# Heading judgement from second-order motion

Mitsuhiko Hanada and Yoshimichi Ejima

Graduate School of Human and Environmental Studies

Kyoto University

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

Abstract

We examined human heading judgement from second-order motion which was generated by random-dots with the contrast polarity determined randomly on each frame. It was found that human observers can judge heading fairly accurately from second-order motion when pure translation is simulated or when self-motion toward a ground plane with gaze rotation is simulated but they cannot when self-motion toward cloud-like random dots with gaze rotations is simulated. It is suggested that the human visual system cannot decompose the flow fields into rotational and translational components by using second-order motion information alone, but it can do in some ways from the flow field of the ground plane.

**Introduction**

Psychophysical studies reveal two mechanisms for visual motion processing. In first-order motion, we detect real or apparent displacement of boundaries defined by luminance differences. In second-order motion, we detect the displacement of boundaries where the amount of luminance on either side is identical, but there is a difference in contrast produced by texture or flicker (Cavanagh & Mather, 1989; Chubb & Sperling, 1988a). Although most studies of motion perception have centered on first-order motion that is defined by spatiotemporal changes in luminance, there has been recent interest in second-order motion. First-order motion can be detected by motion energy detectors like the Reichardt detector (Reichardt, 1961; van Santen & Sperling, 1984) or linear spatiotemporal filtering followed by a squaring non-linearity. (Adelson & Bergen, 1985). van Santen & Sperling's type of Reichardt model and Adelson & Bergen motion energy model are basically identical and computing the same things in different ways (Adelson & Bergen, 1985). Some kinds of second-order motion can be detected by full- or half-wave rectification followed by a motion energy detector (Chubb & Sperling, 1988a). However, an initial linear spatial and/or temporal filtering is required for some second-order motion, such as moving flicker.

Motion on the retina is one of the cues for recovery of self-motion, 3-D structure and depth of objects. Recovery of camera motion and structure from motion have been of central interest in computer vision. Most algorithms recover both structure of the objects and camera

motion. It is relatively simple to recover structure of objects if motion of a camera is known, and vice versa.

Human observers can perceive the structure of moving objects from motion information (e.g., Johansson, 1975). They can also perceive their heading from motion information reasonably accurately when the depth variation is large and the rotation rate is low (Warren & Hannon, 1990; van den Berg, 1993). A number of models of structure from motion and heading recovery in human vision were presented (e.g, Rieger & Lawton, 1985; Perrone, 1992; Hildrerh, 1992; Lappe & Rauschecker, 1993; Beusmans, 1993; Hildreth et al., 1995; Buracas & Albright, 1996; Beintema & van den Berg, 1998). In several models, both self-motion and structure of objects are recovered simultaneously. However, it has not been known whether heading recovery and structure from motion are processed by the same mechanism in the human brain.

Human perception for second-order motion differs from perception for first-order motion in various ways. For example, second-order motion produces weaker motion aftereffects to static test stimuli than first-order motion does, although motion aftereffects to second-order motion are fairly strong when the test pattern is dynamical flickering (Nishida & Sato, 1995). Also, the human visual system can recover shape from first-order motion, but can not from second-order motion (Dosher et al., 1989; Landy et al., 1991). These results suggest that first-order and second-order motion is encoded by different mechanisms.

Gurnsey et al. (1998) examined whether second-order motion is used to process

motion-in-depth in the human visual system by measuring vection and they found that second-order motion contributes to vection. This report suggests that second-order motion is used for human locomotion. It is also important to know the heading direction for locomotion. In psychophysical studies of human heading judgment, first-order motion stimuli have been used. In this paper, we examined whether human observers can judge their heading from second-order motion.

**Experiment**

We used stimuli with random contrast polarity on each frame as second-order motion stimuli as used by Landy et al. (1991) for experiments of shape from motion. We compared human performance of heading judgement from second-order motion with that from first-order motion.

**General methods**

Apparatus: The experiments were conducted on a Silicon Graphics O2 workstation with a color monitor. Computer-simulated motion sequences were monocularly presented at a framerate of 60 Hz. The image on the screen was 34.4 cm wide (1280 pixels) and 27.5 cm (1024 pixels) from top to bottom. At the viewing distance of 40 cm, the screen subtended 46.5 deg horizontally × 38 deg vertically. One red dot, which was always at the center of the screen, served as a fixation point.

The observer was asked to fixate the point and not to move their eyes during the presentation of the stimuli. Apart from the stimuli, the room was dark.

Stimuli: Stimuli were white or black dos of 4×4 pixels on a gray background with a luminance of 29 cd/m$^2$. The contrast of dots was defined as:

{(luminance of dot) – (background luminance) } / (background luminance)

We used microbalanced motion with random contrast polarity as second-order motion stimuli, developed by Chubb & Sperling (1988a). Let I be any visual stimulus, that is, I(x, y, t) is luminance of a point (x, y) in a visual field at time t. Stimulus I is called microbalanced, when, for any (x, y, t), (x', y', t') $\in \mathbf{Z}^3$,

$$E[I(x, y, t)I(x', y', t')] = E[I(x, y, t')I(x', y', t)]$$

E represents expectation. A random-dot stimulus with contrast determined randomly on each frame is microbalanced. Chubb & Sperling (1988a) proved that if a stimulus is microbalanced, the expected output of every first-order motion energy detector will be zero. Thus, motion direction of the random-polarity stimulus cannot be detected by the motion energy detectors.

One might think that unless average luminance of white and black dots is exactly the same as background luminance, first-order motion energy detectors can detect motion direction by artifact of inaccurate setting of background luminance. However, it should be noted that even if the background is perturbed around the average luminance of white and black dots by about

6

10 %, responses of first-order motion energy detectors are little affected and the effects of the perturbation are negligible (See Appendix). For recovery of heading, the speed and the 2-D direction of dots are required. Even if motion direction along a 1-D axis is detected correctly by motion energy detectors due to the small perturbation of average luminance, the velocity cannot be calculated from responses of first-order motion energy detectors to the random-polarity stimuli with background luminance perturbed by about 10 % from average luminance of dots. (See also Appendix.) For our experiment, even if there are errors in luminance measurement, the errors were far less than 10% and obviously the non-linearity of monitors for high spatial-frequency stimuli was not so large. Thus, the artifact due to the error of background luminance setting was negligible, at least for our experiment of heading judgement.

**Preliminary experiment: direction discrimination of first- and second-order motion**

In order to equate the visibility of first- and second-order motion for heading experiments, we measured the effects of contrast on direction discrimination for first-order or second-order motion. The visibility for first-order motion was equated to (or less than) second-order motion stimuli by setting the contrast to an appropriate value from performance of motion discrimination.

**Methods:**

Observers: One of the authors (MH) and one graduate student (SO) in our laboratory participated in the experiment. SO was naïve as to the purpose of the experiment. Both had normal or corrected-to-normal vision.

Stimuli: Dots were randomly distributed in the whole screen. Each dot moved at a rate of 1 [pixel] / 1 [frame] to one of eight neighboring pixels. The drift speed of dots was 2.2 deg/sec for upward, downward, leftward or rightward movement. Since the distance from a position to the diagonal neighboring pixels (i.e. the upper-right, upper-left, lower-left or lower-right pixels) is $2^{1/2}$ times larger than the distance to the vertical or horizontal neighboring pixels, the speed of diagonally moving dots was 3.1 ($\approx 2^{1/2} \times 2.2$) deg/sec. The number of dots was 400. Each dot was moved in a constant direction and displayed for 16 frames (267ms) and disappeared, and then a new dot was displayed at a random position. Total duration of stimulus presentation was 2.0 sec. On a given trial, a certain percentage of dots (target dots) moved in the leftward or rightward three directions (right, upper-right and lower-right when the target direction was right, or left, upper-left and lower-left when the target direction was left). An equal proportion of the target dots moved in each of the three target directions. The direction of motion of the remaining dots (non-target dots) was selected from the other five directions. An equal proportion of the non-target dots moved in one of the five non-target directions. We defined coherency as a proportion of the number of target dots to that of all dots. When the target direction was right and the coherency was 100%, one-third of the dots moved in the right direction, one-third in the

upper-right direction and one-third in the lower-right direction. When the coherency was 75%

and the target direction was right, the moving direction of 25% dots was right, 25% upper-right,

25% lower-right, and the remaining dots moved in one of the other five directions. When the

coherency was 37.5%, the proportion of dots moving in each direction was the same (12.5%).

Procedure: The observers had to indicate in which direction (leftward or rightward) the dots were

moving as a whole by pressing a mouse button. No feedback was given. The contrast conditions

were 100%, 50%, 25%, 12.5% and 6.25% and the coherency conditions were 100%, 87.5%, 75%,

67.5%, 50% and 37.5%. For each condition, 50 trials were conducted. The contrast condition of

dots was varied between sessions and the coherency condition was varied between trials in a

session.

**Results**

Percentages of correct direction discrimination are shown in Fig. 1. There existed no

correct response for 37.5 % coherency because the number of dots moving in each direction was

equal. We defined the performance for 37.5% coherency as 50% correct responses. We focus on

the effects of contrast on motion discrimination. For constant polarity, performance was almost

constant from 100% to 12.5% contrast and was low for 6.25% contrast. In the random-polarity

condition, performance worsened for observer SO as the contrast decreased. For observer MH,

performance was constant up to 25% contrast, and worsened as the contrast decreased to less than 25%.

---

Insert Figure 1 about here

---

Contrast setting for heading experiments

The performance level of MH was almost constant for contrasts over 25% in the random-polarity condition and for contrasts over 12.5% in the constant-polarity condition. Performance of SO at 50% contrast was little different from that at 50 % contrast in the random-polarity condition, and SO's performance was constant for contrasts over 12.5% in the constant-polarity condition. We assumed that the contrast of the random-polarity stimulus was effectively equivalent to one-fourth contrast of the constant-polarity stimulus for SO and half for MH concerning motion visibility. We used 100% contrast for the random-polarity stimulus and 25% contrast for the constant-polarity stimulus for the heading experiments. For SO, both stimuli were effectively equivalent. For MH, first-order motion stimuli with 25% contrast had less visibility than second-order motion stimuli with 100% contrast. Fig. 2 shows the performance of motion discrimination for second-order 100% contrast and first-order 25 % contrast. The percentages of correct responses under the two conditions were almost the same. We also used 100% contrast for constant polarity stimuli to obtain standard performance.

10

---

Insert Figure 2 about here

---

## Experiment 1: heading judgement

We examined human heading judgements from first-order motion and second-order motion in this experiment.

### Methods

Observers: Two observers in the preliminary experiment and one more observer (RM) participated in this experiment. SO and RM were naïve to the purpose of this experiment. SO and RM had not participated in other types of psychophysical experiments on heading judgment before. All had normal or corrected-to-normal acuity. RM could not participate in the preliminary experiment of motion discrimination, but we think that there is no problem for the interpretation of the experimental results described later.

Stimuli: Stimuli were generated by simulating translation toward a ground plane or a 3D cloud which consisted of random dots with fixation to a static point in the environment.

11

If 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 the retinocentric coordinates. The observers may indicate the heading direction at a moment different from the time at which they should respond and the response may be affected by the change of the heading direction. Therefore we did not employ this path. We simulated situations where the observer moved in a fixed direction with respect to the current line-of-sight in the observer-fixed coordinates while fixating a point. The motion path is schematically illustrated in Fig. 3. An example of the actual path is shown Fig. 3(b). For the movement along the path, the heading angle was constant in the retinocentric (observer-fixed) coordinates. In the world-centered coordinates, however, the observer's path was curved. Self-rotation arose from the pursuit. This curved-path paradigm was also used in Hanada & Ejima (2000).

We briefly introduce the mathematical expression of self-motion to explain the observer's simulated movement. We use the notations of 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 with the origin at the observer's viewpoint. 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 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. The 3-D velocity of a point, P(X, Y, Z) relative to the observer is given by:

$$
\begin{aligned}
\dot{X}(t) &= -U(t) - B(t)Z(t) + C(t)Y(t) \\
\dot{Y}(t) &= -V(t) - C(t)X(t) + A(t)Z(t) \\
\dot{Z}(t) &= -W(t) - A(t)Y(t) + B(t)X(t)
\end{aligned} \tag{1}
$$

where t represents 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. (1) becomes:

$$
\begin{aligned}
\dot{X}(t) &= -U - BZ(t) \\
\dot{Y}(t) &= 0 \\
\dot{Z}(t) &= -W + BX(t)
\end{aligned} \tag{2}
$$

The point P (X(t), Y(t), Z(t)) moves on a circle relative to the observer following the differential equations of (2) if we assume that the observer is stationary. If we assume 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. We modified this paradigm to include a constraint of fixation to a static point in the environment.

13

Hanada & Ejima (2000) 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. 3(a). When an observer fixates and pursues a point $P_f(0, 0, Z_f)$, the following equations holds (Lappe & Rauschecker, 1995):

$$U = -BZ_f$$
$$V = AZ_f \qquad (3)$$

In this situation, A(t), C(t) and V(t) in (1) are 0, and U(t) and W(t) are constant in the observer-fixed coordinates. Therefore, from (1) and (3), point P (X, Y, Z) moves relative to the observer according to the following differential equations:

$$\dot{X}(t) = -U - B(t)Z(t)$$
$$\dot{Y}(t) = 0$$
$$\dot{Z}(t) = -W + B(t)X(t) \qquad (4)$$
$$\dot{Z}_f(t) = -W$$
$$B(t) = -U / Z_f(t)$$

We calculated the motion paths of dots using Eqs. (4). It is not required to calculate the path of the observer in the world-centered 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 projection

of the path derived from (4) 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 fixating a static point as shown in Fig. 3 (a) or (b).

We also used stimuli which simulated pure translation with respect to a 3D cloud. In this condition, no self-rotation is simulated. The simulated path is shown in Fig. 3 (c). Pure translation can be regarded as translation with fixation to an infinitely distant point. Thus, concerning environments and self-motion, three conditions were simulated: self-motion toward a ground plane with simulated pursuit ("ground-with-pursuit" condition), self-motion toward cloud with simulated pursuit ("cloud-with-pursuit" condition) and pure translation toward a cloud without self-rotation ("cloud-without-rotation" condition).

---

Insert Figure 3 about here

---

The simulated world extended in depth from 4 to 8 m in front of the observer's eye. In the ground condition, the simulated height of eye was 1.6 m from the ground plane. The depth of the fixation point was 6m in the ground condition and it was randomly determined from 4m to 8m in the cloud-with-pursuit condition. We simulated egomotion in the forward and horizontal directions. Translation in the horizontal direction was randomly chosen from -0.25 to 0.25 m/sec and translation in the direction of the line of sight was chosen from 0.75 m/sec to 1.25 m/sec for

each trial. The rotation rate was less than 10 deg/sec. In almost all trials, however, the rotation rate was less than 5 deg/sec during the presentation. For constant polarity, the contrast was 100% (cp-100%; constant polarity, 100% contrast) or 25% (cp-25%; constant polarity, 25% contrast) and for random contrast polarity, it was 100% (rp-100%; random polarity, 100% contrast). Dot lifetime was 16 frames (267ms). The simulated self-motion was displayed for 2.0 sec, and all dots except the red fixation point disappeared. The number of dots was 400.

Procedure: The observers had a few practice sessions. The direction that was actually simulated was shown after the response in the practice sessions, but not in the main experimental sessions. The observers were informed that forward and horizontal self-motion was simulated. The simulated motion and path were fully explained to the naïve observers. No further strategy or instruction was given to the naïve observers. The observers experienced a clear impression of relative motion between the self and the simulated environment in almost all conditions. They were asked to judge the heading angle relative to the line of sight (that is, $\theta$ in Fig.2). They adjusted the position of a pointer so as to indicate the perceived direction of heading. Each trial was terminated by the observer's response (a mouse click). 100 trials were conducted in a session.

The experiment was run in the order of the ground-with-pursuit condition, the cloud-with-pursuit condition and the cloud–without-rotation condition. The order for the

stimulus conditions was as follows: cp-100%, rp-25% and cp-25%. There might be some effects

of practice, they seemed to be negligible because observers' performance is little different for the

experimental session and the last practice session in the cp-100%, cloud-with-pursuit condition.

**Results**

Ground with pursuit

The results are shown in Fig. 4. Perceived heading is plotted as a function of the

simulated heading. Each data point in the graph indicates the result of one trial. When the points

scatter along a straight line with slope of 1.0, heading perception is unbiased. Observers RM and

SO judged the heading direction fairly accurately in the ground conditions. The bias of observer

MH in perceived heading toward the fixation point was fairly large. The performance level for

rp-100% was almost the same as for cp-25%.

We conducted a linear regression analysis. The slopes of all observers' data were smaller

than 1.0 for the cp-100% condition. The slope was 0.8 for RM, 0.75 for SO and 0.6 for MH. It

indicates that there was bias in perceived heading towards the fixation point. The slope was

smaller for the cp-25% condition than for the cp-100% condition. For the rp-100% condition, the

slope was almost the same as for the cp-25% condition. The y-intercepts were near 0 for all

conditions. The correlation coefficients were higher than 0.9 except that of RM in the rp-100%

condition (R = .88).

---

Insert Figure 4 about here

---

Cloud with pursuit

The results are shown in Fig. 5. The data clearly show that the performance was much better for constant polarity than for random polarity. The observers could not judge heading accurately from the random-polarity stimuli. The performance was better for 100% contrast to some extent than for 25% contrast in the random-polarity condition. We conducted a regression analysis. The y-intercepts were near 0 for all conditions. The slope for the cp-100%"condition was also smaller than 1.0 for all observers. For constant polarity, the slope became smaller when the contrast was reduced. For random contrast polarity, the slope was very small (<0.3 for all observers) and the correlation coefficient for random polarity was much smaller than for constant polarity. It indicates that the observers could not perceive heading from the random-polarity stimuli.

---

Insert Figure 5 about here

---

Cloud without rotation

The results are shown in Fig.5. The performance was high for random and constant polarity. For cp-100%, the slopes of the regression lines for observers RM and SO were near 1 and that for MH was 0.87, nearer to 1 than when translation plus rotation due to pursuit was simulated. When the contrast decreased to 25%, the slope became smaller. The slope was smaller for rp-100% than for cp-100%, and larger than for cp-25%. The correlation coefficient was very high under all conditions (R>0.90).

---

Insert Figure 6 about here

---

**Discussion**

The heading was judged fairly accurately from second-order motion when self-motion across a ground plane with pursuit or pure translation without self-rotation was simulated. On the other hand, it was not judged correctly when self-motion through cloud-like dots with pursuit was simulated.

Non-linearity

If the spatial filters for motion-energy detector are not linear, average outputs for the luminance of the white and black dots differ from outputs for the background luminance and the stimulus is imperfectly microbalanced in the internal representation of the visual system. However, even if the weak non-linearity exits, effects of a small difference between background luminance and average luminance of dots in the representation of the visual system are negligible for experiments of heading judgement (see Appendix). Fairly good heading judgement from second-order motion in the ground condition or in the non-rotation condition is not thus due to the artifact of average luminance from the weak non-linearity of the spatial filters in early vision.

If there is strong non-linearity in the initial filtering process of motion detectors, the motion detector is no longer a first-order motion detector, but it is a second-order motion detector. For example, if we use a threshold, we can make an operator similar to a half-wave rectifier, by which we can detect second-order motion.

Bias in perceived heading

It has been reported that bias in perceived heading occurs toward the center of the screen for simulated pursuit (e.g, van den Berg, 1996) and for pure translation (e.g., Royden & Hildreth, 1996; 1999). In this experiment, all observers showed such bias for simulated pursuit and pure translation with 25% contrast, though two out of three observers did not for pure translation with 100% contrast of constant polarity. The possible causes of the bias are presented

by many researchers (Hildreth, 1992; Beusmans, 1993; Royden, 1994; van den Berg, 1996; Stone & Perrone, 1997).

Contrast effects on heading judgement from first-order motion

When the contrast was reduced from 100% to 25% for first-order motion, bias in perceived heading became larger. On the other hand, performance of direction discrimination changed little when the contrast decreased to 25%. It shows that for accurate heading perception, relatively strong contrast is necessary.

Motion direction can be discriminated by the motion in the central field and it is likely that central dots are mainly used for direction discrimination. Heading may be judged using both central and peripheral motion. Since peripheral contrast sensitivity for motion is low (Smith & Ledgeway, 1998), peripheral motion may not be detected accurately. The contrast effects on the slope of perceived vs. simulated heading function may be explained by peripheral contrast sensitivity for motion.

First-order motion vs. second-order motion

The performance for second-order stimuli with 100% contrast was almost the same as for first-order stimuli with 25% contrast in the ground-with-pursuit condition. In the cloud-with-pursuit condition, however, the performance was much worse for second-order

motion than for first-order motion. The poor performance cannot be explained by the visibility of the motion stimuli. When pure translation without rotation was simulated, performance was high for second-order motion in the cloud condition. It seems that the human visual system cannot decompose the second-order motion field into translational and rotational flow components in the cloud condition.

Differential performance in heading judgement from second-order motion for ground and cloud with simulated pursuit.

Performance of heading judgement with simulated pursuit from second-order motion was very poor when cloud-like dots with random polarity were used, but it was fairly high when dots structured into a ground plane were used. In the ground condition, the position of a dot becomes lower in the visual field as the depth of the dot decreases. The ground plane has a static depth cue, and it was reported that human heading judgement becomes robust to noise by static depth cues (van den Berg & Brenner, 1994a; 1994b). A number of models of human heading judgment presented methods for integration of static depth cues (Lappe, 1996; Perrone &, Stone 1994; Hanada & Ejima, 2000). The difference between the cloud and ground condition can be explained by assuming that the system for second-order motion processing is very noisy. When static depth cues are available, the visual system may recover heading from the noisy flow field of second-order motion.

An alternative explanation is that the flow field is smoother when the scene consists of a ground plane as opposed to a 3D cloud. The spatial acuity of second-order motion is low (Chubb & Sperling, 1988b). Thus, with a mixture of speeds, the motion of each dot might not be discriminated. The reason for differential performance with cloud and ground stimuli remains to be explained.

Differential motion

Longuet-Higgins and Prazdny (1980) showed that heading can be computed by the velocity differences of two elements at different depths in the same visual direction. The difference vectors radiate from the heading point in the image and the heading direction can be specified by computing the center of the radial pattern. Rieger and Lawton (1985) generalized the notion to include relative motion between non-overlapping elements within a neighborhood. A number of researchers have claimed that the algorithm is used for heading recovery in the human visual system (e.g.,Rieger & Toet, 1985; Warren & Hannon, 1990; Hildreth, 1992; Royden, 1997). For the models, accurate difference velocities between neighboring dots are required.

Dosher et al. (1989) showed that regions are segmented better by first-order motion than by second-order motion. Motion segmentation from second-order motion may be poor because the human visual system cannot compute the accurate difference velocity from second-order

motion. When the difference vectors computed from second-order motion are inaccurate, the differential-motion model does not work well. Poor performance for second-order motion in the cloud condition with simulated pursuit can be explained in terms of inaccurate calculation of the difference vectors. On the contrary, fairly good performance of heading judgement in the ground-with-rotation cannot be explained by the inaccurate computing of differential vectors. However, the smooth velocity field for a ground plane may explain it as noted before.

Physiological models

Recently it has been suggested that the heading direction is computed in MST of primate brain (Tanaka et al., 1989; Bradley et al. 1996; Britten & van Wezel, 1998). Perrone (1992), and Lappe & Rauschecker (1993) presented models of heading recovery for the biological visual system, which uses cells in MT and MST. Most cells in MT are motion-direction selective. In their models, cells in MT input to those in MST. Albright (1992) reported that 87% of cells in MT respond to both first-order and second-order motion. On the other hand, O'Keefe and Movshon (1998) reported that only 25% of cells respond to both. They also reported that the response to second-order motion is much weaker than that to first-order motion. Although the effective contrast was equated for first-order and second-order motion, differential performance for first-order and second-order motion was observed in the cloud-with-pursuit condition. It seems that this result is not explained by weakness of response to

second-order motion. O'Keefe and Movshon (1998) also found that temporal frequency selectivity of second-order motion is different from that of first-order motion for the same neuron. It means that the speed tuning of second-order is different from that of first-order motion. For the Perrone (1992) model, and the Lappe & Rauschecker (1993) models, poor performance in the cloud-with-pursuit condition for second-order motion may be explained by inaccurate computation of speed if the accurate estimate of speed is required for heading recovery.

Differential performance for ground and cloud is also explained by the models. Lappe (1994) modified the original Lappe and Rauschecker (1993) model. The modified model showed more robustness to noise when static depth cues are available. The Perrone (1992) model can integrate static depth cues easily. Assuming that the second-order motion detectors are noisy, the difference in performance between ground and cloud is explained by the models.

**Conclusion**

Shape is not accurately recovered from second-order motion by the human visual system (Dosher et al., 1989; Landy et al., 1991). However, Gurnsey et al. (1998) showed that second-order motion contributes to vection. In this paper, we have shown that the heading direction is judged fairly accurately from second-order motion when self-motion across a ground plane with pursuit is simulated, although it is not judged correctly when self-motion through cloud-like dots with pursuit is simulated. It means that second-order motion contributes to human

locomotion perception to some extent. Vaina et al. (1998) reported that a patient, who had the selective impairment of perception of first-order motion, judged the direction of straight-line heading, but could not normally judge heading of translation along a curved path (see also Vaina et al., 1996). This paper has shown that the heading direction is accurately judged from second-order motion when self-motion without rotation is simulated. This suggests that the human visual system cannot decompose the flow fields into rotational and translational components by using pure second-order motion information alone, although it can do in some ways for the ground plane.

Acknowledgement

Appendix: Effects of background luminance on first-order motion energy detectors for random-contrast-polarity stimuli.

Fig. 7 (a) (c) and (e) show the stimulus representation of stimuli used in our experiments. We assume that the right side of the spatial axis is positive. Fig. 7(a) shows the stimulus with constant polarity, (c) shows an example of the stimuli with random polarity. Fig 7 (e) shows the same stimulus as in (c), but background luminance was increased by 20%. (The increase of the background luminance is exaggerated for visibility). First-order motion energy models can be thought of as detecting simple patterns of spectral power in the Fourier domain representation of an image corresponding to first-order motion (Adelson & Bergen, 1985). The power $P(f_x, f_t)$ of each spatio-temporal frequency component for stimulus of (a), (c) and (e) is shown in Fig. 7 (b), (d) and (f), respectively. When the stimulus for constant polarity moves at a constant speed, the spatio-temporal frequency with large power is on a line with slope of $f_x / f_t$ through the origin as shown in Fig. 7(b). Using $f_x$, $f_t$ with large powers, we can calculate the speed by $f_x / f_t$. The power spectrum for random polarity in Fig. 7(d) is quite deteriorated. It is impossible to compute the accurate speed from the power spectrum, that is, outputs of first-order motion energy detectors. Moreover, the effects of the 20 % increase of the background luminance on power spectrum are quite small as shown in Fig. 7(f).

We used a technique similar to Dosher et al. (1989) in order to simulate the detection of motion direction from first-order motion energy detectors. We compute the directional power

$D_P$ :

$$D_p = \sum_{f_x \bullet f_t > 0} P(f_x, f_t) - \sum_{f_x \bullet f_t < 0} P(f_x, f_t)$$

Note that the computation of the directional power was simplified compared with that of Dosher et al. (1989) by removing a spatial and temporal window and a threshold for power. However, this simplification scarcely affects the results. This right-left power difference summed over all frequency is considered to be a score of rightward motion from the whole first-order motion detectors with various spatial and temporal tunings and it is similar to the computation that is carried out by first-order motion models (Dosher et al., 1989). $D_P$ larger than 0 means the rightward motion and $D_P$ smaller than 0 means leftward motion.

We performed simulations to examine the effects of background luminance on the first-order directional power. Background luminance was increased by 0, 10, 20, 40, 80 or 100% from the average luminance of white or black dots. Note that decrease of the background luminance has the same effect on directional power. We generated 100 stimuli with random polarity for each increasing rate of background luminance and then calculated the directional power for every stimulus. Directional powers were normalized by the directional power of the constant-polarity stimulus in Fig. 7(a). The average normalized directional power for each increasing rate of background luminance is shown in Fig. 8. Effects of 10 % increase of

28

background luminance on the directional power were negligible. The percentage such that $D_P > 0$ was 50 % for the 10 % increasing rate, which was almost the same as the 0 % increasing rate. The average directional power increased as the increase of background luminance. It should be noted that the directional power for five out of 100 random-polarity stimuli was less than 0 even when the increasing rate of background luminance was 80%.

We demonstrated that it is almost impossible to detect the 1-D motion direction and the speed for random-polarity stimuli with 10%-perturbed background luminance from first-order motion energy detectors. Moreover, in order to compute the velocity of 2-D motion, more accurate outputs of motion energy detectors are necessary. For those reasons, we can say that the 2-D velocity of the random-polarity stimuli cannot be computed by first-order motion energy detectors even if the average luminance of white and black dots differs from background luminance by 10 %.

References

Adelson, E. H. & Bergen, J. R. (1985) Spatiotemporal energy models for the perception of motion. *Journal of Optical Society of America A, 2,* 284-299.

Albright, T. D. (1992) Form-cue invariant motion processing in primate visual cortex. *Science, 255,* 1141-1143.

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, C.B., Maxwell, M, Andersen, R.A., Banks, M. S. & Shenoy, K.V. (1996) Mechanisms 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.

Buračas, G. T. & Albright, T. D. (1996) Contribution of Area MT to perception of three-dimensional shape: A computational study. *Vision Research, 36,* 869-887.

Cavanagh, P. & Mather, G. (1989) Motion: The long and short of it. Spatial Vision, 5, 103-129.

Chubb, C. & Sperling, G. (1988a) Drift-balanced random stimuli: a general basis for studying non-Fouier motion perception. *Journal of Optical Society of America A, 5,* 1986-2007.

Chubb, C. & Sperling, G. (1988b) Processing stages in non-Fourier motion perception. Investigative Ophthalmology and *Visual Sciences (Suppl.), 29,*266.

Dosher, B. A., Landy, M.S. & Sperling, G. (1989) Kinetic depth effect and optic flow- I. 3D shape from fourier motion. *Vision Research, 29,* 1789-1813.

Gurnsey, R. Fleet, D., Potechin, C. (1998) Second-order motions contributes to vection. *Vision Research, 38,* 2801-2816.

Johansson, G. (1975) Visual motion perception. *Scientific American,232,* 76-88.

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

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

Hildreth, E. C., Ando, H., Andersen, R. A. & Treue, S. (1995) Recovering three-dimensional structure from motion with surface reconstruction. *Vision Research, 35,* 117-137.

Landy, M. S., Dosher, B. A., Sperling, G. & Perkins, M. E. (1991) The kinetic depth effect and

optic flow-II. First- and Second-order motion. *Vision Research, 31,* 859-876.

Lappe, M. (1996) Functional consequences of an integration of motion and stereopsis in area MT of monkey extrastriate cortex.. *Neural Computation. 8,* 1449-1461.

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.

Lu. Z. & Sperling, G. (1995) The functional architecture of human visual motion perception. *Vision Research, 35,* 2697-2722

Nishida, S. & Sato, T. (1995) Motion aftereffect with flickering test patterns reveals higher stages of motion processing. *Vision Research, 35,* 477-490.

Reichardt, W. (1961) Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In Rosenblith, W. A. (Ed.), Sensory Communication. New York: Wiley.

O'Keefe, L. P. & Movshon, J. A. (1998) Processing of first-and second-order motion signals by neurons in area MT of the macaque monkey. *Visual Neuroscience, 15,* 305-317.

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.

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

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. & Hildreth, E. C. (1996) Human heading judgments in the presence of moving objects. *Perception & Psychophysics, 58,* 836-56.

Royden, C. S. & Hildreth, E. C. (1999) Differential effects of shared attention on perception of heading and 3-D object motion. *Perception & Psychophysics, 61,* 120-133.

Smith, A. T. & Ledgeway, T. (1998) Sensitivity to second-order motion as a function of temporal frequency and eccentricity. *Vision Research, 38,* 403-410.

van Santen, J. P. H. & Sperling, G. (1984) Temporal covariance model of human motion perception. *Journal of Optical Society of America A, 1,* 451-473.

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.

Vaina, L. M., Royden, C. S., Bienfang, D.C., Makris, N. & Kennedy, D. (1996) Normal perception of heading in a patient with impaired structure from motion. Investigative Ophthalmology and *Visual Sciences (Suppl.), 37,* 515.

Vaina, L. M., Makris, N., Kennedy, D. & Cowey, A. (1998) The selective impairment of the perception of first-order motion by unilateral cortical brain damage. *Visual Neuroscience, 15,* 333-348.

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

van den Berg, A. V. & Brenner, E. (1994a) Why two eyes are better than one for judgement 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.

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

van Santen, J. P. H. & Sperling, G. (1984) Temporal covariance model of human motion perception. *Journal of Optical Society of America A, 2,* 300-321.

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

Figure Captions

Figure 1.   Contrast sensitivity of first- and second-order motion discrimination.

Moving dots whose contrast polarity was randomly determined on each frame were used as second-order motion. The sensitivities to first- and second-order motion were measured. The coherency and the contrast of the dots were varied. The horizontal axis indicates coherency and the vertical axis indicates percentages of correct direction discrimination. (a) and (b) show the results of SO, and (c) and (d) show those of MH. The results of first-order motion are shown in the left panel, those of second-order motion in the right panel.

Figure 2. Percent correct motion direction discrimination for 100%-contrast random-polarity stimuli and 25%-contrast constant-polarity stimuli. (a) Results of SO are shown. (b) Results of MH are shown.

Figure 3. Schematic diagram of a simulated path of an observer.

The X-axis and the Z-axis are the horizontal axis and the axis along the line of sight respectively.

(a) An observer translated in a fixed direction in the retinocentric coordinate frame while fixating a static point in an environment. The observer moved along the arrows of T from frame to frame. The step shown by the arrow in the diagram is much larger than actual one. In the world-centered

coordinates, the observer moved along a curved path. (b) An example of the actual path is shown. The heading direction was 10 deg. (c) A simulated path for pure translation without self-rotation is illustrated.

Figure 4.    The results for the ground-with-pursuit.

The horizontal axis represents the simulated heading direction and the vertical axis represents the perceived heading direction. When the points scatter along a dashed line with slope of 1.0, heading perception is unbiased. The regression lines are also shown.

Figure 5. The results for the cloud-with-pursuit.

The horizontal axis represents the simulated heading direction and the vertical axis represents the perceived heading direction. The regression lines are also shown.

Figure 6. The results for the cloud-without-rotation.

The horizontal axis represents the simulated heading direction and the vertical axis represents the perceived heading direction. The regression lines are also shown.

Figure 7.    Stimulus representations of moving stimuli. (a) shows the stimulus with constant polarity, (c) shows an example of the stimuli with random polarity, and (e) shows the same

stimulus as in (c), but background luminance was increased by 20%. (The increase of the background brightness is exaggerated for visibility). (b), (d) and (f) show the power of spatio-temporal frequency components for stimulus of (a), (c) and (e), respectively.

Figure 8. Effects of the increase of background luminance on the directional power. The horizontal axis indicates the increasing rate of background luminance from average luminance of dots and the vertical axis indicates the average of the directional powers of 100 random-polarity stimuli. The error bar shows the standard deviation of the directional powers of 100 stimuli.

Figure 1

Figure 2

(a)

the fixation point

Z

Z'

θ

θ

θ

θ

X'

X

(b)

Z

θ=10deg

X

(c)

Z

Z'

θ

θ

X'

X
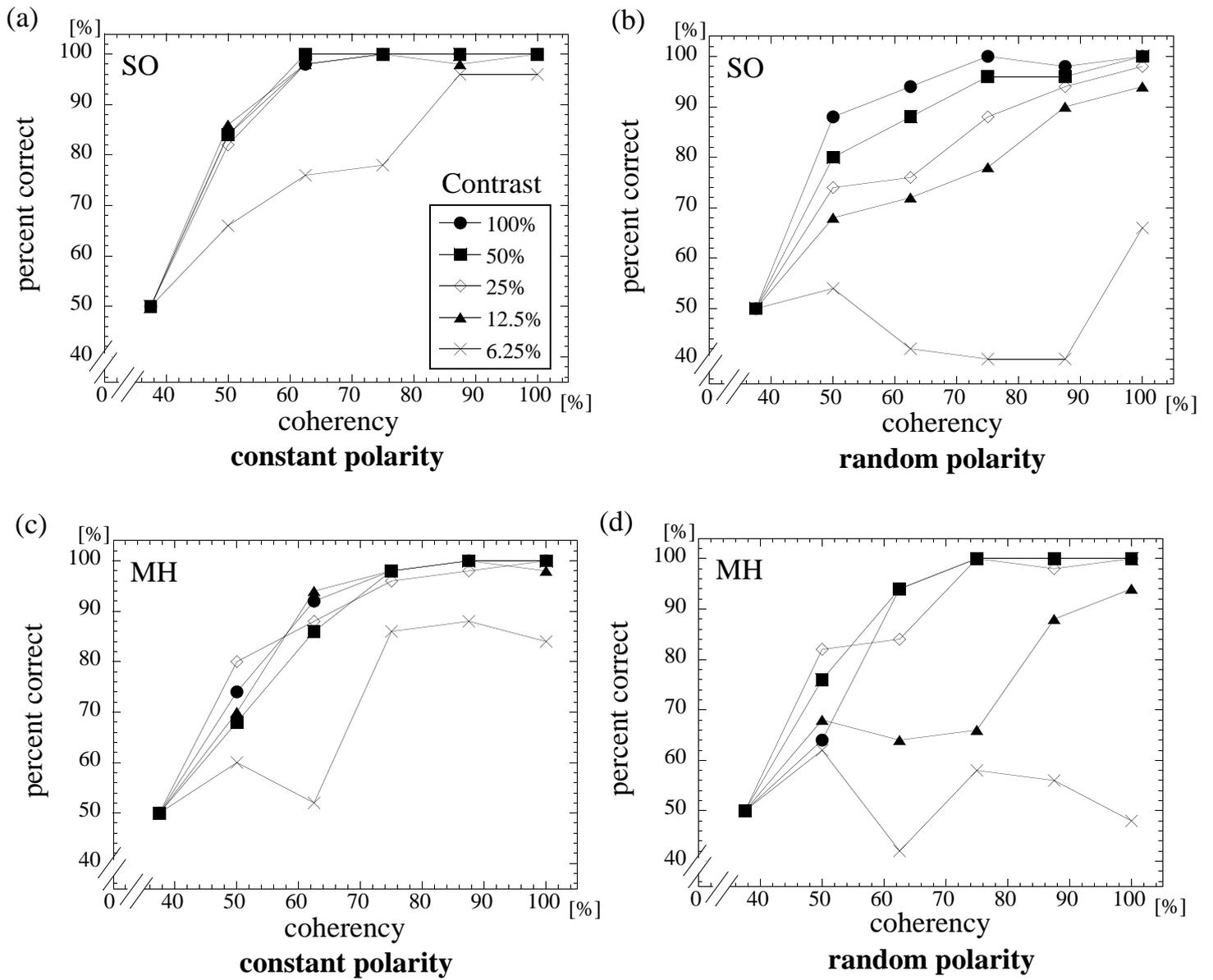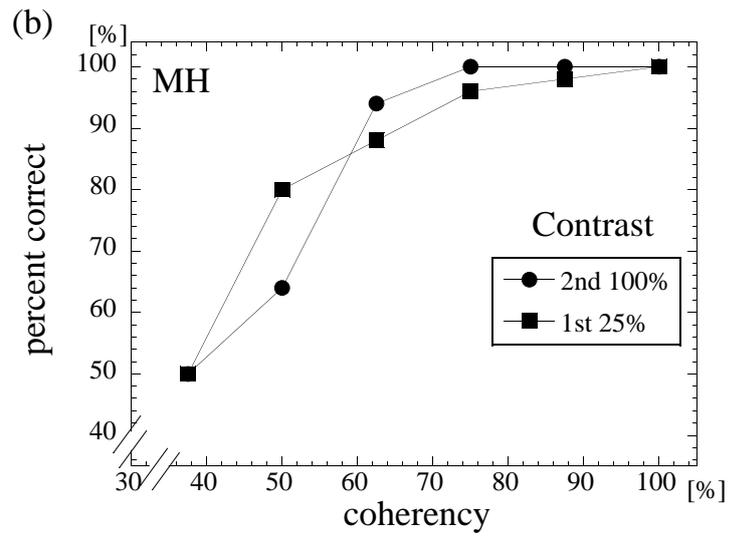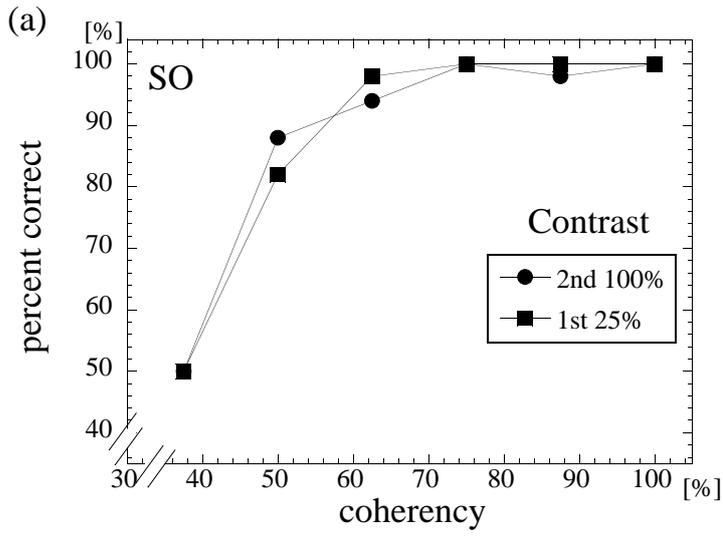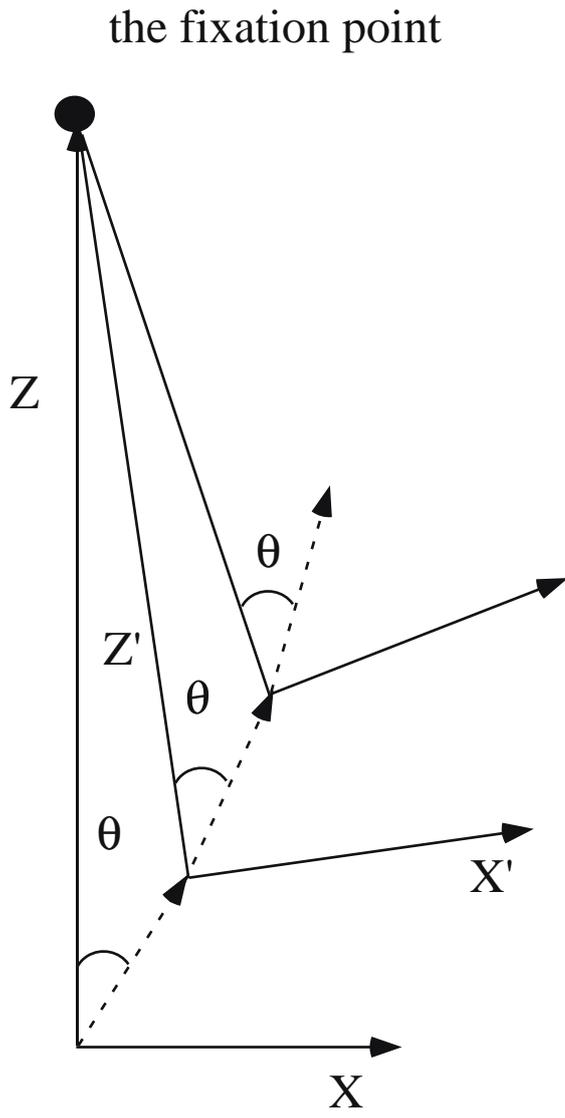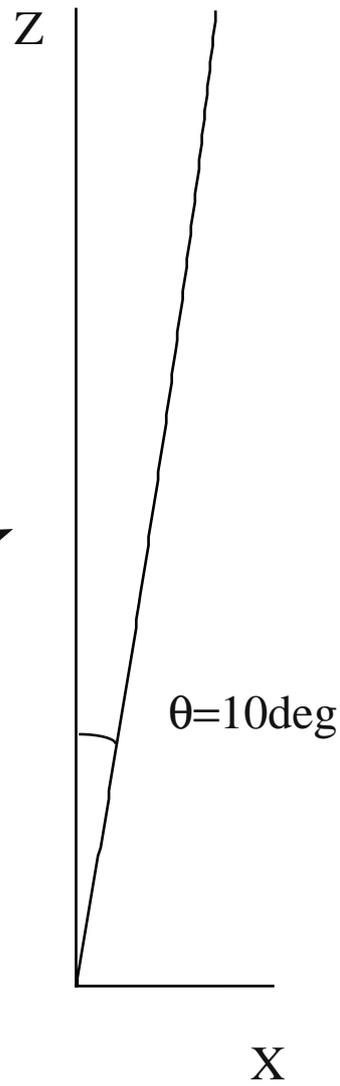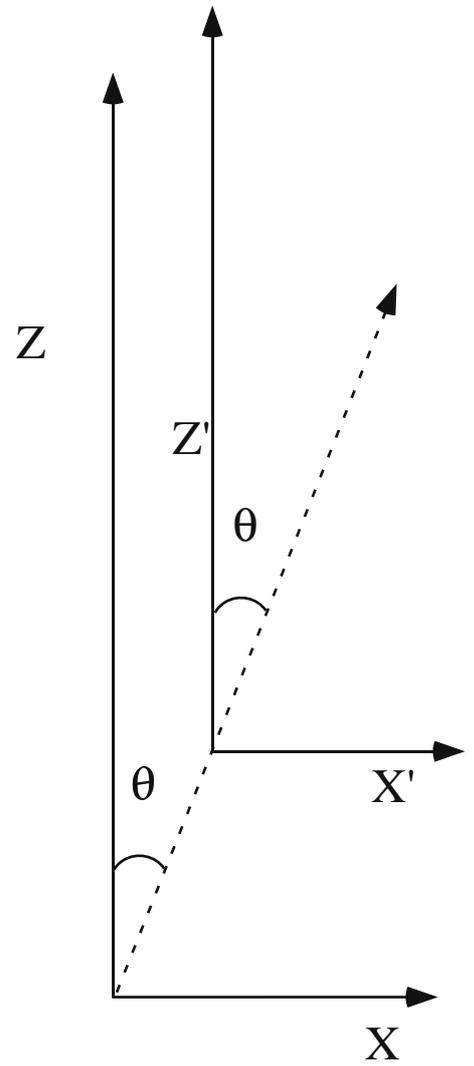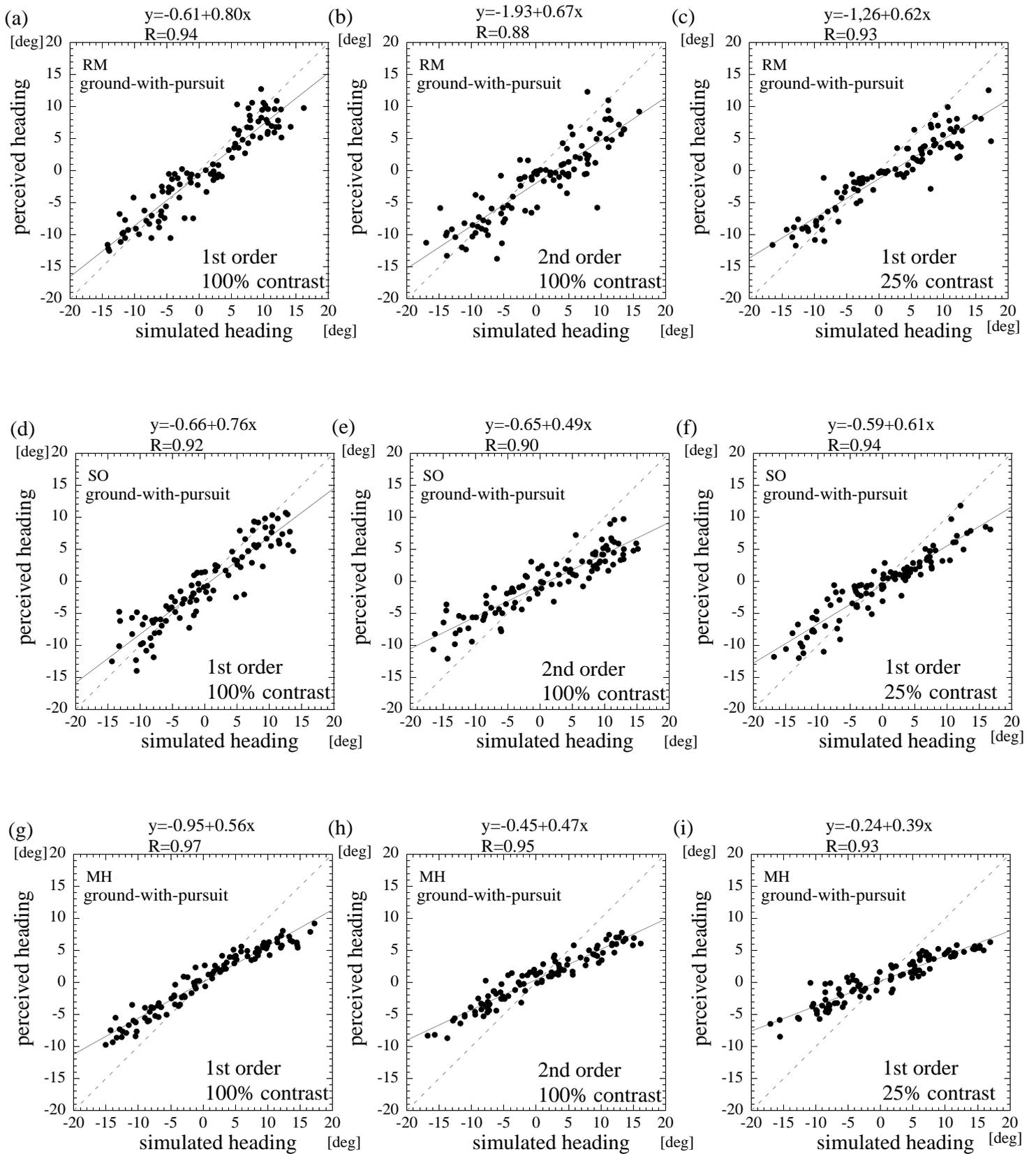
Figure 3

Figure 4

Figure 5

Figure 6

(a)

x

t

(b)

$f_t$

$f_x$

(c)

(d)

20% increase of background luminance

(e)

(f)

Figure 7

Figure 8