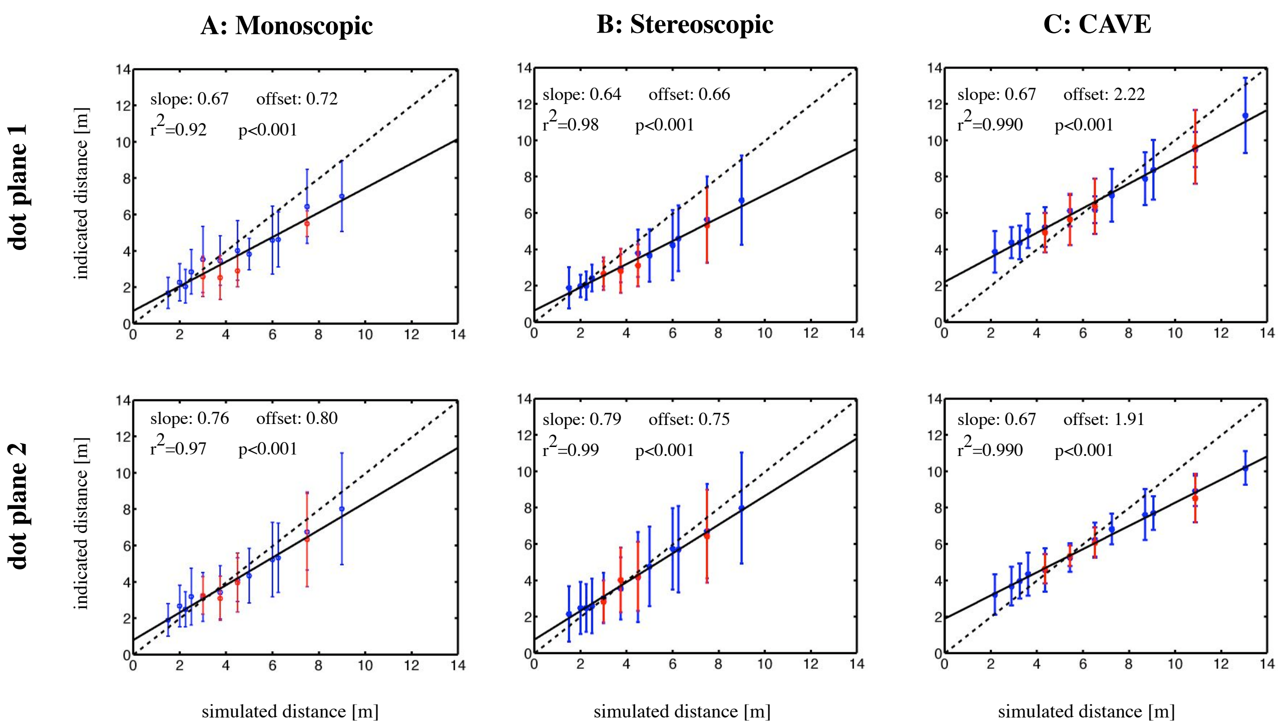

Figure 2. Results obtained with motion simulation on dot plane 1 (top row) and dot plane 2 (bottom row in the single screen experiments either monoscopic (A) or stereoscopic (B) and in the CAVE experiments (C). The sizes of the adjusted intervals are plotted as a function of the simulated travel distances. Each circle corresponds to the mean over subjects. The error bars indicate the standard deviation. Red circles indicate data obtained with higher translation velocity and shorter simulation duration than the corresponding blue circles (travel distances: 3, 3.75, 4.5 and 7.5 m). The solid line corresponds to the fitted regression. The dashed line illustrates hypothetical data of exact distance estimation without errors.

distances were indicated with ground intervals of the same size. From Figure 2 it is clear, that the subjects mostly indicated same travel distances with same sized ground intervals, despite the variation of translation velocity and simulation duration (red vs. corresponding blue markers). This results confirms that subjects predominantly based their distance estimation on the perceived travel distance.

The slopes of the linear regressions to the single subject data obtained with motion simulation on dot plane 1 ranged from 0.23 to 1.1 (Table 1). One subject (*hf*) slightly overestimated the travelled distances, another subject (*jl*) accurately indicated the simulated distances and four subjects underestimated the presented distance of the self-motion For the pooled data, the slope of the fitted linear regression was 0.64 (Figure 2B). This indicates an underestimation of the travelled distance of 36 % on average. The slopes of the linear regressions fitted to the data obtained with motion simulation on dot plane 2 varied between 0.34 and 1.1 (Table 1). Also in this experiment one subject (*hf*) slightly overestimated the traversed distance of the self-motion. Two subjects (*jl* and *lm*) showed accurate distance estimation whereas three participants underestimated the simulated travel distance. The fitted linear regression to all data of all subjects obtained with dot plane 2 had a slope of 0.79 (Figure 2 B) and therefore also indicated an underestimation of the simulated travel distance, here of about 21 %. We tested whether the slopes of the linear regressions fitted to the data of all subjects differed significantly in the two experiments. To normalised the data we subtracted the constant offsets of the regressions (the intercept of the regression line with the y-axis) from the data points. A two-way-ANOVA on the normalised data showed that the difference between the slopes fitted to the data obtained with the different virtual environments was significant ($p < 0.05$). In comparison with the data obtained in the monoscopic condition (Figure 2A) the slopes of the linear regressions were not significantly different (two-way-ANOVA, $p > 0.05$ for dot plane 1, $p = 0.16$ for dot plane 2).

Data obtained with the single screen						
	dot plane 1			dot plane 2		
subject	ρ	slope	r^2	ρ	slope	r^2
hf	0.87	1.1	0.97	0.88	1.1	0.88
jl	0.93	1	0.99	0.89	1.0	0.97
kg	0.77	0.47	0.88	0.81	0.69	0.96
lm	0.64	0.44	0.74	0.55	1.0	0.94
ms	0.72	0.23	0.9	0.72	0.34	0.90
tb	0.83	0.58	0.92	0.84	0.61	0.97
All subjects	0.65	0.64	0.98	0.58	0.79	0.99

Table 1. The parameters of the fitted linear regressions to the data obtained in the single screen stereoscopic experiments. The correlation coefficients (ρ) between the simulated and indicated distances, slopes and r^2 of the linear regressions are listed for each subject and the two used environments. The row labelled "All subjects" corresponds to the fitted linear regression to all data points of all participating subjects.

3.2. Results of the CAVE experiments

The single subject results of the experiments performed in the CAVE are presented in Figure 2C. The upper panel gives the data obtained with dot plane 1, the lower panel the data obtained with dot plane 2. Slopes, offsets, and correlation coefficients for each individual subject are given in table 3.2. Also in these experiments subjects could indicate the perceived distances of the self-motion simulation in terms of a static ground interval: the calculated correlation coefficients between the indicated and the simulated distances varied between 0.71 and 0.83 among subjects when the motion was simulated on dot plane 1. For motion simulation on dot plane 2 the calculated correlation coefficients varied between 0.67 and 0.91. The correlation coefficients calculated between the indicated distances of all subjects and the simulated distances on dot plane 1 (ρ=0.73) and dot plane 2 (ρ=0.76) emphasised the results obtained with the single subject data (see Table 3.2 for a detailed listing of the results). Same travel distances were always indicated with same ground intervals irrespective of speed and duration of the simulation (red and blue symbols in Figure 2C). Thus, subjects based their distance estimation on an integral of the perceived translation velocity over time.

The slopes of the linear regressions fitted to the data represent the error in distance estimation. All regressions obtained in this experiment were accurate descriptions of the data. Correlation coefficients r^2 (listed in Table 3.2) were high and p-values were below 0.05. For the single subject results obtained with motion simulation on dot plane 1 the slopes of the linear regressions varied between 0.48 and 1 (Table 3.2). One subject showed accurate distance estimation whereas four subjects underestimate the traversed distance of the self-motion simulation. The regression fitted to the data of all subjects had a slope of 0.67. Thus, travel distances were generally underestimated by 33 % in this condition. When the self-motion was simulated on dot plane 2, the results indicated an underestimation of the travelled distances for all subjects. The slopes of the regressions varied between 0.46 and 0.85 across subjects. The slope of the regression fitted to the data of all subjects was 0.64 and therefore indicated that the traversed distances were on average underestimated by 36 %. The error in distance estimation for the pooled results of all subjects did not significantly differ between dot plane 1 and dot plane 2 (two-way-ANOVA, $p > 0.05$). Compared to the results obtained without the stereoscopic presentation of the self-motion there was no significant difference when the motion was simulated on dot plane 1 (two-way-ANOVA, $p = 0.96$). The difference of the error in distance estimation obtained with the experiments performed on dot plane 2 with monoscopic presentation (Figure 2 A) and the error obtained in the experiments performed in the CAVE with dot plane 2 was significant (two-way-ANOVA, $p < 0.05$).

In summary, the slopes of the linear regressions fitted to the pooled data of all subjects showed no significant difference between the data obtained with dot plane 1 on the single screen and in the CAVE (two-way-ANOVA, $p > 0.05$). The slopes of the regressions obtained with dot plane 2 showed a significant difference between the two setup: the error in distance estimation was significantly smaller with the single screen setup compared to the CAVE data (two-way-ANOVA, $p < 0.05$).

Data obtained with the CAVE						
	dot plane 1			dot plane 2		
subject	ρ	slope	r^2	ρ	slope	r^2
hf	0.83	1	0.94	0.83	0.8	0.96
jl				0.91	0.85	0.98
kg	0.82	0.67	0.97	0.79	0.59	0.97
lm	0.79	0.48	0.97	0.73	0.57	0.94
ms	0.71	0.57	0.92	0.67	0.56	0.81
tb	0.82	0.62	0.88	0.85	0.46	0.96
All subjects	0.73	0.67	0.99	0.76	0.64	0.99

Table 2. The parameters of the linear regressions fitted to the data obtained with the CAVE. We listed the correlation coefficients (ρ) between the simulated and indicated distances, slopes and r^2 of the linear regressions for each subject and the two tested environments. "All subjects" refers to the results of the linear regression fitted to data points of all participating subjects.

4. DISCUSSION

With the experiments described here and in previous work[3] we showed that human subjects possess an abstract distance gauge derived from optic flow and can indicate this distance gauge in a static scene. However, subjects committed systematic errors in distance estimation as they usually underestimated the travelled distance of a visually simulated self-motion. With the experiments described in this study we investigated whether this underestimation is based on a mis-perception of the optic flow field or whether it can be remedied with more and better information about the 3D layout of the scene.

Palmisano[12] showed that the perceived speed of a self-motion simulation was increased when the motion was simulated with stereoscopic information. Theoretically, there are different possible explanations for this increase of the perceived velocity. First, there is additional *static* information about scene layout that the stereoscopic presentation provides. Optic flow fields are ambiguous with respect to the travelled distance if no scale information in the environment is available. For the ground plane, consistent within-scene scale information is provided by the velocity gradient of the motion simulation or the distances between the observer and environmental objects (Note that the second source of information is not available in dot plane 2). However, with stereoscopic presentation of the virtual environment disparity information is available in addition to determine the distance to objects in the environment and therefore indirectly scale their translation velocity. The second possible explanation for the increased velocity perception is the *dynamic* information provided by the stereoscopic presentation. Dynamic information is provided by the increase of disparity with decreasing distance between observer and environmental object. With this rate of change of disparity, the velocity of the movement could be directly calculated.

If the stereoscopic presentation had improved the perception of the translation velocity, we would have expected that the subjects in our experiments indicated larger travel distances. This should have led to steeper slopes of the fitted linear regressions to the data. This was not the case. The stereoscopic presentation of the self-motion simulation in the present experiments had only a weak effect on the error in distance estimation. Only the results obtained with dot plane 2 performed in the CAVE showed a significant difference between the stereoscopic and non stereoscopic presentation of the stimuli. But in this case, the slopes were even somewhat shallower in the CAVE. All other comparisons between stereoscopic and non-stereoscopic presentation of the self-motion showed no significant effect of the stereoscopic presentation. Taken together the present findings show that stereoscopic presentation does no improved the perception of travel distance from optic flow.

We performed the experiments with two different setups to investigate whether the impression of complete immersion into the virtual environment in the Computer Animated Virtual Environment (CAVE) improved the perception of the travel distance from simulated self-motion. In the CAVE, the orientation of observer's head was tracked and the gaze on the virtual scene rendered accordingly. The subjects could see the whole scene (from the horizon 90° downwards and 180° horizontal) by rotating their heads. But, although there was complete immersion into the virtual environment in the CAVE the ability of subjects to estimate the traversed distances did not improve compared to the single screen experiments. The slopes of the regressions fitted to the data obtained with dot plane 1 showed no significant difference between the setups and the slope of the data obtained with dot plane 2 was even slightly lower in the CAVE. Thus, even the use of the CAVE with the presentation of the whole scene did not improve the subjects' ability to estimate the travelled distances. There are different possible explanations why the subjects' performance in distance estimation did not improve in the CAVE. First, the restricted field of view might have limited the illusion of complete immersion into the virtual environment. Although the gaze on the virtual scene was rendered according to the subjects' head orientation, only 60° x 70° field of view of the scene was visible because of the shutter glasses. Visual flow perception is known to be influenced by the size of the field of view.[16] Grigo and Lappe[17] described that human subjects could judge the heading direction of visually simulated self-motions best when the field of view was large (in their study 90° x 90°). But, although the field of view in the CAVE experiments was reduced compared to the monoscopic experiments (Figure 2A) significant differences to the CAVE experiments occurred only for dot plane 2 of . The reduction of the field of view in the other conditions apparently had no detrimental effect on the distance estimation.

A different explanation for the underestimation may be that the subjects were not adapted to the self-induced optic flow. Subjects did not move through the virtual environment to experience, how their self-motion changed the virtual scene. In the present experiments this adaptation was omitted to make the results comparable to the results obtained with the single screen setup. 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 travelled 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 could clarify together with the results of the present work, if the subjects can calibrate the flow fields with their motion and if this is the case, whether this calibration changes the perception of the simulated self-motion.

5. CONCLUSION

We investigated whether a mis-perception of the optic flow field can explain the underestimation of the travel distance. We used stereoscopic presentation of the self-motion simulation which provided additional depth information based on disparity cues about the structure of the virtual environment. As for monoscopic presentation subjects underestimated the simulated travel distances by about 21 % to 36%. Thus, the addition of disparity cues to the available information for the distance judgement does not solve the problem of the underestimation of travel distance. We think the most promising approach to uncover the reason for the underestimation of the travelled distances of visually simulated self-motions would the adaptation of human subjects to self-induced flow fields in virtual environments. The adaptation of the subjects' real motion to the presented flow field enables the participants to calibrate the flow field based on their own motion. With this calibration, the subjects may perceive the simulated travel distance in its full amount.

ACKNOWLEDGMENTS

ML is supported by the German Science Foundation, the German Federal Ministry of Education and Research BioFuture Prize, and the EC Projects ECoVision and Eurokinesis.

REFERENCES

1. F. Bremmer and M. Lappe, "The use of optic flow for distance discrimination and reproduction during visually simulated self motion," *Exp.Brain Res.* **127**, pp. 33–42, 1999.
2. H. Frenz, F. Bremmer, and M. Lappe, "Discrimination of travel distances from 'situated' optic flow," *Vision Res.* **43**, pp. 2173–2183, 2003.
3. H. Frenz and M. Lappe, "Estimation of absolute travel distances from optic flow," *Vision Res.* , 2005. in press.
4. J. M. Loomis, R. L. Klatzky, R. G. Golledge, J. G. Cicinelli, J. W. Pellegrino, and P. A. Fry, "Nonvisual navigation by blind and sighted: Assessment of path integration ability," *J.Exp.Psych.: Gen.* **122**, pp. 73–91, 1993.

5. B. G. Witmer and P. B. Kline, "Judging perceived and traversed distance in virtual environments," *Presence* **7**, pp. 144–167, 1998.
6. M. J. Kearns, W. H. Warren, A. P. Duchon, and M. J. Tarr, "Path integration from optic flow and body senses in a homing task," *Perception* **31**, pp. 349–374, 2002.
7. H.-J. Sun, J. L. Campos, and G. S. Chan, "Multisensory integration in the estimation of relative path length," *Exp.Brain Res.* **154**, pp. 246–254, 2004.
8. A. Berthoz, I. Israel, P. Georges–Francois, R. Grasso, and T. Tsuzuku, "Spatial memory of body linear displacement: What is being stored?," *Science* **269**, pp. 95–98, 1995.
9. P. Peruch, M. May, and F. Wartenberg, "Homing in virtual environments: Effects of field of view and path layout," *Perception* **26**, pp. 301–311, 1997.
10. B. E. Riecke, H. A. H. C. van Veen, and H. H. Bülthoff, "Visual homing is possible without landmarks: A path integration study in virtual reality," *Presence* **11**, pp. 443–473, 2002.
11. F. P. Redlick, M. Jenkin, and L. R. Harris, "Humans can use optic flow to estimate distance of travel," *Vision Res.* **41**, pp. 213 – 219, 2001.
12. S. Palmisano, "Consistent stereoscopic information increases the perceived speed of vection in depth," *Vision Res.* **31**, pp. 463–480, 2002.
13. A. V. van den Berg and E. Brenner, "Humans combine the optic flow with static depth cues for robust perception of heading," *Vision Res.* **34**, pp. 2153–2167, 1994.
14. A. V. van den Berg and E. Brenner, "Why two eyes are better than one for judgements of heading," *Nature* **371**, pp. 700–702, 1994.
15. A. Grigo and M. Lappe, "Interaction of stereo vision and optic flow processing revealed by an illusory stimulus," *Vision Res.* **38**, pp. 281–290, 1998.
16. J. J. Koenderink and A. J. van Doorn, "Facts on optic flow," *Biol.Cybern.* **56**, pp. 247–254, 1987.
17. A. Grigo and M. Lappe, "Dynamic use of different sources of information in heading judgments from retinal flow," *J.Opt.Soc.Am.* **16**, pp. 2079–2091, 1999.