Computational analyses for illusory transformations in the optic flow field and heading perception in the presence of moving objects

Mitsuhiko Hanada

Department of Media Architecture
Future University-Hakodate
116-2 Kamedanakano-cho, Hakodate, Hokkaido, 041-8655, Japan

Keywords: optic flow, heading perception, focus of expansion

Hanada, M. (2005). Computational analyses for illusory transformations in the optic flow field and heading perception in the presence of moving objects. Vision Research, 45(6), 749-758.

Abstract

When we see a stimulus of a radial flow field (the target flow) overlapped with a lateral flow field or another radial flow field, the focus of expansion (FOE) of the target radial flow appears to be shifted in a direction. Royden and Conti [(2003) *Vision Research*, 43, 2811-26] argued that local motion subtraction is crucial for explanation of this phenomenon. The flow field which causes the illusory displacement of FOE was computationally analyzed. It was shown that the flow field is approximately a rigid-motion flow; the flow can be generated by simulating a situation where an observer moves toward a stationary scene. The heading direction for the observer corresponds to the perceived position of the FOE of the radial flow pattern. It implies that any algorithms which assume rigidity of the scene and recover veridical heading explain the bias in perceived FOE. There is no need for local motion subtraction in order to explain the phenomena. Furthermore, the flow for an observer's translation in the presence of objects moving laterally or in depth was computationally analyzed. It was found that algorithms which minimizes standard error functions with less weights to the independently moving objects show similar biases in recovered heading to the bias of human observers. It implies that local motion subtraction is not necessary for explanation of the bias in perceived heading due to an object moving laterally or in depth, contrary to the argument of Royden [(2002) *Vision Research*, 42, 3043-58].

1 Introduction

When we see a stimulus of a radial flow field overlapped with a lateral flow field, the focus of expansion (FOE) of the radial flow appears to be shifted in the direction of the lateral motion (Duffy & Wurtz, 1993; Grico & Lappe, 1998; Pack & Mingolla, 1998). Two explanations have been proposed for the phenomena. One is that motion opposite to the lateral movement is induced to the radial flow, and the induced motion displaces the FOE (Meese, Smith, & Harris, 1995). The other is that since the flow field of a radial flow plus a lateral flow is very similar to the flow caused by an observer's forward movement toward a front-parallel plane and an extremely distant plane, the visual system compensates for the lateral flow as a flow due to eye movement and recover the true FOE of the head-centered flow (Duffy & Wurtz, 1993; Lappe & Rauschecker, 1995). Recently Royden and Conti (2003) have proposed that Royden's model of heading perception using motion-opponent operators explains the illusory transformation of the optic flow field; the heading direction recovered by the model is consistent with the perceived FOE of the radial flow pattern. Since motion-opponent operators differentiate the velocity field locally, their explanation is similar to the explanation of induced motion by lateral motion in a sense. Also, since the model was developed to model heading perception, their explanation can be regarded as

explanation in terms of heading recovery with compensation for the lateral flow due to eye movement.

Royden and Conti (2003) demonstrated another illusory transformation of the optic flow. When a radial flow is overlapped with another radial flow, the FOE of the first flow appears to be shifted. The direction of the displacement depends on ratio of the simulated depth of the plane for the first flow to the simulated depth of the plane for the second flow. They argued that because the radial flow pattern cannot be generated by eye movement alone, the bias in perceived FOE could not be explained by compensation for the flow due to eye movement. Furthermore, they showed that the heading direction recovered by Royden's (1997) model of heading perception using motion-opponent operators is consistent with the bias in perceived FOE. Royden and Conti (2003) emphasized that local motion subtraction is crucial for explanation of the new illusion.

Indeed, the radial flow cannot be generated by eye movement. However, it does not mean that the flow does not occur when an observer moves in the rigid environment. We will show that the two-FOE flow can be approximately generated by simulating a situation where an observer moves toward two stationary planes with eye movement. In other words, the two-FOE flow is approximately a flow generated by rigid motion. The heading direction is located at a position displaced from the FOE of the first radial flow. The dis-

placement direction is consistent with the bias direction for human observers. Note two facts that Royden's model of heading perception computes heading reliably (Royden, 1997) and that the two-FOE flow field is a rigid-motion flow. They imply that her model simply computes the heading direction for the two-FOE flow. The other heading recovery models which assume rigidity of the scene and find veridical heading can also explain the new illusory transformations of the optic flow field. The purpose of the study is to show that local motion subtraction is not necessary to explain the illusory transformations of the optic flow field, and that the phenomena should be ascribed to heading estimation from two-FOE flows. Thus, heading models which estimate heading reliably can explain them.

The stimulus simulating a situation where an observer translates and objects moves laterally is a radial flow plus a lateral flow within a restricted region. The flow is similar to the flow for the illusory transformation of the optic flow field. The flow generated by simulating a situation where an observer translates and objects moves in depth has the two foci of expansion. It is similar to the flow for the Royden and Conti's (2003) illusory transformation of the optic flow field. Royden and Hildreth (1996) examined heading perception with objects moving laterally or moving in depth. The perceived heading was displaced from the FOE, and the bias direction was consistent with the illusory translation of the optic flow field. Royden

(2002) also showed that her heading perception model using motion-opponent operators (Royden, 1997) explains heading perception with objects moving laterally and moving in depth. Royden (2002) raised a question that other models of heading perception such as a neural network model of Lappe and Rauschecker (1993) and a gain-filed model of Beintema and van den Berg (1998) shows a similar bias to the bias for human observers due to an object moving in depth. We will also address the question in this paper. However, we will not test specific heading models directly. Instead, we will show that algorithms which minimizes standard error functions with less weights to the independently moving objects show similar bias in perceived heading due to moving objects. It implies that the bias in perceived heading due to moving objects should be also attributed to heading estimation from the flow fields and that most of the heading perception models can explain it.

2 Computational analysis of an illusory transformation of the optic flow field

2.1 Computational analyses of a radial flow with a lateral flow

We analyze the flow field for an illusory transformation of the optic flow, and show that the flow can be approximately generated by rigid motion.

We make use of essentially the same notation as Longuet-Higgins and

Prazdny (1980). We use a coordinate system that is fixed with respect to an observer. The translation of the observer in the rigid environment is expressed in terms of translation along three orthogonal directions, which we denote by the vector (U, V, W). U, V and W show translation along the X-axis, Y-axis and Z-axis respectively (Fig. 1). The Z-axis is directed along the optical axis, and the X-axis and Y-axis are horizontal and vertical respectively. The rotation of the observer is expressed in terms of rotation around three orthogonal axes, which we express by the vector (A, B, C). A, B and C indicate rotation around the X-axis, the Y-axis and the Z-axis, respectively (Fig. 1). The 3-D velocity of a point, (X, Y, Z) is given by:

$$\dot{X} = -U - BZ + CY \tag{1}$$

$$\dot{Y} = -V - CX + AZ \tag{2}$$

$$\dot{Z} = -W - AY + BX \tag{3}$$

where $(\dot{X}, \dot{Y}, \dot{Z}) \equiv (dX/dt, dY/dt, dZ/dt)$ (Longuet-Higgins and Prazdny 1980). If we consider perspective projection of the velocity onto the image plane $Z = 1$ for the projection, point P on the image (x, y) is given by

$$x = X/Z \tag{4}$$

$$y = Y/Z \tag{5}$$

The projected velocity $(u, v) \equiv (\dot{x}, \dot{y}) \equiv (dx/dt, dy/dt)$ in the image plane is

Figure 1: Insert the figure about here.

given by (Longuet-Higgins and Prazdny 1980)

$$u = \frac{-U + xW}{Z} - B + Cy + Axy - Bx^2 \quad (6)$$

$$v = \frac{-V + yW}{Z} - Cx + A + Ay^2 - Bxy \quad (7)$$

We consider situations where an observer translates without eye rotation around the line of sight, i.e., $C = 0$. Suppose that the observer moves toward a frontal plane with depth of Z_1 and an infinitely distant plane. For an image point (x_1, y_1) with depth of Z_1, the image velocity (u_1, v_1) is given by

$$u_1 = \frac{-U + x_1 W}{Z_1} - B + Ax_1 y_1 - Bx_1^2 \quad (8)$$

$$v_1 = \frac{-V + y_1 W}{Z_1} + A + Ay_1^2 - Bx_1 y_1 \quad (9)$$

Since the quadratic terms about x_1 and y_1 are small for moderate x_1 and y_1, we neglect them [1].

$$u_1 \approx \frac{-U + x_1 W}{Z_1} - B \quad (10)$$

$$v_1 \approx \frac{-V + y_1 W}{Z_1} + A \quad (11)$$

[1]Humans do not seem to notice the difference between the original and approximate flows at least up to a display size of 45 deg × 35 deg according to the data about heading perception toward a fronto-parallel plane (Grigo & Lappe, 1999; Warren & Hannon, 1990). The display size of the experiments of Royden and Conti (2003) was 25 deg × 25 deg.

The equations imply that the velocity pattern for the plane with depth of Z_1 is a radial pattern, whose center is $(U/W + BZ_1/W,\ \ V/W - AZ_1/W) = (U/W + B\tau_1,\ V/W - A\tau_1)$, where $\tau_1 = Z_1/W$ is time to contact of the plane. The center of the radial flow is displaced from the heading point $(U/W,\ V/W)$ by $(B\tau_1, -A\tau_1)$.

For a point with infinite distance ($Z = \infty$), the velocity (u_2, v_2) on the image point (x_2, y_2) is given by

$$u_2 = -B + Ax_2y_2 - Bx_2^2 \tag{12}$$

$$v_2 = A + Ay_2^2 - Bx_2y_2 \tag{13}$$

Again we neglect the quadratic terms about x_2 and y_2.

$$u_2 \approx -B \tag{14}$$

$$v_2 \approx A \tag{15}$$

The equations imply that the flow field for the infinitely distant plane is uniform motion. The flow field for translation toward a plane with depth of Z_1 and an infinite plane is approximately a radial flow plus a uniform flow.

Conversely, consider a flow field which consists of a uniform velocity (u_2, v_2) and a radial flow pattern with the center of (x_c, y_c) and time to

contact of τ_1. If we built a world with the following translation, rotation and distances, the resulting flow field would have almost the same two components of the flow field.

- Let W be an arbitrary positive real number. $A = v_2$, $B = -u_2$, $C = 0$, $U = (x_c - B\tau_1)W$ and $V = (y_c + A\tau_1)W$. There are two frontal planes. The depth of one plane is $\tau_1 W$. The depth of the other plane is infinite.

Therefore, the flow is a rigid-motion flow corresponding to the above situation. If human observers respond to the heading direction as the FOE of the radial flow (or the rotational flow is first compensated for and then human observers respond to the FOE of the remaining flow), the FOE will be shifted by $(u_2\tau_1, v_2\tau_1)$. This is the original explanation by Duffy and Wurtz (1993) who found the illusion.

2.2 Computational analysis of a flow with two foci of expansion

When a radial flow is overlapped with anther radial flow, the FOE of the first flow appears to be shifted (Royden & Conti, 2003). If the simulated depth of the first plane is larger than that of the second plane, the FOE for the first plane appears to be shifted in the direction of the FOE for the second plane. If the simulated depth of the first plane is smaller than that

of the second plane, the FOE of the first plane appears to be shifted in the direction opposite to the FOE of the second plane. We analyze the two-FOE flow pattern computationally. First we show that the flow with two foci of expansion is approximately a rigid-motion flow as long as the two radial flows have different values of time to contact.

We consider a situation where an observer moves toward two frontal planes with eye movement. Suppose that $C = 0$ and the depth of the two frontal planes are Z_1 and Z_2, respectively. The image velocities are approximately given by

$$u_1 \approx \frac{-U + x_1 W}{Z_1} - B \tag{16}$$

$$v_1 \approx \frac{-V + y_1 W}{Z_1} + A \tag{17}$$

for a point on the first plane and

$$u_2 \approx \frac{-U + x_2 W}{Z_2} - B \tag{18}$$

$$v_2 \approx \frac{-V + y_2 W}{Z_2} + A \tag{19}$$

for a point on the second plane. We neglected the quadratic terms about x_1, y_1, x_2 and y_2 to derive the above equations.

The flow for the first plane is a radial flow with the center of $(U/W + BZ_1/W,\ \ V/W - AZ_1/W) = (U/W + B\tau_1,\ \ V/W - A\tau_1)$, where $\tau_1 = Z_1/W$

is time to contact of the first plane. The flow for the second plane is also a radial flow with the center of $(U/W + B\tau_2,\ \ V/W - A\tau_2)$, where $\tau_2 = Z_2/W$ is time to contact of the second plane. The total flow is two radial flow patterns overlapped with each other.

Conversely, consider two radial flow patterns overlapped with each other such that the centers of the radial patterns are (x_{c1}, y_{c1}) and (x_{c2}, y_{c2}), and time to contact of the two planes is τ_1 and τ_2, respectively. If we built a world with the following translation, rotation and distances, the resulting flow field would have almost the same two components of the flow fields.

- Let W be an arbitrary positive real number. $A = -(y_{c1} - y_{c2})/(\tau_1 - \tau_2)$, $B = (x_{c1} - x_{c2})/(\tau_1 - \tau_2)$, $C = 0$, $U = (x_{c1} - B\tau_1)W$ and $V = (y_{c1} + A\tau_1)W$. There are two frontal planes. The depths of two planes are $\tau_1 W$ and $\tau_2 W$, respectively.

Therefore, the flow is a rigid-motion flow. If human observers respond to the heading direction as the FOE of the first radial flow (or the rotational flow is first compensated for and then human observers respond to the FOE of the remaining flow), the FOE will be displaced by $(-B\tau_1, A\tau_1) = (-(x_{c1} - x_{c2})\tau_1/(\tau_1 - \tau_2),\ \ -(y_{c1} - y_{c2})\tau_1/(\tau_1 - \tau_2))$. The bias direction depends on the sign of the $(\tau_1 - \tau_2) = (Z_1 - Z_2)/W$, and the predicted bias was consistent with the bias in perceived FOEs reported by Royden and Conti (2003). The

magnitude of the predicted bias is inversely proportional to $|\tau_1 - \tau_2|$, or difference of the depths of the two planes. Royden and Conti (2003) also reported that the bias in perceived FOEs increased with decrease of the depth differences.

2.3 Discussion

We have shown that there is at least one approximate solution under the rigidity assumption for the radial plus lateral flow pattern or for the two overlapped radial flow pattern. Is there another solution for the flow fields? It is well known that if a sufficiently large number of points (e.g. eight points) are not on the quadratic surface containing the origin, nor on the two planes with one plane containing the origin, we can uniquely determine the translation from the flow up to a scale factor (e.g., Kanatani, 1993). Neither of the planes for the solutions shown above contains the origin. It implies that the above scene and the observer's movement correspond uniquely to the flow up to a scale factor. (Arbitrary W corresponds to the scale factor.) Strictly speaking, however, we cannot apply this logic to these cases because we neglected some terms in the flow equations. However, it seems that there is no better solution than the one presented above.

Royden and Conti (2003) showed that Royden's (1997) heading perception model explains the illusory transformation of the optic flow field; The

perceived position of the FOE of a flow overlapped with another radial flow or lateral flow corresponds to the heading direction recovered by her model. Her model recovers heading reliably as long as there are enough local depth variations (Royden, 1997). We have shown that the flows for the illusions can be generated by simulating situations where an observer moves toward two stationary planes with different depths while rotating. Hence, Royden's model would recover the heading direction. However, any reliable heading recovery algorithms recover the heading direction. Thus, Royden's model is not unique one that explains the phenomena. Any heading perception models which find veridical heading under the rigidity assumption would explain them. Royden and Conti (2003) emphasized that local motion subtraction is crucial for explanation of the illusions. It is not the case. The key to explanation for the illusory transformation of the optic flow field is the fact that the flows which cause the illusions are approximately rigid-motion flows.

Royden and Conti (2003) reported that the illusory transformation of the optic flow field was larger when each radially moving dot was paired with a dot for another radial or lateral flow within a limited spatial region. They argued that the result supports local motion subtraction rather than global one. However, they did not show that Royden's model using motion-opponent operators explains the difference in the illusion between the matched and unmatched stimuli. It seems that the difference is ascribed to other reasons

than local motion subtraction. The matching of dots would affect primary motion processing and would have effects on motion perception such as motion transparency and computation of self-rotation. It may cause difference in magnitude of the illusory transformations of the optic flow field between the matched and unmatched stimuli.

We treated a radial flow overlapped with another radial flow and a radial flow with another lateral flow separately. If we use projective geometry, however, the distinction is unnecessary (See Kanatani (1991) for image analyses using projective geometry). Any two lines cross a point projectively. Two parallel flow lines cross at a point on an infinite line. The FOE of the lateral flow is located on an infinite line. Time to contact of a lateral flow is also infinite. Thus, in a projective space, we can deal with lateral flows in the same way as with radial flows. A radial flow overlapped with a lateral flow is a special case of two-FOE flows in projective geometry. I analyzed the two cases separately for the readers who are not familiar with projective geometry. It should be noted, however, that the illusory transformation of the optic flow filed due to a lateral flow can be computationally explained in the same way as that due to a radial flow can in projective geometry.

3 Computational analyses of heading judgement in the presence of moving objects

Royden and Hildreth (1996) examined heading perception in the presence of an object moving laterally. They showed that perceived heading was biased in the direction of the object's motion when a moving object crossed the heading point. Although the object spanned a restricted region, the flow was similar to an expansion flow plus a lateral flow for the optic flow illusion reported by Duffy and Wurtz (1993) and the bias in perceived heading in the presence of an object moving laterally was consistent with the bias of perceived FOE for the illusory transformation of the optic flow fields.

Royden and Hildreth (1996) also examined heading perception in the presence of an object moving in depth. They showed that perceived heading was biased in the direction of the object's FOE. The flow had two FOEs, and was similar to the flow with two FOEs for the illusory transformation of the optic flow field used by Royden and Conti (2003).

Royden (2002) reported that her model of heading perception shows a bias similar to the bias for human observers. Royden (2002) argued that local motion subtraction is crucial for explanation of the bias. However, the bias may be ascribed to the property of the flow itself. We analyze the flow fields computationally, and show that the bias in perceived heading due to a

moving object would be ascribed to heading recovery from the flow fields.

Royden and Hildreth (1996) simulated situations where an observer moved toward two static planes in the presence of a moving object. Since the number of the stationary planes were two (and not one), the flows used by Royden and Hildreth (1996) were not rigid-motion flows; they could not be generated by simulating situations where an observer moved in a stationary environment. Hence, computational analyses like those for the illusory transformation of the optic flow field are not possible. Instead, we take a different approach. We recover heading for the flow minimizing standard error functions for heading recovery. We will show that the optimization shows a bias similar to the bias in perceived heading for human observers when weights to the region of the moving object are reduced.

3.1 Error functions

We compute heading from the flow in the presence of a moving object used by Royden and Hildreth (1996) by minimizing error functions. Let (u_i, v_i) be the i-th velocity on the image point (x_i, y_i) $(i = 1, \cdots, N)$. Let $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{U}$, $\hat{V}$ and $\hat{W}$ be estimates of A, B, C, U, V and W, respectively, and let $\hat{Z}_i$ be an estimate of the depth of the i-th image point (Z_i). A natural error function is

$$J_1 = \sum_{i=1}^{N} \omega_i^2 \left((u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2 \right) \tag{20}$$

where

$$\hat{u} = \frac{-\hat{U} + x_i \hat{W}}{\hat{Z}} - \hat{B} + \hat{C} y_i + \hat{A} x_i y_i - \hat{B} x_i^2 \tag{21}$$

$$\hat{v} = \frac{-\hat{V} + y_i \hat{W}}{\hat{Z}} - \hat{C} x_i + \hat{A} + \hat{A} y_i^2 - \hat{B} x_i y_i \tag{22}$$

and ω_i is a weight for the i-th image point. The error function J_1 is minimized over $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{U}$, $\hat{V}$, $\hat{W}$ and $\hat{Z}_i$ $(i = 1, \cdots .N)$. Velocity $(\hat{u}_i, \hat{v}_i)$ indicates the predicted velocity on the i-th image point (x_i, y_i) computed from the estimates. The error function is the sum of the weighted errors for the image velocities.

Another possible error function is the sum of errors for the epipolar constraint. Let $\mathbf{x}_i$ and $\mathbf{u}_i$ be $(x_i, y_i, 1)^t$ and $(u_i, v_i, 0)^t$, respectively, and let $\mathbf{T}$ and $\mathbf{R}$ be $(U, V, W)^t$ and $(A, B, C)^t$, respectively. The following equation holds (Bruss & Horn, 1983; MacLean, Jepson, & Frecker, 1994; Zhuang, Ahuja, & Haralick, 1988).

$$\mathbf{T}^t(\mathbf{u}_i \times \mathbf{x}_i) + (\mathbf{T} \times \mathbf{x}_i)^t(\mathbf{x}_i \times \mathbf{R}) = \mathbf{0} \tag{23}$$

where $\times$ indicates the cross (outer) product. The equation is derived by eliminating Z from Eqs. (6) and (7). It is an instantaneous-time version

of the epipolar constraint (Kanatani, 1993). Hence, we may define an error function as follows.

$$J_2 = \sum_{i=1}^{N} \left| \psi_i \left\{ \hat{\mathbf{T}}^t (\mathbf{u}_i \times \mathbf{x}_i) + (\hat{\mathbf{T}} \times \mathbf{x}_i)^t (\mathbf{x}_i \times \hat{\mathbf{R}}) \right\} \right|^2 \tag{24}$$

where $\hat{\mathbf{T}}$ and $\hat{\mathbf{R}}$ are $(\hat{U}, \hat{V}, \hat{W})^t$ and $(\hat{A}, \hat{B}, \hat{C})^t$, respectively, and ψ_i is a weight for the i-th image point. The error function J_1 is minimized over $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{U}$, $\hat{V}$ and $\hat{W}$.

The two error functions are representatives of the error functions used for recovery of camera motion in computer vision. Since the magnitude of $\mathbf{T}$ cannot be determined in principle, the error function is minimized with a constraint such as $|\mathbf{T}| = 1$.

We minimized the error functions for the flow generated by an observer's translation in the presence of the a moving object by Algorithm II of Ruhe and Wedin (1980) (a version of the Gauss-Newton method) with a constraint of $|\mathbf{T}| = 1$, and examined a bias in recovered heading due to the moving object. We call the method for minimizing J_1 Algorithm A and that for minimizing J_2 Algorithm B. We simulated Experiment 1 and Experiment 8 of Royden and Hildreth (1996). Royden (2002) also performed simulations of her model (Royden, 1997) for the experiments. We show below that the bias in heading recovered by the optimizations due to the moving object is

qualitatively similar to the bias in perceived heading for human observers when the weights (ω_i and ψ_i) are appropriately set.

3.1.1 Simulation of Royden and Hildreth's (1996) Experiment 1

We performed a simulation of Royden and Hildreth's (1996) Experiment 1. In the experiment, observers viewed a simulated scene of an observer moving toward two transparent planes of moving dots with an independently moving object. The object moved laterally relative to the observer. Hence, the size of the object and the distance from the observer did not change during the stimulus presentation. The initial distances from the observer to the two planes were 400 cm and 1000 cm from the observer. The observer's translation speed toward the planes was 200 cm/s. The simulated heading was 4, 5, 6, or 7 deg to the right of the center of the display. The object was a 10 deg $\times$ 10 deg opaque square. It moved to the left or right at a speed of 8.1 deg/s. For a leftward moving object, initial positions of the object were -1.4, 0.6, 4.7, 8.7, 10.7 and 12.7 deg from the center. For a rightward moving object, initial positions of the object were -9.9, -5.9, -1.9, 0.2, 2.2 and 6.3 deg from the center. The viewing window was 30 deg $\times$ 30 deg. For each simulation run, 100 points randomly positioned in the window was generated and 20 points on the moving object was also generated. The number of points was one-fifth of the number of moving dots for the stimuli used in the

psychophysical experiment of Royden and Hildreth (1996). We reduced the number of points because it would take too long time to minimize J_1 if the same number of points as in the psychophysical experiment were employed. However, the reduction of the points would not affect the heading estimates. The positions and the image velocities at the middle of the presentation (i.e., 0.4 s after the start of the presentation) were then computed for each point in the scene. (The presentation time was 0.8 s in their experiment.) The positions and velocities were used as input for the heading recovery. For each object position, 100 runs were performed. The data below show the average over the 100 runs.

We set the weights ω_i and ψ_i as follows.

$$\omega_i = \alpha_\omega \exp\left(-\frac{(x_i - x_c)^2 + (y_i - y_c)^2)}{2\sigma_\omega^2}\right) \tag{25}$$

$$\psi_i = \alpha_\psi \exp\left(-\frac{(x_i - x_c)^2 + (y_i - y_c)^2)}{2\sigma_\psi^2}\right) \tag{26}$$

where $(x_c,\ y_c)$ is the position of the FOE of the static planes, and σ_ω was $5.0 \times \pi/180$ (i.e., 5.0 deg), σ_ψ was $10.0 \times \pi/180$ (i.e., 10.0 deg), and α_ω and α_ψ were 1.0 when the i-th point belonged to either of the two stationary planes, and they were 0.1 when it belonged to the object. The region around the FOE was weighted more heavily by a Gaussian function of the distance between the FOE and the flow point in order to simulate the result that human observers show a larger bias for an object that covers the FOE. We

Figure 2: Insert the figure about here.

used the different weights for the object and for the stationary scene since the bias due to the moving object was too large when the equal weights were used. The FOE for the stationary scene was used as the initial value of $|T|$ for the optimization.

Fig. 2 shows the average bias in the heading recovered by Algorithms A and B. A positive bias means a bias to the right, and negative one indicates a bias to the left. For a leftward moving object, the average bias for both of the algorithms was leftward, and for a rightward moving object, the bias was rightward. The position of the object affected the size of the bias. When the object covered the FOE of the stationary scene, the effect tended to be larger than when it did not, though for Algorithm A the effect of the leftward moving object was largest when the object covered the region right to the FOE. The bias for the algorithms was qualitatively similar to the bias for human observers.

The bias for the algorithms is interpreted as follows. Since the algorithms cannot treat non-rigid-motion flows, the algorithms must explain the lateral flow as eye movement or lateral heading. However, the lateral heading is implausible because the other part of the flow is a radial flow pattern. Hence the algorithms regard the flow of the laterally moving object as the flow

due to eye rotation. If the lateral flow is fully explained by eye movement, however, the other part of the flow (the radial flow) will not be explained well. (Note that if the stationary scene is one frontal plane, the radial pattern can be explained as the illusory transformation of the optic flow can. However, two frontal planes with different depths cannot have the same center of flow unless there is no rotation (i.e. unless $|\mathbf{R}| = 0$)). Hence the observed bias in recovered heading would be some compromise between the two demands from the lateral flow and the radial flow.

3.1.2 Simulation of Royden and Hildreth's (1996) Experiment 8

We performed a simulation of Royden and Hildreth's (1996) Experiment 8. In the experiment, an object moved in depth relative to the observer. The stationary scene and the observer's movement were the same as in the first simulation. The moving object was an opaque square that moved toward the observer at a speed of 300 cm/s, and an angle of motion relative to the observer was 1 or 10 deg to the right of the center of the viewing window. Object starting positions were -1.0, 0.6, 2.25, 3.9, 5.5 and 7.1 deg for 1-deg heading, and 0.6, 2.25, 3.9, 5.5, 7.1, 9.9 deg were for 10-deg heading. The object's starting size was 8 deg $\times$ 8 deg and the final size was 20 deg $\times$ 20 deg. The positions and the image velocities at the middle of the presentation were computed for each point in the scene, and used as input for the heading

Figure 3: Insert the figure about here.

recovery. Algorithms used in these simulations were the same as in the first simulations.

Fig. 3 shows the average bias generated by Algorithms A and B as a function of the starting position of the object. Both of the algorithms showed a leftward bias for the moving object with 1-deg FOE and a rightward bias for the object with 10-deg FOE. The tendency was qualitatively similar to the bias for human observers.

The bias for the algorithms is interpreted as follows. Time to contact of the object was about 1.3 s, and times to contact of the stationary two planes were 2.0 s and 5.0 s. Remember that the computational prediction of the bias for two radial flow pattern is $-(x_{c1} - x_{c2})\tau_1/(\tau_1 - \tau_2)$ (see Section 2.2), where τ_1 is time to contact of the target radial flow, τ_2 is time to contact of the biasing flow, and x_{c1} and x_{c2} are the x coordinates of the center of the target and biasing flows, respectively. For this simulation, τ_2 corresponds to time to contact of the object, that is, $\tau_2 = 1.3$ [s]. We cannot decide τ_1 because the target radial flow is the two overlapped radial flow with different values of time to contact, and we might choose the following three value; $\tau_2 = 2.0$ (minimum), $\tau_2 = 5.0$ (maximum), or $\tau_2 = (5.0 + 2.0)/2$ (average). In either case, $\tau_1 - \tau_2$ is positive. Since $x_{c1} = 4$, 5, 6 or 7 deg, the computational

prediction should be negative for 1-deg heading of the object ($x_{c2} = 1.0$ [deg]), and it should be positive for 10-deg heading of the object ($x_{c2} = 10.0$ [deg]). This predicted bias direction is consistent with the bias directions for Algorithms A and B, Royden's model and human observers.

3.1.3 Discussion

We have shown that the bias in perceived heading due to independently moving objects is qualitatively similar to the bias for the algorithms that recover heading by minimizing weighted sum of square errors of image velocities or weighted sum of epipolar errors. Hence, Royden' (1997) heading perception model using motion-opponent operators is not a unique model to explain the effects of moving objects on heading perception. Since two error functions are representatives used in algorithms for camera motion recovery, many heading algorithms explain the effects of an independently moving object. Also, some models for heading perception were developed on the basis of the error functions. For example, Lappe and Rauschecker (1993, 1995) developed neural network models for heading judgement, which is based on Heeger and Jepson' (1990, 1992) algorithm that minimizes the error function J_1. Their model should also explain the effects of moving objects.

However, we do not mean that the human visual system actually minimizes these error functions. Instead, we intend to show that the key to

explanation of the effects of the moving objects is not local motion subtraction contrary to the argument of Royden (2002). Many algorithms would explain the effect since the rigid-motion flow nearest to the flow with a moving object used by Royden and Hildreth (1996) in a weighted least-square or least-epipolar-error sense is the flow generated by an observer's translation in the biased direction with some eye movement.

There were some discrepancies between the predicted bias by Algorithms A and B and that of human observers. For example, Algorithm A shows a large bias even when the object did not cover the observer's path. It seems that Algorithm B explains the human bias better. However, the peak for the bias generated by Algorithm B was shifted slightly from that of the bias of human observers. Also, the magnitude of the bias for the two algorithms was larger than that for human observers, although we used fairly small weights for the independently moving object. On the other hand, Royden's model predicts the position of the peak bias well, and explains the bias not only qualitatively but also quantitatively. We did not try to find the best weights to explain the human performance because the purpose of this research is not to develop good models to explain the human performance, but to analyze the flow computationally. There may be better weights for the error functions J_1 and J_2. Also, it is implausible that the human visual system minimizes the error functions directly as we did. Especially, direct minimiza-

tion of J_1 is computationally costly. Indirect minimization used by Lappe and Rauschecker (1993, 1995) is needed to implement the algorithm actually. Furthermore, computation for heading recovery should be performed by neurons in the brain. If some neural constraints are included in the computation, better prediction might be obtained.

We used Gaussian-function weighting around the FOE for the stationary scene. The weighting was also adopted by Royden (2002) for weights from motion-opponent operators to heading template cells. The weighting was required to explain the fact that human observers show the bias in perceived heading due to a moving object only when the object crosses the observer's path. However, a question arises; how does the visual system weight the points as a Gaussian centered on the FOE before the heading estimation? The weights might not be constant during the computation of heading. Perhaps the visual system changes the weights dynamically so that the points around the current estimate of heading would be weighted more heavily. It is also possible that heading might be estimated by a template method using template units with larger weights for a region around the heading to which the units are tuned. If the output of the template units reflect the error function of Eq. (20) or (24), the template method may be regarded as approximate minimization of the error function.

Furthermore, we used different weights for the object and for the sta-

tionary scene, though Royden (2002) did not use the different weights. The flows for the psychophysical experiments of Royden and Hildreth (1996) are non-rigid-motion flows. The visual system might notice the object moving independently, and might attempt to neglect the object's movement for the estimation of the observer's movement. The small weighting value reflects the degree of the neglect, though it seems that the visual system cannot neglect it completely. However, detection of objects moving independently is computationally hard. Can the visual system detect independently moving objects?

The observers who participated in their experiments might know that the object moves independently. Two out of the five observers who participate in their experiments were the authors themselves, who should know it. The other observers participated in some practice sessions before the experimental session (perhaps without moving objects). The moving object in experimental sessions should draw attention, and the observers might notice that the object moves independently. Also, there may be a mechanism which detects objects moving independently. In fact, there are many independently moving objects when we walk in everyday life (e.g., automobiles, animals, people and so on.), but we usually know which objects are moving independently of our own movement. Although the computation of the detection of objects moving independently during self-motion is fairly hard, some algorithms for

it have been presented (Adiv, 1985; Hildreth, 1992; MacLean, Jepson, & Frecker, 1994; Tian & Shah, 1997). These algorithms improve estimation of self-motion by discarding motion for the objects moving independently as we decreased the weights for the points on the objects. Thus, it seems reasonable to assume that the visual system uses different weighting for the object and for the stationary scene.

Warren and Saunders (1995) also examined effects of an object moving in depth. Their stimuli were generated by simulating situations where an observer moves towards a plane with an object moving independently in depth. The situations were very similar to those that Royden and Hildreth (1996) used, and the direction of a bias in perceived heading of human observers reported by Warren and Saunders (1995) was similar to the bias observed in Experiment 8 of Royden and Hildreth (1996). Warren and Saunders (1995) simulated a single plane as a stationary scene and the flow pattern was a two-FOE flow pattern as was analyzed for the illusory transformation of the optic flow above. Therefore, we can find a stationary scene corresponding to the flow pattern. The bias direction reported by Warren and Saunders (1995) is also consistent with the bias predicted computationally.

4 General discussion

We analyzed the flow for the illusory transformation of the optic flow field and heading perception in the presence of independently moving objects. We showed that there exists a flow that is generated by an observer's movement in a stationary scene, similar to a radial flow pattern overlapped with another lateral flow or radial flow, which causes an illusory shift of the FOE. We found that the heading direction for the observer's movement corresponds to perceived FOE. We also showed algorithms which minimized the weighted sum of square errors of image velocities or errors of the epipolar constraint explains the bias in perceived heading due to independently moving objects. The computational analyses imply that the bias in perceived heading and FOE should be ascribed to heading estimation from the flows.

Marr (1982) proposed three levels of explanation. One is a computational level of explanation, the second is an algorithmic and expressive level of explanation, and the third is an implementational level of explanation. Furthermore, he suggested that for each visual phenomenon there exists an appropriate level of explanation. He gave an example. When we view a Necker cube, the depth direction appears to reverse. He suggested that the phenomena should be explained in the computational level. The depth direction cannot be uniquely determined from the image because an image of an

object generated by orthographic projection is the same as an image of the object with mirror-reflected depth. In general, structure can be determined from orthographic projection only up to reflection; a pair of solutions exists for an image viewed from orthographic projection. The unstable perception reflects the two possible solutions. The computational level would be an appropriate level of explanation. In this paper, we presented computational explanation of the illusory transformations of the optic flow field reported by Duffy and Wurtz (1993) and Royden and Conti (2003), and heading perception in the presence of moving objects, and show that the phenomena are well explained computationally. Royden (2002) and Royden and Conti (2003) argued that local motion subtraction or motion-opponent operators play a crucial role in the phenomena. Their explanation is algorithmic or implementational. However, we have shown that neither local motion subtraction nor motion-opponent operator is required to explain the phenomena. The appropriate level of explanation for the phenomena is computational, and not algorithmic nor implementational. It should be noted, however, that the computational analyses do not deny explanation of Royden and Royden and Conti (2002). Royden's (1997) model predicts the biases in perceived heading and FOE due to another flow pattern quite well (Royden, 2002; Royden & Conti, 2003). However, the computational analyses indicate most of heading models can explain the biases.

References

Adiv, G. (1985). Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **7**, 384-401.

Beintema, J. A., & van den Berg, A. V. (1998) Heading detection using motion templates and eye velocity gain fields. *Vision Research*, **38**, 2155-79.

Bruss, A. R., & Horn, B. K. (1983). Passive navigation. *Computer vision, graphics, and image processing*, **21**, 3-20.

Duffy, C. J., & Wurtz, R. H. (1993). An illusory transformation of optic flow fields. *Vision Research*, **33**, 1481-90.

Grigo, A., & Lappe, M. (1998). Interaction of stereo vision and optic flow processing revealed by an illusory stimulus. *Vision Research*, **38**, 281-90.

Grigo, A., & Lappe, M. (1999). Dynamical use of different sources of information in heading judgments from retinal flow. *Journal of the Optical Society of America A*, **16**, 2079-91.

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 I: Algorithm and implementation. *International Journal of Computer Vision*, **7**, 95-117.

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

Lappe, M., & Rauschecker, J. P. (1995). An illusory transformation in a model of optic flow processing. *Vision Research*, **35**, 1619-31.

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

Kanatani, K. (1991). Computational Projective Geometry. *Computer vision, graphics, and image processing. Image understanding*, **54**, 333-348.

Kanatani, K. (1993). 3-d interpretation of optical flow by renormalization. *International journal of computer vision*, **11**, 267-282.

MacLean, W. J., Jepson, A. D., & Frecker, R. C. (1994). Recovery of ego-motion and segmentation of independent object motion using the EM

algorithm. In *Proceedings of the 5th British Machine Vision Conference*, York, UK, pp. 13-16.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman & Co.

Meese, T. S., Smith, V., & Harris, M. G. (1995). Induced motion may account for the illusory transformation of optic flow fields found by Duffy and Wurtz. *Vision Research*, **35**, 981-5.

Pack, C., Mingolla, E. (1998). Global induced motion and visual stability in an optic flow illusion. *Vision Research*, **38**, 3083-93.

Royden C. S. (1997). Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *Journal of the Optical Society of America A*, **14**, 2128-43.

Royden C. S. (2002). Computing heading in the presence of moving objects: a model that uses motion-opponent operators. *Vision Research*, **42**, 3043-58.

Royden, C. S., & Conti, D. M. (2003). A model using MT-like motion-opponent operators explains an illusory transformation in the optic flow field. *Vision Research*, **43**, 2811-26.

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

Ruhe, A., & Wedin, P. A. (1980). Algorithms for separable nonlinear least squares problems. *SIAM Review*, **22**, 318-337.

Tian, Y. T., & Shah, M. (1997). Recovering 3D motion of multiple objects using adaptive Hough transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**, 1178-83.

Warren, W. H. Jr., & Saunders, J. A. (1995). Perceiving heading in the presence of moving objects. *Perception*, **24**, 315-31.

Warren, W. H. Jr., & Hannon D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America A*, **7**, 160-9.

Zhuang, X., Huang, T. S., Ahuja, N., & Haralick, R. M. (1988) A simplified linear optic flow-motion algorithm. *Computer vision, graphics, and image processing*, **42**, 334-344.

Figure captions

- Figure 1. An external coordinate system moving with the observer who is located at the origin and the corresponding image coordinates. The observer translates by (U, V, W) and rotates by (A, B, C). A point $P = (X, Y, Z)$ is projected on a image plane ($Z = 1$). The coordinates of the projected point p is (x, y).

- Figure 2. Biases in heading estimates caused by a laterally moving object. The bias (difference between the heading estimate and the actual simulated heading) is plotted as a function of the starting position of the object. A positive bias indicates a bias to the right and a negative bias indicates a bias to the left. Also, positive and negative object positions indicate starting positions to the right and to the left of the center, respectively. Circles indicate the average bias for Algorithm A (which minimizes the error function J_1) and squares indicate the average bias for Algorithm B (which minimizes the error function J_2). (a) Biases for a leftward moving object. (b) Biases for a rightward moving object.

- Figure 3. Biases in heading estimates caused by an object moving in depth. The bias (difference between the heading estimate and the actually simulated heading is plotted as a function of the starting po-

sition of the object. A positive bias indicates a bias to the right and a negative bias indicates a bias to the left. Also, positive and negative object positions indicate starting positions to the right and to left of the center, respectively. Circles indicate the average bias for Algorithm A and squares indicate the average bias for Algorithm B. (a) Biases for an object with FOE at 1 deg to right of the center. (b) Biases for an object with FOE at 10 deg to right of the center.

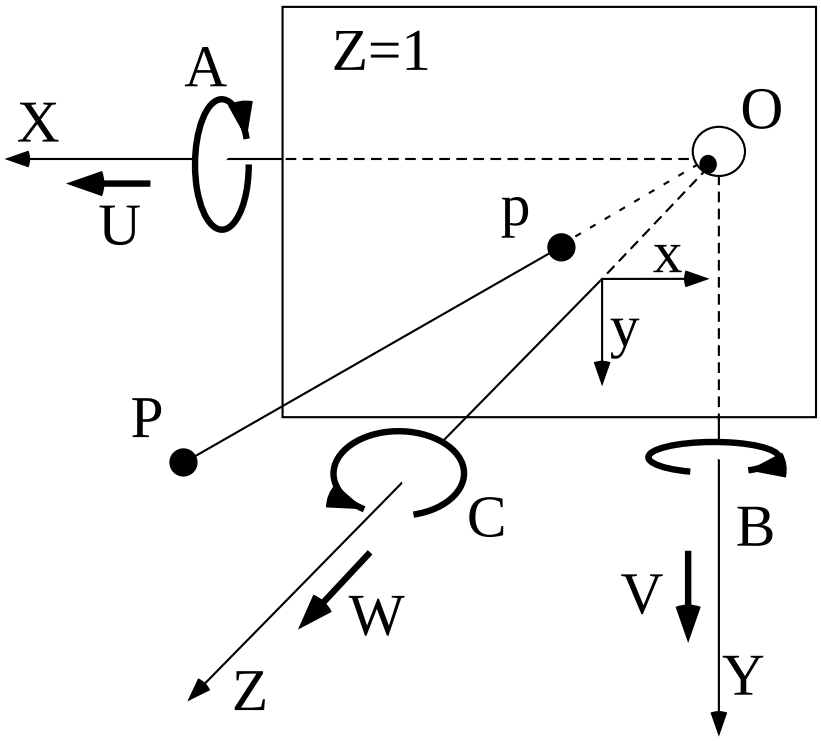


Figure 1

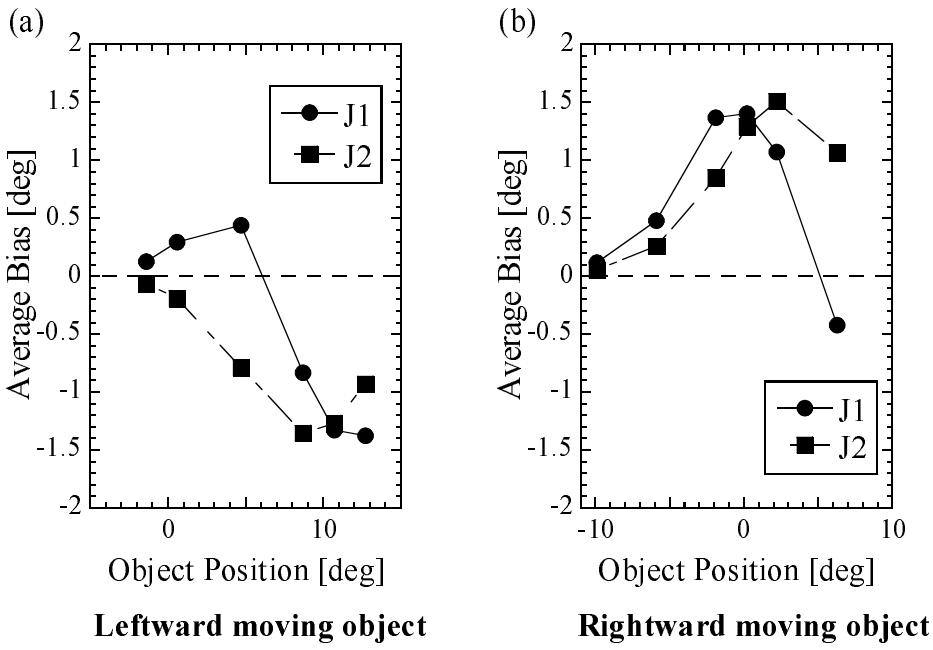
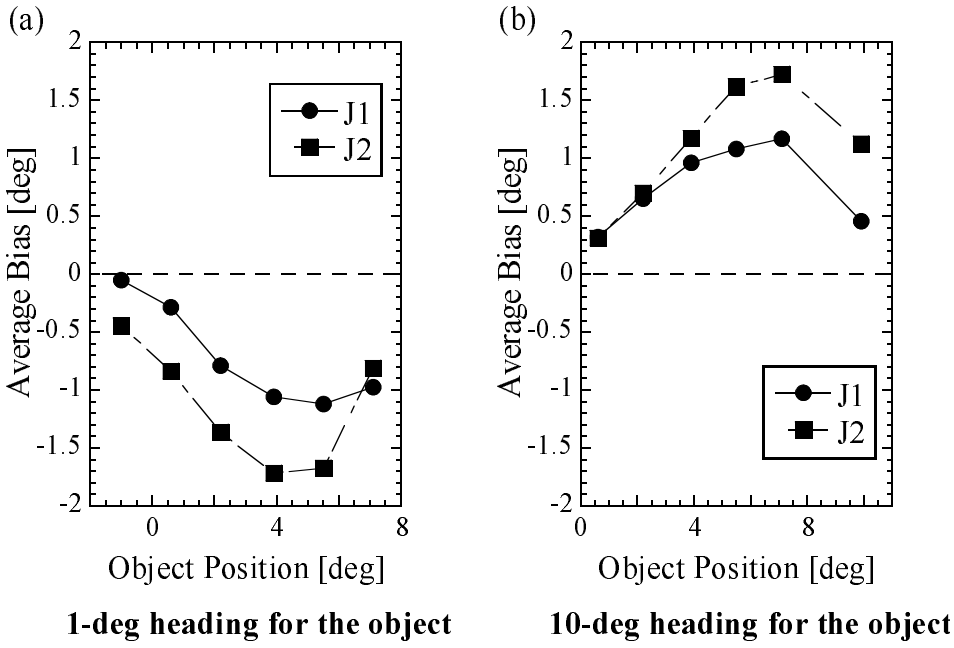

(a)
(b)
Average Bias [deg]
Object Position [deg]
J1
J2
Leftward moving object
Rightward moving object

Figure 2



1-deg heading for the object **10-deg heading for the object**

Figure 3