A model of human heading judgement in forward motion

MITSUHIKO HANADA, YOSHIMICHI EJIMA

Graduate School of Human and Environmental Studies

Kyoto University

Yoshida-nihonmatsu-cho, Sakyo-ku, Kyoto 606-8501, Japan

Hanada, M., & Ejima, Y. (2000). A model of human heading judgement in forward motion. Vision Research, 40 (2), 243-263.

Running Title: Modeling of human heading judgement

Key Words: Optic flow, Motion, Heading, Modeling

ABSTRACT

We developed a new computational model of human heading judgement from retinal flow. The model uses two assumptions: a large number of sampling points in the flow field and a symmetric sampling region around the origin. The algorithm estimates self-rotation parameters by calculating statistics whose expectations correspond to the rotation parameters. After the rotational components are removed from the retinal flow, the heading direction is recovered from the flow field. Performance of the model was compared with human data in three psychophysical experiments. In the first experiment, we generated stimuli which simulated self-motion toward the ground, a cloud or a frontoparallel plane and found that the simulation results of the model were consistent with human performance. In the second and third experiment, we measured the slope of the perceived vs. simulated heading function when a perturbation velocity weighted according to the distance relative to the fixation distance was added to the vertical velocity component under the cloud condition. It was found that as the magnitude of the perturbation was increased, the slope of the function increased. The characteristics observed in the experiments can be explained well by the proposed model.

Introduction

Motion pattern is induced on the retina by self-motion. The motion in the visual image is a source of information about the structure of the environment and about the way the observer is moving through. It is important for navigation tasks (e.g., walking, driving, and so on) to know the heading direction of the moving observer. Human observers can achieve an accuracy of about 1-2 deg in judging their heading direction from optical flow when rotational components due to self-rotation are not included in the retinal flow (Warren et al. , 1988).

Translation of an observer through a stationary environment without eye movements generates a radial pattern of optical flow, in which the focus of outflow specifies the heading direction. The focus no longer corresponds to heading when the flow pattern includes rotational components due to eye movements. Warren and Hannon (1988, 1990) examined whether human observers can decouple rotational and translational components of retinal flow. They compared performance of heading judgement under two conditions: (a) the observer tracked a moving point, introducing a rotational component of motion (“moving” condition). (b) the observer maintained stationary fixation while the display contained both translational and rotational components of motion (“simulated” condition). The same flow pattern appeared on the retina for conditions (a) and (b), though the rotation information could be derived from extraretinal sources in condition (a). For the case of movement toward a cloud of random dots or a ground plane, observers performed heading judgement accurately and there was essentially no difference in performance between the conditions. However, when simulating translation toward a frontal plane, the performance reached a high level of accuracy in condition (a), but was at chance in condition (b). This suggests that the decoupling of the rotational and translational components of motion only from visual input requires motion parallax produced by elements at different depths, and that extraretinal information is also used for the recovery of the heading direction.

Many algorithms for the recovery of heading from motion have been presented for computer vision (e.g., Bruss & Horn, 1983; Prazdny, 1980; Longuet-Higgins & Prazdny, 1981; Tsai & Huang, 1984; Weng, et al., 1989; Kanatani, 1993; Tomasi & Shi, 1993; Zhang, 1997; Nagel et al., 1997; Soatto & Perona,

1998) and for human computational modeling (e.g, Cutting, 1986; Heeger & Jepson, 1990; 1992; Rieger & Lawton, 1985; Hildreth, 1992). Some neural network models were also presented (e.g., Lappe & Rauschecker, 1993; Perrone, 1992; Perrone & Stone, 1994; Hatsopoulos & Warren, 1991; ; Warren & Saunders, 1995; Royden, 1997; Zemel & Sejnowski, 1998; Beintema & van den Berg, 1998) .

In this paper, we propose a new method to recover the heading, which is useful in our daily life or experimental situations in psychophysics. Our method recovers the heading direction by using the deviation from a radial flow pattern. The validity of the proposed method as a human model is tested in three psychophysical experiments. The experimental results show that the data in the experiment can be explained by our model.

Recovery of heading

We present a new algorithm to estimate heading parameters from retinal flow. The algorithm estimates self-rotation parameters calculating statistics whose expectations correspond to the rotation parameters when the velocities in the image are sampled in many positions randomly. After removing the rotational components from the flow field, the algorithm estimates the heading direction using the flow field without the rotational components.

We make use of essentially the same notation as Longuet-Higgins and Prazdny (1980). We use a coordinate system that is fixed with respect to an observer, with the Z-axis directed along the optical axis. The X-axis and Y-axis are horizontal and vertical respectively. The translation of the observer in the rigid environment can be expressed in terms of translation along three orthogonal directions, which we denote by the vector (U, V, W). U, V and W are the translation along the X-axis, Y-axis and Z-axis respectively. The rotation of the observer can be expressed in terms of rotation around three orthogonal axes, which we express by the vector (A, B, C). A, B and C are rotation around the X-axis, Y-axis and Z-axis, and are called pitch, yaw and roll respectively.

We use the equations by Longuet-Higgins and Prazdny (1980) to obtain the projected velocity of a point in 3-D space. The 3-D velocity of a point, P(X,Y,Z) is given by:

$$\dot{X} = -U - BZ + CY$$
$$\dot{Y} = -V - CX + AZ \qquad (1)$$
$$\dot{Z} = -W - AY + BX$$

If we consider perspective projection of the velocity onto the image plane, with a focal length for the projection of 1, the point P on the image (x,y) is given by:

$$x = \frac{X}{Z}$$
$$y = \frac{Y}{Z} \qquad (2)$$

The projected velocity (u,v) in the image plane is given by:

$$u = \frac{(-U + xW)}{Z} - B + Cy + Axy - Bx^2$$
$$v = \frac{(-V + yW)}{Z} - Cx + A + Ay^2 - Bxy \qquad (3)$$

The first term represents the component of image velocity due to translation of the observer and depends on the depth Z. The remaining terms represent the component of velocity due to rotation of the observer and do not depend on the depth Z.

Here we propose a new model of human heading judgement based on the following assumptions:

(a) Yaw(B) and pitch(A) are small.

(b) There is a large number of sampling points which are randomly chosen in a symmetrical region around the origin in the image.

(c) An observer pursues a static point smoothly.

(d) The fixation point is on a smooth surface, or the depth of the fixation point is the average of sampling points' depth.

Only when eye movement velocity is slow (less than 1.5 deg/sec) , heading judgement can be performed well by human observers using visual information alone (e.g., Royden, et al.; Banks, et al., 1996). Therefore, we can assume (a) for the model of human heading perception. As for assumption (b), one can

use a large number of sampling points in real life. Assumptions (c) and (d) are used temporarily, and we will introduce an extended algorithm later which does not use (c) and (d).

When an observer fixates and pursues a point $P_f(0, 0, Z_f)$, the velocity of P_f in the image plane is (0, 0). Therefore, the following equations hold :

$$\begin{aligned} U &= -BZ_f \\ V &= AZ_f \end{aligned} \qquad (4)$$

Such fixation constraints have been used for computer vision (Fermüller & Aloimonos, 1992; Taalebinezhaad, 1992; Radiv & Herman 1994; Lappe & Rauschecker, 1995; Soatto & Perona, 1998). Substituting U and V of Equation (1) by the equations above, we obtain:

$$\begin{aligned} u &= \frac{BZ_f}{Z} - B + x\frac{W}{Z} + Axy - Bx^2 + Cy \\ v &= -\frac{AZ_f}{Z} + A + y\frac{W}{Z} + Ay^2 - Bxy - Cx \end{aligned} \qquad (5)$$

Let (x_i,y_i), (u_i,v_i) and Z_i be the projected position, the velocity and the depth of the i- th sampling point respectively.

First we derive the following equation from (5) to estimate C:

$$u_i y_i - v_i x_i = \frac{BZ_f y_i}{Z_i} - By_i + Cy_i^2 + \frac{AZ_f x_i}{Z_i} - Ax_i + Cx_i^2 \qquad (6)$$

with the inner product of $(y_i, -x_i)$ and (u_i, v_i). $(y_i, -x_i)$ is a vector which is orthogonal to the radial direction, (x_i, y_i).

If many sampled points are randomly located in a symmetrical region, C can be estimated by the following equation from (6):

$$C_e = \frac{1}{N_c} \sum_{\substack{|x_i|>T_{cx} \\ or \\ |y_i|>T_{cy}}} \frac{u_i y_i - v_i x_i}{x_i^2 + y_i^2} \qquad (7)$$

where C_e is the estimation value of C, and N_c is the number of points which satisfy the condition, $|x_i|>T_{cx}$ or $|y_i|>T_{cy}$. Because the expectation of the terms except Cx_i^2 and Cy_i^2 in Eq. (6) is 0, C_e corresponds to C

according to a statistical theorem, the law of large numbers[1]. T_{cx} and T_{cy} are thresholds. They are introduced so that the estimate may not be easily upset by noise when we sample dots near the origin.

For the case of movement on the ground plane, this estimation is inappropriate because 1/Z is correlated with y and the expectation of $BZ_f y_i / Z_i$ in Equation (6) is not 0. The following equation is preferable:

$$C_e = \frac{1}{N_{cv}} \sum_{|x_i|>T_{cx}} \frac{-v_i}{x_i} \qquad (8)$$

where N_{cv} is the number of points which satisfy the condition, $|x_i|>T_{cx}$. The sampling region should be determined carefully because we cannot sample dots in the region above the horizon. The sampling region should be symmetric around the origin.

After the estimation of C, we remove the velocity components of C by redefining v_i as $v_i+C_e x_i$, and u_i as $u_i-C_e y_i$. Thus we can regard C as 0. Therefore, if x_i and y_i are not 0, we obtain the following equation from Eqs. (5):

$$\frac{v_i}{y_i} - \frac{u_i}{x_i} = \left(1 - \frac{Z_f}{Z_i}\right)\left(\frac{B}{x_i} + \frac{A}{y_i}\right) \qquad (9)$$

We rewrite the equation:

$$B = \frac{\frac{x_i}{y_i} v_i - u_i}{1 - \frac{Z_f}{Z_i}} - \frac{x_i}{y_i} A \qquad (10)$$

Z_f / Z_i is unknown. If |B| and |A| are small compared with W/Z, and the fixation point is set at the average depth of the dots, it is estimated as:

$$\frac{Z_f}{Z_i} = \frac{\frac{Z_f}{W}}{\frac{Z_i}{W}} \approx \frac{\tau_f}{\tau_i} \qquad (11)$$

[1] Z, x and y in Eq. (6) are regarded as random variables.

where

$$\tau_f = \frac{1}{N_z} \sum_{\substack{\sqrt{u_i^2+v_i^2}>T_{zf} \\ \sqrt{x_i^2+y_i^2}<T_c}} \tau_i \approx \frac{Z_f}{W} \qquad (12)$$

$$\tau_i = \frac{\sqrt{x_i^2+y_i^2}}{\sqrt{u_i^2+v_i^2}} = \sqrt{\frac{x_i^2+y_i^2}{(x_i^2+y_i^2)\frac{W}{Z_i}+O(A,B)}} \approx \frac{Z_i}{W} \qquad (13)$$

O(A,B) shows the terms including A or B and N_z is the number of points which satisfy the conditions:

$$\sqrt{u_i^2+v_i^2} > T_{zf} \text{ and } \sqrt{x_i^2+y_i^2} < T_c$$

The thresholds, T_{zf} and T_c are used to avoid using $x_i^2 + y_i^2$ around 0. Equation (12) means the eccentricity divided by the dot speed. If U and V are 0, Z/W represents the time to contact. Z/W can be regarded as an index of relative depth. The approximation of (12) is valid when the retinal flow are dominated by the translational components. When the fixation point is on a continuous surface, we should select points on the surface to estimate τ_f (=Z_f/W) from Eq. (13), and T_c in (13) should be small to use points near the fixation point because Z_f /W is near Z_i /W for points around the fixation point.

B can be estimated from Eq. (10) by the following:

$$B_e = \frac{1}{N_b} \sum_{\substack{|y_i|>T_{by} \\ T_{zl}<|1-\frac{\tau_f}{\tau_i}|<T_{zh}}} \frac{\frac{x_i}{y_i}v_i - u_i}{1-\frac{\tau_f}{\tau_i}} \qquad (14)$$

where N_b is the number of dots which satisfy the conditions, $|y_i|>T_{by}$ and $T_{zl}< |1 -\tau_f/\tau_i | < T_{zh}$. T_{by}, T_{zl} and T_{zh} are thresholds. As the expectation of x_iA/y_i in Eq. (10) is 0, B_e corresponds to B if many points are sampled according to the law of large numbers.

We rewrite Eq. (10) to estimate A as follows:

$$A = \frac{v_i - \frac{y_i}{x_i}u_i}{1-\frac{Z_f}{Z_i}} - \frac{y_i}{x_i}B \qquad (15)$$

From the equation, A can be estimated in the same way as the estimation of B because the expectation of y_iB/x_i in (15) is 0.

$$A_e = \frac{1}{N_a} \sum_{\substack{|x_i|>T_{ax} \\ T_{zl}<|1-\frac{\tau_f}{\tau_i}|<T_{zh}}} \frac{v_i - \frac{y_i}{x_i}u_i}{1-\frac{\tau_f}{\tau_i}} \qquad (16)$$

where N_a is the number of dots which satisfy the conditions, $|x_i|>T_{ax}$ and $T_{zl}< |1- \tau_f/\tau_i| < T_{zh}$. T_{ax} is a threshold. x_iv_i/y_i-u_i （$v_i-y_iu_i/x_i$）represents the degree of deviation from a radial pattern in the x (y)-axis direction (Fig. 1). Parameters B and A are recovered by averaging the values of the deviations weighted by $1- \tau_f/\tau_i$.

Insert Figure 1 about here

The absolute translation parameters cannot be recovered from the flow information alone. We can obtain only U/W and V/W, which represent the heading direction. We estimate their values using Eqs. (4) and (12) as:

$$U_e = -B_e\tau_f$$
$$V_e = A_e\tau_f \qquad (17)$$

where U_e and V_e are the estimation values of U/W and V/W respectively. To obtain better estimates of U/W and V/W, we can use the methods for pure translation after removing the rotational component from the retinal flow.

If an observer moves horizontally, V is about 0. If V is 0, U/W is given in the following way. First we replace A, C and V in Eq. (3) with 0. Then we obtain the following equation by eliminating Z using Eqs. (3):

$$U = -u_iW\alpha_i - BW\alpha_i + Wx_i - Bx_i^2W\alpha_i \qquad (18)$$

where,

$$\alpha_i = \frac{y_i}{v_i + Bx_i y_i} \qquad (19)$$

From (18):

$$\frac{U}{W} = -u_i\alpha_i + x_i - B\,\alpha_i(1 + x_i^2) \qquad (20)$$

Therefore U/W is estimated as:

$$\frac{U}{W} \approx U_e = \frac{1}{N_u} \sum_{|v_i + Bx_i y_i| > T_u} \left\{ -u_i\alpha_i + x_i - B_e\alpha_i(1 + x_i^2) \right\} \qquad (21)$$

where N_u is the number of dots which satisfy the condition, $|v_i+Bx_iy_i| > T_u$.

We repeat the same procedure after removing the rotational components of the estimated A and B, and setting the estimated (U/W, V/W) to the new origin in the image plane. We can obtain better estimates of translation and rotation parameters by iterating the procedure several times, though they are adequate without iteration when A and B are small. Setting a new origin interfere with the assumption that the points are symmetrically distributed around the origin. However, the shift of the origin is small for small A and B. If the new origin is in the image, we can address the problem by decreasing the size of the sampling region with the iterations.

The method has some limitations. First A and B cannot be estimated accurately if |B| and |A| are very large. However, this limitation is also observed for humans. Royden et al. (1994) reported that heading judgement could be performed well by human observers from visual input alone only when the eye movement velocity was 1 deg/sec or less. But van den Berg (1993; 1996) reported that human observers could perform heading judgement accurately by visual information when eye movement of 5 deg/sec was simulated. However, 5 deg/sec is rather small. Poor performance for stimuli that simulated

fast eye movement is also reported in other studies (Royden et al. 1992; Banks et al., 1996). These reports support the conclusion that human observers can not judge heading accurately with fast eye rotation. However, Stone & Perrone (1997) reported that experienced observers can perform the judgement of heading precisely even for rotation rates as high as 16 deg/sec, without extraretinal information when the simulated translation was fast. They suggested that the ratio of rotation to translation rates rather than the absolute rotation rate should limit precision[2]. We present the dependence of our model on the ratio in the simulations later.

Second, our method needs the value of Z_f / W. It can be easily obtained by Eqs. (12) and (13) if the fixation point is on a surface. Eq. (12) is essentially the same equation $\theta/(d\theta/dt)$ to calculate Z_f/W used by Regan and Kashal (1994) (θ is the eccentricity angle). If binocular disparities are available or if an observer moves toward the ground while looking toward the horizon, the calculation Z_f/W in Eq. (11) is not required because one can know the relative depth between dots and calculate the value of Z_f/Z_i in Eqs. (10) and (11) . We represent the ground plane as:

$$Y = \frac{\beta Z}{Z_f} - \beta$$
$$\beta = \frac{Z_f h}{\sqrt{Z_f{}^2 - h^2}} \qquad (22)$$

where h is the height of the observer's eye from the ground plane. The y-axis value of a point in the image plane is:

$$y \equiv \frac{Y}{Z} = \frac{\beta}{Z_f} - \frac{\beta}{Z} \qquad (23)$$

The y-axis value of the horizon is obtained by calculating the limit of y in (23) with infinite Z:

$$y_\infty = \frac{\beta}{Z_f} \qquad (24)$$

[2] If the translation and depth are multiplied by two, the same flow appears in a retina. The ratio is meaningless if different depth configurations are used.

Z_f/Z_i is given by:

$$\frac{Z_f}{Z_i} = \frac{y_i - y_\infty}{-y_\infty} \qquad (25)$$

Thus we can obtain relative depth by (25).

We assume random sampling in a region symmetrical about the origin. When a large number of dots are available, we can select a symmetrical region by limiting sampling points. The limitation due to the assumption is not so serious. The major limitation of the above algorithm is the gaze stability constraint. In the next section we extend the algorithm in case of gaze-unstabilized situations. In the extended algorithm, assumption (d) is not necessary either.

Gaze unstability algorithm

The model so far assumes conditions of opposite directions of rotation and the heading direction to the center of the image. This may not be the case for gaze non-stabilized condition. Then we propose to search a transformation of the image that will guarantee that such opposite directions occur after the transformation and proceed with the derivation as presented in the current text.

We assume that C is 0 (see Appendix A for the method to estimate C and to remove the components from the flow field in gaze-unstabilized situations). Let the center of flow be a point that minimizes the square sum of the distance (d in Fig. 2) between the point and the line which passes through the velocity flow vector.

Insert Figure 2 about here

We use a new coordinate system; The axes are transformed so that the center of the outflow is defined on the Z-axis. Then we define the depth of the center point (Z_f) as -U/B or V/A in the new coordinate system:

$$Z_f \equiv -\frac{U_{new}}{B_{new}}$$
$$or$$
$$Z_f \equiv \frac{V_{new}}{A_{new}} \qquad (26)$$

A_{new}, B_{new}, U_{new} and V_{new} are transformed A, B, U and V in the new coordinate system respectively. Comparing (26) with (4), we can think that at the moment the observer is tracking the center of outflow with the depth of Z_f in the new coordinates. We found that Z_f in (26) nearly equals the average depth of other sampling points. We present the proof in Appendix A.

It follows that if the center of outflow of point (x_c, y_c) is found, we can rotate the axes so that the following equation holds:

$$\frac{-U_{new}}{B_{new}} = Z_f \approx Z_0$$
$$\frac{-V_{new}}{A_{new}} = Z_f \approx Z_0 \qquad (27)$$

where Z_0 represents the average depth.

Comparing (27) with (4), one can say that the observer is tracking the center of outflow in the new coordinates, whose depth(Z_f) is near the average of other sampling points. It means that the gaze-stability algorithm can be used in gaze-unstabilized cases if the coordinates are transformed first. But we should change Eq. (13) by removing a threshold T_c as follows:

$$\tau_f = \frac{1}{N_z} \sum_{\sqrt{u_i^2+v_i^2} > T_{zf}} \tau_i \approx \frac{Z_f}{W} \qquad (28)$$

Perrone and Stone's model (1994) predicts better performance for stimuli with a direction of simulated eye rotation that is opposite to the retinal horizontal heading direction than for stimuli with identical directions of heading and simulated eye rotation. Their prediction does not correspond to human performance (Crowell, 1996). Crowell (1996) found that the above prediction resulted from the gaze stability constraint. Our gaze-unstability algorithm does not predict it because the coordinate transform is interpreted as the procedure that ensures that the direction of the eye rotation is opposite to the horizontal heading direction (see Eqs. (27)). This technique can be used for other models that use the gaze stability constraint.

Lappe and Rauschecker (1994) stated that "in situations where humans succeed in heading judgement, the retinal flow field is overall centrifugal in structure, whereas in situations where humans cannot correctly detect the direction of heading, it is not". We believe that the centrifugal structure contains the center of outflow and it is essential for human heading detection.

Simulations

We performed simulations to test the new model using three environments; a ground, a cloud and a frontoparallel plane. They were composed of discrete points whose image motion was determined by translation and rotation of an observer relative to a random dot surface or a cloud in space. We assumed that the observer gazed at a static point. The motions of the dots on the image plane were computed and these velocities formed the input for heading recovery.

The image subtended 40 deg horizontal × 30 deg vertical. We used 1000 dots. Noise was added to each dot. The horizontal and vertical components of the noise velocity were randomly set within 10% of the original dot speed. The horizontal and the vertical component was determined independently.

We focused on the horizontal movement of the observer, which means V=0 and A=0 because most psychophysical experiments have been conducted for the horizontal movement. We assumed that C (roll) was 0.

100 trials were conducted. We set the thresholds as follows: T_{ax}=0.05 [1/sec], T_{by}=0.05 [1/sec], T_{zl}=0.2 [1/sec], T_{zh}=5.0 [1/sec] and T_u=0.05 [1/sec], T_c=∞. In this simulation, we did not use the iterative procedure. We used Equations (20), but not (17) for the estimation of U/W and did not use Equation (25) . We did not use the gaze-unstability algorithm, but the results do not change significantly if we use the gaze-unstability algorithm because the center of outflow for the stimuli in the simulations is always the center of the display.

Ground plane

The following conditions were simulated here:

*Observer's translation: U was randomly set to a value between –0.25 m/sec and 0.25 m/sec, and W was set between 0.5 m/sec and 1.5 m/sec for each trial.

*Distance of fixation point: 9.5 m on the ground

*3-D structure: the observer's simulated eye height was 1.6 m and points covered a plane extending from 1 m to 18 m in front of the observer.

*Rotation parameters: A=0, |B|<1.5 deg/sec and C=0.

Figure 3(a) shows the results of the simulation of our model. The horizontal axis represents the simulated heading direction, and the vertical axis represents the direction estimated by our model. Each point denotes the result of each trial. If the points are scattered along a straight line with slope 1, heading judgement is unbiased. In the figure, the points lie along a line with slope about 1. We conducted a linear regression analysis. The slope of the fitting line is 1.08. The proposed method can estimate the direction of heading precisely in this condition.

Insert Figure 3 about here

Cloud

The following conditions were simulated:

*Observer's translation: U was set to a value between –0.25m/sec and 0.25m/sec, and W was set between 0.75 m/sec and 1.25 m/sec for each trial randomly .

* Distance of fixation point: 4 m

*3-D structure: Points were placed randomly within a depth range of 3 m – 5 m.

*Rotation parameters: A=0, |B|< 3.6 deg/sec and C=0.

Figure 3(b) shows the result of the simulation. The slope of the fitting line is 0.74. The eccentricity of the direction of heading was underestimated. In other words, the estimate was closer to the direction straight ahead relative to the simulated direction. The underestimation in the cloud condition may be ascribed to the smaller depth variation of scattered dots relative to the fixation point or to the difference in the scene configuration between cloud and ground.

We performed another simulation with more depth variations. Points were placed randomly within a depth range of 1 m – 7 m. In this condition, the heading direction was not underestimated. The underestimation in the cloud condition is ascribed to the smaller depth variation, but not to the scene configuration. The proposed model obtained fairly good estimates of heading in this condition.

Frontoparallel plane

The conditions were the same as in the cloud condition with one exception: The simulated plane and the simulated fixation point were both placed at a distance of 4 m away from the observer. The condition represented an extreme case in small depth variations.

Figure 3(c) shows the result of the simulation. The estimation was inaccurate because $1-Z_f/Z$ is too small when the observer moves toward a frontoparallel plane. Our model cannot estimate the direction of heading for the frontoparallel-plane condition accurately. Poor performance of human observers in the frontoparallel-plane condition was also reported (Rieger and Toet, 1985; Warren & Hannon, 1990; Royden et al., 1994). The simulation results of the model are similar to performance of human observers.

The number of sampling points

We assume that the number of dots is large. Here we examine how sensitive the estimation by our model is to the assumption. We made use of the law of large number in Eqs. (7), (8), (13), (14) and (16) in

the gaze-stabilized algorithm. In most psychophysical studies, roll (C) has been 0 and pitch(A) has been negligible. If C and A are 0, the law of large number is not used in Eqs. (7), (8), (14) and (16) for recovery of heading by our model. When static depth information is available, (13) is also unnecessary. It means that a few dots are sufficient for the estimation in situations with static depth cues such as translation on the ground. In cloud conditions without static depth cues, more dots are necessary. In the gaze-unstabilized situations , many dots are necessary because the center of outflow must be calculated. Our model needs a few dots to obtain accurate estimates in some situations. However, the accuracy of our models depends on other factors (e.g., simulated environments, thresholds, the rotation axis and the display size).

We performed simulations in the same way as the former cloud conditions. We used thresholds different from those in the previous simulations because the thresholds are inappropriate when the number of dots is small. The thresholds were as follows: T_{ax}=0.01 [1/sec], T_{by}=0.01 [1/sec], T_{zl}=0.1 [1/sec], T_{zh}=10.0 [1/sec] and T_{u}=0.01 [1/sec]. We conducted the linear regression analyses between the simulated heading and the estimated heading. Fig. 4 shows the results. For 1000 dots conditions, the slope was smaller than that of the previous simulations. This was ascribed to the differences of the thresholds. In the simulations we used the iterative procedure. As the number of iterations increased, the slope became nearer to 1. We can say that the iterative procedure eliminates the underestimation and improves the estimate. For 100 dots conditions, the results were almost the same as for 1000 dots conditions. For 10 dots conditions, the slope was about 0.6 after two iterations and the correlation coefficients were larger than 0.8. The model can judge heading fairly accurately for 10 dots in this condition though the performance became worse than for 100 or 1000 dots.

Insert Figure 4 about here

Warren et al. (1988) reported that accurate heading judgements can be obtained from very few dots in motion for pure translations. However, Warren & Hannon(1990) reported that performance of heading

judgement became poor for six dots when translation and rotation were simulated. Human visual system appears to need a large number of dots to decompose retinal flow into rotational and translational components in retinal flow.

yaw rate

Here we examine how sensitive our model is to the assumption of small rotation. In gaze stabilized situations, the yaw rate is confounded with the heading direction (see (4)) and a yaw rate can not varied independently. Therefore we simulated gaze-unstabilized situations. We simulated situations where an observer moves towards 100 dots cloud with the depth range of 3 – 5 m. Parameter W was 1.25 m/sec. The heading direction was 4.0 deg from the line of sight. We used the gaze-unstability algorithm. The results are shown in Fig. 5. The estimate was accurate when the yaw rate was smaller than 4.0 deg/sec. The performance became worse when it was larger than 6.0 deg/sec. The bias in the estimated heading to the center of outflow, namely, to the direction of rotation occurred. The tendency was also observed for human observers (Royden et al. 1994)..

Insert Figure 5 about here

We made use of the assumption of small yaw and pitch in Eq. (12). The terms of A^2, B^2, Z_f/Z_i $(1-Z_f/Z_i)A^2$ and Z_f/Z_i $(1-Z_f/Z_i)B^2$ in O(A,B) of (12) which do not include x or y, are much larger than other terms. Z_f/Z_i $(1-Z_f/Z_i)A^2$ is less than 0.25 A^2 and Z_f/Z_i $(1-Z_f/Z_i)B^2$ less than 0.25 B^2. Roughly the accuracy of the approximation of (12) depends on the ratio of $(W/Z_i)^2$ to A^2+B^2. Therefore more accurate estimate is expected when W is increased.

Small pitch and yaw are also necessary for the symmetric sampling around the origin after the transformation of the Z-axis so that it passes through the center of outflow. The deviation from symmetric sampling also depends on the ratio of W/Z_i to the root of A^2+B^2.

We performed simulations using W of 1.875 and 2.5 m/sec. Fig. 5(b) shows the results. The fairly accurate estimates were obtained up to the yaw rate of 8 deg/sec for W of 2.5 m/sec. For 1.875 m/sec, the

effect of yaw was smaller than for 1.5 msec, though the performance was worse than for 2.5 msec. Stone & Perrone (1997) reported that human performance depends on the ratio of rotation to translation. The simulation results are consistent with human performance qualitatively.

The effects of a yaw rate depends on many factors such as the depth range, the display size and the number of dots. One need to be careful for the test of our model.

Next, we present results of psychophysical experiments and compare performance of the model with human heading judgement to test the proposed model.

Psychophysical experiments of human heading judgement

Experiment 1

To test the validity of our method as human model, we compared the performance of human heading judgement under the conditions used in the simulations with the performance of our proposed model. We simulated the case where the observer fixated a static point while translating in a fixed direction with respect to the current line-of-sight as shown schematically in Fig. 6(a). The heading direction did not change during the presentation in the egocentric(eye-center) coordinate system from one frame to another, while the direction of the path changed in the exocentric (world-center) coordinate system. Thus the observer's path was like a curve in Fig. 6 (a) in the exocentric coordinates. One example of the actual path is shown Fig. 6(b). It means that we must rewrite Eqs. (1) and (4) taking time into account when we assume V=0, C=0 and A=0.

$$\begin{aligned}
\dot{X}(t) &= -U - B(t)Z(t) \\
\dot{Y}(t) &= 0 \\
\dot{Z}(t) &= -W + B(t)X(t) \\
\dot{Z}_f(t) &= -W + B(t)X(t) \\
B(t) &= -U(t)/Z_f(t)
\end{aligned} \qquad (29)$$

where t shows time, and U and W are constants. We adopted this paradigm because U and V are constant in the retinocentric coordinate frame (fig 4a).

Insert Figure 6 about here

When we use the following equations instead of (29), not assuming the gaze-stability constraint but assuming constant B, the trajectory is a circular path that Stone & Perrone (1997) used in their psychophysical experiments, though we did not use the paradigm.

$$\begin{aligned} \dot{X}(t) &= -U - BZ(t) \\ \dot{Y}(t) &= 0 \\ \dot{Z}(t) &= -W + X(t) \end{aligned} \qquad (30)$$

where B is a constant. Mathematical derivation of the trajectory is presented by Royden (1997).

If the observer moves along a straight path in the environment while (s)he fixates a fixation point (Fig. 6(c)), we must use the following equations instead of (29).

$$\begin{aligned} \overset{*}{X}(t) &= -U(t) - B(t)Z(t) \\ \dot{Y}(t) &= 0 \\ \dot{Z}(t) &= -W(t) + X(t) \end{aligned} \qquad (31)$$

Generally U(t) W(t) and U(t)/W(t) are not constants in the retinocentric coordinate system as shown in Fig. 6(c), (which means the direction of heading is not constant in retinocentric coordinates), though the heading direction is constant in exocentric coordinates. It is problematic when the fixation point is the only reliable landmark because the direction of heading relative to the fixation point changes during the stimulus presentation. Therefore we did not use this paradigm.

Methods

Observers: Two authors (MH and YE) and three naïve observers (HI, HT and YN) participated in the experiment.

Apparatus: Experiments were conducted using a Silicon Graphics O2 workstation with a color monitor. The image on the screen was 34.4cm wide (1280 pixels) and 27.5cm from top to bottom (1024 pixels). The observers sat with their heads stabilized in a chin-rest at 40 cm from the screen, resulting in an image of 46.5 deg × 38.0 deg of visual angle. The stimuli were presented at a frame rate of 60Hz. Observers viewed them monocularly. Apart from the stimuli, the room was dark.

Stimuli: One red point served as a fixation point. It was located at the center of the screen and remained stationary so that no eye movement occurred during stimulus presentation. Translation with eye movements was simulated for the visual stimuli. We used white dots, which were 2 × 2 pixels with a luminance of 70 cd/m^2 on a black background. The simulated environment consisted of 1000 randomly located dots that were configured in a horizontal plane, in a cloud, or in a frontoparallel plane. Simulated horizontal ego-motion was displayed for 2 sec and then all dots except the fixation point disappeared. When dots went out of the screen, new dots appeared at randomly determined positions in the screen to keep the number of dots on the screen constant. No noise was added. The depth range of dots was determined so that it was very similar to that of the model simulation at the end of the presentation for 2 sec.

Procedure: 100 trials were conducted in a session. Observers had to make retinocentric heading judgements after the stimulus presentation: They had to judge the direction of heading relative to the fixation point (θ in Fig. 6(a)) by setting a pointer and pressing a mouse button. The simulated motion was explained well to all observers. The observers did not receive feedback although they had some practice with feedback.

Results

Ground condition

Simulated translation in the world was determined in the same way as in the simulation for each trial: Parameter U was randomly set to a value between –0.25 m/sec and 0.25 m/sec, and W was set

between 0.5 m/sec and 1.5 m/sec for each trial. The simulated world extended in depth from 3 to 20 m in front of the observer's eye and the simulated distance of the fixation point was 11.5 m at the beginning of the presentation. The absolute value of parameter B (yaw) was less than 1.25 deg/sec at the beginning of the presentation and less than 1.5 deg/sec at the end. Parameters A and C were 0.

Figure 7(a) shows an observer's result in the ground condition. The results of the other observers were similar. In this figure, most of the data points lie near the regression line. We can say that there is linear relationship between the perceived heading direction and the simulated heading direction. The slope of the regression line is 0.82. The slopes and correlation coefficients between the data points and the regression line for all observers are shown in Table 1. The slope was smaller than 0.85 for all observers. It means that all observers showed underestimation of the heading direction in the ground condition, despite individual differences in the degree of underestimation. The correlation coefficients were larger than 0.9 for all observers.

Insert Figure 7 and Table 1 about here

Cloud condition

Simulated translation in the world was determined in the same way as in the simulation for each trial: Parameter U was randomly set to a value between –0.25 m/sec and 0.25 m/sec, and W was set between 0.75 m/sec and 1.25 m/sec for each trial. The simulated world extended in depth from 5 to 7 m in front of the observer's eye and the simulated distance of the fixation point was 6 m at the beginning of the presentation. The absolute value of parameter B was less than 2.9 deg/sec at the beginning of the presentation and less than 3.6 deg/sec at the end.

Figure 7(b) shows an observer's result under the cloud condition. The slope of the regression line is 0.76. The slopes and correlation coefficients for all observers are shown in Table 1. All observers showed some underestimation in the cloud condition, though there were individual differences about the degree. The perceived heading was more underestimated in the cloud condition than in the ground condition for

four out of the five observers. The correlation coefficient was high (> 0.7), although it was smaller than in the ground condition for all observers.

Frontoparallel plane

The stimuli were the same as in the cloud condition with one exception: The simulated plane was placed at a depth of 6 m in front of the observer's eye and the simulated distance of the fixation point was 6 m at the beginning of the presentation.

Figure 7(c) shows an observer's result under the frontoparallel-plane condition. The data points are scattered around the horizontal line. Results of the regression analyses for all observers are shown in Table 1. The slope was less than 0.2 and the correlation coefficient less than 0.5 for all observers. It means that heading could not be judged accurately by all observers in the frontoparallel-plane condition. The results were consistent with the simulation results of the model.

Discussion

Comparison with the model

Fairly good performance in the ground and cloud condition and poor performance in the frontoparallel-plane condition are consistent with the model's performance. However, the bias in perceived heading is a little different from the performance of our model. The model without the iteration process showed some underestimation of the heading direction in the cloud condition, but not in the ground condition, while human observers showed underestimation in both conditions and the individual differences about the degree were large. The bias in perceived heading has often been explained as follows: A flow pattern that simulated translation with eye movement results in perceived ego-motion on a curved path because of contradicted extraretinal information, and hence the bias occurs (van den Berg, 1996; Royden, 1997; Crowell, 1997). In our model, the underestimation occurred in at least one condition and might therefore be explained partly by our model.

Comparison with other studies

Warren and Hannon (1990) reported that when simulated eye-rotation was small(<1.5 deg/sec), their observers judged heading accurately when there were sufficient depth variations in the simulated scenes. Some researchers reported inaccurate heading judgement from visual information alone in situations where egomotion with fast eye movement (> 1.0 deg/sec) was simulated (Royden et al., 1994; Banks et. al., 1996). van den Berg (1993), however, reported that human observers judged heading with a relatively fast rotation rate of 5 deg/sec accurately. In our psychophysical experiment where the rotation rate was rather low (<1.5deg/sec in the cloud conditions and <3.6deg/sec in the ground condition), performance was good in cloud and ground conditions although biased.

In the research of Royden et al. (1994), it seems that inaccurate heading perception occurred when there was no center of outflow in the display. In the absence of the center of outflow, our model also does not work well because the symmetric sampling about the center of outflow is difficult. On the other hand, human observers judged heading accurately with a relatively fast rotation rate in the results of van den Berg (1993; 1994). The display that he used included the center of outflow. In our experiment, the center of outflow was at the center of the display and human observers showed fairly good performance in the cloud and ground conditions. Differences of the studies may be explained in terms of the center of outflow.

Bias in heading perception to the line of sight in some conditions was reported in various studies (van den Berg, 1996; van den Berg & Brenner, 1994; Cutting et al., 1997). This was also observed in our experiment. The results of van den Berg (1996) showed that there were some individual differences in the magnitude of the bias as in our experiment. The bias observed in the results of van den Berg (1994) and ours can also be interpreted as bias to the center of outflow. In the research of Royden et al. (1994), the bias toward the center of outflow caused an error in heading judgement. When egomotion toward a frontoparallel plane was simulated, a bias towards the center of outflow (or the singularity point) was observed in many investigations (e.g., Warren & Hannon, 1990; Royden et at., 1994; Stone & Perrone, 1997). Human observers may perceive heading biased toward the center of outflow under some conditions.

In the ground condition of our experiment, correlation coefficients between the perceived directions and the regression line were very high though some underestimation occurred. We can say that under some conditions, human observers can discriminate the heading direction correctly although there is a bias in the perceived direction and that absolute heading judgement from a retinal flow field is more difficult than discrimination of the heading direction.

Experiment 2

In Experiment 2, we introduce a result consistent with our model. In the experiment, we use flow fields with perturbed velocities of dots, and we examine the relationship between the performance of our model and human performance with the stimuli. For the perturbation, another velocity component, $gxyB(1-Z_f/Z_i)$, is added to the vertical component of the velocity of each dot. The perturbed stimuli do not occur in real situations[3]. The perturbation velocity has specific effects on our model. We compare the effects of the perturbation on the estimations of our model with those on human performance. To show that all heading recovery algorithms do not predict the same results, we also consider the performance of another algorithm, the differential motion algorithm by Rieger and Lawton (1985). The algorithm and its concept were used to model human heading judgement (Hildreth, 1992; Royden, 1997). Details of the implementation are described in Appendix B.

Performance of the models

For the perturbation, we added a vertical velocity component, $gxyB(1-Z_f/Z_i)$ to the original velocity of each dot assuming that viewing distance was 1. Parameter g indicates the magnitude of the added velocity. When g was 0, no velocity component was added to the original simulated velocity.

The extra velocity component affects the estimation of our model in a specific way. Assuming that A and C are 0, and $v_i' = v_i + gx_iy_iB(1-Z_f/Z_i)$, we obtain the following equation by replacing v_i in (10) with v_i':

[3] A least-square algorithm (Heeger & Jepson, 1990) did not achieve the solution without error for the perturbed stimuli even when no other noise was added. Therefore we can say

$$\frac{\frac{x_i}{y_i}v_i' - u_i}{1-\frac{Z_f}{Z_i}} = \frac{\frac{x_i}{y_i}v_i - u_i}{1-\frac{Z_f}{Z_i}} + gx_i^2 B = B + gx_i^2 B \qquad (32)$$

From (11) and (32), B_e is :

$$\begin{aligned} B_e &= \frac{1}{N_b}\sum \frac{\frac{x_i}{y_i}v_i' - u_i}{1-\frac{\tau_f}{\tau_i}} \\ &\approx \frac{1}{N_b}\sum \frac{\frac{x_i}{y_i}v_i - u_i}{1-\frac{\tau_f}{\tau_i}} + \frac{1}{N_b}\sum gx_i^2 B \qquad (33) \\ &\approx B + \frac{1}{N_b}\sum gx_i^2 B \end{aligned}$$

From (17) and (33), the horizontal direction of heading is computed as:

$$U_e \approx B_e\tau_f \approx B\tau_f + \frac{B\tau_f}{N_b}\sum gx_i^2 \approx \frac{U}{W}\left(1+\frac{g}{N_b}\sum x_i^2\right) \qquad (34)$$

The above equations show that the horizontal heading direction is overestimated when g is larger than 0. The magnitude of the overestimation depends on $(1/N_b)\sum x_i^2$, or the width of the sampling range along the x-axis direction.

We performed simulations for the perturbed stimuli. Noise was also added as in the previous simulations. The cloud condition was used in this experiment. Four g values were used. 100 trials were conducted for each g value. Figures 8 and 9 show the simulation results of our model (Fig. 8) and Rieger and Lawton's algorithm (Fig. 9) respectively. The same parameters as in the previous simulations were

---

that this operation is the introduction of noise.

used for our model. The estimates of V/W by Rieger and Lawton's algorithm are near 0. Thus we showed only the estimated horizontal direction.

Insert Figure 8 and Figure 9 about here

As the g value is increased, the performance of our model and Rieger and Lawton's algorithm changes. For our model, the slope of perceived vs. estimated heading function increases as g is increased and the shape holds nearly linearly. For Rieger and Lawton's algorithm, however, the functional shape remarkably changes. As the g value is increased, the shape changes from linear shape to an inclined S-shape. The heading direction is estimated around the line of sight when |U/W| is large. But we do not know the reason for the performance of Rieger and Lawton's algorithm and perhaps it depends on the different implementations and parameters. It seems that algorithm performance in the case of our model does not arise from an aspect inherent to the heading recovery problem, but rather from the properties of the specific algorithm used.

Human performance

Methods

Both authors and one naïve observer participated in the experiments. We used the same experimental apparatus and stimuli as in Experiment 1, except that another velocity component, $gxyB(1-Z_f/Z_i)$ for the perturbation was added to dot velocities assuming that viewing distance is 1. The experiment was carried out under the cloud condition as in the simulations. We also used four g values. Because the stimuli sometimes make observers feel themselves moving in a non-rigid environment when g is 60 and |U/W| is large, the observers were asked to report the direction in which they felt themselves moving most. 100 trials were conducted for each value of g. One session included 400 trials. The observers viewed randomly ordered sequences of 400 trials. The observers had much practice with unperturbed stimuli so that estimated slope parameters could be reliable.

Results and discussion

We conducted regression analyses as in Experiment 1. We found that no systematic deviations from the line occurred at any g value for all observers. The correlation coefficients of the regression analyses were greater than 0.9 for all sessions but one. One exception occurred in a session in the condition of g=0 for observer YE in which the correlation coefficient was 0.78.

Insert Figure 10 about here

To examine how the slope depends on the value of g, we plotted the slope of the perceived vs. simulated heading function against g, as shown in Figure 10(a). In Figure 10(b), the slope is normalized at g=0. At g=0, there is a little difference between the results of Experiment 1 in Figure 3 and those of Experiment 2 under the same condition. This difference may be due to a training effect because the observers had practice between the experiments. As the g value is increased from 0 to 60, the slope increased from 0.77 to 0.90 for HI, from 0.65 to 0.79 for MH and from 0.73 to 0.88 for YE. The increasing rate is 17% for HI, 22% for MH and 21% for YE.

Systematic deviations from the line did not occur in the experiment. The findings are consistent with the prediction by our model, but not with the predictions by our implementation of Rieger and Lawton's algorithm. There is, however, a difference in the rate of increase between the psychophysical experiment and our model prediction. The observed rate is about 20 % for observers, whereas the rate of increase predicted by our model is about 200% as seen in Fig. 8.

This discrepancy may be ascribed to the following reasons. First, we employed a screen size of about 46.5 deg wide, disabling the observers from pointing out peripheral heading directions over 24 deg from the center. This might cause compression of the response span. Second, the width of the sampling region along the x-axis affects the slopes obtained by the model in the simulation. We used a 46 deg-wide sampling region, but human observers might use a narrower sampling region. When a narrower sampling region is used, $(1/N_b) \sum x_i^2$ in Eq. (34) becomes smaller and the effects of parameter g become smaller. But why does the visual system use a narrower region?

In our experiment, parameter A (pitch) was 0. Otherwise, the term, Ax_i/y_i in Eq. (10) is a disturbance for the estimation of B (Yaw) by Eq. (14) and our model removes the components according to the law of large numbers. If x_i/y_i is small, the effect of Ax_i/y_i is small and a better estimate is obtained. Accordingly the visual system may use a narrower sampling region.

Experiment 3

In Experiment 3, we used a narrower display and examined the effects of display width under the same conditions as in Experiment 2. In the cloud condition in Experiment 1 and 2, heading judgement seemed to be difficult for some observers and it appeared impossible to obtain an accurate estimate of the slope of the perceived vs. simulated heading function. However, we found that heading judgement was easier when we used a larger depth range for dots in the simulated environment (see also Rieger & Toet, 1985). We used a larger depth range in the experiment.

Method

The conditions are the same as in Experiment 2 with two exceptions. First, the simulated world extended in depth from 4.5 to 7.5 m in front of the observer's eye and the simulated distance of the fixation point was 6m at the beginning of the presentation. Second, we used two display sizes, 46.5deg × 38.0 deg (1280 × 1024 pixels) and 19 deg × 38.0 deg(500 × 1024 pixels). 1000 dots were presented for the wide display and 390 dots were presented for the narrow display.

Both authors (MH, and YE) and three naïve observers (HI, HT and YN) participated in the experiment. We used a wider depth range, and the average and maximum of $|gxyB(1-Z_f/Z_i)|$ were larger than in Experiment 2. The stimuli make observers feel themselves translating in a non-rigid environment more than those in Experiment 2, or feel themselves rolling. They were asked to judge the direction in which they felt themselves moving most. Each observer participated in only one session with 400 trials for each display condition.

Results and discussion

The results for the wide-display condition are shown in Fig. 11. Fig. 11 (a) shows the slopes of perceived vs. simulated direction functions against g. The slope of the regression line at g=0 was different between individuals as in Experiment 1. Fig. 11(b) shows the slope normalized at g=0 and Fig. 11(c) shows the correlation coefficients. There appears to be two types of observers. For observers HI, MH and YN, correlation coefficients were high at all g values and the slope increased as g was increased, while for observers HT and YE, the correlation coefficient decreased as g increased and the slope decreased as g changed from 20 to 60. Observer YE reported that he sometimes felt himself rolling when observing the stimuli and the feeling might affect the results. It seems that HT and YE could not perform heading judgement when g was large because the perturbation was too large for them.

Insert Figure 11 about here

Observers HI, MH and YN could maintain the performance when g was large and the slope estimates of the perceived vs. simulated heading function were reliable. The slopes for the three observers increased as g was increased as shown in Fig. 11 (a) and (b), though the degree was different between individuals. No systematic deviation from the regression line was observed. The results were similar to those in Experiment 2 and consistent with the prediction of our model.

Results for the narrow-display condition are shown in Fig. 12. Fig.12 (a) shows the slopes of the perceived vs. simulated heading function against g. Fig.12b shows the increasing rate of the slope and Fig. 12c shows the correlation coefficients. The slope of the regression line at g=0 was different between individuals. The slope at g=0 was a little smaller in the narrow-display condition than in the wide-display condition. The correlation coefficient at g=0 was high (R>0.95 for HI, MH and YN, R=0.94 for YE, R=0. 88 for HT) and almost the same as in the wide-display condition. It shows that heading judgement was fairly accurate in the narrow-display condition and narrowing the field did not have much effect on heading judgement. However, the slope of the perceived vs. simulated heading function did not increase as g was increased for all observers except observer MH as shown in Fig. 12(a) and (b), and the rate of

increase of MH in the narrow-display condition was smaller than in the wide-display condition. The resulting increase of slope is seemingly inconsistent with the model's performance, but we present an extended model to explain the results.

Insert Figure 12 about here

It is suggested from the results that although the role of the periphery is small, the increase of the slope by the perturbation is mainly due to periphery dots. Next we present a model taking the findings of Experiment 2 and 3 into account.

Weighted Average Model

Computationally it is suggested that sampling points with small x_i/y_i should be used to estimate B(Yaw) by Eq. (14) as we noted in the discussion of Experiment 2 and it is suggested in the psychophysical experiment that the central and narrow region play an important role in judging heading.

Therefore we use the weighted average obtained with a Gauss function to estimate B instead of Eq. (14) so that the central and narrow region might be used effectively and periphery dots might also be used to a limited extent.

$$B_e = \frac{\omega_i}{\kappa} \sum \frac{\frac{x_i}{y_i} v_i - u_i}{1 - \frac{\tau_f}{\tau_i}} \qquad (35)$$

where

$$\omega_i = \exp\left(\frac{-\left(\frac{x_i}{y_i}\right)^2}{\sigma^2} \right), \quad \kappa = \sum \omega_i \qquad (36)$$

κ is a normalization parameter. If σ becomes infinite, Eq. (35) becomes the same as (14).We rewrite (16) along the same reasoning.

$$A_e = \frac{\psi_i}{\mu} \sum \frac{v_i - \frac{y_i}{x_i} u_i}{1 - \frac{\tau_f}{\tau_i}} \qquad (37)$$

where,

$$\psi_i = \exp\left(\frac{-\left(\frac{y_i}{x_i}\right)^2}{\sigma^2} \right), \quad \mu = \sum \psi_i \qquad (38)$$

We performed simulations using (35). In the simulations, we used an iterative procedure. The number of iterations was two. Because the degree of underestimation depends on observers and the iteration process eliminates the underestimation of the heading direction, we focus on the increasing rate of the slope as parameter g. σ was set to 0.3. Other parameters were the same as in previous simulations.

Results of the simulations are shown in Fig. 10 (b), 11(b) and 12(b) with human data. The rate of increase of the model in Fig. 10 (b) is similar to the rate of the human observers. Human performance in Experiment 2 is consistent with the model's performance.

Because it seems that observers YE and HT had trouble in performing the task in Experiment 3 when g was large, we do not discuss their results here. The rate of increase in Fig. 11 (b) obtained from the model is similar to the rate of two observers, HI and YN. The increasing rate of MH is larger than the modeled rate. The rate of increase largely depends on σ. As σ becomes larger, the rate of increase becomes larger. The rate of increase also depends on the number of iterations. Possibly the results of observer MH in this condition are explained by selecting appropriate parameters.

In the narrow-field simulations, the slope at g=0 is larger than in conditions where g>=20 as shown in Fig 12(b). There were large individual differences in the results under the narrow-field conditions for human observers and the individual differences might be explained by parameter differences.

Qualitatively the results in Experiments 2 and 3 correspond to the modeled results. Quantitative differences between human performance and the model may be explained by selected parameters.

We adopted Eqs. (35) and (36) as a weighting function in the sampling region due to a computational motivation, but there might be another valid function. Rieger and Toet (1985) reported that a reduction of the field of view from 20 × 20 deg to 10 × 10 deg did not affect the judgements significantly. The report indicates that the central small region plays an important role in heading judgement, and this was confirmed by other studies (Cutting et al., 1997; Warren and Kurtz, 1992). Crowell & Banks (1993), however, reported that for translational heading, a small region with a radial configuration was important for heading judgement. Warren & Saunders (1995) used the center-weighted expansion unit by a Gaussian function. The results in Experiment 3 suggest that the narrow region is important for human heading judgement and the narrow region about the center of outflow plays an important role in our model. Further research on the weight of regions is required.

General discussion

We developed a new model of human heading judgement that uses the deviation from a radial retinal flow pattern. The proposed model showed similar performance to that of human observers; fairly good performance in the ground and cloud condition, poor performance in the frontoparallel-plane condition. Our model also showed a tendency similar to human observers when $gxyB(1-Z_f/Z)$ was added to the vertical velocity component for the perturbation. Thus we can say that our model is a candidate for a model of human heading judgement.

Other models

It seems that our model makes use of information similar to that used in the algorithm of Cutting (1986). What he refers to as differential motion parallax (or differential parallactic displacement) is similar

to the concept of deviation from a radial pattern in our model if the retinal velocity is limited to the horizontal component and the observer's translation is limited within a horizontal plane (Fig 13). But our model differs from Cutting's method in various aspects. First, Cutting's algorithm uses horizontal *displacements* of a few points, but our model uses a *velocity field*. Second, Cutting's algorithm needs multiple fixations or relative depth information from sources other than retinal flow, for accurate absolute heading judgement. Our model does not need it in principle. Third, Cutting's algorithm is applicable to a limited range of situations where multiple fixations or other depth information, gaze stability, horizontal translation in the 3-D environment occur, but without roll. Fourth, our algorithm decouples rotational and translational components in retinal flow, but Cutting's algorithm does not. The mathematical framework of our algorithm is different from Cutting's algorithm. In the limited situation as Cutting and his colleagues used in psychophysical experiments (Cutting et al., 1992; Vishton & Cutting, 1995; Cutting et al., 1997), however, both algorithms use similar visual information (differential motion parallax, or deviations from a radial pattern). Cutting and his colleagues doubted the feasibility of decomposing a velocity field into rotational and translational components by an optic flow algorithm representative of human wayfinding, but it appears that many results of their psychophysical experiments can be explained well not only by Cutting algorithm, but also by our decomposition model.

Insert Figure 13 about here

Our model is one of the candidates for human model, but it has not yet been tested sufficiently. Since our model has some limitations (slow eye rotation and a large number of sampling dots), we should test whether humans share the limitations of the model. Moreover the results of Experiment 2 do not deny Rieger and Lawton's algorithm as a candidate for a human model because the performance might depend on the algorithm implementations. For example, Hildreth (1992) modified Rieger and Lawton's algorithm to achieve accurate heading estimation by removing noisy dots incrementally in the calculation. Possibly

her method reduces the effects of perturbation in the stimuli of Experiment 2. It is necessary to determine which of these models accurately describes human heading judgement by further studies.

Depth cues and extraretinal information

van den Berg and Brenner (1994a, b) reported that human observers show large noise tolerance when stimuli include binocular disparity or static depth cues. Our model integrated the depth cue easily as shown previously. It is expected that depth cues improve the performance of our model if noise is large because it is not needed to use approximate equations, (11), (12) and (13) .

Extraretinal information about eye movements is used for the recovery of heading direction (Warren et al., 1990; Royden, et al., 1994). Our method recovers the rotation parameters first and next the direction of heading. Because our algorithm is a rotation-first approach, extraretinal information about eye rotation could be easily integrated.

Physiological bases

Our modes calculates Z_f /W, which can be obtained in the same way as calculating time to contact. Humans can judge time to contact (Regan and Vincent, 1995; Gray and Regan, 1997). There are cells which respond to expansion/contraction or roll pattern in the brain of monkey (Saito et al., 1986). Perhaps time to contact is calculated by the cells in MST which respond to an expansion patterns. Our model searches the center of outflow. It may also be calculated by the expansion cells in MST. The magnitude of the roll (C) may be estimated by the cells in MST which respond to roll patterns. But our model has not yet been implemented as a neural network.

As far as we know, the existence of cells that respond to a deviation from a radial pattern for heading judgement has not been reported, though cells that respond to spiral motion were found (Graziano et al., 1994). We suggest that physiological substrates of heading judgement may be found by searching for cells responding to a deviation from a radial flow pattern.

Appendix A: The proof of Eqs. (27)

Z_f defined in (26) nearly equals the average depth of other sampling points. Here we present the proof. First we eliminate the roll component in the gaze-unstability condition. From (3) we obtain:

$$y_i u_i - x_i v_i = (x_i^2 + y_i^2)C - y_i \frac{U}{Z_i} + x_i \frac{V}{Z_i} - By_i - Ax_i \qquad (39)$$

In the random-cloud condition, Eq. (7) can also be used to estimate C because the expectation of the terms, y_iU/Z_i, x_iV/Z_i, $-By_i$, $-Ax_i$ in (39) is 0. In the ground condition, Eq. (8) can also be used. Then we remove the roll components from the retinal flow field using the estimate. Therefore Eqs. (3) become:

$$\begin{aligned} u &= \frac{(-U + xW)}{Z} - B + Axy - Bx^2 \\ v &= \frac{(-V + yW)}{Z} + A + Ay^2 - Bxy \end{aligned} \qquad (40)$$

Assuming that A and B are small, and x and y are also small, we neglect the terms, Axy, $-Bx^2$, Ay^2 and –Bxy because these terms are much smaller than the other terms. We obtain:

$$\begin{aligned} u &= \frac{(-U + xW)}{Z} - B \\ v &= \frac{(-V + yW)}{Z} + A \end{aligned} \qquad (41)$$

We assume that Z_0 is the average depth of sampling points. We choose a point with a depth of Z_0. (x_0, y_0), and (u_0, v_0) shows the position and velocity projected to the image plane respectively. From Eqs. (41), we obtain:

$$\begin{aligned} u_0 &= \frac{(-U + x_0W)}{Z_0} - B \\ v_0 &= \frac{(-V + y_0W)}{Z_0} + A \end{aligned} \qquad (42)$$

The equation of the line passing through (x_0, y_0) and orienting to (u_0, v_0) is:

$$(y - y_0)(\frac{-U + x_0 W}{Z_0} - B) = (x - x_0)(\frac{-V - y_0 W}{Z_0} + A) \qquad (43)$$

Regardless of any x_0, y_0, the line always passes through a point (x_c, y_c). x_c and y_c are:

$$x_c = \frac{BZ_0 + U}{W}$$
$$y_c = \frac{-AZ_0 + V}{W} \qquad (44)$$

Every line through the velocity vector of points with the average depth passes through the point (x_c, y_c). Next we show that (x_c, y_c) nearly coincides with the center of outflow.

If we select a point whose depth is different from the average depth (Z_0), the line passing through the velocity vector in the image plane is located on the left or right of (x_c,y_c) and above or below (x_c,y_c) depending on Z-Z_0. When we calculate the center of outflow, the effect of the points nearer than the average depth is cancelled by the effect of the points further away if the observer moves toward a random cloud, a plane, an ellipsoid and other 3D shapes[4]. It implies that the point (x_c, y_c) nearly equals the center of outflow.

We rotate the axes so that the Z-axis is through (x_c, y_c). First the X- and Z- axes are rotated around the Y-axis by arctan((BZ_0+X)/W). Next the Y- and Z- axis are rotated around the X-axis by[5]:

$$\arctan\left(\frac{-AZ_0 + V}{W\sqrt{1 + \left(\frac{BZ + U}{W}\right)^2}}\right) \qquad (45)$$

Roll components arise from the transform of the axes because the rotation axis are also transformed. If the transform is small, it is negligible. Otherwise, the roll components should be estimated by Eq. (7) or (8).

In the new coordinates, (x_c, y_c) is (0, 0). From Eqs. (44), we obtain:

---

[4] Note that it is not true for all stimuli. For example, if the following stimulus is used, it is not true: $B \neq 0$, $A=C=0$. $V=0$. $U \neq 0$, $W>0$, $x_c = (B \times 4 + U)/W$ and $y_c = 0$. if ($X > x_c$ and $Y > y_c$) or ($X < x_c$ and $Y < y_c$) then $Z=3$. If ($X < x_c$ and $Y > y_c$) or($X > x_c$ and $Y < y_c$), then $Z=5$. The y-axis value of the center of outflow is different from y_c in the situation. But it appears that such stimuli are exceptional.

[5] Practically it is sufficient that the image plane translates only by($-x_c$, $-y_c$) to set (x_c, y_c) to (0, 0).

$$0 = \frac{B_{new} Z_0 + U_{new}}{W_{new}}$$

$$0 = \frac{-A_{new} Z_0 + V_{new}}{W_{new}} \qquad (46)$$

Therefore we get:

$$B_{new} = \frac{-U_{new}}{Z_0}$$

$$A_{new} = \frac{V_{new}}{Z_0} \qquad (47)$$

Compared with Eqs. (26), we obtain Eqs. (27):

Appendix B: Differential motion algorithm by Rieger and Lawton(1985)

The algorithm is based on the observation that at a depth discontinuity in the visual field, the translation component of the image velocity field will be discontinuous due to the dependence of this component on depth, while the rotational component will be roughly constant across the discontinuity. Rieger and Lawton (1985) presented an algorithm, which is based on the algorithm by Longuet-Higgins and Prazdny (1981). We estimated the heading direction using the algorithm in the following way. First, the differences between each local image velocity and other velocities measured within a restricted neighborhood were computed. To compute the difference between velocities, we set the neighborhood size to 5.7 deg. From the resulting distribution of velocity difference vectors, the dominant orientation of the vectors was computed by PCA (principal component analysis) and preserved only at locations where the distribution of velocity differences was strongly anisotropic in some direction. In the next stage, we used only the results for which the first factor explained the difference vectors over 90%. Such points typically arise where there is a strong depth variation. The result from this stage was a set of distributions at a number of points in the image that were all roughly aligned with the translational field lines. The heading direction was then calculated as the best-fit intersection point for all the resulting vector directions.

Acknowledgements

We thanks G. van Tonder for reading and making extensive comments on the manuscript. M. Hanada was supported by Fellowships from the Japan Society for the Promotion of Science for Japanese Junior Scientists, and Y. Ejima was supported by Grants in Aid for Scientific Research (No. 10551003, 10164223 and 08301007) from the Ministry of Education, Science, Sports and Culture.

References

Banks, M., Ehrlich, S. M., Backus, B. T., Crowell, J. A. (1996) Estimating heading during real and simulated eye movements. *Vision Research, 36,* 431-443.

Beintema, J. & van den Berg, A.V. (1998) Heading detection using motion templates and eye velocity gain fields, *Vision Research, 38,* 2155-2179.

Bruss, A. R & Horn, B. K. P. (1983) Passive navigation. *Computer Vision, Graphics and Image Processing, 21,* 3-20.

Crowell, J. A. (1997) Testing the Perrone and Stone(1994) model of heading estimation. *Vision Research, 37,* 1653-1671.

Crowell, J. A. & Banks, M. S. (1993) Perceiving heading with different retinal regions and types of optic flow. *Perception & Psychophysics, 53,* 325-337

Cutting, J. E., (1986) Perception with an eye for motion. The MIT Press, Cambridge, MA.

Cutting, J. E. (1992) Wayfinding on foot from information in retinal, not optical flow. *Journal of Experimental Psychology: General,* 121, 41-72.

Cutting, J. E. & Vishton, P. M. (1995) Wayfinding, Displacement, and Mental Maps: Velocity fields are not typically used to determine one's aimpoint. Journal of Experimental Psychology: *Human Perception and Performance, 21,* 978-995.

Cutting, J. E., Vishton, P. M. & Grendt, J. D. (1997) Heading and path information from retinal flow in naturalistic environments. *Perception & Psychophysics, 59,* 426-441.

Fermüller, C. & Aloimonos, Y. (1992) Tracking facilitates 3-D motion estimation. *Biological Cybernetics, 67,* 259-268.

Graziano, M.S.A., Andersen, R.A. & Snowden, R.J. (1994) Tuning of MST neurons to spiral stimuli. *Journal of Neuroscience, 14,* 54-67.

Gray, R. & Regan, D. (1998) Accuracy of estimating time to collision using binocular and monocular information. *Vision Research, 38,* 499-512.

Hatsopoulos, N. G. & Warren, W. H. (1991) Visual navigation with a neural network. *Neural Networks, 4,* 303-317.

Heeger, D. J. & Jepson, A. (1990) Visual perception of three-dimensional motion. *Neural Computation, 2,* 129-137.

Heeger, D.J. & Jepson, A. (1992) Subspace methods for recovering rigid motion 1: Algorithm and implementation. *International Journal of Computer Vision, 7,* 95-117.

Hildreth, E. C. (1992) Recovering heading for visually-guided navigation. *Vision Research, 32,* 1177-1192.

Kanatani, K. (1993) 3-d interpretation of optical flow by renormalization. *International Journal of Computer Vision, 11,* 267-282.

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

Lappe, M. & Rauschecker, J.P. (1993) A neural network for the processing of optic flow from ego-motion in higher mammals. *Neural Computation. 5,* 374-391.

Lappe, M. & Rauschecker, J.P. (1994) Heading detection from optic flow. *Nature, 369,* 712-713.

Prazdny, K. (1980) Egomotion and relative depth map from optical flow. *Biological Cybernetics, 36,* 87-102.

Lappe, M & Rauschecker, J.P. (1995) Motion anisotropies and heading detection. Biological Cybernetics, 72, 261-277.

Nagel, M.G., Srinivasan, M. V. & Wilson, D. L. (1997) Image interpretation technique for measurement of ego motion in 6 degrees of freedom. *Journal of Optical Society of America A, 14,* 3233-3241.

Perrone, J. A. (1992) Model for the computation of self-motion in biological systems. *Journal of the Optical Society of America A, 9,* 177-194

Perrone, J. A. & Stone, L. S. (1994) A model of self-motion estimation within primate extrastriate visual cortex, *Vision Research. 34,* 2917-2938.

Raviv, D. & Herman, M. (1994) A unified approach to camera fixation and vision-based road following. *IEEE Transactions on System, Man, and Cybernetics, 24,* 1125-1141.

Regan, D. & Kashal, S. (1994) Monocular discrimination of the direction of motion in depth. *Vision Research, 34,* 163-177.

Regan D. & Vincent, A. (1995) Visual processing of looming and time to contact throughout the visual field. *Vision Research, 35,* 1845-1857.

Rieger, J. H. & Lawton, D. T. (1985) Processing differential image motion. *Journal of Optical Society of America A, 2,* 354-360.

Rieger, J.H. & Toet, L. (1985) Human visual navigation in the presence of 3D rotations. *Biological Cybernetics, 52,* 377-381.

Royden, C. S. (1994) Analysis of misperceived observer motion during simulated eye rotations. *Vision Research, 34,* 3215-3222.

Royden, C. S. (1997) Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *Journal of Optical Society of America A, 14,* 2128-2143.

Royden, C. S., Banks, M. S. and Crowell, J. A. (1992) The perception of heading during eye movements. *Nature, 360,* 583-585

Royden, C. S., Crowell, J. A. and Banks, M. S. (1994) Estimating heading during eye movements. *Vision Research, 34,* 3197-3214.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1986) Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience, 6,* 145-157.

Soatto, S. & Perona, P. (1998) Reducing "structure from motion": A general framework for dynamic vision Part 1: Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20,* 933-942.

Stone, L. S. & Perrone, J. A. (1997) Human heading estimation during visually simulated curvilinear motion. *Vision Research, 37,* 573-590.

Taalebinezhaad, M. A. (1992) Direct recovery of motion and shape in the general case by fixation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14,* 847-853.

Tomasi, C. & Shi, J. (1993) Direction of heading from image deformations. *In Proceedings of IEEE Computer Vision and Pattern Recognition,* 422 427.

Tsai, R.Y. & Huang, T. S. (1984) Uniqueness and estimation of 3-D motion parameters of rigid bodies with curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6,* 13-27.

van den Berg, A. V. (1993) Perception of heading. *Nature, 365,* 497-498

van den Berg, A. V. (1996) Judgement of heading. *Vision Research, 36,* 2337-2350.

van den Berg, A. V. & Brenner, E. (1994a) Why two eyes are better than one for judgemnt of heading. *Nature 371,* 700-702.

van den Berg, A. V. & Brenner (1994b) Human combine the optic flow with static depth cues for robust perception of heading. *Vision Research, 34,* 2153-2167.

Warren, W. H. & Kurtz, K. J. (1992) The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics, 51,* 443-454

Warren, W. H., Morris, W. H. & Kalish, M. (1988) Perception of translational heading from optic flow. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 646-660.

Warren, W. H. & Hannon, D. (1988) Direction of self-motion is perceived from optical flow. *Nature, 336,* 162-163.

Warren, W. H. & Hannon, D. (1990) Eye movement and optical flow. *Journal of Optical Society of America, 7,* 160-169.

Warren , W. H., Morris, M. W. & Kalish, M. (1988) Perception of translational heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 646-660

Weng, J., Huang, T. S. & Ahuja, N. (1989) Motion and structure from two perspective views: Algorithm, error analysis, and error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11,*

451-476.

Zemel, R. S. & Sejnowski, T. J. (1998) A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. *The Journal of Neuroscience, 18,* 531-547.

Zhang, Z. (1997) Motion and structure from two perspective views: from essential parameters to Euclidean motion through the fundamental matrix. *Journal of Optical Society of America, A, 14,* 2938-2950.

Figure and Table Legends

Table 1　Slopes of perceived vs. simulated heading function and the correlation coefficients (R) between the perceived direction and the regression line in Experiment 1.

Figure 1　Deviation from a radial optic flow pattern.

The arrow indicates the velocity of the point, p(x,y) in the image plane. |v-xu/y| and |yv/x-u| represent the deviations from a radial pattern.

Figure 2　The center of outflow.

The center of outflow is defined as a point that minimizes the square sum of d in the figure.

Figure 3　Results of simulations.

The horizontal axis represents the horizontal simulated direction and the vertical axis represents the estimated direction. (a) the ground condition (b) the cloud condition (c) the frontoparallel-plane condition. If the points are scattered along a straight line with slope 1.0, heading perception is unbiased. A linear regression analysis was conducted. The equation and the correlation coefficient (R) are shown above the figures.

Figure 4　Effects of the number of dots.

The slope between simulated vs. estimated heading function is plotted against the number of dots in (a) and The correlation coefficient in (b).

Figure 5　Effects of a yaw rate

The estimated heading is plotted against the simulated yaw rate. The simulated heading angle was 4.0 deg, which is indicated by the dotted line. Each point denotes the average of the estimates of 10 trials. (a)

Simulated forward translation (W) was 1.25 m/sec. (b) Simulated forward translation was varied. The number of iteration was two.

Figure 6　Simulated path for the experiments.

(a) The schematic diagram of the simulated self-motion in our experiments. We simulated a static point while translating in a fixed direction (θ) with respect to the current line-of-sight. (b) An example of the simulated path that is actually used in the experiments is shown. The heading direction is 10 deg. (c) If the observer translates in a fixed direction in the exocentric coordinates,　the heading direction changes with time in the retinocentric coordinates.

Figure 7　Results of Experiment 1.

The results of observer HI are shown. The horizontal axis represents the simulated heading direction. The vertical axis represents the perceived heading direction. (a) the ground condition (b) the cloud condition (c) the frontoparallel-plane condition.

Figure 8　Results of our model’s simulations.

The estimated heading is shown as a function of the simulated heading. We used four g values, which represent the magnitude of the additional vertical velocity component for the perturbation.(a) g=0 (b) g=20 (c) g=40 (d) g=60

Figure 9　Results of the simulations of Rieger and Lawton’s algorithm.

The estimated heading is plotted as a function of the heading. We used four g values, which represent the magnitude of the additional vertical velocity component. (a) g=0 (b) g=20 (c) g=40 (d) g=60

Figure 10　The slopes of perceived vs. simulated heading function in Experiment 2.

Parameter g represents the magnitude of the perturbation. (a) The circle(observer HI), square (observer MH) and triangle (observer YE) symbols show the results in experiment 2. Each point shows the average slopes of perceived vs. simulated heading function obtained in four sessions. The error bars indicate ± one standard error about the slopes obtained in the four sessions. (b) Slopes normalized at g=0 were shown. The dotted lines show the results of the simulations of the weighted average model.

Figure 11　The slopes of perceived vs. simulated heading function under the wide-display condition in Experiment 3.

Parameter g represents the magnitude of the perturbation. (a) The results under the wide-display condition in Experiment 3 are shown. Each point shows the slopes of perceived vs. simulated heading function at each g. (b) Slopes normalized at g=0 were shown. The dotted lines show the results of the simulations of the weighted average model. (c) The correlation coefficients between the perceived heading and the regression line are shown.

Figure 12　The slopes of perceived vs. simulated heading function under the narrow-display condition in Experiment 3.

Parameter g represents the magnitude of the perturbation. (a) The results under the narrow-display condition in Experiment 3 are shown. Each point shows the slopes of perceived vs. simulated heading function at each g. (b) Slopes normalized at g=0 were shown. The dotted lines show the results of the simulations of the weighted average model. (c) The correlation coefficients between the perceived direction and the regression line are shown.

Figure 13　Differential motion parallax.

Differential motion parallax is one of the source for recovery of heading. The highest retinal velocity across the line of fixation is opposite the direction of self-motion and indicates whether the heading is to

the left or to the right of the fixation point. The information can regarded as the deviation from the radial displacement.

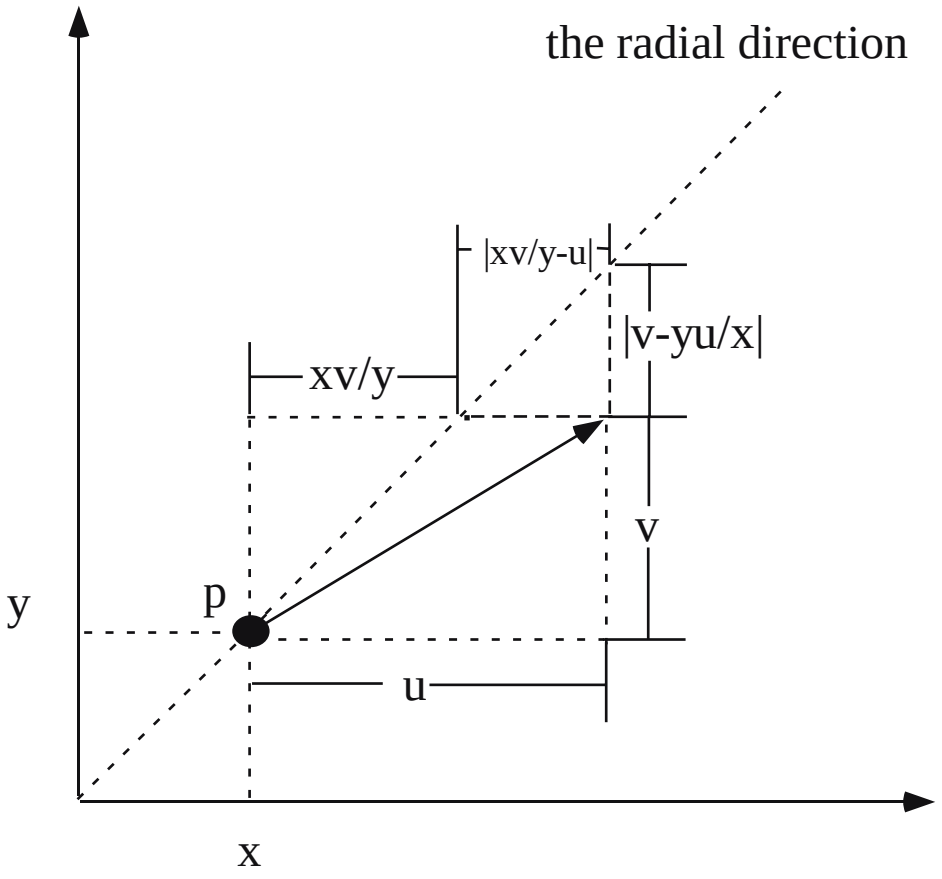


Figure 1

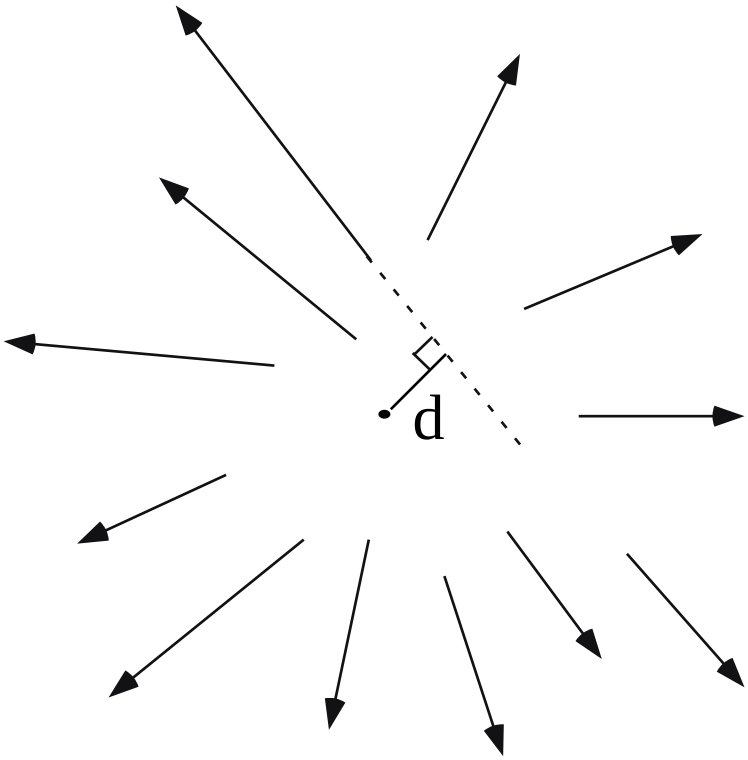
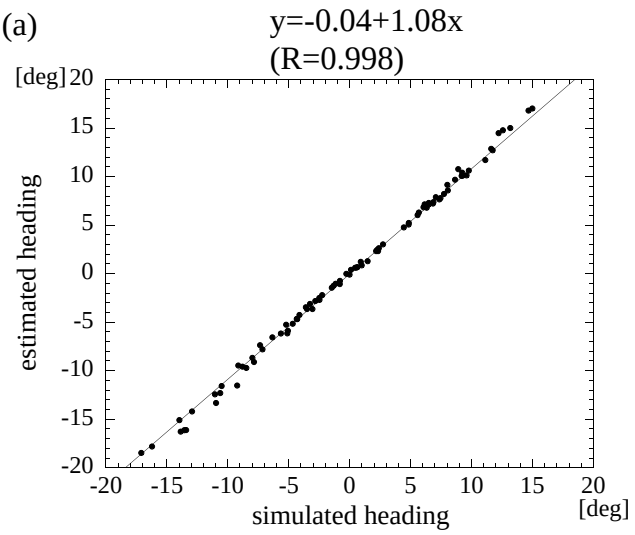


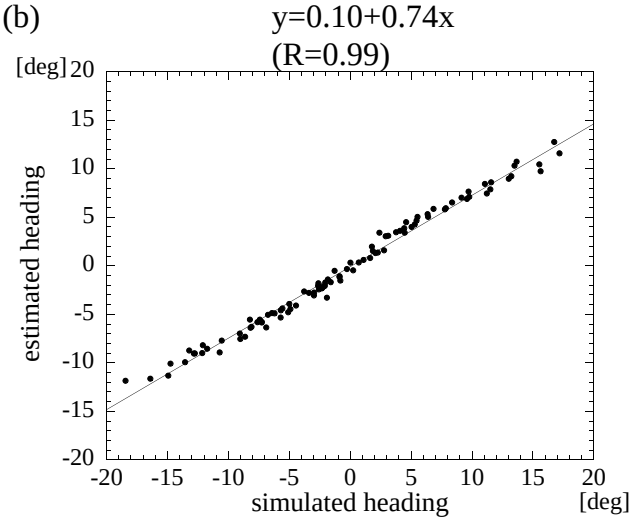
Figure 2

(a)



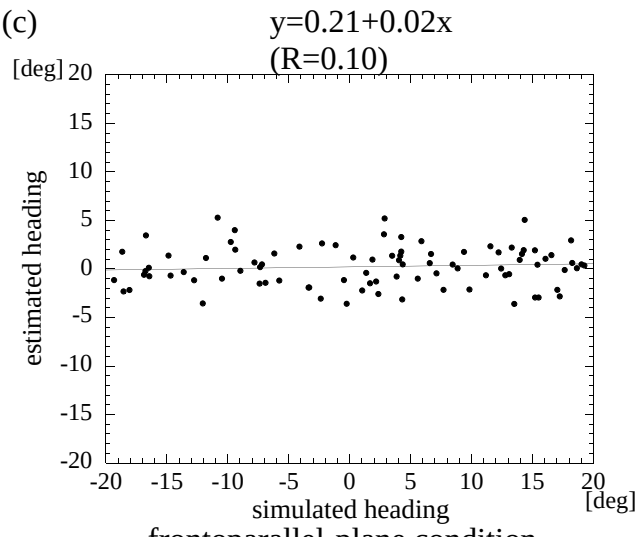
ground condition

(b)



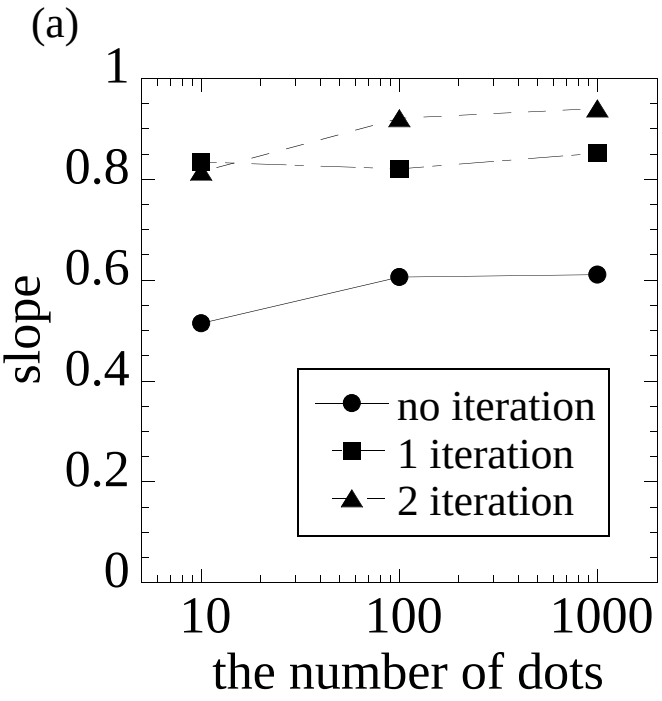
cloud condition

(c)



frontoparallel-plane condition

Figure 3


(a)
slope
1
0.8
0.6
0.4
0.2
0
10
100
1000
the number of dots
no iteration
1 iteration
2 iteration

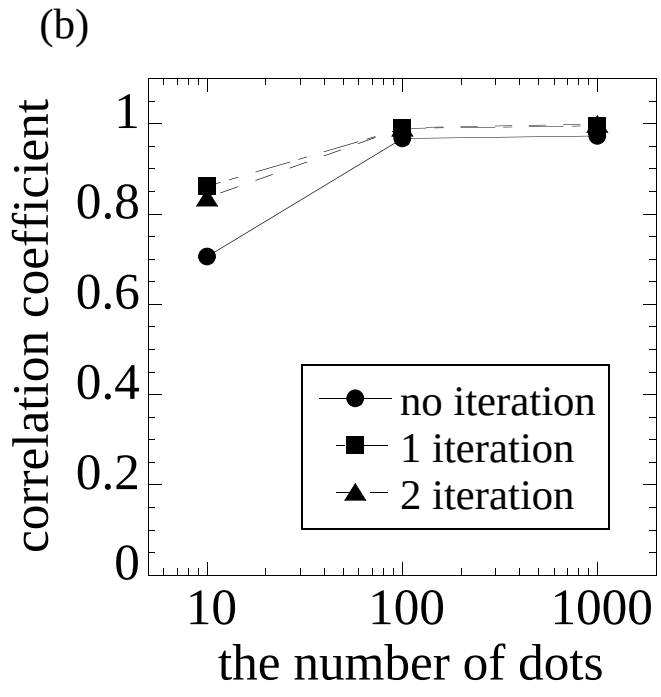
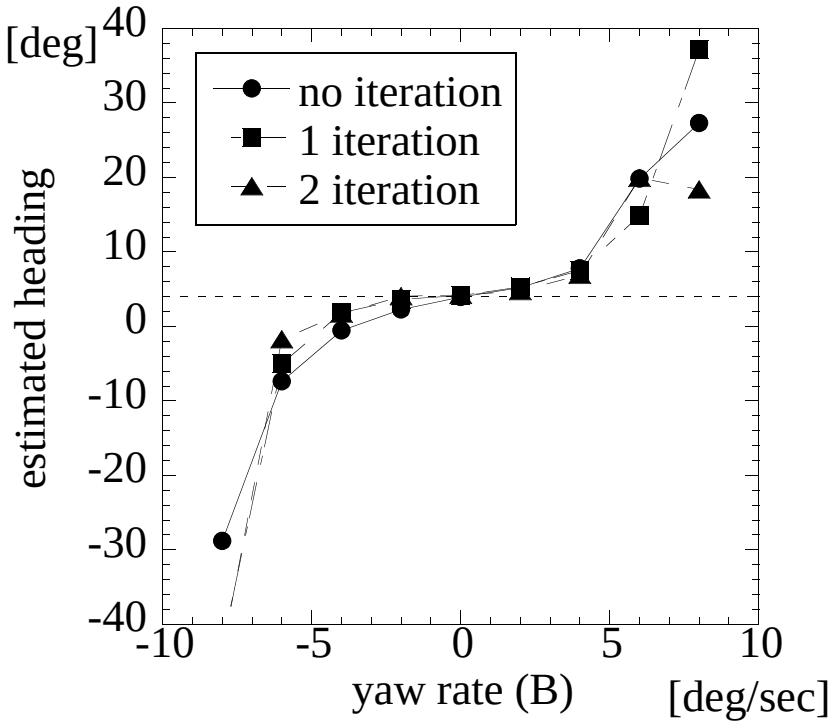

(b)
correlation coefficient
1
0.8
0.6
0.4
0.2
0
10
100
1000
the number of dots
no iteration
1 iteration
2 iteration

Figure 4

(a)


[deg]
40
30
20
10
0
-10
-20
-30
-40
estimated heading
no iteration
1 iteration
2 iteration
-10
-5
0
5
10
yaw rate (B)
[deg/sec]

(b)

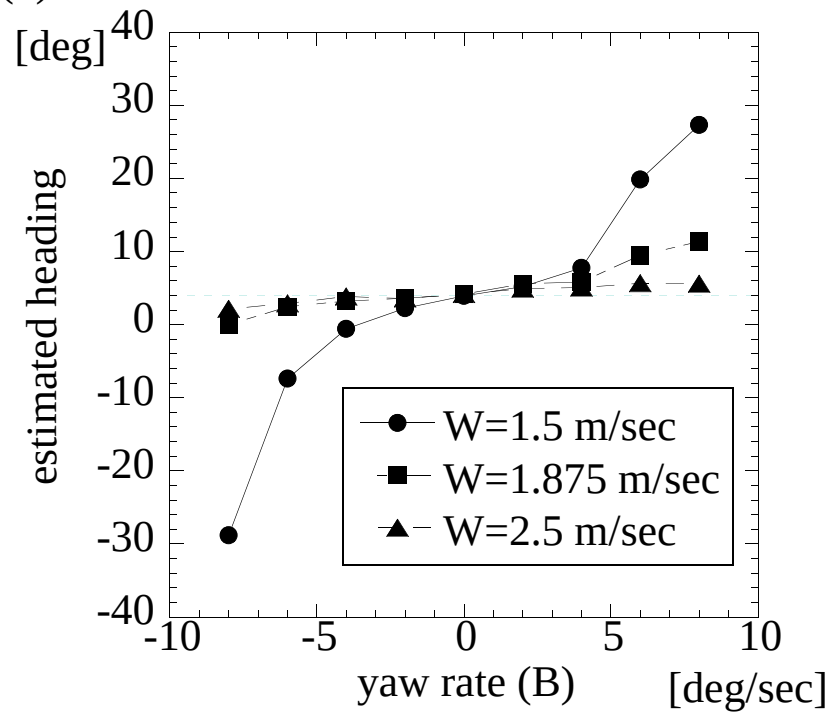
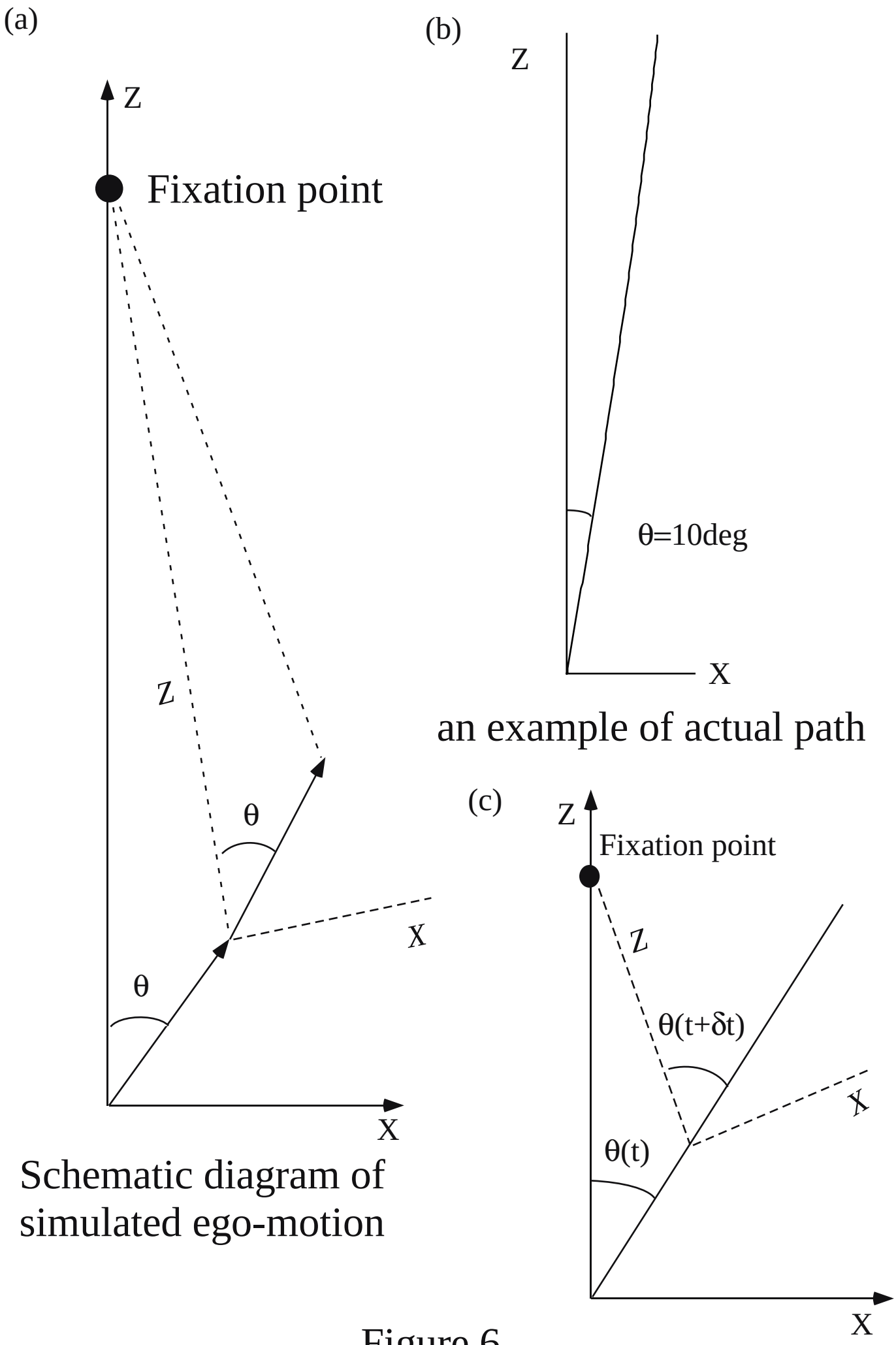

[deg]
40
30
20
10
0
-10
-20
-30
-40
estimated heading
W=1.5 m/sec
W=1.875 m/sec
W=2.5 m/sec
-10
-5
0
5
10
yaw rate (B)
[deg/sec]

Figure 5



Schematic diagram of simulated ego-motion

Figure 6

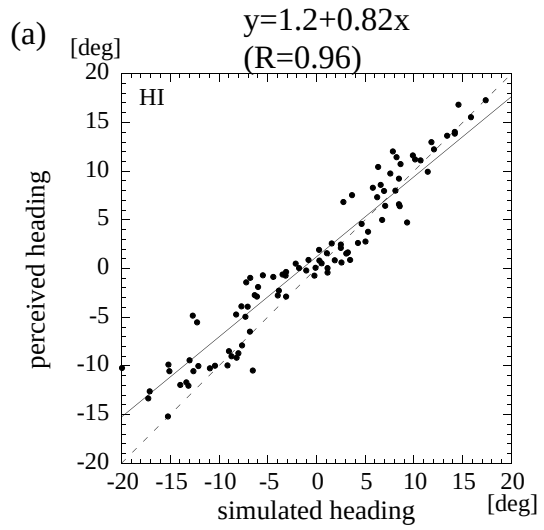

(a)
y=1.2+0.82x
(R=0.96)
[deg]
HI
perceived heading
simulated heading
[deg]
ground condition
(b)
y=0.54+0.76x
(R=0.87)
[deg]
HI
perceived heading
simulated heading
[deg]
cloud condition
(c)
y=-0.39+0.023x
(R=0.079)
[deg]
HI
perceived heading
simulated heading
[deg]
frontoparallel plane condition

Figure 7

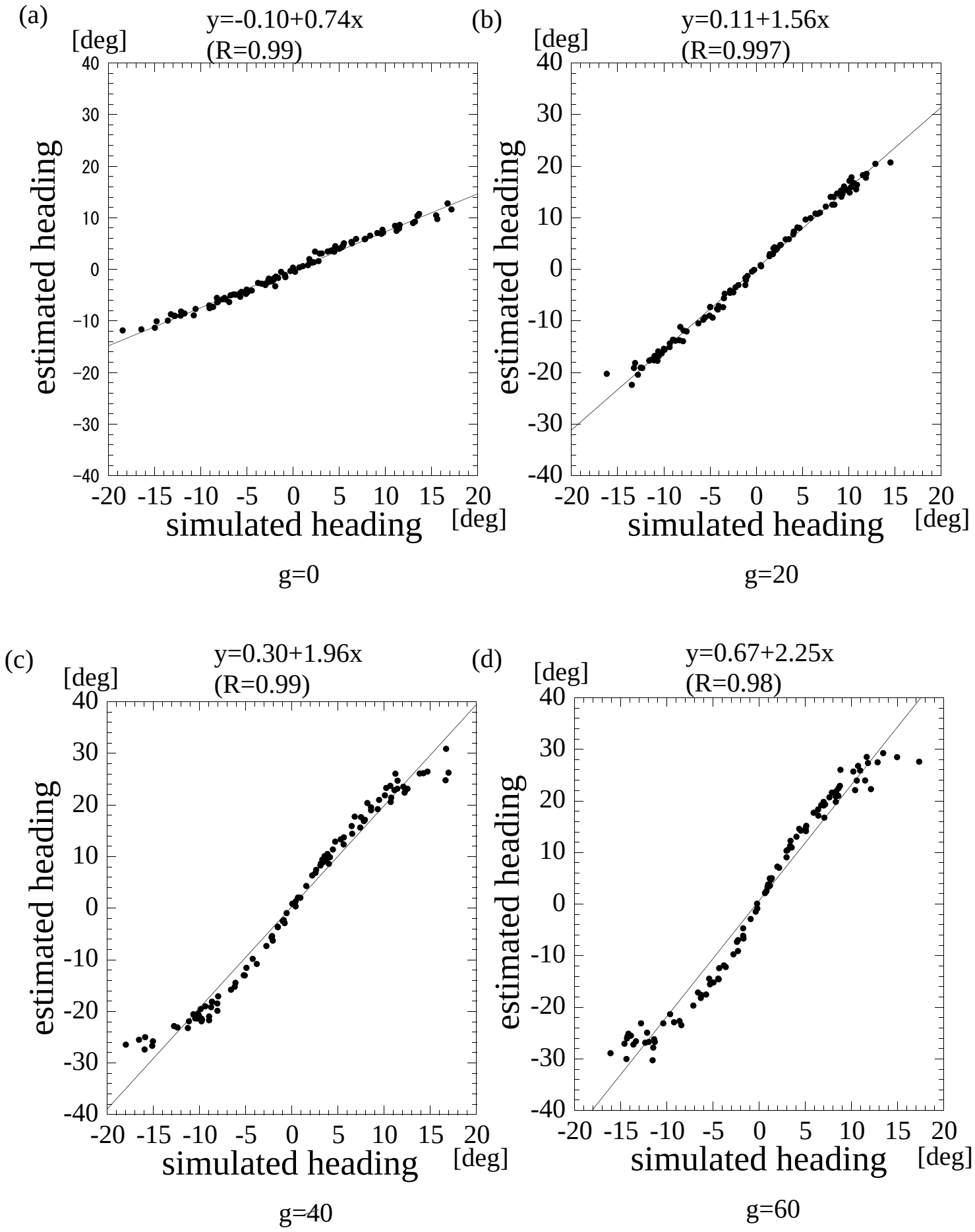

(a)
y=-0.10+0.74x
(R=0.99)
[deg]
estimated heading
simulated heading
[deg]
g=0
(b)
y=0.11+1.56x
(R=0.997)
[deg]
estimated heading
simulated heading
[deg]
g=20
(c)
y=0.30+1.96x
(R=0.99)
[deg]
estimated heading
simulated heading
[deg]
g=40
(d)
y=0.67+2.25x
(R=0.98)
[deg]
estimated heading
simulated heading
[deg]
g=60

Figure 8

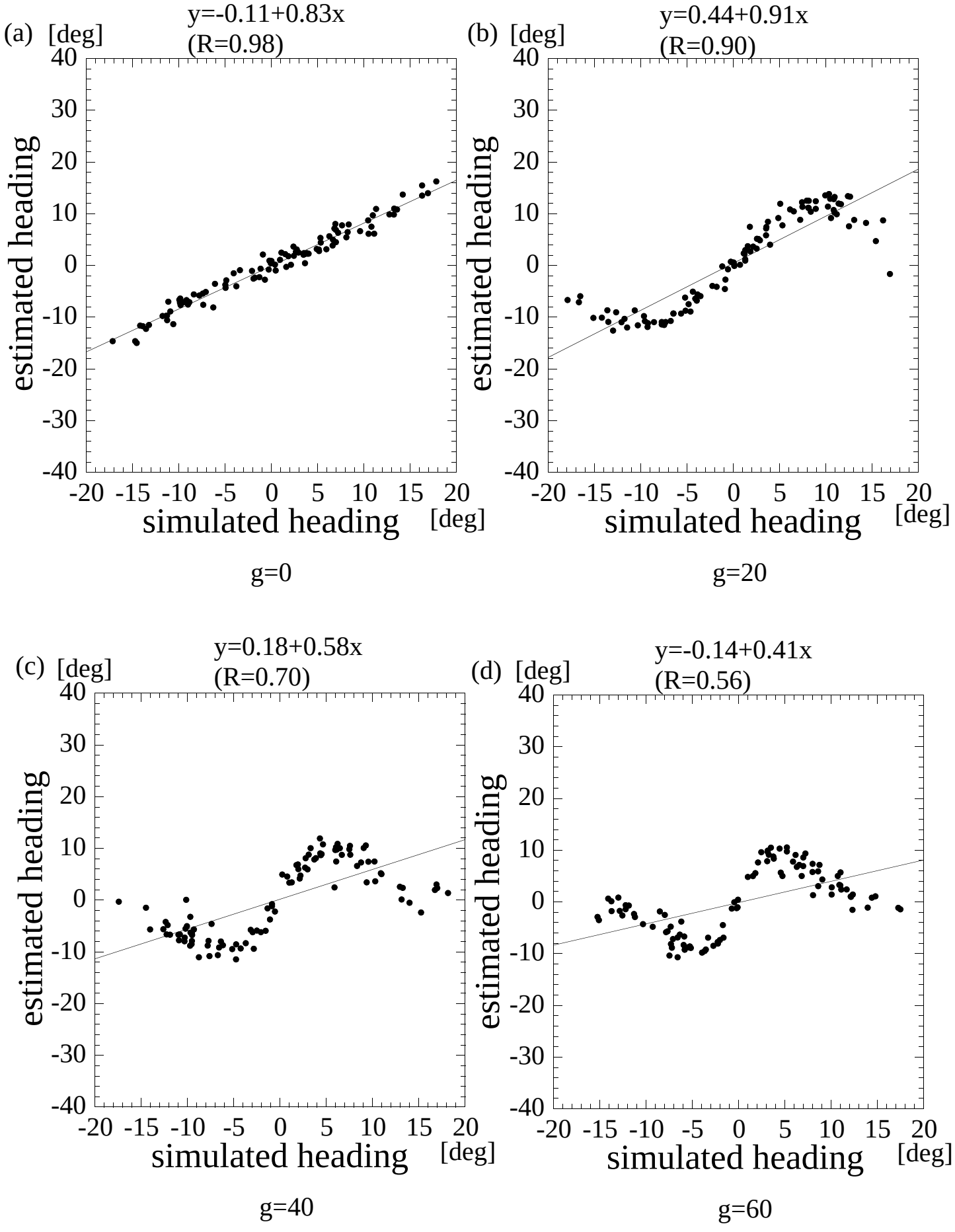

(a)
y=-0.11+0.83x
(R=0.98)
[deg]
estimated heading
simulated heading
[deg]
g=0
(b)
y=0.44+0.91x
(R=0.90)
[deg]
estimated heading
simulated heading
[deg]
g=20
(c)
y=0.18+0.58x
(R=0.70)
[deg]
estimated heading
simulated heading
[deg]
g=40
(d)
y=-0.14+0.41x
(R=0.56)
[deg]
estimated heading
simulated heading
[deg]
g=60

Figure 9

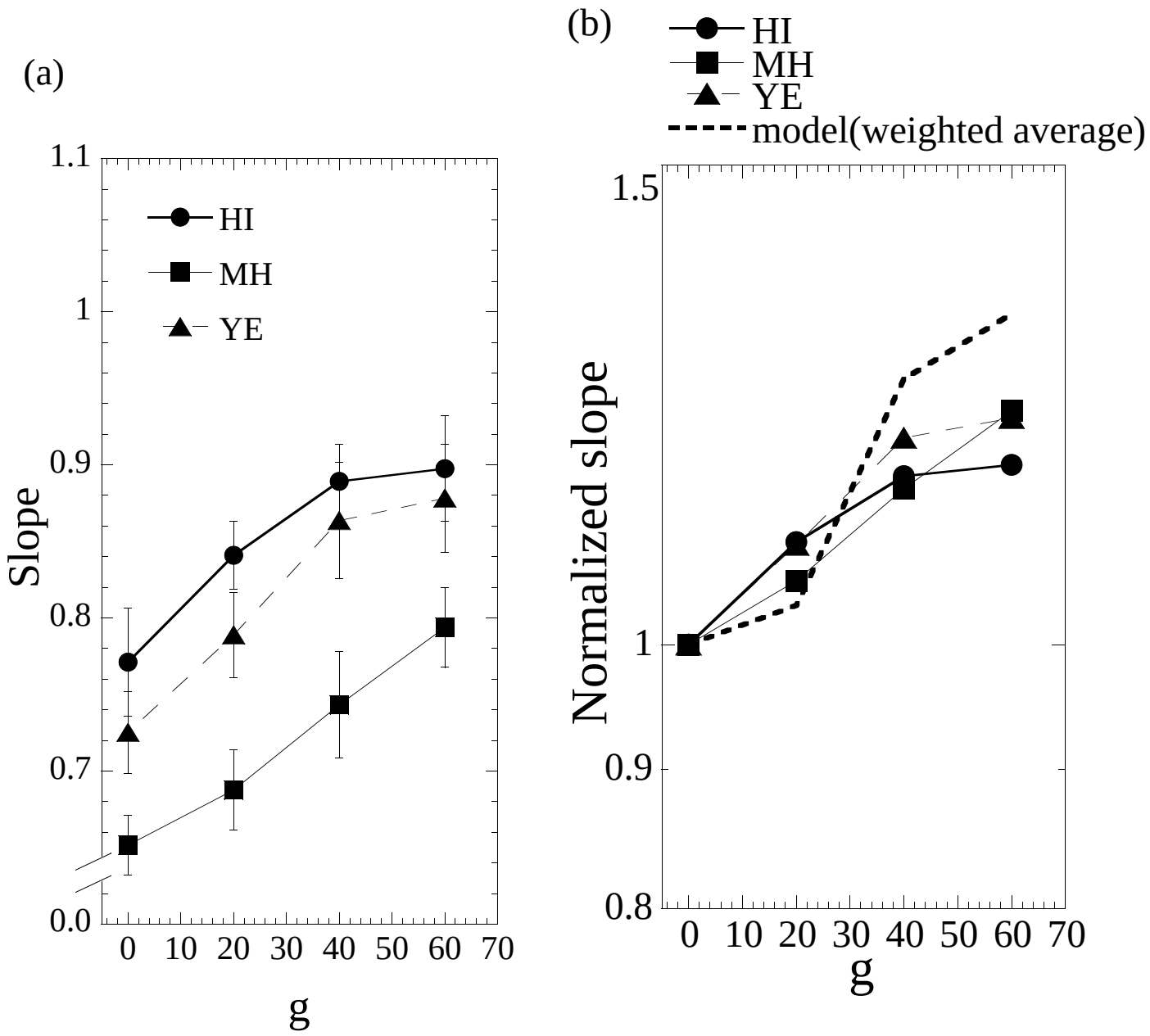

(a)
1.1
1
0.9
0.8
0.7
0.0
Slope
HI
MH
YE
0
10
20
30
40
50
60
70
g
(b)
HI
MH
YE
model(weighted average)
1.5
1
0.9
0.8
Normalized slope
0
10
20
30
40
50
60
70
g

Figure 10

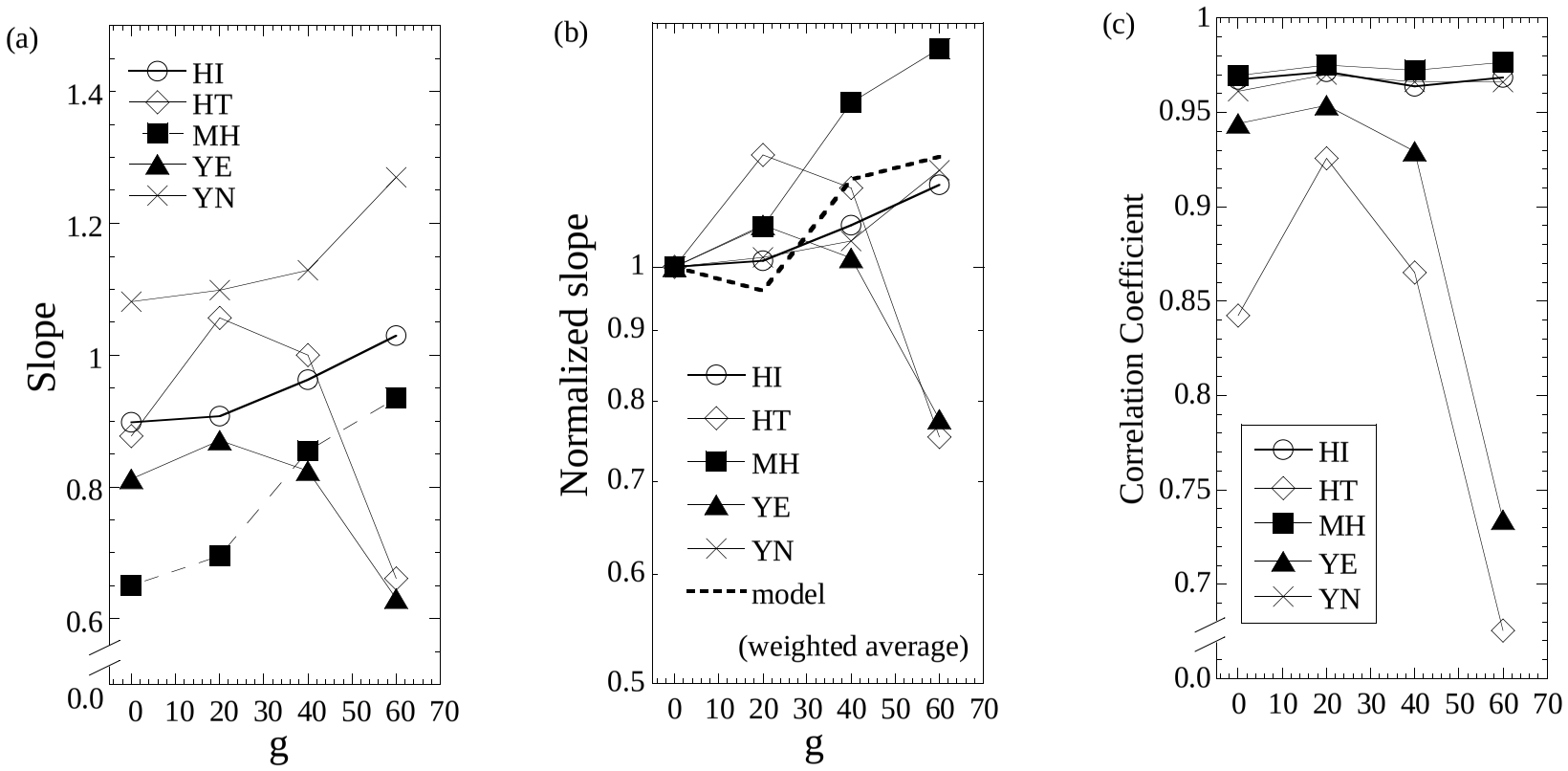

(a)
HI
HT
MH
YE
YN
Slope
g
(b)
Normalized slope
HI
HT
MH
YE
YN
model
(weighted average)
g
(c)
Correlation Coefficient
HI
HT
MH
YE
YN

Figure 11

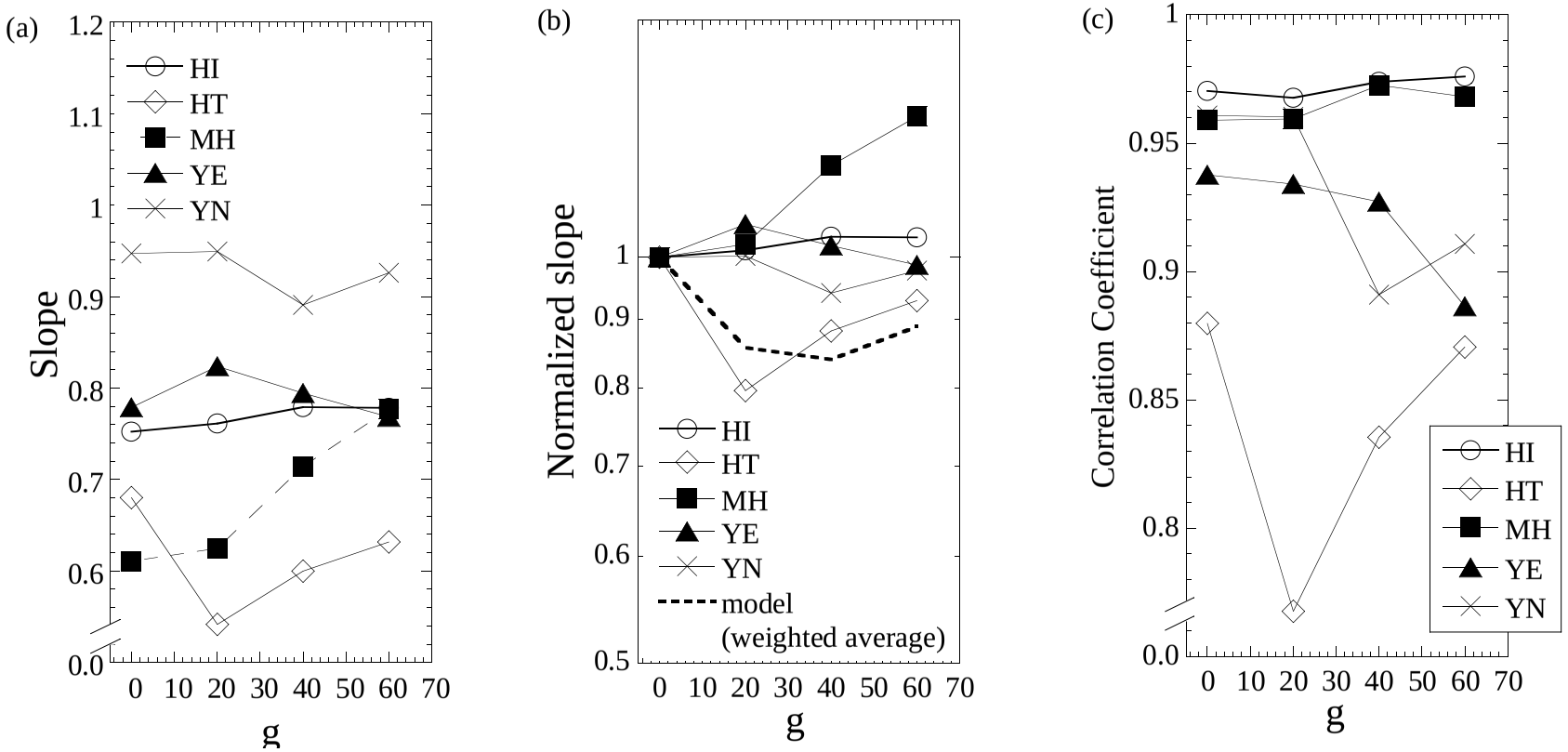

(a)
Slope
g
HI
HT
MH
YE
YN
(b)
Normalized slope
g
HI
HT
MH
YE
YN
model
(weighted average)
(c)
Correlation Coefficient
g
HI
HT
MH
YE
YN

Figure 12

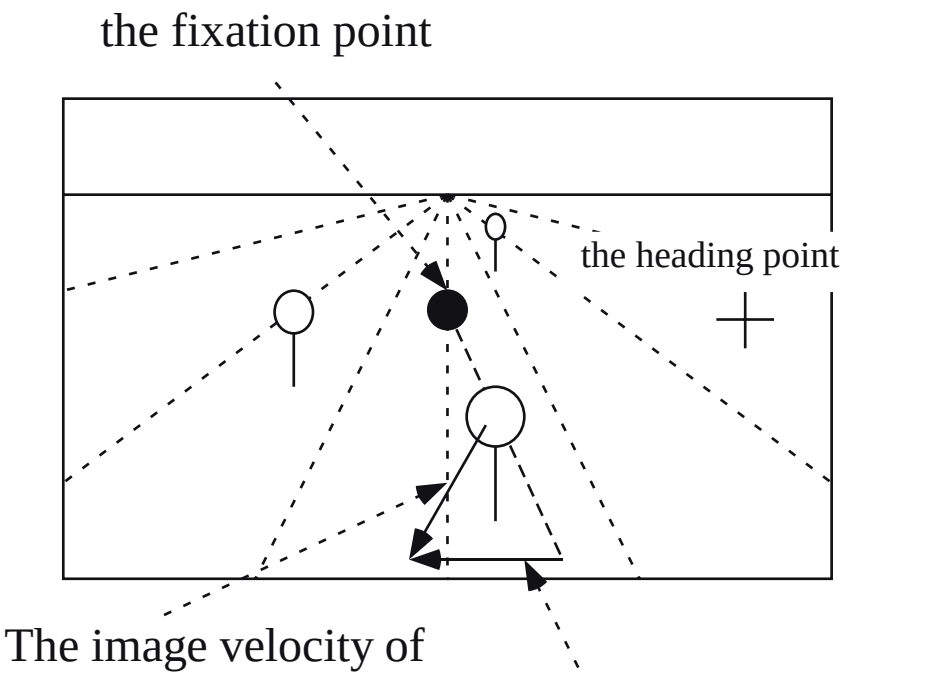

the fixation point
the heading point
The image velocity of
the fastest object
Differential motion parallax

Figure 13