Effects of roll and pitch components in retinal flow on heading judgement

Mitsuhiko Hanada and 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). Effects of roll and pitch components in retinal flow on heading judgement. Vision Research, 40 (14), 1827-1838

Keyword: heading, motion, optic flow, roll

Running Title: Heading Judgement with Roll and Pitch

Abstract

We investigated effects of roll (rotation around line of sight) and pitch (rotation around the horizontal axis) components of retinal flow on heading judgement from visual motion information. It was found that performance level of human observers for yaw (rotation around the vertical axis) plus pitch is little different from that for only yaw although there is bias in perceived heading toward the fixation point, and that heading judgement is fairly robust with respect to roll. It was also found that there are some observers who can perceive heading with pitch, yaw and roll at a roll rate of 11.5 deg/sec without extraretinal information. It suggests that there exist compensation mechanisms for roll in the human visual system.

Introduction

When an observer moves through the 3-dimentional environment, the visual image changes continuously. 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. In this paper, we will focus on the problem of recovering the heading direction from motion information.

Human observers can judge heading from a retinal flow pattern without extraretinal information such as efference copies of eye movement when a simulated rotation rate is small (< 1.5 deg/sec) (Warren & Hannon, 1990). When a rotation rate is high (> 1.5 deg/sec), human observers can judge heading accurately in some conditions (van den Berg, 1993; Stone & Perrone, 1997), but they show large bias in perceived heading in other conditions (Royden et al., 1992). The reason for the differences remains unclear. The necessary and sufficient conditions for accurate heading perception are still unknown.

When an observer does not rotate, the focus of expansion (FOE) corresponds to the heading direction (Gibson, 1950). When an observer rotates due to head, body or eye movement, it is not trivial to recover heading from the retinal flow alone. A large number of algorithms to recover heading have been presented for computer vision (e.g., Bruss and Horn, 1983; Louguet-Higgins and Prazdny, 1980; Kanatani, 1993) and several models of human heading perception have been proposed (e.g, Rieger & Lawton, 1985; Hildreth, 1992; Beusmans, 1993; Lappe & Rauschecker, 1993; Perrone, 1992; Royden, 1997; Beintema & van den Berg, 1998).

Rotation around the light of sight is called roll, rotation around the horizontally directed axis pitch and rotation around the vertically directed axis yaw (Fig. 1). Some models recover heading assuming no roll because it is very small when human observers usually move on the ground. Perrone and Stone (1994) presented a template-matching model to recover heading, which uses many templates of motion patterns to estimate the heading direction. No roll is assumed in order to reduce the templates. Lappe and Rauschecker (1995) presented an algorithm using the non-roll assumption based on subspace algorithm of Heeger and Jepson (1990; 1992). Royden (1997) presented a model of heading judgement using motion-opponent mechanisms in brain. Her model also assumes no roll. Roll components make heading estimates of the model unreliable.

Insert Figure 1 about here

However, a velocity field includes roll components when human observers incline their head. When a pilot navigates an airplane, roll often occurs. Kaiser & Hecht (1995) reported that human judgement of the time until an approaching object will pass the observer, is fairly robust with respect to roll. Freeman, Harris & Tyler (1994) reported that estimates of time-to-collision are unaffected by the addition of circular motion which is generally induced by roll. These reports suggest that the human visual system has a compensation mechanism for roll. However, it has not been examined whether human observers can perceive heading from the flow field with roll components. It is

necessary to test the validity of the non-roll assumption for the models psychophysically.

Rieger and Toet (1985) used stimuli with pitch components and reported that human observers accurately judged the direction of heading relative to the fixation mark when the rotation rate in their psychophysical experiment was small (< 2.0 deg/sec). However, human heading judgement has not been examined so much using stimuli which include large pitch components. In this study, we examined effects of pitch as well as roll components in retinal flow on heading judgement.

Experiment 1

Two types of self-motion paths have been used in psychophysical experiments to examine effects of self-rotation (yaw), a) straight paths in the environment in which the instantaneous heading changes during the presentation with respect to the line of sight (Warren & Hannon, 1990; van den Berg, 1993, 1996; Royden, 1994), and b) curved paths in which the instantaneous heading is a fixed angle with respect to the line of sight (Stone & Perrone, 1997; Hanada & Ejima, in press). In this study, we examined effects of roll components in the flow field on human heading judgement using paths of b), which was generated by simulating horizontal and forward translation in a fixed direction with respect to the line of sight while fixating a static point and rolling.

We briefly introduce the mathematical expression of self-motion to explain the simulated observer's movement for the stimulus generation. We use a method by Longuet-Higgins and Prazdny (1980) to calculate the velocity of dots in the image plane. 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 (Fig. 1). The rotation of the observer can be expressed by rotation around three orthogonal axes, which we express by the vector (A, B, C). A, B and C, which are rotation around the X-axis, Y-axis and Z-axis, indicate pitch, yaw and roll respectively (Fig. 1). The 3-D velocity of a point, P(X,Y,Z) relative to the observer is given by:

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

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

$$\begin{aligned} x &= \frac{X}{Z} \\ y &= \frac{Y}{Z} \end{aligned} \qquad (2)$$

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

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

When we consider situations where the observer moves for a period, we must rewrite (1) taking time into account as follows.

$$\begin{aligned}\dot{X}(t) &= -U(t) - B(t)Z(t) + C(t)Y(t)\\ \dot{Y}(t) &= -V(t) - C(t)X(t) + A(t)Z(t)\\ \dot{Z}(t) &= -W(t) - A(t)Y(t) + B(t)X(t)\end{aligned} \qquad (4)$$

where t shows time. In most psychophysical experiments, stimuli were generated by simulating forward and horizontal translation without pitch and roll. It means that A(t), C(t) and V(t) are 0. If B(t), U(t) and W(t) are constant in the observer-fixed coordinates, Eq. (4) becomes:

$$\begin{aligned}\dot{X}(t) &= -U - BZ(t)\\ \dot{Y}(t) &= 0\\ \dot{Z}(t) &= -W + BX(t)\end{aligned} \qquad (5)$$

The point P (X(t), Y(t), Z(t)) moves on a circle relative to the observer following the differential equations (5) if we consider that the observer is stationary. If we consider that the observer is moving

and the point P is stationary, the observer moves on a circular path in the environment (Royden, 1994). Stone & Perrone (1997) used this circular path in their psychophysical experiments. The subjects for their experiments were required to judge the heading direction relative to the line of sight, which was constant in the retinocentric (observer-fixed) coordinates, and the judgment was accurate for a yaw rate of 16 deg/sec. These results show that human observers can judge the retinocentric heading direction. We modified this paradigm in this study.

Hanada & Ejima (in press) generated stimuli by simulating a situation where the observer translated in a fixed direction with respect to the current line-of-sight while fixating a static point as shown schematically in Fig. 2(a). An example of the actual path is shown Fig. 2(b). 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) (Lappe & Rauschecker, 1995). Therefore we obtain the following equations from (3):

$$\begin{aligned} U &= -BZ_f \\ V &= AZ_f \end{aligned} \qquad (6)$$

In this situation, A(t), C(t) and V(t) in (4) are 0, and U(t) and W(t) are constant in the observer-fixed coordinates. Therefore, from (4) and (6), the point P (X, Y, Z) moves relative to the observer according to the following differential equations:

$$\begin{aligned} \dot{X}(t) &= -U - B(t)Z(t) \\ \dot{Y}(t) &= 0 \\ \dot{Z}(t) &= -W + B(t)X(t) \\ \dot{Z}_f(t) &= -W \\ B(t) &= -U / Z_f(t) \end{aligned} \qquad (7)$$

Insert Figure 2 about here

The path is not a circle because the curvature radius is $(U^2+W^2)^{1/2}/B(t) = (U^2+W^2)^{1/2}U/Z_f(t)$ (Royden, 1997), and changes with time as $Z_f(t)$ decreases with time.

Here we add roll to this observer's motion. We show a schematic diagram of this observer's motion in the world-centered coordinates in Fig. 3 (a). If C(t) is constant, but not 0, we obtain the followings from (4) and (6) instead of (7):

$$\begin{aligned}
\dot{X}(t) &= -U - B(t)Z(t) + CY(t) \\
\dot{Y}(t) &= -CX(t) \\
\dot{Z}(t) &= -W + B(t)X(t) \\
\dot{Z}_f(t) &= -W \\
B(t) &= -U / Z_f(t)
\end{aligned} \qquad (8)$$

We used Eqs. (8) to generate stimuli in this experiment. From (8), we can calculate the movement of the point relative to the observer in the observer-fixed coordinates. Therefore we can obtain the projected velocities of P using (2) and (8). It means that we can generate the stimuli by calculating movements of the objects in the observer-fixed (retinocentric) coordinates. It is not required to calculate the path in the world-center coordinates for the generation of the images, though it is possible

to calculate the path in the environment numerically. Note that the stimuli generated by (2) and (8) were the same as the display generated by simulating the situation where the observer translated in a fixed direction with respect to the current line of sight while rolling and fixating a static point. An example of the actual paths for no roll (C=0) is shown in Fig. 2(b). If C is not 0, the observer moves along a complex curved path in the 3-D environment as schematically shown in fig. 3 (a). Roll(C) in (8) induced torsion because $\dot{Y}(t)$ is not 0, which means that the path is not on a plane. For the movement along the path, the heading angle is constant in the retinocentric (observer-fixed) coordinates. We generated stimuli using (8) and (2) in this experiment.

Insert Figure 3 about here

Straight paths in the world-center environment have been used in many psychophysical experiments (Warren & Hannon, 1990; Royden et al., 1994; van den Berg, 1993; 1996). When the observer moves along a straight path in the environment while (s)he fixates a static point, the direction of heading is not constant in retinocentric (observer-fixed) coordinates (Fig. 2(c)). It is problematic for the observer's judgement because the direction of heading relative to the fixation point changes during the stimulus presentation. We cannot neglect effects of the change of the heading angle in the retinocentric coordinates when roll is large. Therefore we did not use the path like Fig. 2(c).

Methods

Observers: One of the authors (MH) and an undergraduate student (HI) participated in this experiment. The observers wore corrective spectacles or contact lenses during the experiment to achieve normal acuity. HI was naïve as to the hypothesis under study. Both observers had participated in other types of psychophysical experiments on heading judgement. HI had not seen simulated roll plus translation display in experimental situations before.

Apparatus: The observer was seated and his head stabilized with a chinrest. Experiments were conducted using a Silicon Graphics O2 workstation with a color monitor. Computer-simulated motion sequences were monocularly presented at a framerate of 60Hz. The image on the screen was 34.4cm wide (1280 pixels) and 27.5cm (1024 pixels) from top to bottom. At the viewing distance of 40cm, the screen subtended 46.5deg horizontally × 38 deg vertically. Apart from the stimuli, the room was dark.

Stimuli: The simulated environment consisted of 100 randomly located white dots of 58 cd/m^2 which were configured in a cloud. The dots size was 2 × 2 pixels. The simulated world extended in depth from 4 to 8m in front of the observer's eye. Dots were randomly located in the viewing frustum. 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. One red dot served as a fixation point. The point was always at the center of the screen. The observer was asked to fixate the point and not to move their eye

during the presentation of simulated self-motion. However, the observer saw stimuli that simulated translation and self-rotation. Simulated self-motion was presented for 2.0 sec. After the presentation of simulated self-motion, all dots but the fixation point disappeared.

We generated images using (8) and (2). Point paths in a condition are shown in Fig. 3(b). Translation component U was randomly set to a value between -0.25 and 0.25 m/sec, and W was randomly set between 0.75 m/sec and 1.25 m/sec for each trial. V was 0. The fixation point was a point among cloud dots and its initial distance (Z_f) was determined between 4 and 8m. The yaw rate, which depends on Z_f and U as shown in Eq. (2), varied for each trial. Yaw (B) was less than 10 deg/sec. In almost all trials, however, the yaw rate was less than 5 deg/sec during the presentation. Nine values of roll were simulated; 0, ±5.7, ±11.5, ±17.2 and ±22.9 deg/sec. Fifty trials were conducted for each roll condition and there were 450 trials in a session. The session lasted about 40 min at most.

Procedures

The observers had some training sessions with stimuli which simulated egomotion without roll. The actually simulated direction was shown after the response in the training sessions, but not in the proper sessions. The simulated motion and path were explained to the naïve observer well.

The observers experienced a clear impression of relative motion between the self and the simulated environment and did not perceive changes of the heading angle relative to the line of sight during the stimulus presentation. They were asked to adjust the position of a pointer horizontally by

moving a mouse so as to indicate the perceived heading angle relative to the line of sight (θ in Fig. 3). The trial was terminated by the observer's response (a mouse click). The observers performed the heading judgement relative to the fixation point well after a few practice sessions.

Results

Pointing responses of the observers are shown in Fig 4 (observer HI) and Fig. 5(observer MH). The results of negative roll (counterclockwise rotation) are not shown here. They were similar to those of the positive roll with the same absolute value. Perceived heading is plotted as a function of the simulated heading direction. Thus, a point at the origin indicates that the observer perceived heading towards the fixation point. Each data point indicates the result of one trial. When the points are on the line with slope of 1.0, heading perception is unbiased.

Insert Figure 4 and 5 about here

The observer HI fairly correctly judged the heading direction when no roll was simulated, although he showed small bias toward the fixation point. MH showed relatively large bias toward the fixation point for no roll. However, there was little difference in performance between 0, 5.7 and 11.5 deg/sec roll for both observers. For the rate of 22.9 deg/sec, HI accurately judged whether the heading

was leftward or rightward relative to the line of sight. The bias of MH increased as the roll increased to 22.9 deg/sec.

We conducted a linear regression analysis. Deviation from slope of 1 shows the bias, and a low correlation coefficient between the regression line and the data points indicates variability of data. The correlation coefficients of HI were high and the slopes of HI were more than 0.78 for all roll rates. The correlation coefficients of MH were fairly high, for all roll rate (>0.7). The slopes of the regression line decreased for MH as a roll rate increased. It indicates that bias in perceived heading for MH increased with the increasing roll rate though the variability was rather small.

Discussion

Performance of heading judgement changes little up to 11.5 deg/sec roll. This result is inconsistent with the assumption of no roll used in several models of heading perception. It indicates that the human visual system recovers heading from visual information assuming some roll. One of the observers (HI) precisely judged whether the heading direction was leftward or rightward from his line of sight for a roll rate of 22.9 deg/sec as seen in Fig. 4, and clearly the performance of the other observer (MH) was significantly above chance. The results in this experiment show that humans can judge the heading direction reasonably accurately at a relatively high rotation rate of roll.

Both observers showed some underestimation of the heading direction even when no roll was simulated. The bias toward the center of screen or the fixation point was often reported in the

simulated eye-movement condition (van den Berg, 1996; Cutting et al, 1997; Ehrlich et al, 1998). Some researchers argued that the bias occurs because human observers perceive a curved path in the simulated eye-movement condition and predict the future path even when a straight path in the world-center coordinates is simulated (Royden, 1994; van den Berg, 1996). In our experiment, the heading direction relative to the line of sight was constant during the stimulus presentation. The observers judged heading relative to the line of sight and were not required to predict the future path for the task. Since the perception of the curved path has little effect on the judgement, the bias observed in our experiment cannot be explained by the perceived curved path. Several models of human heading recovery successfully explain the bias with stimuli which simulate a small depth range (Hildreth, 1992; Beusmans, 1993; Hanada & Ejima, in press). The bias appears to arise from the limitation of computation in the human visual system.

Experiment 2

Effects of yaw components have been studied well on human heading judgement although effects of pitch components have not been examined sufficiently. Rieger and Toet (1985) reported that human observers judged accurately whether the heading direction was left or right, and up or down relative to the line of sight from the retinal flow with pitch components. However, the flow pattern with no pitch or with small pitch components has been used in most psychophysical studies. In this experiment, we examined the effects of pitch components on human heading judgement from the flow

pattern. If an observer fixates a static point in the environment, Eq. (6) holds. Pitch (A) is not 0 unless V is 0. Thus, flow fields were generated by simulating situations where an observer translates in a forward plus horizontal plus vertical direction, (i.e., an observer translates toward a certain position in the screen), while fixating a static point.

Methods

The same apparatus and procedures as in Experiment 1 were used. Stimuli were generated by simulating situations where an observer translated in a fixed direction in the screen while fixating a point without rolling. In this situation, the point P(X(t), Y(t), Z(t)) moves relative to the observer according to the following the differential equations:

$$\begin{aligned}
\dot{X}(t) &= -U - B(t)Z(t) \\
\dot{Y}(t) &= -V + A(t)Z(t) \\
\dot{Z}(t) &= -W - A(t)Y(t) + B(t)X(t) \\
\dot{Z}_f(t) &= -W \\
A(t) &= V / Z_f(t) \\
B(t) &= -U / Z_f(t)
\end{aligned} \qquad (9)$$

We used Eqs. (2) and (9) to generate stimuli. In the world-center coordinates, we simulated situations where an observer moved along a path like in Fig. 2(b) in a plane of Y= (V/U)X. The observers did not

perceive change of the heading direction relative to the line of sight during the stimulus presentation.

U was set a value between -0.25 m/sec and 0.25 m/sec, and V was set between -0.20 m/sec and 0.20 m/sec for each trial. W was chosen from 0.75m/sec to 1.25m/sec. A narrower range of V was used than that of U because the screen does not sometimes include the heading point if the range of –0.25 to 0.25 m/sec is used. Pitch was less than 8 deg/sec. For almost all trials, however, the pitch rate was less than 4 deg/sec during the presentation. The observers were asked to indicate the perceived heading angle relative to the line of sight. 100 trials were conducted in a session.

Results and discussion.

The results of pointing responses are shown in Fig. 6. Observer HI judged heading fairly accurately except for one trial. Observer MH showed relatively large bias toward the center of the screen although the variability is fairly small. When the vertical component of heading was smaller than 2.0 deg, the response of MH was drawn to the x-axis. Concerning the horizontal component of heading, the performance of both observers was little different from that for no roll in Experiment 1.

We conducted the regression analysis. The correlation coefficients are very high (>0.85). for vertical and horizontal components of heading. The slopes for HI were larger than 0.8 . The observer judged the heading direction fairly accurately.

Insert Figure 6 about here

These results indicate that performance of human observers does not worsen when pitch components are added. One of the observers (HI) judged heading fairly accurately from the flow field with pitch as well as yaw without retinal information, while the other (MH) showed fairly large bias. It shows that there are fairly large individual differences in bias for human heading judgment.

Experiment 3

In this experiment, we examined effects of roll on judgement of the horizontal plus vertical heading. Stimuli were designed to include all components of rotations.

Methods

We added roll components to the flow field used in Experiment 2. We generated stimuli by simulating the situation where the observer translated in a fixed forward plus horizontal plus vertical direction with respect to the current line-of-sight while fixating a point and rolling. Therefore we used the following differential equations to generate the stimuli:

$$\begin{aligned}
\dot{X}(t) &= -U - B(t)Z(t) + CY(t) \\
\dot{Y}(t) &= -V - CX(t) + A(t)Z(t) \\
\dot{Z}(t) &= -W - A(t)Y(t) + B(t)X(t) \\
\dot{Z}_f(t) &= -W \\
A(t) &= V / Z_f(t) \\
B(t) &= -U / Z_f(t)
\end{aligned} \tag{10}$$

The apparatus and procedure were the same as in Experiment 1. The observers were asked to indicate the perceived heading angle relative to the line of sight.

Results and discussion

Three naïve observers and one of the authors participated in the experiments. We show the results of only two naïve observers. The results of the other observers had tendencies similar to the two observers'. Results are shown in Fig 7 and Fig. 8. The results of HI for no roll in this experiment are not shown, because the condition was exactly the same as that for Experiment 2 and they were very similar to the results in Experiment 2 shown in Fig. 6. The results of negative roll (counterclockwise rotation) were not shown here because they were similar.

Observer HI judged the horizontal component of heading fairly accurately up to 17.2 deg/sec roll (Fig. 7) . For the vertical component of heading, the variance of the data was large when the roll rate was larger than 5.7 deg/sec. The correlation coefficient in the regression analysis was very high ($R>0.8$) and the slope was larger than 0.69 up to 17.2deg/sec roll. It implies that the performance was maintained up to 17.2 deg/sec roll.

The results of YN are shown in Fig. 8. The observer judged the heading direction reliably though the variability of the response for the observer was relatively high. The performance was fairly high up to 11.5 deg/sec roll though the bias for YN was larger than for HI. The observer could not perceive heading accurately when the roll rate was larger than 17.2 deg/sec. The correlation coefficient

between the regression line and the data points was high up to a roll rate of 11.5 deg/sec (R>0.75). The slopes were more than 0.67 up to 11.5 deg/sec roll.

Insert Figure 7 and 8 about here

These results indicate human observers can judge their heading fairly accurately from the flow including pitch, yaw and roll components and their judgement is robust to roll, though there were individual differences in bias of perceived heading toward the line of sight. The results seem to be inconsistent with the non-roll assumption for models of human heading recovery from motion.

General discussion

We examined the effects of pitch and roll on human heading judgement from retinal flow. We have found that performance level of human observers for yaw plus pitch is little different from that for only yaw and that performance changes little up to 11.5 deg/sec roll. We have also found that there are some observers who can perceive heading with pitch, yaw and roll at a roll rate of 11.5 deg/sec without extraretinal information.

Validity of non- roll assumption

Roll is generally small when humans translate on the ground. However, roll components

enter the flow pattern easily when the observers incline their head. For that reason, it is unlikely that the human visual system assumes no roll for recovery of self-motion. However, several models of human heading perception assume no roll. Lappe and Rauchecker (1995) presented a subspace algorithm with the non-roll assumption, based on the least-square algorithm of Heeger and Jepson (1990; 1992). We performed simulations of the algorithm with the non-roll assumption and examined performance when self-motion with roll was simulated (see Appendix for detail of the algorithm and our implementation). Input to the calculation was the position and velocity of 100 dots at a depth range of 4 – 8m. No noise was added. We computed input velocities by simulating the same situations as at the beginning of the stimulus presentation in Experiment 3. We used the simulated heading direction as initial value for the minimization of the residual function. We adopted the initial value to avoid a local minimum of the residual function. Fig. 9 shows the results. When no roll was simulated, recovery of the heading direction was indeed perfect. When 5.7 deg/sec roll was simulated, the estimates became worse. On the other hand, 5.7 deg/sec roll had little effect on performance of human observes in our experiment. For a roll rate of 11.5 deg/sec, the algorithm could not estimate the heading direction though human observers can perceive heading clearly at the roll rate. It is evident that human performance for 11.5 deg/sec roll was much higher than performance of the algorithm with the non-roll constraint. The correlation coefficients between the regression line and the data points obtained by the algorithm were lower than human data for roll of 11.5 deg/sec. Therefore this least-square algorithm with the non-roll assumption is inappropriate for human model.

Insert Figure 9 about here

Perrone (1992) presented a model of template matching. Perrone and Stone (1994) presented a biological neural model based on Perrone's original model, which recovers heading using no roll and gaze stability constraints. The constraints were used to reduce the number of templates. Crowell (1997) compared estimation of the model and human performance, and showed that the performance of the model with the gaze-stability assumption was not consistent with human data. We showed that the human visual system does not use the non-roll assumption. It means that the model needs templates for roll.

Royden (1997) presented a model which recovers heading using motion-opponent mechanisms in MT of primate brain. The model also assumes no roll. Since humans can recover heading with roll over 10 deg/sec from motion information alone, the model needs some mechanisms to compensate roll components, although the model appears to show relatively large robustness to roll.

If roll is estimated accurately, the components can be removed from a flow field. We presented a simple method for estimating roll without knowing the heading direction (Hanada & Ejima, in press). If we remove roll components after the estimation, we can assume no roll. This technique can be used for the models using the non-roll assumption.

Physiological mechanisms to estimate roll

Saito et al. (1986) found cells in MST of monkey brain which respond to circular motion patterns. Tanaka et al. (1989) suggested that the cells responding to circular motion play an important role in analyzing self-motion from motion because they have large receptive fields. Bradley et al. (1996) and Britten & Wezel (1998) reported cells that are related to heading perception and eye movement. From these reports, it is plausible that roll is estimated by the cells in MST. Actual computation of heading in MST, however, has not been known sufficiently. Perrone (1992), and Lappe & Rauschecker (1993) presented neural models of MST for recovery of heading. The cells in their neural networks show response similar to those of cells in MST. Zemel and Sejnowski (1998) presented a neural network which was trained with an unsupervised learning rule and the units showed many of the known properties of MST cells. Hence neural mechanisms of heading recovery in MST are still controversial.

Effects of yaw, pitch and roll on heading judgement.

Previous studies have examined the effects of yaw on heading judgement in the simulated eye movement conditions. It has not been clearly known up to how much rotation rate humans can judge heading from retinal flow alone. Under some conditions, humans perceive heading accurately for a simulated yaw rate over 5.0 deg/sec (van den Berg, 1993; Stone and Perrone, 1997), but under other conditions they do not (Royden et al, 1994; Ehrlich et al., 1998). The bias toward the fixation

point is usually observed when fixation to a static point is simulated. van den Berg and Brenner (1994) showed that fairly large bias occurred as the depth range was reduced. In the case of the least depth range, namely, in the case of egomotion toward a frontoparallel plane, heading judgement only from visual information is almost impossible and the strong bias toward the center of a centrifugal structure arises (Warren & Hannon, 1990; Royden et al, 1994; Stone & Perrone, 1997). The bias in perceived heading seems to be affected by the following factors.

(a) Simulated depth range.

(b) Ratio of the rotation rate to the translation speed (Stone & Perrone, 1997).

(c) Cause of rotation: Rotation caused by fixation to a self-moving object, or fixation to a static point in 3-D environment.

(c) Configurations of simulated environments (e.g., ground, cloud or frontoparallel plane)

(d) Observer's task (egocentric or exocentric judgement) (van den Berg, 1996; Stone & Perrone, 1997).

No current model of human heading judgement can explain the effects of all the factors above. Further studies are needed about the bias in perceived heading.

We found that in gaze-stabilized situations, human observers can judge the heading direction reasonably accurately up to a roll rate of 11.5 deg/sec. The effect of roll appears to be smaller than yaw or pitch. One possible reason is that the retinal velocity due to roll is smaller than due to yaw or pitch in the central visual field because the velocity due to yaw and pitch is roughly constant within

the limited visual field, while the retinal velocity due to roll is almost in proportionate with the eccentricity.

Coordinate system in human visual system

We can use different axes as a coordinate system. Even if the coordinates are transformed by rotation around the line of sight and the transformed coordinates are used to describe self-rotation, the visual system can process information in the same way. It is possible that the human visual system uses axes different from the vertical and horizontal axes to define self-rotation. If different axes are used, the concept of pitch, yaw and roll is meaningless. The representation of self-rotation in human brain has not been clearly known yet. It is needed to know the representation of self-rotation for the analysis of visual function.

Translation with various types of rotation and in the upward direction

Rotation was caused by eye movement in almost studies of human heading judgement, but rotation often arises not only from eye or body rotation which can be detected by efference copies or feedback signals, but also from vehicle rotation which cannot be detected by those signals. Only the acceleration or gravitation detectors in the middle ear and visual information are available in such situations. Previous studies have focused on heading in the horizontal and forward direction. Today we can enjoy simulation games of airplane flight. In such games, visual information alone is used for

perception of egomotion. An airplane's path usually curves with roll, but we perceive the translation clearly. It implies that visual information plays an important role in detecting the various heading directions and various types of self-rotation. Because there are situations where efference copies or feedback signals are not available to detect rotation, it is important to study perception of various kinds of rotation from visual information. The study on perception of heading in the upward direction is also important.

We investigated heading recovery from motion. Visual information different from motion such as dynamical changing binocular disparities may be used. Further studies will be required on the integration of various types of information.

We examined retinocentric (observer-center) heading perception in this paper. However, it remains unclear whether human observers can judge the destination in the environment from the flow with roll or pitch components in real-life situations. For the navigation task, transform of the self-motion path in the observer-center coordinates to the path in the world-center coordinates is required. Recently the transform has been discussed in the context of heading detection (Royden, 1994; van den Berg, 1996; Stone & Perrone, 1997). Fast calculation is necessary for the navigation task because the world-center coordinates are always changing relative to the egocentric coordinates. Investigations about human ability of the transform will be more important in the study of human heading judgement.

ACKONWLEDGEMENT

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 Special Coordination Funds for Promotion of Science and Technology Agency of the Japanese Government.

Appendix: Subspace algorithm with non-roll constraint

We use the subspace algorithm of Heeger and Jepson (1990; 1992) to obtain a least-squares solution. First we introduce the notation. Each point in a scene has an associated position vector, **P**=(X, Y, Z)^t relative to a viewer-centered coordinate frame. (^{t} denotes transpose of a matrix or vector). Under perspective projection, the point projects to a point in the image plane of 1/Z, (x, y) = (X/Z, Y/Z). Every point of a rigid body shares the same six motion parameters relative to the viewer-centered coordinate frame. Due to the motion of the observer the relative motion of the point is:

$$\mathbf{V} = (\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt}) = -(\Omega \times \mathbf{P} + \mathbf{T}) \qquad (A-1)$$

Where **T**=(U, V, W)^t and **Ω**=(A, B, C)^t denote, respectively, the translational and rotational velocities.

Image velocity, θ(x, y), is defined as the time derivatives of (x,y)^t. From (3), we obtain:

$$(\frac{dx}{dt}, \frac{dy}{dt})^t = \vartheta(x, y) = p(x, y)\mathbf{K}(x, y)\mathbf{T} + \mathbf{M}(x, y)\Omega \qquad (A-2)$$

where p(x,y)=1/Z is the inverse depth, and **K** and **M** are:

$$\mathbf{K}(x, y) = \begin{bmatrix} -1 & 0 & x \\ 0 & 1 & y \end{bmatrix}$$
$$\mathbf{M}(x, y) = \begin{bmatrix} -xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{bmatrix} \qquad (A-3)$$

The subspace algorithm uses flow vectors of N image points with a minimization method in the following way. The N separate equations are combined into one matrix equation:

$$\boldsymbol{\Theta} = \mathbf{L}(\mathbf{T})\mathbf{q} \qquad (A-4)$$

where

$$\boldsymbol{\Theta} = (\vartheta_1, \ldots, \vartheta_N)^t$$

$$\mathbf{q} = [p(x_1, y_1), \ldots p(x_N, y_N), \Omega_x, \Omega_u, \Omega_z]^t$$

$$\mathbf{L}(\mathbf{T}) = \begin{pmatrix} \mathbf{K}(x_1, y_1) & & \mathbf{M}(x_1, y_1) \\ & \ddots & \vdots \\ & & \mathbf{K}(x_N, y_N)\,\mathbf{M}(x_N, y_N) \end{pmatrix} \qquad (A-5)$$

Then we minimize the residual function R(**T**) to recover the translation direction.

$$R(\mathbf{T}) = \left\| \boldsymbol{\Theta} - \mathbf{L}(\mathbf{T})\mathbf{q} \right\|^2 \qquad (A-6)$$

When **T** is fixed, (A-4) is linear about **q**. When **q** has the following value, R(**T**) is minimized with a fixed **T**:

$$\mathbf{q} = (\mathbf{L}(\mathbf{T})^t \mathbf{L}(\mathbf{T}))^{-1} \mathbf{L}(\mathbf{T})\boldsymbol{\Theta} \qquad (A-7)$$

It is rare that $\mathbf{L}(\mathbf{T})^t\ \mathbf{L}(\mathbf{T})$ is singular. R(**T**) can be minimized by the minimization of the following equation.

$$R(\mathbf{T}) = \left\| \Theta - \mathbf{L}(\mathbf{T})(\mathbf{L}(\mathbf{T})^t \mathbf{L}(\mathbf{T}))^{-1} \mathbf{L}(\mathbf{T}) \Theta \right\|^2 \qquad (A-8)$$

The residual function R(**T**) has the same value as R(a**T**) for any constant a. It means that we can obtain only the ratio of U, V and W, namely, only the direction of translation. Therefore we used another constraint, W=1 for convenience of the calculation. We minimized the R(**T**) of (9) by Fletcher-Reeves-Polak-Ribiere method with an initial value of **T**.

Lappe & Rauschecker (1995) showed a subspace algorithm with non-roll constraint (C=0). The same procedure is used with the following changes:

$$\mathbf{q} = [\, p(x_1, y_1), \ldots p(x_N, y_N), \Omega_x, \Omega_u \,]^t$$
$$\mathbf{M}(x, y) = \begin{bmatrix} -xy & -(1+x^2) \\ 1+y^2 & -xy \end{bmatrix} \qquad (A-9)$$

The remaining points are the same as the algorithm that does not use the non-roll constraint.

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

Beusmans, J.M.H. (1993) Computing the direction of heading from affine image flow. *Biological Cybernetics,70,* 123-136.

Bradley, D. C., Maxwell, M., Anderson, R. A., Banks, M. S., Shenoy, K. V. (1996) Mechanism of heading perception in primate visual cortex. Science, 273, 1544-1547.

Britten, K.H. & van Wezel, R.J.A. (1998) Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nature neuroscience, 1,* 59-63.

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

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

Ehrlich, S. M., Beck, D. M., Crowell, J.A., Freeman, T. C. A. & Banks, M.S. (1998) Depth information and perceived self-motion during simulated gaze rotations. *Vision Research, 38,* 3129-3145.

Freeman, T. C., Harris, M. G. & Tyler, P. A. Human sensitivity to temporal proximity: The role of spatial and temporal speed gradients. *Perception & Psychophisics, 55,* 689-699.

Gibson, J. J. (1950) The perception of the visual world. Houghton Mifflin, Boston.

Hanada, M & Ejima, Y. (in press) A model of human heading judgement in forward motion. *Vision Research*

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.

Kaiser, M. K. & Hecht, H. (1995) Time-to-passage judgements in nonconstant optical flow fields. *Perception & Psychophysics, 57,* 817-825.

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

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. (1995) Motion anisotropies and heading detection. Biological *Cybernetics, 72,* 261-277.

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.

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.

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. (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., Crowell, J.A. (1992) The perception of heading during eye movements. *Nature, 360,* 583-5.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1989) Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience, 6,* 145-157.

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

Tanaka, K & Saito, H. (1989) Analysis of motion of the visual field by direction, Expansion/Contraction, and Rotation cells clustered in the dorsal part of the medial superior temporal

area of the macaque monkey. *Journal of Neurophysiology, 62,* 626-641.

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.

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

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.

Figure captions

Figure 1: Rotation and translation of an observer

Pitch, yaw and roll are rotation around the horizontal axis, the vertical axis and the axis along the line of sight, respectively. Rotation can be expressed by the three components. A, B and C indicate the rotation rate of pitch, yaw and roll respectively. U, V and W show translation in the horizontal direction, in the vertical direction and in the direction of the line of sight, respectively.

Figure 2: Path without roll.

(a) The schematic diagram of the simulated self-motion in our experiments. We simulated translating in a fixed direction (θ) with respect to the current line of sight (the Z axis) while fixating a static point. The observer moved stepwise on the arrows from frame to frame. The step shown by the arrow in this diagram is much larger than actual one. (b) An example of the simulated path that was actually used in the experiments is shown. The heading direction was 10 deg. (c) If the observer translates in a fixed direction in the world-center coordinates, the instantaneous heading direction changes with time in the retinocentric (observer-fixed) coordinates.

Figure 3: Schematic diagram of a simulated path of an observer while rolling, and the flow field.

(a) Schematic diagram of the simulated self-motion in the experiments is shown. We simulated

translating in a fixed direction (θ) with respect to the current line of sight (the Z axis) while fixating a static point and rolling. The observer moved along the arrows of T from frame to frame, and rolls around the line of sight (the Z axis). The step shown by the arrow in the diagram is much larger than actual one. The coordinates fixed to the observer's view (shown by (X, Y, Z), (X', Y', Z') or (X'', Y'', Z'')) rotated with time due to fixation and roll. (b) Point paths are shown for a condition of forward and horizontal translation with 10.5 deg/sec roll.

Figure 4: Results of observer HI in Experiment 1.

Results of observer HI in Experiment 1 are shown. Perceived heading is indicated as a function of the simulated heading. Each point indicates the result of one trial. When the judgement is correct, the point is on a line with slope of 1. The results for five positive roll values (22.9, 17.2, 11.5, 5.7 and 0.0 deg/sec) are shown. The results of negative roll are similar.

Figure 5: Results of observer MH in Experiment 1.

Results of observer MH in Experiment 1 are shown.

Figure 6: Results of Experiment 2.

Results of observer HI and MH in Experiment 2 are shown. Simulated heading and perceived one were divided into horizontal and vertical components. Perceived heading is indicated as a function of

the simulated heading. The results of the horizontal component of heading are shown in the upper panel, those of the vertical component in the lower panel.

Figure 7: Results of observer HI in Experiment 3.

Results of observer HI in Experiment 3 are shown. Simulated heading and perceived one were divided into horizontal and vertical components. Perceived heading is indicated as a function of the simulated heading. The results of the horizontal component of heading are shown in the left panel, those of vertical component of heading in the right panel. The results for nine four values (22.9, 17.2, 11.5 and 5.7 deg/sec) are shown.

Figure 8: Results of observer MH in Experiment 3.

Results of observer YN in Experiment 3 are shown. The results for five roll values (22.9, 17.2, 11.5, 5.7 and 0 deg/sec) are shown.

Figure 9: Estimation by subspace algorithm with non-roll constraint

Estimated heading is plotted as a function of the simulated heading. Each point indicates the result of one trial. When the judgement is correct, the point is on a line with slope of 1. Estimation results for 5.7 deg/sec roll are shown in the upper panel, and results for 11.5 deg/sec roll in the lower panel. Input velocities to the algorithm were computed by simulating the same conditions as in Experiment 3.

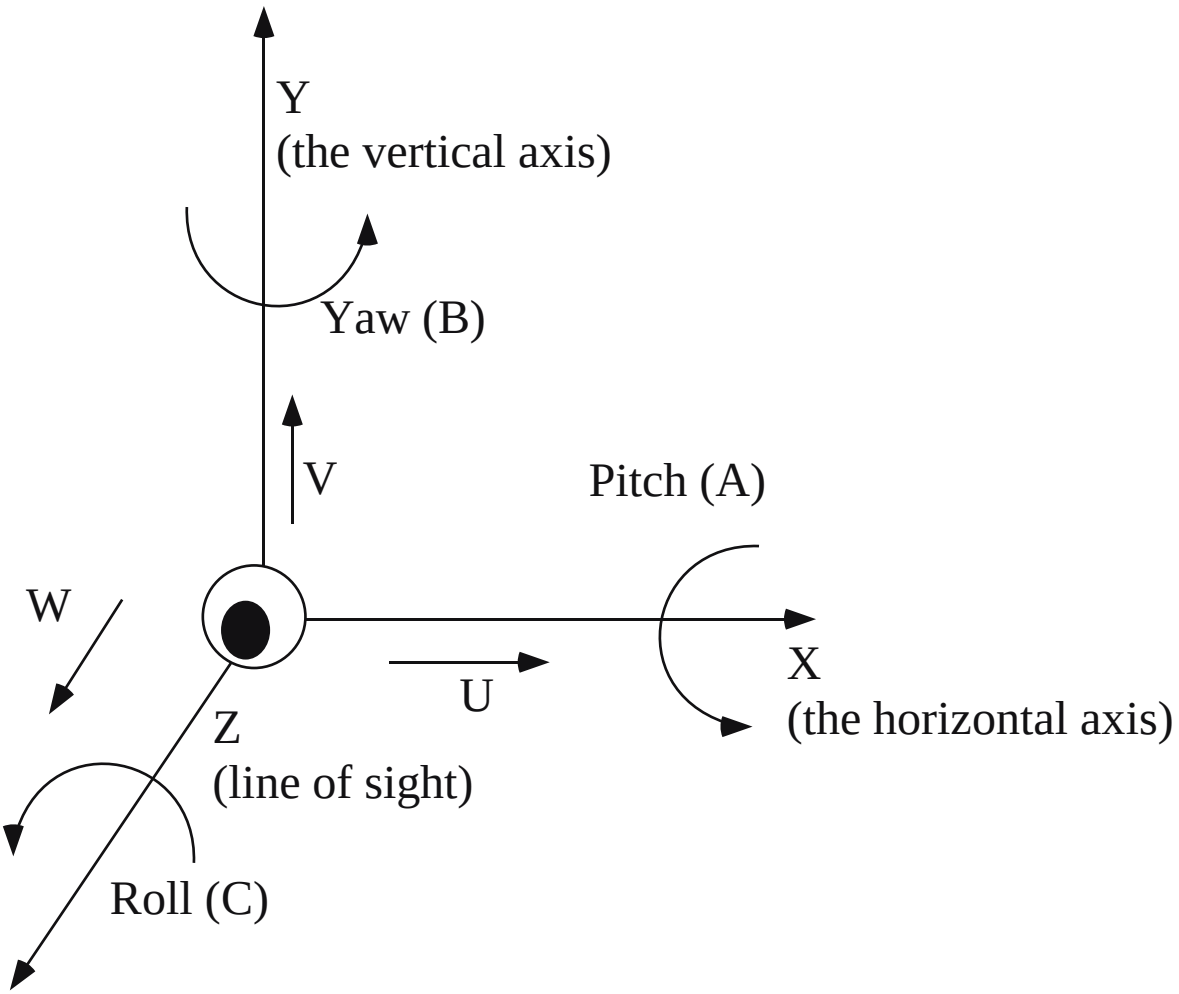

Y
(the vertical axis)
Yaw (B)
V
Pitch (A)
W
X
(the horizontal axis)
U
Z
(line of sight)
Roll (C)

Figure 1

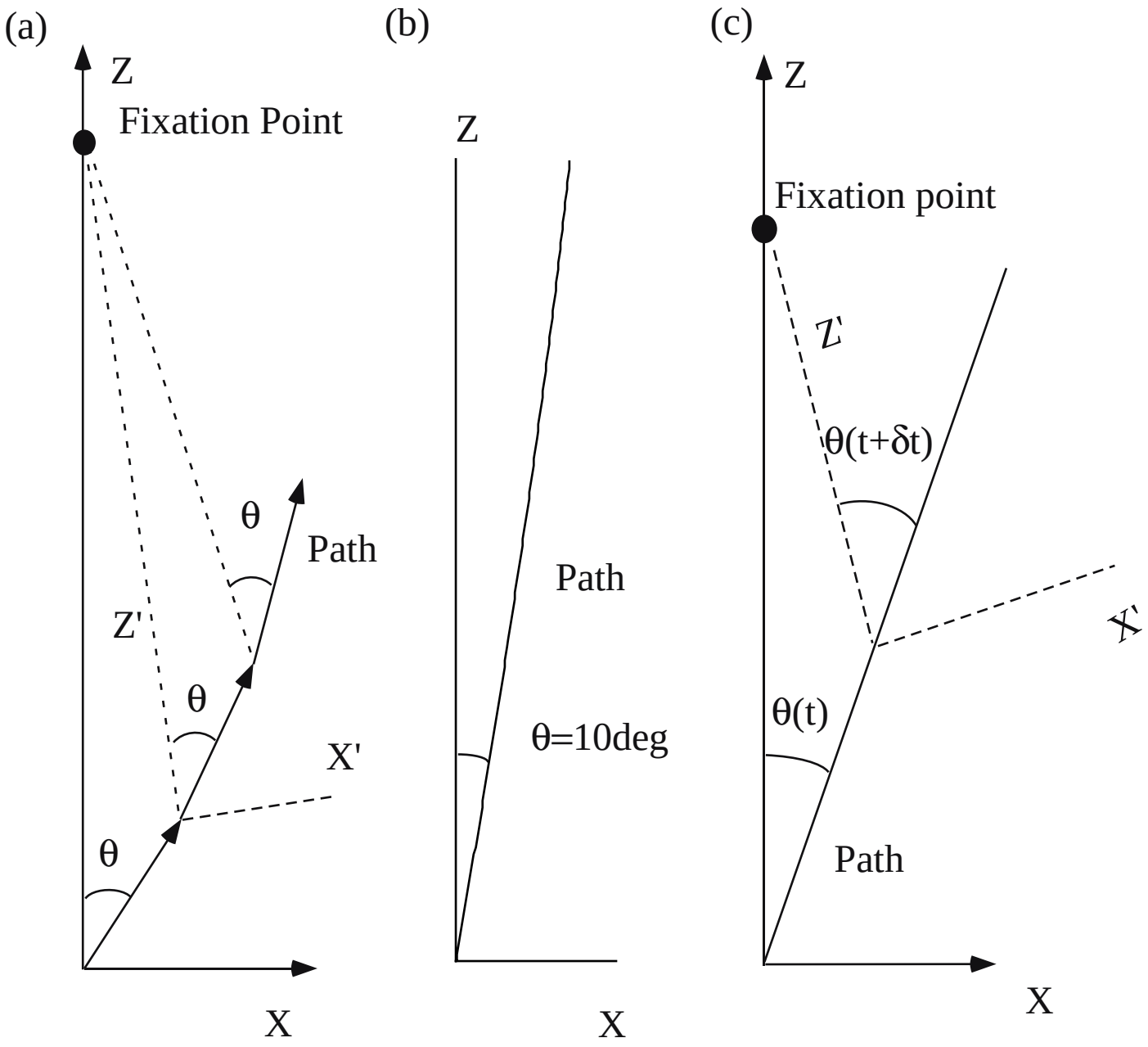

(a)
Z
Fixation Point
θ
Path
Z'
θ
X'
θ
X
(b)
Z
Path
θ=10deg
X
(c)
Z
Fixation point
Z'
θ(t+δt)
X'
θ(t)
Path
X

Figure 2

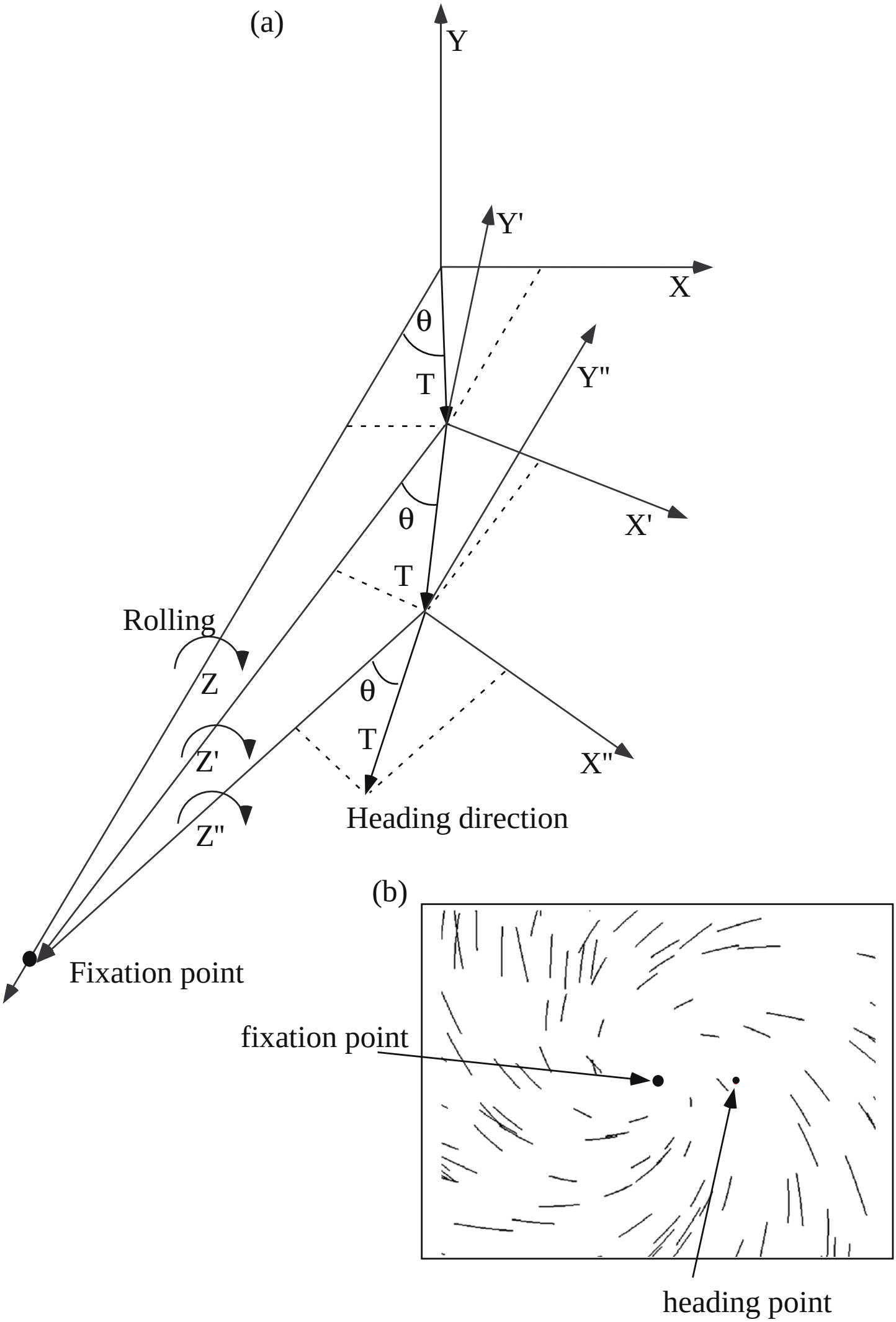

(a)
Y
Y'
X
θ
T
Y"
θ
X'
T
Rolling
Z
θ
T
Z'
X"
Z"
Heading direction
Fixation point
(b)
fixation point
heading point

Figure 3

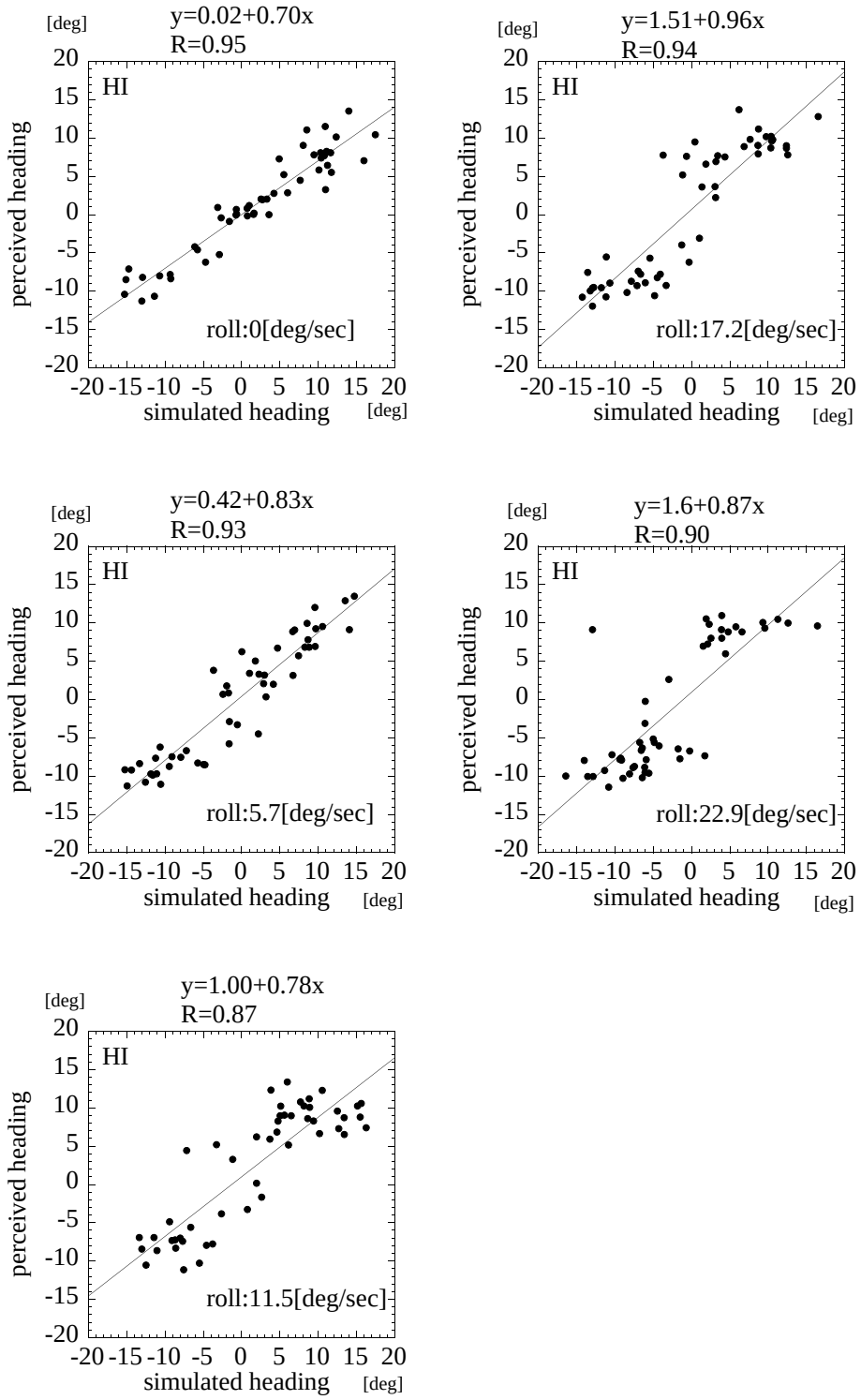

y=0.02+0.70x
R=0.95
HI
roll:0[deg/sec]
y=1.51+0.96x
R=0.94
HI
roll:17.2[deg/sec]
y=0.42+0.83x
R=0.93
HI
roll:5.7[deg/sec]
y=1.6+0.87x
R=0.90
HI
roll:22.9[deg/sec]
y=1.00+0.78x
R=0.87
HI
roll:11.5[deg/sec]
perceived heading
simulated heading
[deg]

Figure 4

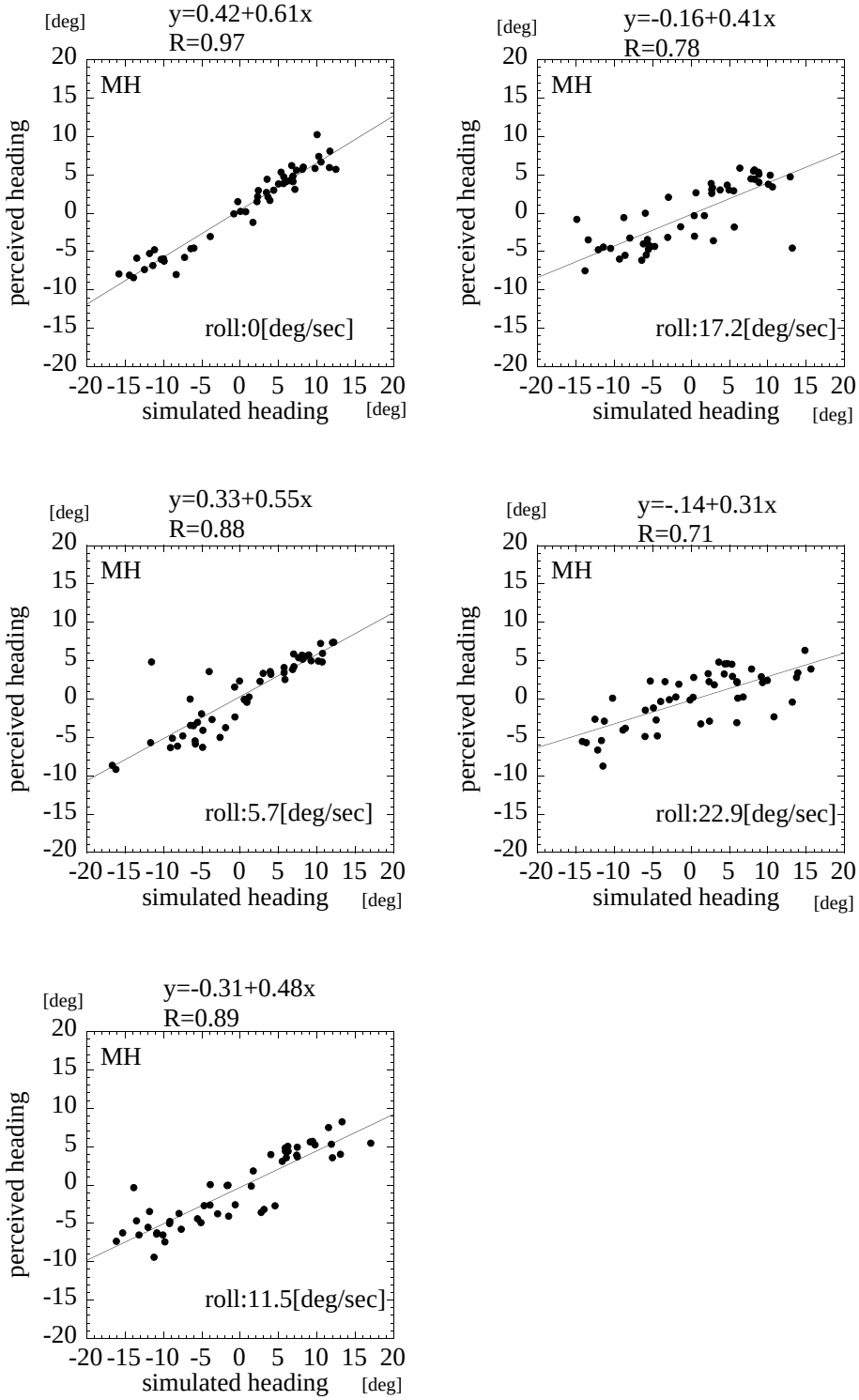

y=0.42+0.61x
R=0.97
MH
roll:0[deg/sec]
y=-0.16+0.41x
R=0.78
MH
roll:17.2[deg/sec]
y=0.33+0.55x
R=0.88
MH
roll:5.7[deg/sec]
y=-.14+0.31x
R=0.71
MH
roll:22.9[deg/sec]
y=-0.31+0.48x
R=0.89
MH
roll:11.5[deg/sec]
perceived heading
simulated heading
[deg]

Figure 5

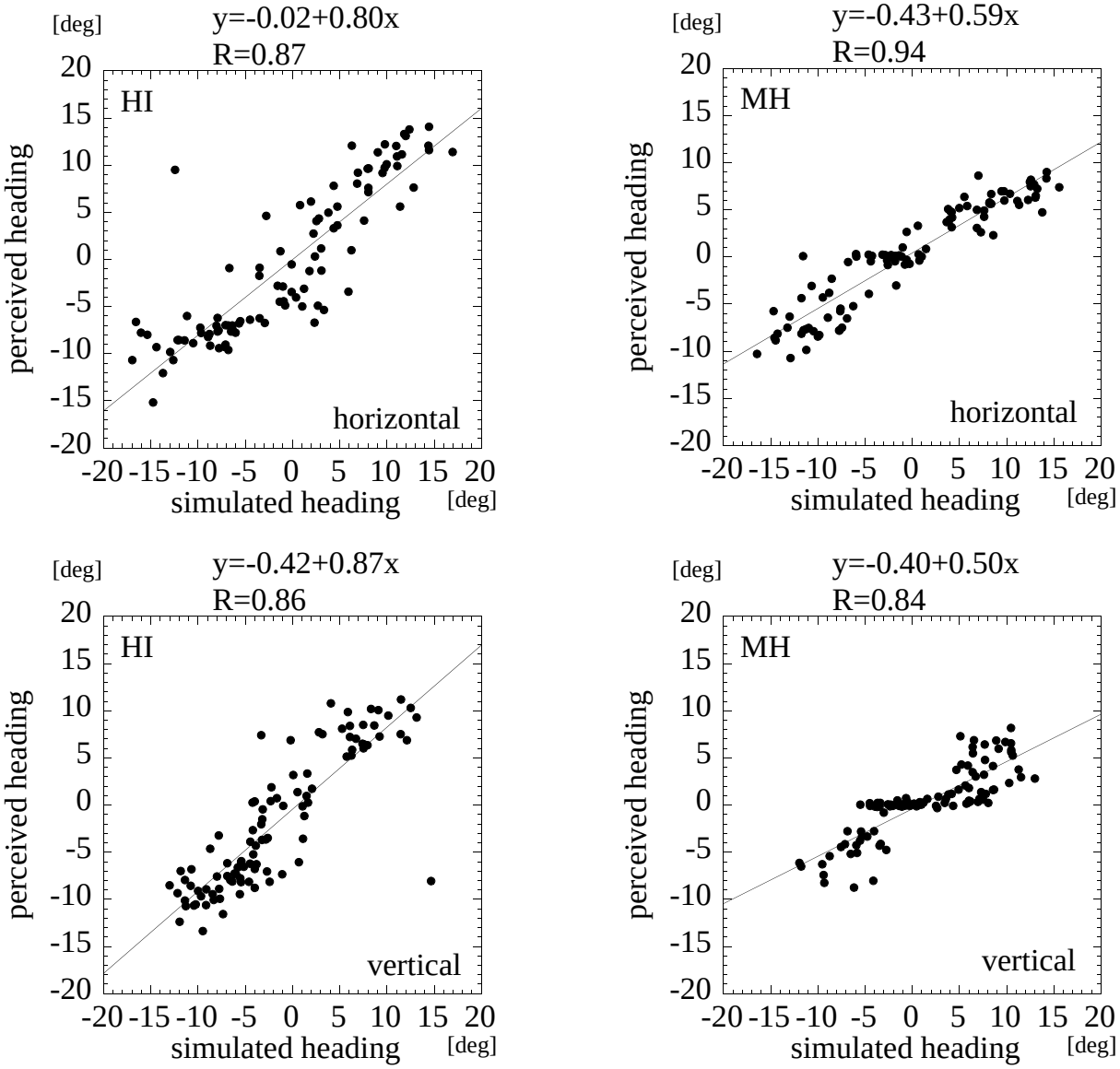

[deg]
y=-0.02+0.80x
R=0.87
HI
horizontal
perceived heading
simulated heading
[deg]
[deg]
y=-0.43+0.59x
R=0.94
MH
horizontal
perceived heading
simulated heading
[deg]
[deg]
y=-0.42+0.87x
R=0.86
HI
vertical
perceived heading
simulated heading
[deg]
[deg]
y=-0.40+0.50x
R=0.84
MH
vertical
perceived heading
simulated heading
[deg]

Figure 6

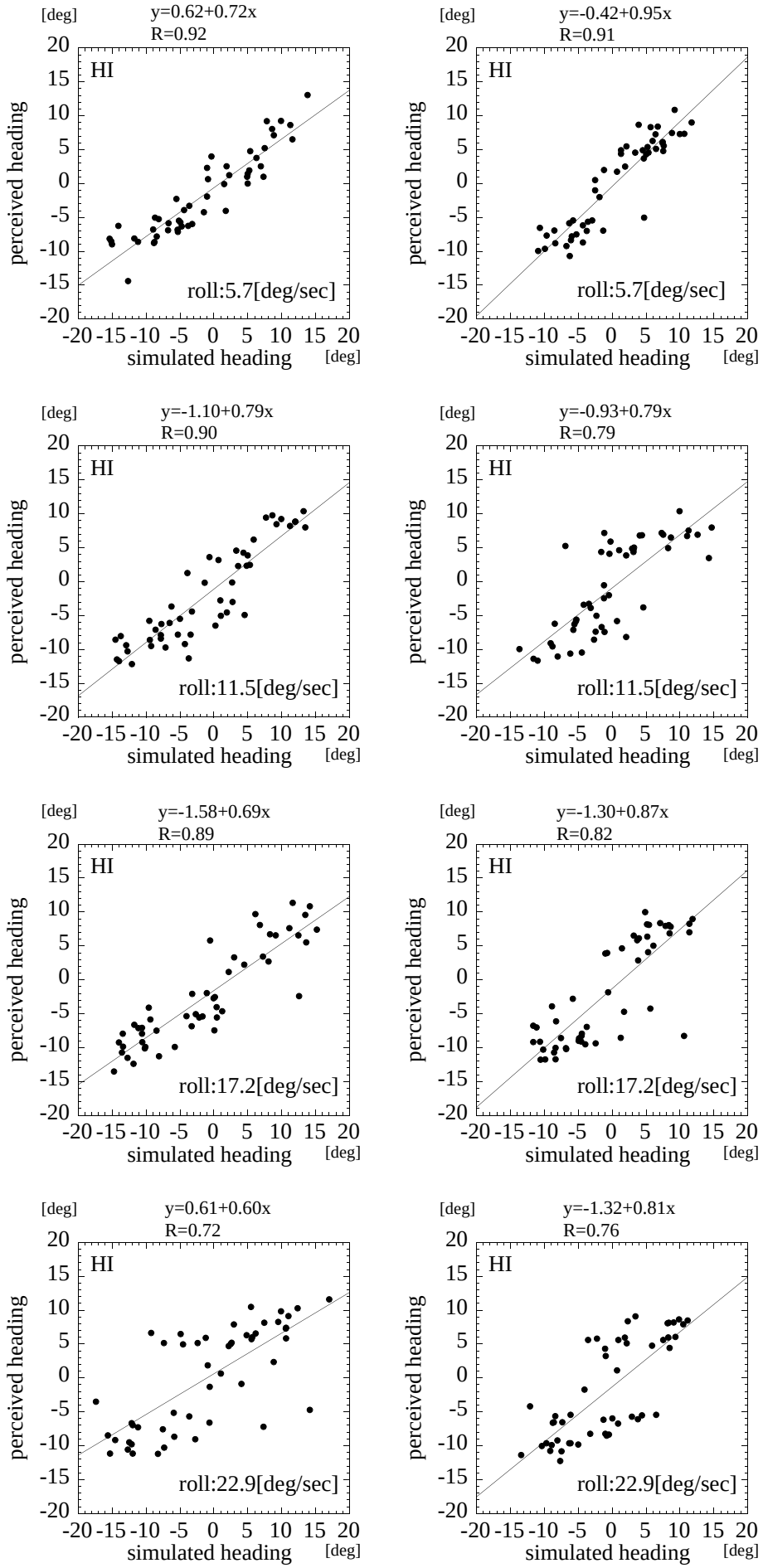
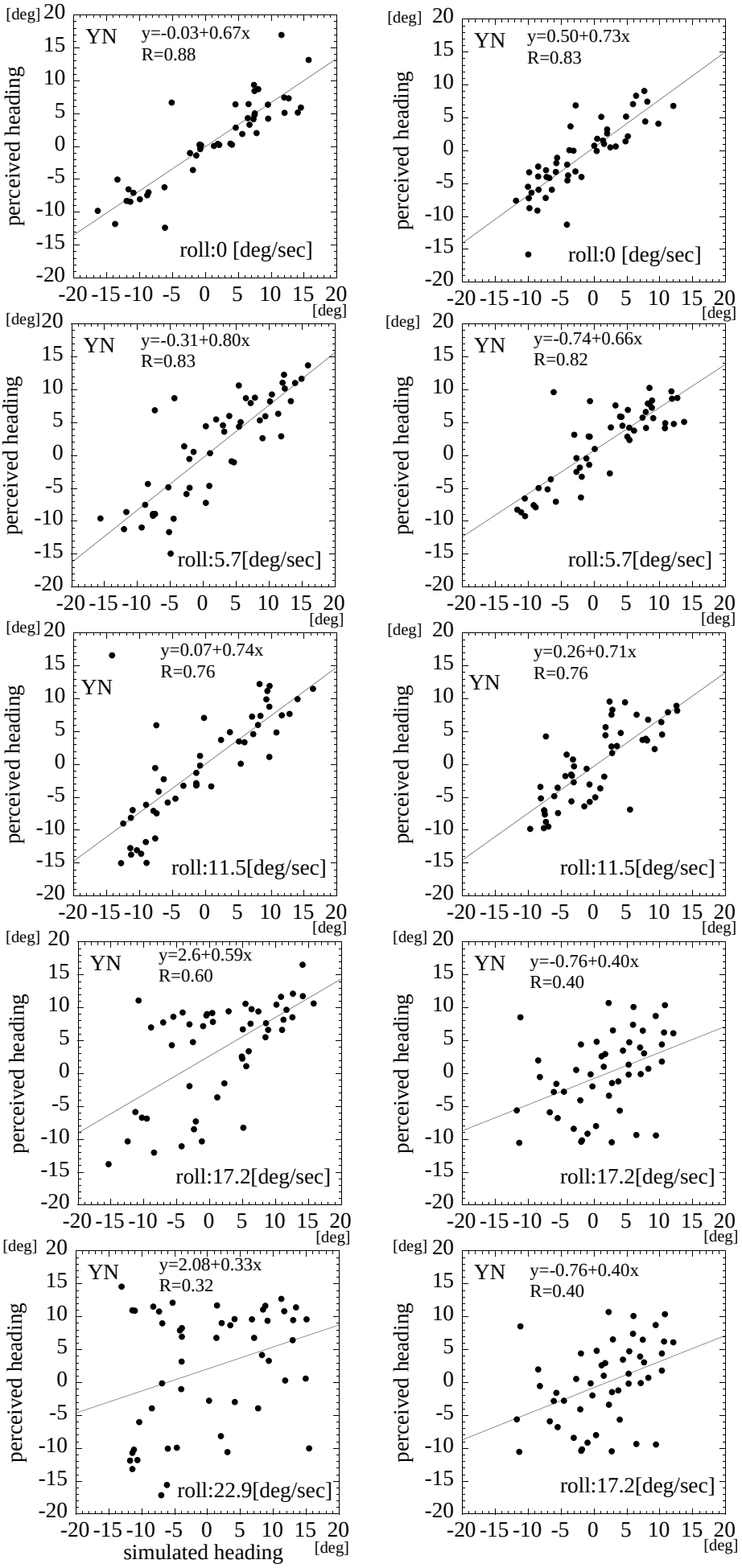

[deg]
y=0.62+0.72x
R=0.92
HI
perceived heading
roll:5.7[deg/sec]
simulated heading
[deg]
[deg]
y=-0.42+0.95x
R=0.91
HI
perceived heading
roll:5.7[deg/sec]
simulated heading
[deg]
[deg]
y=-1.10+0.79x
R=0.90
HI
perceived heading
roll:11.5[deg/sec]
simulated heading
[deg]
[deg]
y=-0.93+0.79x
R=0.79
HI
perceived heading
roll:11.5[deg/sec]
simulated heading
[deg]
[deg]
y=-1.58+0.69x
R=0.89
HI
perceived heading
roll:17.2[deg/sec]
simulated heading
[deg]
[deg]
y=-1.30+0.87x
R=0.82
HI
perceived heading
roll:17.2[deg/sec]
simulated heading
[deg]
[deg]
y=0.61+0.60x
R=0.72
HI
perceived heading
roll:22.9[deg/sec]
simulated heading
[deg]
[deg]
y=-1.32+0.81x
R=0.76
HI
perceived heading
roll:22.9[deg/sec]
simulated heading
[deg]
horizontal
vertical

Figure 7



horizontal vertical

Figure 8

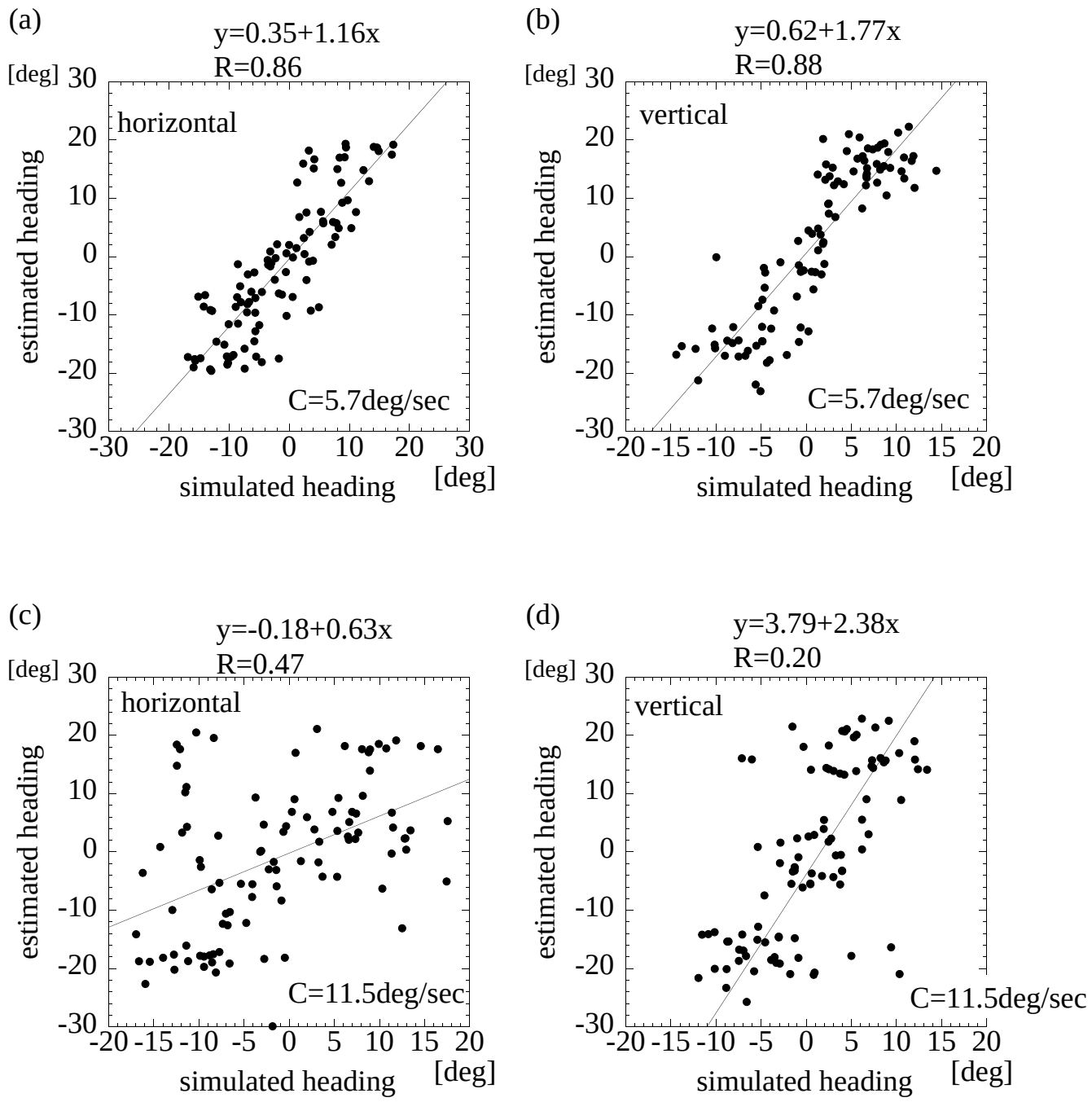

(a)
y=0.35+1.16x
R=0.86
horizontal
C=5.7deg/sec
estimated heading
simulated heading
[deg]
(b)
y=0.62+1.77x
R=0.88
vertical
C=5.7deg/sec
estimated heading
simulated heading
[deg]
(c)
y=-0.18+0.63x
R=0.47
horizontal
C=11.5deg/sec
estimated heading
simulated heading
[deg]
(d)
y=3.79+2.38x
R=0.20
vertical
C=11.5deg/sec
estimated heading
simulated heading
[deg]

Figure 9