PERGAMON

Vision Research 42 (2002) 1991–2003

Vision Research

www.elsevier.com/locate/visres

Absolute distance perception during in-depth head movement: calibrating optic flow with extra-retinal information

Chin-Hwee Peh [a,*], Francesco Panerai [b], Jacques Droulez [b], Valérie Cornilleau-Pérès [c,d], Loong-Fah Cheong [a]

[a] *Department of Electrical and Computer Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore*
[b] *Laboratoire de Physiologie de la Perception et de l'Action, CNRS-Collège de France, 11 pl. M. Berthelot, 75005 Paris Cedex 06, France*
[c] *Image Processing and Application Laboratory (IPAL-CNRS), KDRL, 21 Heng Mui Keng Terrace, Singapore 119613*
[d] *Singapore Eye Research Institute, 11 Third Hospital Avenue, Singapore 168751*

Received 11 June 2001; received in revised form 4 March 2002

Abstract

We investigated the ability of monocular human observer to scale absolute distance during sagittal head motion in the presence of pure optic flow information. Subjects were presented at eye-level computer-generated spheres (covered with randomly distributed dots) placed at several distances. We compared the condition of self-motion (SM) versus object-motion (OM) using equivalent optic flow field. When the amplitude of head movement was relatively constant, subjects estimated absolute distance rather accurately in both the SM and OM conditions. However, when the amplitude changed on a trial-to-trial basis, subjects' performance deteriorated only in the OM condition. We found that distance judgment in OM condition correlated strongly with optic flow divergence, and that non-visual cues served as important factors for scaling distances in SM condition. Absolute distance also seemed to be better scaled with sagittal head movement when compared with lateral head translation.

Keywords: Self-motion; Object-motion; Absolute distance perception; Optic flow; Extra-retinal cue

1. Introduction

The perception of absolute distances, as opposed to relative distances between objects or distance changes along visible surface, is essential for planning goal directed movements. This capability is known to rely on many cues such as accommodation, vergence and motion parallax (Cutting & Vishton, 1995; Johansson, 1973). A recent study shows that multiple visual cues are combined accordingly based on their reliability in estimating the object spatial attributes (Landy, Maloney, Johnston, & Young, 1995). In terms of absolute distance perception, cues elicited by lateral head movement have been studied extensively (Dees, 1966; Eriksson, 1974; Ferris, 1972; Gogel, 1973; Johansson, 1973; Panerai, Cornilleau-Pérès, & Droulez, 2002). In fact, this way of generating distance cues is important for an enucleated observer who has to rely on monocular vision to perceive 3D space (Gonzalez, Steinbach, Ono, & Wolf, 1989; Marotta, Perrot, Nicolle, & Goodale, 1995). Animals have also been known to use head movements to plan their actions (Collett, 1978; Ellard, Goodale, & Timney, 1984; Poteser & Kral, 1995; Russell, 1931; Wallace, 1959).

Although a large proportion of our head movement is along the in-depth (or sagittal) direction, distance cues generated from such head movement are seldom discussed in the literature. It was found in gerbils that forward motion seemed to play a role in absolute distance perception (Ellard et al., 1984). Bingham and Stassen (1994) also suggested that if information of an oscillatory motion was available, egocentric distance could be scaled. Their simulation was based solely on a simple noiseless mechanical model and no performance on human subject was reported. In a separate work, Bingham and Pagano (1998) showed that subjects were able to perceive egocentric distance through isolated optic flow while performing sagittal head motion (HM)

---

[*] Corresponding author. Tel.: +65-6322-4512; fax: +65-6322-4598. *E-mail address:* elepehch@nus.edu.sg (C.-H. Peh).

0042-6989/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0042-6989(02)00120-7

in their reaching task. However, their responses were not as stable as those obtained when other monocular cues were present. In yet another related work, Pagano and Bingham (1998) reported that verbal estimates of distance were unstable and unreliable compared to reaching responses. However, both studies revealed that the perceived distance correlated well with the object distance. Due to physical limits, the distance studied in both works was limited to armlength measures and the subjects' verbal responses were also in proportional of armlength. Whilst the main objective of the latter work was at comparing the accurateness in verbal and motor responses, we are interested in comparing the judgement of egocentric distance in self-motion (SM) and object-motion (OM). We are also keen to assess this distance judgement at a larger distance: beyond the limit of one's armlength.

The analysis of optic flow developed by Longuet-Higgins and Prazdny (1980) demonstrates that an in-depth translation differs substantially from a fronto-parallel translation in terms of its induced flow structure. For instance, a depth discontinuity in the central image point can be readily resolved from lateral movement but not from in-depth movement. Similarly, the recovery of absolute distance from SM (i.e. scaling visual motion by the amplitude of the SM) is likely to be influenced critically by the motion direction. Firstly, the visual input is a diverging optic flow for in-depth movement but a lamellar one for lateral movement (Koenderink, 1986). If the visual scene were to be reduced to the central point, the retinal motion cannot be used for scaling absolute distance in the former case. Secondly, lateral eye movements for gaze stabilization are mainly involved in a lateral translation. Conversely, convergence is mainly involved in an in-depth translation. While the vestibular–ocular reflex constitutes an effective distance cue to lateral head translation, vergence and accommodation cues are important factors in distance perception during sagittal axis translation (Fisher & Ciuffreda, 1988; Gogel, 1961; Mon-Williams & Tresilian, 2000; Owens & Liebowitz, 1980; Tresilian, Mon-Williams, & Kelly, 1999; Viguier, Clement, & Trotter, 2001). This difference pertaining to the oculomotor behavior governs the strategy that an observer adopts to estimate distance. Interestingly, a recent paper (Wickelgren, McConnell, & Bingham, 2000) reveals that subjects are more accurate in performing reaching task in forward movement than in side-to-side movement experiment with isolated optic flow cue. Given that both the visual and non-visual information are different in lateral and in-depth HM, we anticipate that the ability of an observer to recover distance information from the two motion directions also differs.

Besides related work on distance perception, studies involving time-to-collision measure are also closely associated with our study. These time-to-collsion studies usually involve the passive observation of an approaching object by a stationary subject (e.g. Gray & Regan 1999a,b, 2000; Schiff & Detwiler, 1979; Steeves, Gray, Steinbach, & Regan, 2000). In this situation, the visual information about absolute distance is in principle ambiguous under monocular condition. Therefore, the question arises as to whether the dependency on absolute distance will occur when the observer is actively moving. Interestingly, Gray and Regan (2000) showed that the percept of time to collision is modified in the presence of a large field of optic flow. Hence, for true SM, we question whether the processing of the expansional flow is coupled with non-visual SM information so as to enable the perception of absolute distance.

2. Experiment 1

In this experiment, we compared the ability for a monocular subject to estimate the egocentric distance under two conditions: (a) during sagittal head movement toward/away from a stationary sphere (i.e. the SM condition); and (b) passive observation of a sphere moving along a sagittal plane trajectory (i.e. the OM condition).

2.1. Methods

2.1.1. Subjects

Four subjects (aged 18–25) were recruited for this experiment and paid an hourly rate. All subjects had normal vision and were familiar with the centimeter scale. One subject was one of the authors. The remaining three were naïve as to the goal of the experiment. The non-dominant eye of the subject was covered with a patch to achieve monocular viewing.

2.1.2. Apparatus

The visual stimuli were displayed on a large screen (189 cm in height and 243 cm in width) through a high-resolution (1024 × 1280) BARCO video projector connected to the graphic output of a computer (Pentium II 600 MHz with 3D graphic accelerator). The subject stood facing the screen at a distance of 60 cm away, while the display was projected from the other side of the large screen.

Images displaced on the screen were calculated as seen from the subject's eye position. The eye position was tracked in real-time, using the mechanical head tracker described by Panerai, Hanneton, Droulez, and Cornilleau-Pérès (1999). This device tracks the 3D head position in real-time at a high sampling rate of 85 Hz. It resolves the three translation and three rotation parameters of the head movements with a spatial resolution better than 1 mm and a latency of 0.011 s.

2.1.3. Stimulus

A computer-generated sphere located in a virtual 3D space was used as the main visual stimulus. The black opaque sphere was placed over a black background. Only the front surface of the sphere, covered with 200 ± 10 randomly distributed green dots, was visible. Without any motion cue, the sphere appeared as a flat 2D circle with random dots scattered over its surface. The stimulus luminance was in the range of 0.2–0.3 cd/m^2 while the background luminance was in the range of 0.002–0.003 cd/m^2. Each new stimulus view was refreshed at a rate of 85 Hz and presented to the subject at his eye-level.

2.1.4. Design

The sphere was presented at one of the nine simulated distances ranging from 30 to 238 cm (at a regular interval of 26 cm). The above range was chosen for its effectiveness in providing distance cues while making natural sagittal head movement with amplitude of 10 cm and frequency of 0.3 Hz. A simple calculation showed that for a 6° object at 100, 200 and 250 cm, the subject observed a movement of 1.17, 0.58 and 0.44 cm respectively on the screen. The subject can, thus, readily pick up the motion experienced in our chosen distance range.

Spheres of two apparent sizes (6° and 15° angular radius) were used. The apparent size was maintained by co-varying the 3D diameter of the sphere with the simulated distance. With these two apparent sizes and the nine object distances, we had a total of 18 distinct stimuli. Each of these stimuli was viewed 15 times, resulting in 270 experimental trials. The trials were randomly divided into six sessions of 45 trials. Hence, within each session, trials with spheres belonging to any of the two apparent sizes and any of the nine simulated distances could be found.

2.1.5. Procedure

The experiment was conducted under two motion conditions: SM and OM. With six sessions each for the SM and OM condition, we staggered sessions of different motion conditions. In other words, after each session of SM condition, the corresponding (recorded) session for the OM condition followed. The experiment was conducted in a dark room. Verbal reports were collected from the subjects at the end of each session.

During the SM session, the subject performed sagittal head movement while fixating the center of the dot distribution on the screen. Upon reaching 9 cm away from the initial 3D point in either forward or backward direction, a soft tone was emitted by the computer. This signaled the subject that the maximum amplitude of oscillation was about to reach and that he had to reverse his motion direction.

The subject was instructed to estimate his distance from the simulated sphere while performing the sagittal head movement. Since the distance was not constant during the movement, the subject was asked to report the distance from the center of oscillations to the surface of the sphere. After three periods of oscillations, the stimulus disappeared and the following set of four options for the perceived stimulus distance was displayed: (1) below 50 cm; (2) between 50 and 100 cm; (3) between 100 and 200 cm; and (4) above 200 cm. He was required to do a forced choice among these four options by browsing through a graphical menu with a joystick.

After the completion of each SM session, the subject was given a break before proceeding to the corresponding OM session. In the OM session, the subject stood stationary in front of the large screen while observing the simulated sphere moving back and forth in the 3D virtual space. Since the 3D movement of the sphere was a replay of the 3D movement in the earlier SM session, the subject would observe three periods of oscillation along the sagittal axis. Upon disappearance of the stimulus, the subject made his four forced choice decision.

2.2. Data analysis

2.2.1. Grouping of data

The four experimental conditions were abbreviated as follows:

1. SM_small—SM and small apparent size;
2. OM_small—OM and small apparent size;
3. SM_large—SM and large apparent size; and
4. OM_large—OM and large apparent size.

2.2.2. Psychometric curves and threshold

We derived three psychometric curves corresponding to the three distance thresholds (50, 100 and 200 cm) defined in the interval of simulated distances (30–238 cm) for each experimental condition. These curves establish the correspondence between the simulated distance and the distance perceived by the subject at the three threshold values (the values identifying this correspondence will be referred with the symbols D50, D100 and D200). The psychometric curves were obtained by regrouping the four possible responses (i.e. their associated probabilities p_1, p_2, p_3 and p_4, where $p_1 + p_2 + p_3 + p_4 = 1$) with respect to the distance thresholds. Hence, the probability of responses in estimating that the sphere distance exceeds a given distance are computed as follows:

$$P(>\mathrm{D50}) = p_2 + p_3 + p_4;$$
$$P(>\mathrm{D100}) = p_3 + p_4; \quad \text{and}$$
$$P(>\mathrm{D200}) = p_4.$$

The data obtained from all subjects were combined together to obtain the "averaged" psychometric curves. We adopted a bootstrap method (Foster & Walter, 1997) to derive the psychometric functions. This method fitted the data with a normal cumulative distribution function using weighted linear regression. It provided an estimate of the threshold, as well as the slope and spread of the fitted function at the threshold of any criterion level. In addition, the 95% confidence limits of these estimates were computed. We ran 1000 iterations on the program and observed the threshold at the criterion level of 50%.

For each of the average psychometric curve, the threshold, the 95% confidence limit interval (CLI) of the threshold and the inverse width were computed. While the threshold determined the perceptual level of the subjects, the 95% CLI and the inverse gradient of the threshold for each psychometric curve reflected the discriminatory power of the subjects, as well as the variability of their responses (both inter- and intra-individual) at each perceptual level.

2.3. Results

2.3.1. Psychometric curves

Fig. 1 shows the curves obtained by fitting all subjects' responses with psychometric functions under the four experimental conditions. Each subplot presents the averaged psychometric curves for the three distance thresholds (50, 100 and 200 cm) obtained by grouping the individual data according to the criteria "motion condition" and "apparent size".

Fig. 1 shows the monotonic widening of the CLI across curves of increasing perceptual level (i.e. from D50 to D200). There is also a progressive increment in the inverse gradients of the curves. In fact, there exists a strong correlation between the width of the CLI and the inverse gradient (Spearman $R = 0.951$, $p < 0.05$).

2.3.2. Threshold plots

Fig. 2 plots the physical distance against the perceived distance derived from the thresholds of the psychometric curves. The subplots were systematically grouped so that the effects across experimental conditions could be readily compared. The 95% CLIs were included as error bars in the graphs.

We observed that the distance estimates co-varied strongly with the stimulus distance under all conditions. The estimates were also rather accurate (median error of 7.72%, 7.56%, 8.93% and 22.63% in conditions SM_small, SM_large, OM_small and OM_large respectively). SM_large condition presented the best estimate whereas OM_large bore the largest error.

The distance estimates given in condition OM were relatively small as compared to condition SM (see the top two graphs of Fig. 2). However, this difference was significant only in two of the six cases. In these two cases, condition OM yielded larger errors in distance estimates than condition SM, in terms of an underestimation. Under the same motion condition (see the two

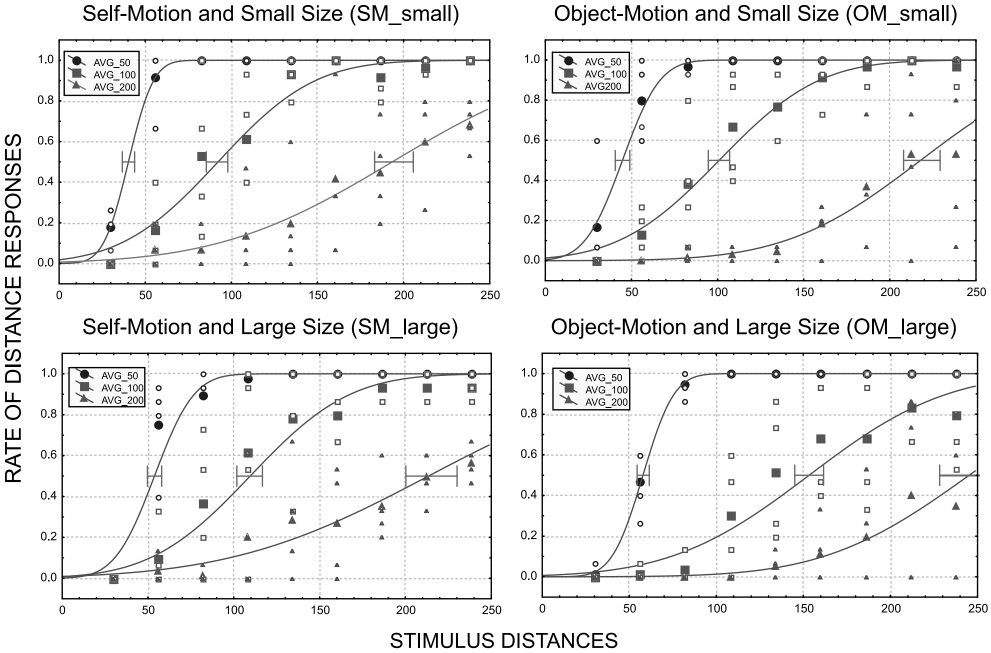
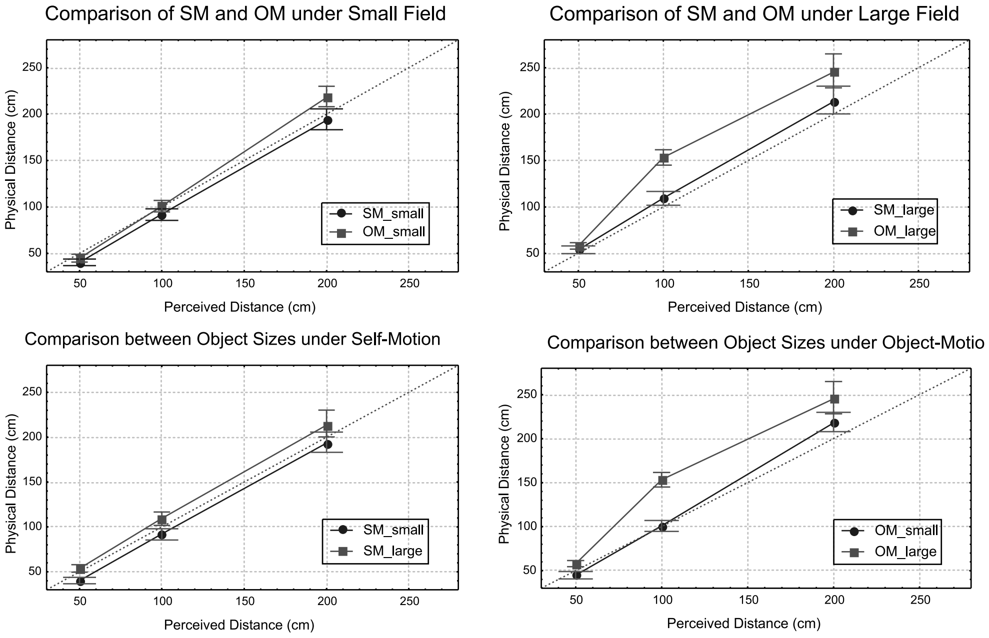


Fig. 1. Psychometric curves for four different categories of 2 stimuli, 2 conditions (open marker: individual data; closed marker: averaged data; ordinates: rate of distance responses; abscissae: stimulus distance).


Comparison of SM and OM under Small Field

Comparison of SM and OM under Large Field

Comparison between Object Sizes under Self-Motion

Comparison between Object Sizes under Object-Motio

Fig. 2. The comparisons of thresholds under different conditions.

bottom plots of Fig. 2), object of large apparent size tended to be perceived as nearer than that of small apparent size. However, the influence by the size was not significant for two of the six cases. In these two non-significant cases, the object was situated at the furthest distance.

2.4. Discussion

The steeper psychometric curves for nearer distance perceptual level suggest that subjects are better at judging near distances. The observed drop in subjects' confidence and ability to estimate far object distances has been compatible with their verbal reports. Indeed, all subjects claimed that they perceived the sphere to be closer to them when the diverging flow of the moving dots was strong. Conversely, they faced greater difficulties in estimating distance when this diverging motion was weak.

An extension to the computational model developed in (Panerai et al., (in press)) is presented here to account for the case of sagittal HM. This model expresses the estimated distance (D_{est}) as a function of the stimulus distance (D), head motion amplitude (HMA, T_Z), the angular size of the object (ρ) and the error terms, ε_1 and ε_2 associated with 3D translation and 2D velocity estimation respectively. We document the elaborated derivation of this model in the Appendix A and quote only the final expressions here:

$$D_{\text{est}} \cong D + \varepsilon_1 \frac{D}{T_Z} - \frac{\varepsilon_2}{\rho} \frac{D^2}{T_Z^2} \tag{1}$$

$$\sigma(D_{\text{est}}) \cong |\varepsilon_1| \left(\frac{D}{T_Z} \right) + \frac{|\varepsilon_2|}{\rho} \left(\frac{D^2}{T_Z^2} \right) \tag{2}$$

The model predicts that for a fixed HMA, the error in distance estimation increases linearly with distance when ε_1 is much larger than ε_2/ρ (i.e. $\rho(\varepsilon_1/\varepsilon_2) \gg 1$). On the contrary, if the reverse is true (i.e. ε_1 is much smaller than ε_2/ρ), then the distance error increases quadratically with distance.

In Eq. (2), only the intra-subject variability is considered in its derivation. Since CLI describes both the inter- and intra-subject variability found in our experimental data, a component σ_k, independent of experimental parameters and accounting for any inter-individual variability, has to be included before fitting CLI into the model. That is,

$$\text{CLI} = \sigma(D_{\text{est}}) + \sigma_k$$

Table 1 shows the coefficients obtained using a non-linear fitting of CLI with distances for the different categories of data. A 100% of variance was accounted for in the fitting of each data set. As predicted by the

Table 1
Non-linear estimation of CLI with distance

Category	Linear coeff.	Quadratic coeff.	$\lvert\sigma_k\rvert$
SM_small	0.11	0	1.58
OM_small	0.06	0	5.16
SM_large	0.12	0	2.24
OM_large	0.18	0	2.29

Second column shows the linear coefficient, third column shows the coefficient in the quadratic term and fourth column shows the constant term.

model, results in Section 2 showed a linear increase in CLI with distance for a fixed apparent size. As the linear coefficient was closely related to the error in the estimation of 3D translation, we observed that this parameter was a constant for the two SM-related categories (i.e. SM_small and SM_large). This revealed that subjects were consistent in their estimates in 3D translation, and hence leading to the same error characteristics, independent of the apparent size conditions.

From $|\sigma_k|$ in Table 1 which accounts for any inter-individual variability, we found that smaller measures had been obtained for the SM-related categories. This suggested that SM could lead to a better calibration of distance or resulting in a more consistent distance estimates across all subjects. This inter-individual variability, however, increases drastically in OM condition for a small apparent size object indicating a larger inconsistency in distance estimation with this object.

As the model predicts both an overestimation and underestimation of distance depending on the sign of ε_1 and ε_2, it is difficult for us to have a conclusive inference from the perceived distance. Nevertheless, if we assume the error committed in the 3D motion to be much larger than the error induced by 2D visual motion, the model described adequately the trend of variability found in our psychophysical experiments. Strictly speaking, the last quadratic terms of Eqs. (1) and (2) should not be neglected if ε_1 is sufficiently small. This is seldom the case. In all cases, the contribution by the last quadratic term is always there except that the effect was not found in our experimental results.

The comparison between OM and SM reveals an underestimation of distance in the OM condition. This is by no means surprising and can be linked to the underestimation of time-to-collision repeatedly reported in several psychophysical experiments (Gray & Regan, 1999b; Schiff & Detwiler, 1979; Steeves et al., 2000). Subjects tend to give a more conservative estimate of an approaching object to avoid any potential head-on collision. Similarly, a larger distance allowance is given to the oscillating sphere to prevent any collision. However, when the subject has full control over his own motion as in the SM condition, a less underestimation or even an overestimation of distance is expected. This prediction tallied with our experimental data.

We were surprised to find such a small difference between the SM and OM conditions and the strong correlation between response and distance in OM condition. The subject had in principle no cue to absolute distance in OM condition. Indeed, the diverging and converging cue received in the OM condition alone could not serve to help subjects differentiate the 3D distance and in-depth translation. One possible interpretation is that the subject could have memorized some parameters of the head movement and used this information to scale for distance in OM condition. However, this could not explain why they performed worst in the OM condition as compared to the SM condition for a given distance especially at larger distances. Another possibility was that they learnt to make association with visual motion and the perceived distance (e.g. small apparent dot velocity as large distance). If they did succeed in linking such association, the main reason why they performed worst in OM condition was the lack of non-visual information. Hence, it was hypothesized that in the absence of SM information, optic flow alone would not enable the subject to estimate distance. In order to test this hypothesis, we conducted a second experiment by varying the head movement amplitude between trials to remove any effects of head movement learning by the subject. This also served to vary the flow divergence of an object at a fixed distance in the OM condition (i.e. generating different optic flow divergences for stimulus at each distance).

3. Experiment 2

The purpose of Section 3 was to dissociate the amplitude of head movement from the amount of expansion and contraction experienced. In the SM trials, subjects performed head movement of three different amplitudes while determining objects positioned at one of three absolute distances.

We recruited nine normal subjects for this experiment. Two of the subjects were from Section 2. The apparatus setup and type of stimulus were similar to that of Section 2. However, the head movement amplitude in the SM trials was randomly varied between three values.

3.1. Design

The subjects were asked to perform sagittal head translations of three amplitudes (HMA): 5, 10 and 15 cm, in random order. The stimulus distance was randomly chosen to be at 50, 100 or 150 cm. The apparent size of the sphere was fixed at 15°.

For each HMA and viewing distance, subjects performed 10 trials. Hence, we collected a total of 180 trials: 90 for SM and 90 for OM condition. These trials were randomly divided into two sessions of 45 trials for each motion condition. The varying parameters in each session were the HMA and stimulus distance.

3.2. Procedure

The subject moved his head along the sagittal axis in one of the three HMA while maintaining a frequency of 0.3 Hz. Before presenting the stimulus in each trial, the subject was instructed by a message on the screen on the HMA to be executed. If the motion frequency of the subject were not kept at a tolerable level of ±0.05 Hz, the trial would be discarded and the subject would have to repeat it. Before starting the experiment, the subject was given a training session to familiarize him on making the correct HMA. The mean and standard deviation of the RMS head movement recorded in the experiment for the 5, 10 and 15 cm HMA were 4.8 ($\sigma = 2.25$), 9.4 ($\sigma = 4.18$) and 13.6 ($\sigma = 6.08$) cm respectively.

In the SM condition, the subject was asked to estimate his distance from the surface of a sphere while performing sagittal head movement. After three periods of oscillations, the subject, with the help of a joystick, indicated if this distance was smaller or larger than 100 cm. After a session on the SM condition, the subject was given some rest before proceeding to a session in the OM condition. During the OM session, the subject remained stationary while observing a sphere oscillating back and forth along the sagittal axis in the virtual 3D space. As in Section 2, the 3D motion of the sphere was the exact replay of the subject's motion in the SM session.

3.3. Data analysis

We computed the responses with "distance larger than 100 cm" for each distance under each motion condition. The experimental data were arranged and analyzed from three different perspectives:

(1) We analyzed all the trials together and evaluated the effects of each independent factor using the standard MANOVA test.
(2) We selected only trials that shared similar divergence (i.e. identical ratio between HMA and object distance) from the entire set of data and grouped them according to their motion conditions. These constant divergence trials exhibited similar visual input in both SM and OM conditions. In this case, the only difference between the SM and OM condition within this group was related to the non-visual input.
(3) We chose only trials with stimulus distance at 100 cm from the entire data set and grouped them according to their motion conditions. In effect, we analyzed the dependency of the subjects' response on HMA with these constant distance trials.

3.4. Results

3.4.1. All trials

In this experiment, there were three independent factors: motion (SM, OM), HMA (5, 10, 15 cm) and stimulus distances (50, 100, 150 cm). The effects of these three independent factors were reported in Table 2. The MANOVA test showed significant differences in responses for all of the three factors.

Table 2
Results of MANOVA with entire set of data

Condition	Df	*F*	*p*
SM, OM	(1, 8)	29.28	0.0006
HMA	(2, 16)	9.26	0.0021
Distance	(2, 16)	51.47	<0.0001

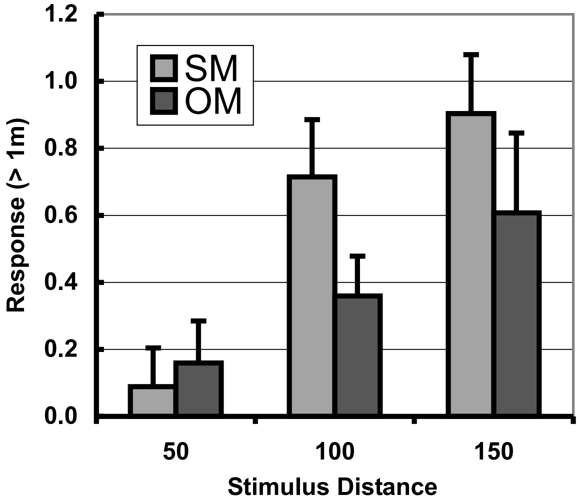


Fig. 3. Comparison of response between SM and OM with entire set of data.

Fig. 3 shows the comparison of responses for condition SM and OM. In both conditions, the responses correlated well with object distance (Spearman correlation $R = 0.842$; $p < 0.05$ and $R = 0.553$; $p < 0.05$ for SM and OM respectively). In the OM condition, it was observed that there was a strong underestimation of distances for large object distances (100 and 150 cm).

3.4.2. Constant divergence trials

Fig. 4 shows a significant difference in response for SM and OM conditions (ANOVA, $F_{1,8} = 12.71$, $p < 0.05$) when trials with a constant divergence (i.e. constant ratio between stimulus distance and HMA of 0.1) were being selected. The ability to discriminate distances was still observable in the SM condition. In other words, the response correlates strongly with increasing distance (Spearman correlation: $R = 0.87$, $p < 0.05$). Comparatively, the distribution of response was flat in the OM condition (Spearman correlation: $R = 0.05$, N.S.). Therefore, the distance judgments were not affected by the constant divergence for SM, as responses co-varied in the same way with distance on Figs. 3 and 4. On the contrary, when the same visual information (or optic flow with constant divergence) was presented to the subjects in condition OM, they gave a constant distance estimates regardless of the stimulus distance.

A further analysis shown in Fig. 5 indicated that in the OM condition, subject responses depended strictly

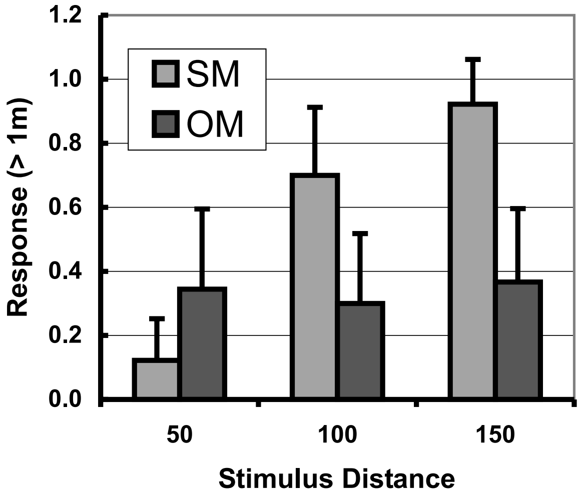


Fig. 4. Comparison of response between SM and OM under constant divergence.

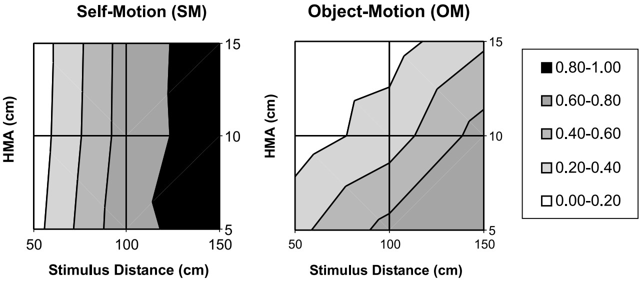


Fig. 5. The variations of response with distance and HMA.

on the pure visual information (i.e. the divergence). In each plot of Fig. 5, we monitored the changes in response (tagged with a different gray intensity) as we simultaneously varied the two independent variables: stimulus distance and HMA. The surface was obtained through the interpolation of the available experimental data along the two axes. While a line that cut the plots horizontally at a constant amplitude level determined the relationship between the response and the stimulus distance for that particular HMA level, a vertical line checked the dependency of the response on HMA when the stimulus was at a fixed distance. As divergence was proportional to the ratio between HMA and stimulus distance, a family of lines that was inclined at 45° to the horizontal axis corresponded to the iso-divergence lines (lines along which divergence remains constant).

The vertical bands in Fig. 5 of the SM condition show that the percentages of the response increase with increasing stimulus distance but remain relatively constant with the changing HMA (i.e. variations of response in the horizontal direction but not in the vertical direction). This effect is not observed in the OM condition. Instead, the response bands in the OM plot are in parallel to the iso-divergence contours (inclined at an angle near 45°). The tendency for distance underestimation in the OM condition is also evidenced in the plot whereby no response in the 0.8–1 range could be identified (black areas).

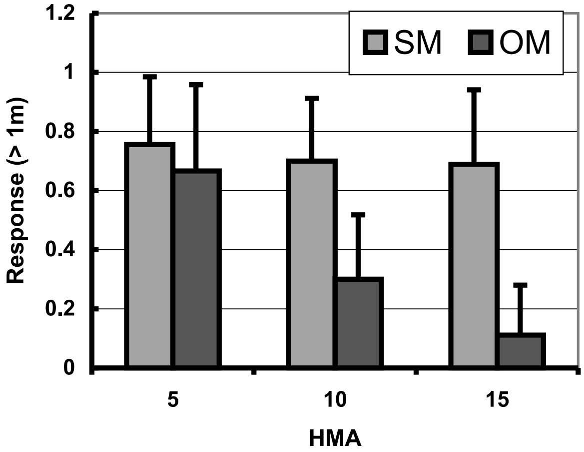


Fig. 6. Responses for object at 100 cm.

3.4.3. Constant distance trials (at 100 cm)

When the object was placed at 100 cm, the subjects' responses did not co-vary with the HMA in the SM condition (Spearman $R = -0.12$, N.S.). Instead, it hovered at an almost constant ($F_{2,24} = 0.21$; $p = 0.8093$) overestimated value of mean $= 0.72 \pm 0.23$ regardless of the change in HMA (see Fig. 6). In comparison, the response decreases with increasing head movement amplitude (Spearman $R = -0.71$, $p < 0.05$) in the OM condition, confirming that perceived distance decreased with divergence under this condition. Recall that divergence increases with HMA for constant distance.

The almost constant response in SM condition for all HMA when the object was at 100 cm showed that the overestimation of object distance was not dependent on the HMA. Even when the diverging flow was significantly strong (i.e. HMA = 15 cm), the subjects did not perceive the object to be of any nearer. In fact, they perceived it to be of the same distance as any other object that had a relatively weaker diverging flow but was placed at the same distance.

Our results confirmed the tendency for one to underestimate object distance in OM condition. This observation was also reported in Section 2. However, the slight underestimation of distance in SM condition for object at 100 cm in Section 2 was not revealed in Section 3. There was, in fact, an overestimation of distance in the second experiment. One explanation for this discrepancy was the use of only one HMA in Section 2 that has led to a better scaling of space with a minimal error.

3.5. Discussion

In Section 3, we introduced three different HMA and reduced the number of simulated distances to only three.

With such a modified design, we tested the hypothesis that subjects were able to judge egocentric distance by weighting (i.e. scaling), on a trial by trial basis, the visual cues produced by their voluntary displacement with the signals related to momentary SM. The prediction was that the subject in the SM condition would be able to discriminate the three simulated distances (i.e. 50, 100 and 150 cm). On the contrary, in the OM condition and without any prior knowledge on the type of movement, the subject would fail to recover the object distance.

The only possible way in the OM condition would be to judge an object to be near if its 2D flow was large, since a near object would naturally be perceived as moving faster and hence resulted in a stronger optical flow field. This makes sense ecologically since no other motion information was available then. The fact that subjects, in the OM condition, failed to recover the object distance was unanimously depicted in Fig. 4 where chance level performance were observed for the three simulated distances. Moreover, from Fig. 5 (right-hand plot), it is clear that the behavior of responses (i.e. the estimated distance) in the OM condition corresponded to a criterion of constant divergence. This means that similar judgment of distance is performed for trials presenting a constant ratio between simulated distance and HMA.

The divergence of flow was in fact proportional to the average 2D velocity of the moving dots. The use of average 2D velocity to perform psychophysical tasks in OM condition is not uncommon, and has been reported in at least two studies (Cornilleau-Pérès, Wong, Cheong, & Droulez, 2000; Domini & Caudek, 1999). In a recent work studying the effects of orthographic and perspective projection on slant perception, Cornilleau-Pérès et al. (2000) noted the dependency of perceived slant on the 2D velocity of the visual stimulus. Also, despite the presence of slant information in large visual field condition, this dependency was as significant as that for small visual field.

Although the divergence was also found under the SM condition, this information was not exclusively used in the estimation task. Despite the constant diverging flow across trials (i.e. constant ratio between distance and HMA), the subject was still capable of making satisfactory distance discrimination. This suggested that some additional non-visual cues, apart from the visual information, were exclusively available under the SM condition to enable such discrimination.

4. General discussion

The primary visual cue for estimating distances in both SM and OM conditions stemmed from the diverging flow of the moving dots. Since the flow field in the OM session was essentially a replayed version of that experienced in SM session, the only difference was whether the flow was actively generated or passively observed. When the flow was actively generated, subjects received extra-retinal signals from several non-visual sources (e.g. efference copies of motor commands, proprioceptive and vestibular systems) and coupled these signals together with the optic flow information to scale distances. In contrast, when the flow was passively experienced, subjects had no distance cues other than the visual ones.

Results in Section 2 suggest a rather good estimation of absolute distance in both the SM and OM conditions. A strong correlation was found between the perceptual distance and the physical distance, in accordance with several recent studies (Bingham & Pagano, 1998; Pagano & Bingham, 1998; Wickelgren et al., 2000). Our result shows that a naïve subject can scale distance ranges relatively well with his internal representation of distance. As mentioned earlier, the subject has various motion cues available to help him estimates the absolute distance from the diverging flow. On the other hand, these cues were absent in OM condition. Therefore, it was indeed a surprise to find such a good performance in the OM condition.

The findings can be explained by the ability to correlate the observed optic flow information with the constant HM parameters. In order to test this hypothesis, a second experiment was conceived based on the idea of randomizing the amplitudes of head movement across trials. The results show that in the OM condition, subjects did not perform as well as in Section 2. In the OM condition, a strong dependency of response on flow divergence was found, but none was observed in the SM condition. This supports the conclusion that in the SM condition, non-visual cues were present and effectively used for scaling distances. Note that when we compared the subsets of results in Sections 2 and 3, no effect of memorization or *a priori* learning of head parameters (Fig. 7) was evidenced. In the OM condition of the two

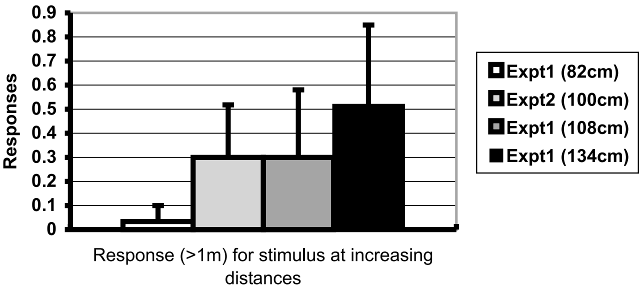


Fig. 7. Comparison of subjects' response (>1 m) for OM condition between Sections 2 and 3. The angular size of stimulus for both experiments was 15°. In Section 2, four subjects were considered and the stimulus was at 82 cm, 108 cm and 134 cm respectively (first, second and the last bars). In Section 3, nine subjects were considered and the stimulus was at 100 cm (fourth bar). No significant difference in response was found for responses at 100 cm and 108 cm (the middle two bars).

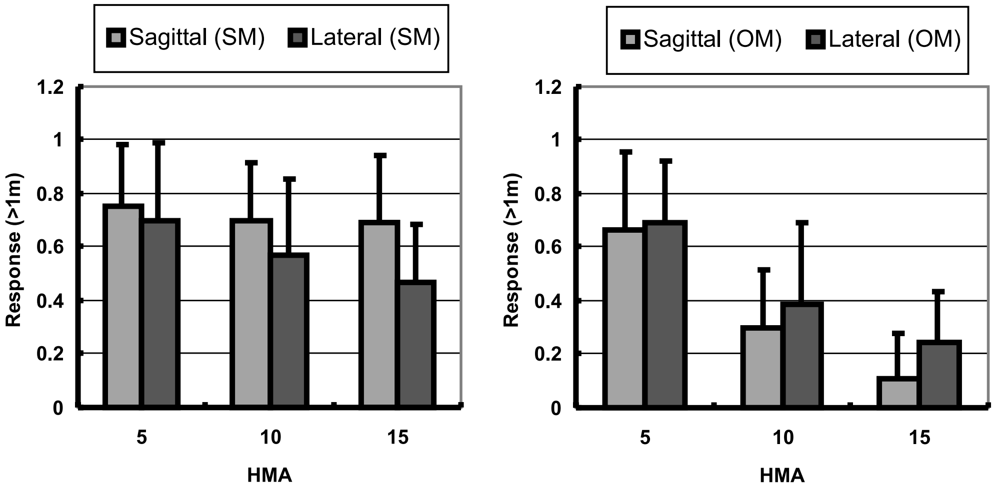


Fig. 8. Comparison of responses in sagittal and lateral experiments for object at 100 cm.

experiments, no significant difference was found in the subjects' response at 108 cm and 100 cm respectively. This suggests that although the learning of HM parameters in the SM condition is probable, it has no effect on the subjects' response in Section 2 since subjects in Section 3 had no *a priori* HM information and yet similar responses were obtained. Thus, it was likely that the subject classified objects into the four distance ranges based on the relative strength or divergence of the optic flow in Section 2. This reliance on divergence was also clearly demonstrated in Section 3 where objects positioned at the same distance but producing different divergence (by changing HMA) were not correctly identified for the OM condition.

The importance of dissociating the visual information from HM parameters has often been overlooked in psychophysical experiments that compare SM with passive observation of moving object. For example, in quantifying the effects of lateral head movement on curved quadratic surface percept, van Damme and van de Grind (1996) used a constant amplitude and frequency throughout HM condition. This motion was then mimicked in the non head motion (NHM) condition. Only two out of three subjects showed significant differences in their results between HM and NHM. Though not explicitly reported, there seems to be a strong correlation between the two motion conditions. The good results in the NHM condition may be a direct consequence of the subjects' ability to correlate the visual information with the constant motion parameters. Hence, it is more appropriate or conclusive to assess the performance between the two motion conditions by decorrelating the visual information from the HM parameters. Only the use of such methodological approach can determine correctly the contributions and effects of non-visual information during SM. We argue that in general, any psychophysical study that compares SM and OM condition with respect to tasks pertaining to absolute attributes or tasks allowing subjects to group responses (e.g. forced choice experiments) needs to decorrelate the visual and the non-visual information to better assess the effect of SM.

An interesting outcome is obtained when we make a comparison of our results with that obtained in a lateral head translation experiment (Panerai et al., 2002). Fig. 8 compares the responses for object at 100 cm in the two experiments. Both studies employed nine subjects and essentially adopted the same methodology in gathering responses. As the subjects were forced to make a choice between "more than 100 cm" and "less than 100 cm", the response should be at around chance level for object situated at 100 cm.

In the SM condition of the sagittal experiment (Fig. 8, left), the responses show that the optic flow information was well scaled for reporting absolute distances as reflected by an almost constant level for all HMA. This indicates that subjects did not depend solely on visual information to estimate distance. The strong optic flow divergence elicited by large HMA was not perceived to be near. In fact, there was an overall overestimation of distances regardless of the strength of divergence. On the contrary, a downward trend with increasing HMA (or 2D optic flow velocity) is observed for the lateral HM experiment. While the estimation of absolute distance tended to be influenced by the 2D optic flow, this effect was less felt by subjects in sagittal motion than in lateral translation.

In OM condition of the sagittal experiment, responses changed significantly with the HMA (i.e. absolute 2D velocity). For the lateral experiment, the rate of change of response with HMA, though present, was a less drastic one. One could explain the significant difference in responses for the two experiments in terms of the average dot velocity experienced. To verify this claim, we computed the average velocity of 200 dots resulting from a sagittal and a lateral translation while fixating the center of a sphere. We modeled a 3D sphere at a distance of 100 cm with 200 random dots scattering over its 15° viewing surface. These parameters corresponded to the settings of Section 3. The angular velocities of the flow vectors projected onto the retina with a translation of 10 cm/s in both sagittal and lateral directions at the initial viewing position were derived. We observed that a great reduction in the flow velocity actually resulted from the gaze stabilization during lateral translation. The average angular flow velocities computed for sagittal and lateral translation were respectively 1.09 and 0.57 rad/s. This corresponded to a ratio of 1.91. Hence, the average angular flow in sagittal translation was, in general, significantly higher than that in lateral translation. This explained why subjects responded more often to the "less than 100 cm" category in the OM condition of sagittal experiment. The larger difference in responses found for larger flows (i.e. at HMA = 15 cm) might be explained by the psychological effect that a shorter distance estimate is usually assigned to a more rapidly approaching object. Since in the case of lateral translation, the object was not approaching the subjects, no such phenomenon was found.

The better scaling of optic flow with SM in sagittal experiment may be closely related to the way we perform our daily activities in the 3D physical space. When we walk, run or maneuver, it is often along the sagittal axis that the motion is carried out. Hence, the brain might perform better correlation of this type of movement with the corresponding visual information experienced. Furthermore, eye movement is usually involved in lateral head translation for stabilization purposes. This might induce an error that varies with the changing HMA, which is not found in sagittal HM. In a recent study, Wickelgren et al. (2000) reported that monocular reaching performed with haptic feedback and pure optic flow cues to distance, was better during sagittal than lateral HM. When haptic feedback was removed, performance deteriorated greater for lateral head movement. Our results, together with theirs, provide strong evidence to support the conclusion that sagittal head movement might lead to a more calibrated perception of egocentric distance than lateral head translation.

5. Conclusion

As we perform SM, we may acquire motion information from several sources. This information can then be used to scale for absolute distance from the optical flow information. In the absence of SM information, the subjects could only rely on the optic flow and their responses co-vary with the image velocity. When the 2D flow was large, the subject perceived the object to be near; when the 2D flow was small, the subject perceived it to be far. This corresponded positively to the perception that near object moves faster than a far object. The reliance on 2D velocity to judge distance was nevertheless not observed in SM. This shows that distance estimation from the 2D optic flow requires the integration of non-visual information for scaling.

When we compare our results with that obtained in lateral head translation experiment, we found that estimation of absolute distances was less dependent on HMA when subject moved along the sagittal axis. This suggested that a better scaling of optic flow by the non-visual cues had been found in sagittal HM as compared to lateral HM. This better coordination can be closely associated with the movement we performed in our daily activities.

While it has been established in our findings that absolute distance perception can possibly be achieved with sagittal head movement, it remains to show how the various mechanisms from the multiple sensory systems interact and work together to allow the perception of absolute distance.

Acknowledgements

The major part of this work has been carried out during the first author's stay at LPPA, CNRS-Collège de France.

Appendix A

We develop a simple model to describe the relationship between estimated distance D_{est} and real distance D as a function of 3D motion estimation errors and the error found in 2D visual motion. If we consider only 3D translation components, from Longuet-Higgins and Prazdny (1980) we get:

$$\vec{U}(x,y) = \frac{1}{D}[T_X - xT_Z, T_Y - yT_Z] \tag{A.1}$$

where $\vec{U}(x,y)$ is the optical flow vector at retinal point (x, y) and (T_X, T_Y, T_Z) is the 3D translation with respect to the 3D world coordinate (X,Y,Z).

If we further assume that the movement of the subject is restricted to be along the sagittal plane and there is no error associated with T_X and T_Y (i.e. $T_X = T_Y = 0$), we get

$$\vec{U}(x,y) = \frac{T_Z}{D}\rho\vec{r} \tag{A.2}$$

where $\rho = |\sqrt{x^2 + y^2}|$ and $\vec{r}$ is the unit radial vector from the retinal center. Here, we assume that there is no directional error corresponding to the optic flow. This assumption is valid when we consider the errors in measuring the horizontal and vertical component of the optic flow are identical.

Now, if the brain estimates the amplitude of the movement T_Z with an intrinsic error ε_1 and measures the retinal motion velocity with an intrinsic error ε_2, by considering only the optic flow magnitude, we may rewrite Eq. (A.2) as

$$\begin{aligned} D_{\mathrm{est}} &= \frac{[T_Z + \varepsilon_1]}{|\vec{U}| + \varepsilon_2}\rho|\vec{r}| \\ &= \frac{[T_Z + \varepsilon_1]}{\left[\frac{T_Z}{D}\rho|\vec{r}| + \varepsilon_2\right]}\rho|\vec{r}| \\ &= \frac{[T_Z + \varepsilon_1]\rho}{\left[\frac{T_Z\rho}{D}\left(\left[1 + \varepsilon_2\frac{D}{T_Z\rho}\right]\right)\right]} \end{aligned}$$

since $|\vec{r}| = 1$.

If $\varepsilon_2 D/T_Z\rho = \varepsilon_2/|\vec{U}|$ is much smaller than 1, the following approximation holds true:

$$\begin{aligned} D_{\mathrm{est}} &\cong \frac{D[T_Z + \varepsilon_1][1 - \varepsilon_2 D/T_Z\rho]}{T_Z} \\ &= D\left[1 + \frac{\varepsilon_1}{T_Z}\right]\left[1 - \frac{\varepsilon_2 D}{T_Z\rho}\right] \\ &= D + \varepsilon_1\frac{D}{T_Z} - \frac{\varepsilon_2}{\rho}\frac{D^2}{T_Z^2} \end{aligned}$$

Hence, the variance of the first order approximation is

$$\mathrm{var}(D_{\mathrm{est}}) \cong |\varepsilon_1|^2\left(\frac{D}{T_Z}\right)^2 + \frac{|\varepsilon_2|^2}{\rho^2}\left(\frac{D^2}{T_Z^2}\right)^2$$

$$\sigma(D_{\mathrm{est}}) \cong |\varepsilon_1|\left(\frac{D}{T_Z}\right) + \frac{|\varepsilon_2|}{\rho}\left(\frac{D^2}{T_Z^2}\right)$$

References

Bingham, G. P., & Stassen, M. G. (1994). Monocular egocentric distance information generated by head movement. *Ecological Psychology, 6*(3), 219–238.

Bingham, G. P., & Pagano, C. C. (1998). The necessity of a perception-action approach to definite distance perception: Monocular distance perception to guide reaching. *Journal of Experimental Psychology: Human Perception & Performance, 24*(1), 145–168.

Collett, T. (1978). Peering: a locust behavior pattern for obtaining motion parallax information. *Journal of Experimental Biology, 76*, 237–241.

Cornilleau-Pérès, V., Wong, T. K., Cheong, L. F., & Droulez, J. (2000). Visual perception of slant from optic flow under orthographic and perspective projection. *Investigative Ophthalmology and Visual Science, 41*(4), S718 (Abstract).

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein, & S. Rogers (Eds.), *Perception of Space and Motion* (pp. 69–117). San Diego, CA: Academic Press, Chap. 3.

Dees, J. W. (1966). Accuracy of absolute visual distance and size estimation in space as a function of stereopsis and motion parallax. *Journal of Experimental Psychology, 72*(3), 466–476.

Domini, F., & Caudek, C. (1999). Perceiving surface slant from deformation of optic flow. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 426–444.

Ellard, C. G., Goodale, M. A., & Timney, B. (1984). Distance estimation in the mongolian gerbil: The role of dynamic depth cues. *Behavioral Brain Research, 14*, 29–39.

Eriksson, E. S. (1974). Motion parallax during locomotion. *Perception & Psychophysics, 16*, 197–200.

Ferris, S. H. (1972). Motion parallax and absolute distance. *Journal of Experimental Psychology, 95*(2), 258–263.

Fisher, S. K., & Ciuffreda, K. J. (1988). Accommodation and apparent distance. *Perception, 17*(5), 609–621.

Foster, D. H., & Walter, F. B. (1997). Bootstrap estimates of the statistical accuracy of thresholds obtained from psychometric functions. *Spatial Vision, 11*(1), 135–139.

Gogel, W. C. (1961). Convergence as a cue to absolute distance. *Journal of Psychology, 52*, 287–301.

Gogel, W. C. (1973). Absolute motion parallax and the specific distance tendency. *Perception and Psychophysics, 13*(2), 284–292.

Gonzalez, E. G., Steinbach, M. J., Ono, H., & Wolf, M. E. (1989). Depth perception in children enucleated at an early age. *Clinical Vision Sciences, 4*, 173–177.

Gray, R., & Regan, D. (1999a). Do monocular time to collision estimates necessarily involve perceived distance? *Perception, 28*, 1257–1264.

Gray, R., & Regan, D. (1999b). Adapting to expansion increases perceived time-to-collision. *Vision Research, 39*(21), 3602–3607.

Gray, R., & Regan, D. (2000). Simulated self-motion alters perceived time to collision. *Current Biology, 10*(10), 587–590.

Johansson, G. (1973). Monocular movement parallax and near-space perception. *Perception, 2*, 135–146.

Koenderink, J. J. (1986). Optic flow. *Vision Research, 26*, 161–180.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35*, 389–412.

Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B, 208*, 385–397.

Marotta, J. J., Perrot, T. S., Nicolle, D., & Goodale, M. A. (1995). Adapting to monocular vision—Grasping with one eye. *Experimental Brain Research, 104*, 107–114.

Mon-Williams, M., & Tresilian, J. R. (2000). Ordinal depth information from accommodation? *Ergonomics, 43*(3), 391–404.

Owens, D. A., & Liebowitz, H. W. (1980). Accommodation, convergence, and distance perception in low illumination. *American Journal of Optometry & Physiological Optics, 57*(9), 540–550.

Pagano, C. C., & Bingham, G. P. (1998). Comparing measures of monocular distance perception: Verbal and reaching errors are not correlated. *Journal of Experimental Psychology: Human Perception & Performance, 24*(4), 1037–1051.

Panerai, F., Hanneton, S., Droulez, J., & Cornilleau-Pérès, V. (1999). A 6-dof device to measure head movements in active vision experiments: geometric modeling and metric accuracy. *Journal of Neuroscience Methods, 90*, 97–106.

Panerai, F., Cornilleau-Pérès, V., & Droulez, J. (2002). Contribution of extra-retinal signals to the scaling of object distance during self-motion. *Perception* & *Psychophysics*, *64*(5), in press.

Poteser, M., & Kral, K. (1995). Visual distance discrimination between stationary target in praying mantis: An index of the use of motion parallax. *Journal of Experimental Biology, 198*, 2127–2137.

Russell, J. T. (1931). Depth discrimination in the rat. *Journal of Genetic Psychology, 40*, 136–159.

Schiff, W., & Detwiler, M. L. (1979). Information used in judging impending collision. *Perception, 8*, 647–658.

Steeves, J. K., Gray, R., Steinbach, M. J., & Regan, D. (2000). Accuracy of estimating time to collision using only monocular information in unilaterally enucleated observers and monocularly viewing normal controls. *Vision Research, 40*(27), 3783–3789.

Tresilian, J. R., Mon-Williams, M., & Kelly, B. M. (1999). Increasing confidence in vergence as a cue to distance. *Proceedings: Biological Sciences, 266*(1414), 39.

van Damme, W. J. M., & van de Grind, W. A. (1996). Non-visual information in structure from-motion. *Vision Research, 36*(19), 3119–3127.

Viguier, A., Clement, G., & Trotter, Y. (2001). Distance perception within near visual space. *Perception, 30*(1), 115–124.

Wallace, G. K. (1959). Visual scanning in the desert locust *Schistocerca Gregaria Forskål*. *Journal of Experimental Biology, 36*, 512–525.

Wickelgren, E. A., McConnell, D. S., & Bingham, G. P. (2000). Reaching measures of monocular distance perception: Forward versus side-to-side head movements and haptic feedback. *Perception & Psychophysics, 62*(5), 1051–1059.