Title
The Perceived Motion of a Sterokinetic Stimulus

Authors
Rokers, Bas
Yuille, Alan L
Liu, Zili

Publication Date
2005-05-21

Available online at www.sciencedirect.com
SCIENCE DIRECT
Vision Research 46 (2006) 2375–2387
Vision Research
www.elsevier.com/locate/visres

The perceived motion of a stereokinetic stimulus

Bas Rokers [a], Alan Yuille [a,b], Zili Liu [a,*]

[a] Department of Psychology, University of California, Los Angeles, USA
[b] Department of Statistics, University of California, Los Angeles, USA

Received 21 May 2005; received in revised form 2 December 2005

Abstract

An ellipse rotating in the image plane can produce several different percepts. The two-dimensional (2D) percepts are either a rotating rigid ellipse or a constantly deforming non-rigid ellipse. The 3D percept is a rotating rigid circular disk that is tilted relative to the image plane. Stimuli that generate 3D percepts based on purely 2D rotational motion are known as stereokinetic stimuli. We examined the 3D percepts generated by the rotating ellipse stimulus. In theory, the motion of the 3D percept cannot be reliably inferred based on the 2D stimulus. When we quantitatively estimated observers' perceived motion, however, we found that the perceived motion was nearly identical across observers. These results suggest that all observers had similar 3D percepts. We assumed that given the 2D rotating ellipse stimulus the visual system generates a rigid 3D percept that is as slow and smooth as possible. The percepts predicted by these assumptions closely matched the experimental data. These findings suggest that perceptual ambiguity in stereokinetic stimuli is resolved using slow and smooth motion assumptions.

Keywords: Stereokinetic; Structure from motion; Slow and smooth; Bayesian inference; Rigidity

1. Introduction

The visual system infers the properties of an object, such as its shape, distance, and motion, based on the object's projection onto the retina. The object itself is called the distal stimulus, while the projection on the retina is called the proximal stimulus. Since the proximal stimulus is a projection of the distal stimulus, some information about the distal stimulus is inevitably lost. The visual system attempts to recover this information through a process called perceptual inference (Rock, 1983; von Helmholtz, 1924). When properties of the distal stimulus cannot be determined based on the proximal stimulus, the visual system relies on prior assumptions that reflect the properties objects tend to have. Together, the proximal stimulus and prior assumptions produce a percept that is an estimate of the properties of the distal stimulus.

The use of prior assumptions in perceptual inference can explain many visual phenomena. For example, assumptions of viewpoint have been used to account for the perceived surface shape of objects (Freeman, 1994; Nakayama & Shimojo, 1992). Assumptions of object distance have been used to account for the fact that near objects are perceived to be slightly farther away, while far away objects are perceived to be slightly nearer (Yang & Purves, 2003). Assumptions of object shape and orientation have been used to account for the perceived 3D shape of 2D line drawings (Mamassian & Landy, 1998). We used shape and motion assumptions to account for the perceived 3D objects generated by 2D stimuli. If all observers used the same assumptions, we should find that observers achieved the same 3D percepts given these 2D stimuli.

The assumption of minimal shape change initially proposed by Jansson and Johansson (1973) has been refined by Ullman (1979) into the rigidity assumption. The rigidity assumption posits that a set of 2D elements undergoing 2D transformations that has a rigid 3D object interpretation tends to be perceived as such. Given an ambiguous

* Corresponding author. Fax: +1 310 206 5895.
E-mail address: zili@psych.ucla.edu (Z. Liu).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.01.032

stimulus human observers tend to prefer a single rigid percept or a percept where several rigid components are moving independently over a non-rigid percept (Kersten, Bülthoff, Schwartz, & Kurtz, 1991; Shiffrar & Pavel, 1991; Ullman, 1984).

Theories that focus on minimal motion and minimal motion change (Hildreth, 1984; Yuille & Grzywacz, 1988) have suggested that the visual system favors percepts that have small and smoothly varying velocities across space. In accordance with the rigidity assumption, a 'skinny' ellipse (i.e., an ellipse with a small aspect ratio) rotating in the fronto-parallel plane appears to rotate rigidly. However, observers report that a 'fat' ellipse (i.e., an ellipse with a large aspect ratio) appears to deform non-rigidly (Wallach, Weisz, & Adams, 1956). Hildreth (1984) showed that the use of a prior assumption of motion smoothness predicts percepts that qualitatively match those reported by observers. More recently, Weiss and Adelson (1998) presented a model that incorporated both slow and smooth motion assumptions that accounted for a range of 2D percepts, including percepts generated by a rotating ellipse.

Structure from motion stimuli are 2D stimuli that generate the illusion of a 3D percept based on motion alone. A stereokinetic stimulus is a structure from motion stimulus that generates a 3D percept based on rotation in the image plane.[1] Such stimuli initially generate 2D percepts, but after longer observation produce vivid 3D percepts (Musatti, 1924). Since both the 2D and 3D percepts are compatible with the stimulus, prior assumptions must play a key role in the generation of these percepts. After some observation, a rotating ellipse stimulus generates the percept of a 3D rigid circular disk slanted relative to the image plane. Once an observer experiences the rigid 3D percept it becomes difficult to switch back to the initial 2D percept, suggesting that the 3D percept is preferred over the 2D percept. Despite nearly a century of research, the mechanisms underlying the stereokinetic effect are not well understood.

Shape-based assumptions, such as a rigidity assumption are not sufficient to account for the percepts generated by stereokinetic stimuli. Stereokinetic stimuli rely on the ambiguity in the perceived trajectory of points on curved contours. We will show that for stereokinetic stimuli determining the shape of the percept is not sufficient to determine the motion of the percept.

An ellipse rotating in the image plane can be the projection of an infinite number of shapes, including non-rigid shapes, non-planar shapes, and tilted elliptical disks. From a generic viewpoint consideration (Freeman, 1994; Nakayama & Shimojo, 1992), planar shapes are most likely. Following the rigidity assumption, we assumed that the rotating ellipse is perceived as the projection of a rigid circular disk. However, this assumption is not sufficient to determine the perceived motion of the circular disk.

[1] For an example of the stereokinetic stimuli we used and the percepts they generate, see http://zililab.psych.ucla.edu/demo/vr05.

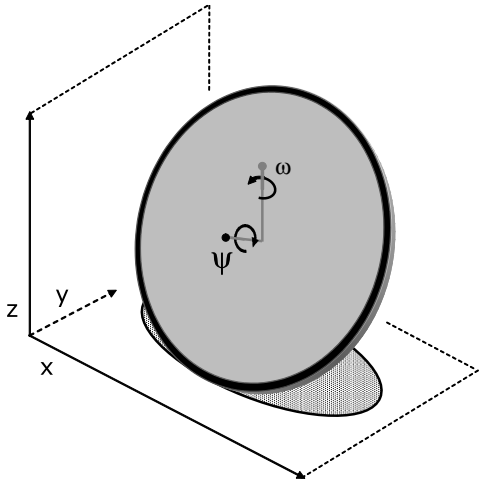

Fig. 1. Schematic illustration of the percept generated by an ellipse rotating in the image plane (Oxy). The stimulus is rotated around its center with an angular velocity ω. The percept generated by this stimulus is a rotating 3D circular disk that is also spinning around its surface normal with angular velocity ψ. In our experiments, participants adjusted the direction and magnitude of the rotation around the surface normal on a computer display so that the initially perceived rotation ψ was nulled. This nulling provided a measure of perceived rotation around the surface normal ψ. The stimulus contained a dot that helped participants achieve a 3D percept and also served as a tracking point.

When we decompose the circular disk's motion it can be characterized by two rotations: ω around the z-axis orthogonal to the image plane and ψ around the disk's surface normal (see Fig. 1). Although the rotation ω around the z-axis can be determined based on the projection of the stimulus, the rotation ψ around the disk's surface normal cannot. Any rotation ψ will be compatible with both the projected stimulus and the rigid 3D percept of a circular disk.

Minimum motion assumptions have been used to account for the 2D percepts generated by visual stimuli (Weiss, Simoncelli, & Adelson, 2002; Yuille & Grzywacz, 1988). However, to our knowledge, minimum motion assumptions have not been used to account for 3D percepts. We hypothesized that given the rotating ellipse stimulus the visual system generates a rigid 3D percept of a circular disk that moves as slowly and smoothly as possible. In the following we will derive the predicted rotation ψ around the surface normal of a rigid circular disk that minimizes motion. Subsequently, we will compare the predictions to empirical results.

2. Theoretical predictions

In cases where motion assumptions have been used to predict the perceived 2D motion generated by ambiguous stimuli the percept has minimized the perceived motion in the image plane. However, there are multiple ways minimal motion assumptions may be applied when a stimulus generates a 3D percept. The visual system may minimize the perceived 2D motion in the image plane or alternatively the perceived motion in 3D space.

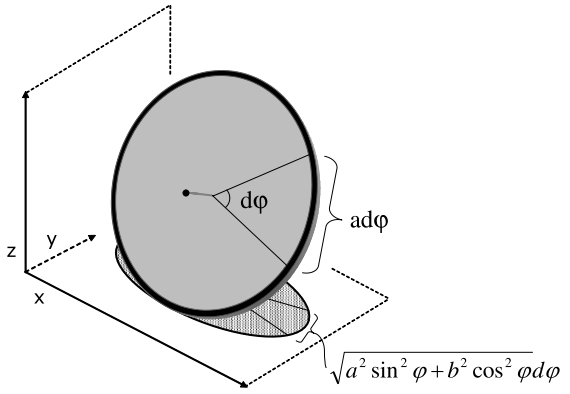

Fig. 2. Projection of a tilted 3D circle onto the image plane (Oxy). An infinitesimal contour segment $a\mathrm{d}\varphi$ in 3D is projected to $\mathrm{d}\varphi\sqrt{a^2\sin^2\varphi + b^2\cos^2\varphi}$ in the image plane. As a result, the minimum motion solution depended on whether velocity was integrated in the image plane or across 3D space.

Additionally, it is unknown whether the visual system combines motion estimates based on the 3D shape or based on the 2D projection. Points that are close in the image plane may be far away in 3D space. Since motion smoothness stipulates that nearby points have similar velocities, considering smoothness in the image plane may lead to different results than considering smoothness in 3D space (see Fig. 2 for a visual illustration).

We considered four possible scenarios: 3D motion may be computed across 3D space, 3D motion may be computed in the 2D image plane, 2D image plane motion may be computed across 3D space, or 2D image plane motion may be computed in the 2D image plane.

First, we will set up a general framework to describe the motion of a rotating ellipse, after which we will derive the predictions resulting from the different scenarios. We assumed that an ellipse rotating in the image plane is the orthogonal projection of a circular disk tilted in depth. The time dependent x and y coordinates of this disk project onto the rotating ellipse in the image plane

$$\begin{bmatrix} x(\phi,\psi,t) \\ y(\phi,\psi,t) \end{bmatrix} = \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix} \begin{bmatrix} a\cos(\phi+\psi t) \\ b\sin(\phi+\psi t) \end{bmatrix}, \tag{1}$$

where ϕ is the parameterized angle of the ellipse, ψ is an unknown rotation around the surface normal of the disk, ω is the rotation of the disk in the image plane (i.e., rotation around the image plane normal), and a and b are the major and minor radii of the ellipse, respectively. The depth component z of the disk is then given by

$$z(\phi,\psi,t) = \sqrt{a^2-b^2}\sin(\varphi+\psi t), \quad \text{with } a \geqslant b \tag{2}$$

so that the x, y, and z coordinates define a circle with radius a.

The velocity of points on the contour of the disk can be obtained by differentiation with respect to time t.

$$v_x = \frac{\mathrm{d}x}{\mathrm{d}t} = -(\psi a+\omega b)\sin(\phi+\psi t)\cos\omega t - (\omega a+\psi b) \times \cos(\phi+\psi t)\sin\omega t, \tag{3}$$

$$v_y = \frac{\mathrm{d}y}{\mathrm{d}t} = -(\psi a+\omega b)\sin(\phi+\psi t)\sin\omega t + (\omega a+\psi b) \times \cos(\phi+\psi t)\cos\omega t, \tag{4}$$

$$v_z = \frac{\mathrm{d}z}{\mathrm{d}t} = \psi\sqrt{a^2-b^2}\cos(\phi+\psi t). \tag{5}$$

If the circular disk remains rigid while moving, ψ is constant with respect to φ so that we can substitute θ for $\phi + \psi t$.

Models that use a motion minimization approach generally constrain both motion slowness and smoothness. However, for the rotating ellipse stimulus motion slowness and motion smoothness constraints predict similar solutions. For simplicity we provide the derivation based on motion slowness alone in the main body of the paper. For derivations incorporating both motion slowness and smoothness, please see Appendix A. To obtain the slowest motion solution we minimize

$$F(\psi,t) = \int |\vec{v}(\psi,t)|^2\,\mathrm{d}s, \tag{6}$$

where s is the arc length of the stimulus.

2.1. Predictions for experiment 1

2.1.1. Case i: Minimizing 3D motion along the 3D contour

We first assume that the visual system integrates 3D motion along the 3D contour. We called this the $3D_v$–$3D_\theta$ solution. The $3D_v$ component reflects that velocity is considered in 3D space and the $3D_\theta$ component reflects that we integrate motion along the disk's contour in 3D space. Then

$$|\vec{v}|^2 = v_x^2 + v_y^2 + v_z^2 \tag{7}$$

and

$$\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2} = a\mathrm{d}\theta \tag{8}$$

so that

$$F_{3D-3D}(\psi,t) = \int_0^{2\pi} \left(a^2\psi^2 + 2ab\psi\omega + \omega^2 \times (a^2\cos^2\theta + b^2\sin^2\theta)\right) a\mathrm{d}\theta. \tag{9}$$

Solving the integral

$$F_{3D-3D}(\psi) = 2\pi a^3\psi^2 + 4\pi a^2 b\psi\omega + \pi a\omega^2(a^2+b^2). \tag{10}$$

When we solve for $\psi_{\min}$ using $\frac{\mathrm{d}F_{3D-3D}(\psi)}{\mathrm{d}\psi} = 0$ we find

$$\psi_{\min} = -\frac{b}{a}\omega. \tag{11}$$

For a physical intuition of this result, imagine looking down at a coin with radius a lying on a table. Now imagine that we tilt the coin up so that its rim touches the table at a

single point and its projection on the table's surface is an ellipse of aspect ratio b/a (Fig. 1). Let the coin roll around while making sure that: (1) the center of the coin maintains its position in space and (2) the point on the coin that touches the table moves without slipping. The coin's motion can be formally decomposed into two simultaneous rotations: a rotation ω around the viewing direction and a rotation ψ around the coin's surface normal. Consider the velocity of the point on the coin's rim that touches the table. Its speed due to the rotation ω is ωb since this point's distance to the center of the ellipse projected onto the table is b. The speed due to the rotation ψ is ψa. When the instantaneous speed of this point is zero so that the coin is rolling on the table without slipping, we have

$$\psi a + \omega b = 0, \tag{12}$$

which provides the same solution as Eq. (11).

2.1.2. Case ii: Minimizing 2D motion along the 3D contour

Rather than minimizing the perceived 3D velocity, the visual system may only minimize the motion components in the image plane. We called this the $2D_v$–$3D_\theta$ scenario. The $2D_v$ component reflects that velocity is only considered along the disk's contour in the image plane, whereas the $3D_\theta$ component reflects that we integrated motion along the disk's contour as it is defined in 3D space. To find the predicted contour rotation $\psi_{\min}$ that minimizes motion in the image plane we use

$$|\vec{v}|^2 = v_x^2 + v_y^2 \tag{13}$$

while other equations are identical to those used in Case i above. It follows:

$$F_{2D-3D}(\psi,t) = \int_0^{2\pi} \left(\psi^2(a^2\sin^2\theta + b^2\cos^2\theta) + 2ab\omega\psi + \omega^2(a^2\cos^2\theta + b^2\sin^2\theta)\right)a\,d\theta. \tag{14}$$

Solving the integral

$$F_{2D-3D}(\psi) = a\psi^2(a^2\pi + b^2\pi) + 4\pi a^2 b\omega\psi + a\omega^2(a^2\pi + b^2\pi). \tag{15}$$

When we solve for $\psi_{\min}$ using $\frac{dF_{2D-3D}(\psi)}{d\psi} = 0$ we find

$$\psi_{\min} = -\frac{2ab}{a^2+b^2}\omega. \tag{16}$$

2.1.3. Case iii: Minimizing 3D velocity along the 2D contour

Thus far we have presented solutions of the two scenarios that integrate motion across the stimulus contour in 3D space. The visual system may integrate motion along the projection of the disk's contour onto the image plane instead. Since we originally parameterized the stimulus in 3D space, an infinitesimal segment of the stimulus contour in the image plane becomes

$$ds = \sqrt{(dx)^2 + (dy)^2} = \sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\,d\theta. \tag{17}$$

The motion flow in the $3D_v$–$2D_\theta$ scenario becomes

$$F_{3D-2D}(\psi) = \int_0^{2\pi} \left(a^2\psi^2 + 2ab\omega\psi + \omega^2(a^2\cos^2\theta + b^2\sin^2\theta)\right) \times \sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\,d\theta. \tag{18}$$

Let

$$S = \int_0^{2\pi} \sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\,d\theta \tag{19}$$

and

$$T = \int_0^{2\pi} \omega^2(a^2\cos^2\theta + b^2\sin^2\theta)\sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\,d\theta. \tag{20}$$

Observe that S and T are constant with respect to ψ. Then

$$F_{3D-2D}(\psi) = a^2\psi^2 S + 2ab\omega\psi S + T. \tag{21}$$

When we solve for $\psi_{\min}$ using $\frac{dF_{3D-2D}(\psi)}{d\psi} = 0$ we find

$$\psi_{\min} = -\frac{b}{a}\omega, \tag{22}$$

which is identical to Eq. (11). The $3D_v$–$2D_\theta$ scenario and the $3D_v$–$3D_\theta$ scenario of Case i thus predict the same $\psi_{\min}$.

2.1.4. Case iv: Minimizing 2D velocity along the 2D contour

In the $2D_v$–$2D_\theta$ scenario, 2D velocity is integrated along the contour of the ellipse in the 2D image plane. The derivation is similar to the $2D_v$–$3D_\theta$ scenario derivation in Case ii with the exception that we used Eq. (17) to integrate across the stimulus contour in the image plane, so that

$$F_{2D-2D}(\psi) = \int_0^{2\pi} \left(\psi^2\left(\frac{ds}{d\theta}\right)^2 + 2ab\omega\psi + \omega^2\left(\frac{ds}{d\theta}\right)^2\right) \times \sqrt{\left(\frac{ds}{d\theta}\right)^2}\,d\theta, \tag{23}$$

where $\left(\frac{ds}{d\theta}\right)^2 = a^2\sin^2\theta + b^2\cos^2\theta$.

Let

$$U = \int_0^{2\pi} \sin^2\theta\sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\,d\theta \tag{24}$$

and

$$V = \int_0^{2\pi} \cos^2\theta\sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\,d\theta. \tag{25}$$

Then using Eq. (20) it follows:

$$F_{2D-2D}(\psi) = \psi^2(a^2U + b^2V) + 2ab\omega\psi S + T. \tag{26}$$

When we solve for $\psi_{\min}$ using $\frac{dF_{2D-2D}(\psi)}{d\psi} = 0$ we find

$$\psi_{\min} = -\frac{abS}{a^2U + b^2V}\omega. \tag{27}$$

We can then numerically approximate S, U, and V. The solution we obtained for $\psi_{\min}$ lies in between the $2D_v$–$3D_\theta$ and $3D_v$ predictions, thus

$$-\frac{2ab}{a^2+b^2}\omega < \psi_{\min} < -\frac{b}{a}\omega. \tag{28}$$

Instead of minimizing motion flow along the contour of the stimulus, observers may minimize motion across the entire surface of the stimulus. Although we do not provide a proof, it can be shown that the same results would have been obtained, since all four scenarios provided $\psi_{\min}$ predictions that are a function of the aspect ratio of the rotating ellipse b/a, rather than the major radius a or minor radius b themselves. The surface of the stimulus can be considered as an infinite set of stimulus contours with different lengths of the major radius a, but with constant aspect ratio b/a.

The scenarios above provide predictions for the magnitude of the perceived rotation ψ around the surface normal. The scenarios predict that the perceived rotation ψ around the surface normal is a function of aspect ratio b/a of the stimulus and rotation ω in the image plane based on the assumption that the percept is a rigid 3D disk that moves as slowly as possible. In all scenarios a linear relationship between rotation ω in the image plane and perceived rotation ψ around the surface normal is predicted. In addition in all scenarios the predicted ψ monotonically increases with the aspect ratio b/a of the ellipse. These predictions are parameter free, and thus provide strong constraints on the percepts we expected observers to report.

2.2. Predictions for experiment 2

We conducted an second experiment in which the viewing direction was changed to be along the x-axis. The Oyz plane thus became the image plane. We again derived the predicted motion of the 3D percept. Naturally, the $\psi_{\min}$ prediction based on the $3D_v$–$3D_\theta$ scenario is independent of the viewing direction. However, in the $3D_v$–$2D_\theta$ scenario velocity is integrated along the stimulus contour in the Oyz plane so that

$$\begin{aligned} ds &= \sqrt{(dy)^2 + (dz)^2} \\ &= d\theta\sqrt{(-a\sin\theta\sin\omega t + b\cos\theta\cos\omega t)^2 + (a^2 - b^2)\cos^2\theta}. \end{aligned} \tag{29}$$

In this case, and the cases that follow, we derived the predictions numerically. When the stimulus is viewed from the side the motion flow became a function of time. To derive the solutions we integrated both over time t and along the stimulus contour s. Motion flow F was defined as

$$F(\psi, t) = \int\int |\vec{v}|^2\, ds\, dt. \tag{30}$$

Like in the derivations for experiment 1, numerical analysis suggested that the ψ that minimized motion flow in the $3D_v$–$2D_\theta$ scenario was the same as in the $3D_v$–$3D_\theta$ scenario.

In the $2D_v$–$3D_\theta$ and $2D_v$–$2D_\theta$ scenarios the motion flow in the image plane is defined as

$$|\vec{v}|^2 = v_y^2 + v_z^2. \tag{31}$$

In the $2D_v$–$3D_\theta$ scenario we used $ds = a\,d\theta$ as obtained in Eq. (8) to integrate along the stimulus contour in 3D space. Numerical analysis suggested that the ψ that minimized motion flow was a linear function of ω and a monotonically increasing function of aspect ratio b/a.

Finally in the $2D_v$–$2D_\theta$ scenario we used Eq. (29) to integrate along the stimulus contour in the image plane. Numerical analysis suggested again that ψ was linear function of ω and a monotonically increasing function of aspect ratio b/a. In the next section, we present experiments to estimate the perceived motion generated by the stereokinetic stimuli.

3. Experimental methods

3.1. Apparatus

We performed two experiments. In both experiments stimuli were displayed on a 17 in. Dell Trinitron CRT display at a resolution of 1600 by 1200 pixels with a 75 Hz refresh rate. The stimuli were rendered on a 1.8 GHz Pentium 4 computer using a NVIDIA GeForce2 GTS video card. The OpenGL graphics library was used for stimulus visualization. Viewing distance was 70 cm. Observers used a headrest to stabilize head position. A viewing tube was placed between the headrest and the monitor. During the experiment the only light source in the room was the monitor which had a background luminance of 0.01 cd/m^2.

3.2. Stimuli

Stimuli were specified in 3D space and rendered under orthographic projection. The absence of perspective cues might have affected participant performance, but was required to eliminate any depth cues in the stimuli. Before projection onto the image plane, the circular stimulus was composed of 20 segments of 15° arc length and 20 gaps of 3° arc length. The number and size of the gaps were chosen such that participants readily perceived an additional rotation around the surface normal of the circular disk. In pilot experiments we did not find evidence to suggest that the results varied with different numbers and sizes of these gaps. Due to the projection of the 3D stimuli on the computer display, segments were not physically identical in length in the image plane.

The projected ellipses had one of three aspect ratios (0.5, 0.7, and 0.9). All stimuli subtended a visual angle of approximately 8°. Stimuli were displayed using low luminance (3 cd/m^2) to counteract the effects of motion blur on the CRT screen.

In both the familiarization and experimental phase of the experiments a dot was displayed to make the 3D percept more vivid. This dot was positioned on the surface normal that passed through the origin of the circular disk. To observers the dot appeared as the tip of a rod that was oriented perpendicular to the surface of the circular disk or the apex of a 3D cone with the disk forming the base. Due

to the placement of the dot on the perceived axis of rotation we did not expect the dot to affect the perceived rotation of the disk.

The speed of rotation of the stimulus in the image plane ω was one of 112.5, 75, or 37.5 deg/s in either a clockwise or counter-clockwise direction. At the start of each trial, the rotation around the ellipse contour was randomly set to be faster or slower (0 or -2ω) than any of the theoretically predicted rotations.

For each observer two sessions of 36 stimuli were shown using a full factorial design of ellipse aspect ratios, initial starting velocities, rotation speeds, and rotation directions. The stimulus sequence was displayed in random order without replacement in each session.

We covered the observers' non-dominant eye with an eye-patch so that stimuli were viewed monocularly. Observers were familiarized with the stereokinetic phenomenon by viewing a number of rotating ellipse stimuli at varying speeds and with varying aspect ratios. To prevent observers from tracking contour motion explicitly we asked observers to fixate on the dot during presentation of the stimuli. No time limit was imposed. On average observers took about 45 min to complete the experiment.

3.3. Procedure of Experiment 1

The stereokinetic stimuli generated the percept of a 3D disk slanted in depth that rotated both around an axis perpendicular to the image plane and around the surface normal of the disk. We dashed the line of the ellipse so that the rotation around the surface normal of the ellipse became visible to the observer (see Fig. 3). In Experiment 1 observers were asked to adjust the perceived rotation around the surface normal of the disk until it disappeared. We believe this adjustment was equal in size, but opposite in sign, to the perceived rotation around the surface normal of the disk. This method is similar to motion nulling procedures used in a number of psychophysical experiments (Cavanagh & Anstis, 1991; Chichilnisky, Heeger, & Wandell, 1993; Gregory, 1985). However, unlike in these experiments, the final percept in our procedure is not stationary.

Observers were asked to adjust the rotation around the surface normal of the 3D disk to make it appear as if there was "no twirling motion of the circle around its own center". The step-size of the adjustment was 0.1 times the speed of rotation ω in the image plane. In order to achieve a percept with no rotation around the surface normal, observers were actually adding a rotational motion to the stimulus.

3.4. Observers of Experiment 1

Eleven observers participated in this experiment. One observer was familiar and the remaining 10 participants were naïve with regard to the purpose of the study.

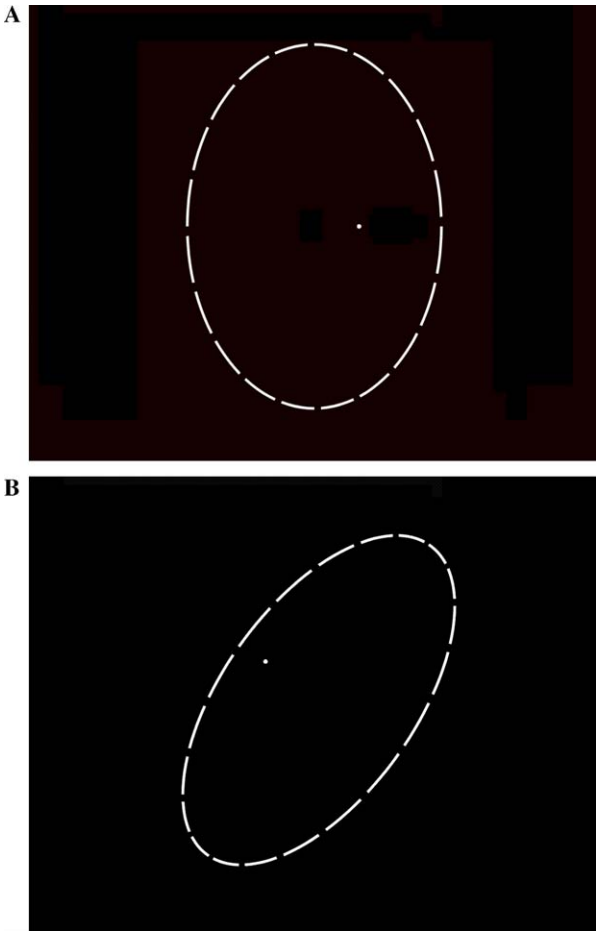

Fig. 3. High contrast versions of the experimental stimuli. The stimuli were presented in low contrast on a dark background. (A) The stimulus in Experiment 1 was rotated around its center in the image plane. This stimulus generated the percept of a 3D disk tilted in depth that rotated around an axis perpendicular to the image plane and around the surface normal of the disk. Observers were asked to adjust motion around the contour of the disk so that the rotation around the surface normal of the disk perceptually disappeared. (B) The stimulus in Experiment 2 was the same as the stimulus in Experiment 1, except that it was viewed 'from the side'. In this experiment observers were asked to adjust the dashed line stimulus so that the percept that it generated resembled the percept generated by the corresponding solid line stimulus. In both experiments the adjustments provided quantitative estimates of the perceived contour rotation ψ around the surface normal of the 3D percept. The dot in the display helped observers achieve a 3D percept. The stimuli shown here have an aspect ratio of 0.7.

3.5. Results of Experiment 1

Figs. 4A–C show the rotation around the surface normal ψ added by participants as a function of ellipse rotation ω for three different aspect ratios (b/a) of the ellipse. Figs. 4A–C show that the estimated perceived rotation ψ around the surface normal varied little across observers. Averaged across conditions the standard error of the perceived rotation ψ around the surface normal was 2.43 deg/s or 3.39% of the rotation ω in the image plane. This suggests that for each stimulus all observers had similar percepts.

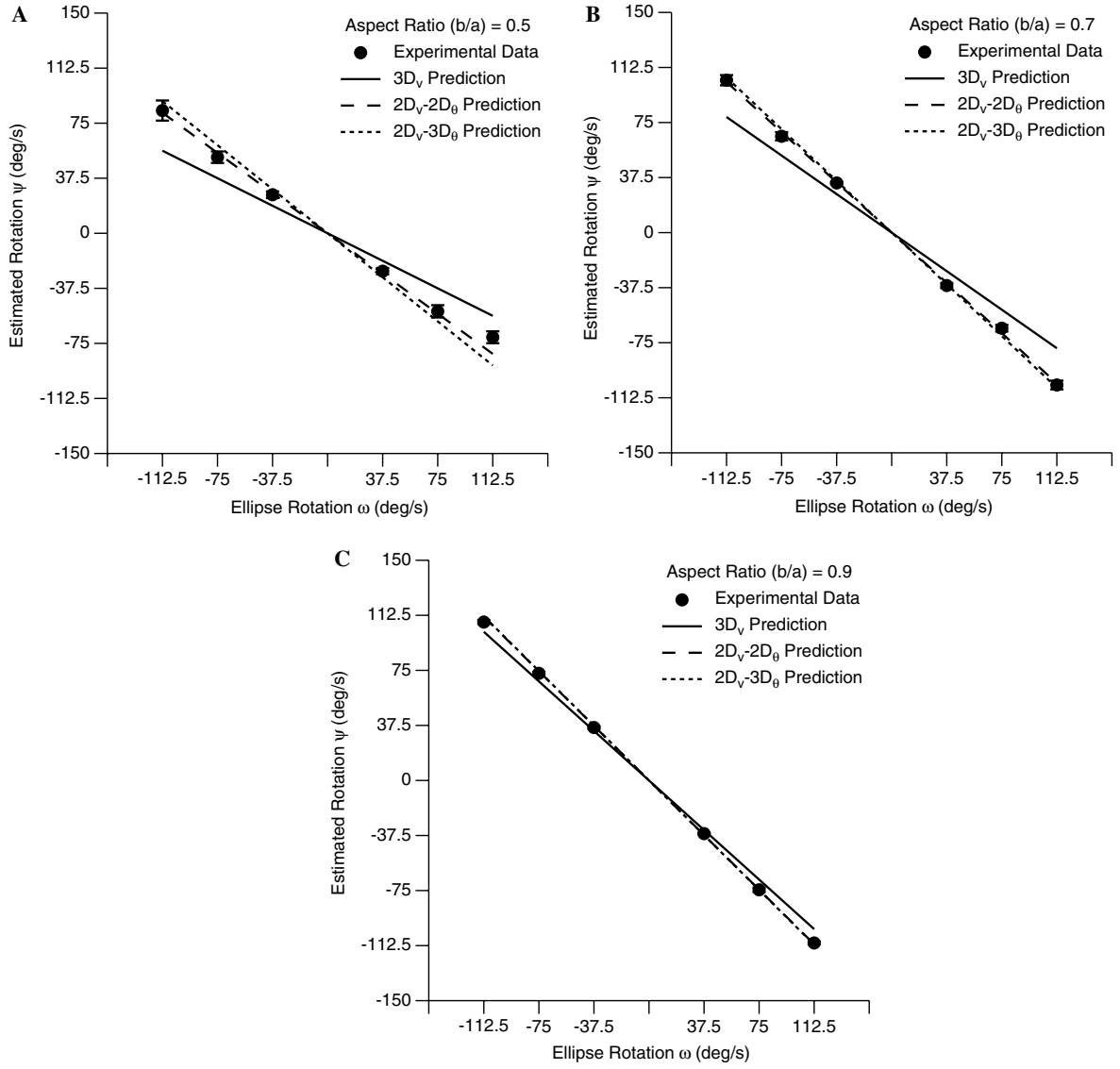

Fig. 4. Empirical results and theoretical predictions for Experiment 1. An ellipse was rotated with speed ω in the image plane. Observers perceived a circular disk tilted in depth. Observers were asked to adjust motion around the contour of the disk so that the rotation around the surface normal of the disk perceptually disappeared. The actual rotation ψ around the surface normal (black dots) was revealed when observers nulled the perceived rotation around the surface normal. The theoretical data (dashed and solid lines) show the predicted rotation ψ around the surface normal if the visual system minimized the motion of a 3D rigid percept. Theoretical predictions closely matched the data. Experimental data and predictions were a function of rotation speed in the image plane (ω) and aspect ratio (b/a) of the ellipse. Aspect ratio of the ellipse in the image plane equaled (A) 0.5, (B) 0.7, and (C) 0.9. Error bars depict standard error of the mean ($n = 11$).

The perceived rotation ψ around the surface normal appeared to be a linear function of the rotational velocities ω of the stimulus in the image plane. The apparent relationship was tested using a linear regression analysis. We found a high correlation coefficient r^2 across aspect ratios (for $b/a = 0.5$, $r^2 = .94$, for $b/a = 0.7$, $r^2 = .99$, and for $b/a = 0.9$, $r^2 = 1.00$).

In Section 2, we predicted the perceived rotation ψ around the surface normal of a 3D circular disk slanted in depth. The predictions were based on the assumption that observers generate percepts that minimize the velocity of the percept. We presented four different scenarios ($3D_v$–$3D_\theta$, $3D_v$–$2D_\theta$, $2D_v$–$2D_\theta$, and $2D_v$–$3D_\theta$). The $3D_v$–$3D_\theta$ and $3D_v$–$2D_\theta$ scenarios gave identical predictions, so we presented these predictions as a single prediction ($3D_v$) in Fig. 4. In all four scenarios it was predicted that rotation ψ around the surface normal would be a linear function of stimulus rotation ω in the image plane. The experimental data showed the predicted linear relationship between the estimated rotation ψ around the surface normal and the rotation ω of the stimulus in the image plane for each aspect ratio (b/a). In addition, the experimental data were bounded by the predictions derived under the $3D_v$ scenario and the predictions derived under the $2D_v$–$2D_\theta$, and $2D_v$–$3D_\theta$ scenarios.

We believe that the consistency of the data is a result of the observers having similar percepts. However, since we did not objectively verify the observer's percepts, the

consistency of the results could have been the result of a different common strategy used by observers. We can think of two strategies that would have resulted in consistent results. Observers could have selected a perceived rotation around the surface normal of a circular disk that was opposite in sign but equal in size to the rotation of the ellipse in the image plane. In such a case the line segments that make up the stimulus maintain their radial location over time. Alternatively, participants could have adjusted the rotation around the surface normal of the circular disk to zero. In that case, the stimulus is compatible with the 2D percept of a rigid ellipse rotating in the image plane. Neither strategy accounted for the experimental data. Estimates of perceived contour motion fall in between these two settings. Given our data we do not believe that the rotation around the surface normal could have been adjusted consistently unless observers adjusted the stimulus based on a common percept.

3.6. Procedure of Experiment 2

In the second experiment we changed the viewing direction so that observers perceived the rotating ellipse 'from the side', rather than from the front. In other words the viewing direction was along the x axis rather than the z axis. In this case the stimulus is not compatible with a rigid 2D interpretation, but the perceived rotation around the surface normal remains ambiguous. If observers minimized perceived 3D motion across the 3D shape, the results of this experiment should be identical to the results reported in Experiment 1. In Experiment 2 the projected stimulus changes shape over time. If observers minimized perceived 2D motion in the image plane the perceived rotation becomes a function of time. In the predictions below we assumed that observers integrated motion across time as well as across space.

When viewed 'from the side' the stimuli generated the percept of a 3D disk slanted in depth that rotated around an axis perpendicular to the Oxy plane and around the surface normal of the disk. We again dashed the line of the ellipse so that the rotation around the surface normal of the ellipse became visible to the observer.

In the results reported here the method of adjustment used by the observers was also modified. Similar results were obtained based on the method used in Experiment 1. Recall that in Experiment 1 observers adjusted the rotation around the surface normal so that it perceptually disappeared. It could be argued that in Experiment 1 observers may not have adjusted the stimulus to resemble the 3D percept generated by the solid line stereokinetic stimulus. Therefore observers in Experiment 2 were instructed to adjust the dashed line ellipse to best resemble the percept generated by the corresponding solid line stimulus. Observers used two buttons to switch between the display of a dashed line stimulus and the solid line stimulus while making the stimulus adjustment.

3.7. Observers of Experiment 2

Twenty-eight observers participated in this experiment. Two observers were familiar with the purpose of the study and also participated in Experiment 1. The remaining 26 participants did not take part in Experiment 1 and were naïve with regard to the purpose of the study.

3.8. Results of Experiment 2

Figs. 5A–C show the rotation around the surface normal ψ added by participants as a function of ellipse rotation ω for three different aspect ratios (b/a) of the ellipse. Averaged across conditions the standard error of the perceived rotation ψ around the surface normal was 3.24 deg/s or 4.28% of the rotation ω in the image plane. We used 28 participants in Experiment 2 compared to 11 participants in Experiment 1 because the method of adjustment used in Experiment 2 produced somewhat more variable data. When using the same method of adjustment as in Experiment 1, the variance was closer to that found in Experiment 1.

The perceived rotation ψ around the surface normal appeared to be a linear function of the rotational velocities ω of the stimulus in the image plane. The apparent relationship was tested using a linear regression analysis. We found high correlation coefficients r^2 across aspect ratios (for $b/a = 0.5$, $r^2 = .72$, for $b/a = 0.7$, $r^2 = .91$, and for $b/a = 0.9$, $r^2 = .97$).

We presented the theoretical predictions alongside the experimental data in Fig. 5. We predicted that rotation ψ around the surface normal would be a linear function of stimulus rotation ω in the image plane. The experimental data showed the predicted linear relationship between the estimated rotation ψ around the surface normal and the rotation ω of the stimulus in the image plane for each aspect ratio (b/a). In addition, the experimental data were bound by the predictions derived under the $3D_v$ scenarios and the predictions derived under the $2D_v$–$2D_\theta$ and $2D_v$–$3D_\theta$ scenarios.

4. Discussion

We investigated the 3D percepts generated by 2D stereokinetic stimuli. While there was no a priori reason to assume so, all observers reported similar results. Since the 2D stimuli did not determine the motion of the 3D percepts, these results suggest that when given an ambiguous stimulus all observers used similar assumptions to generate a 3D percept.

We showed that if motion is minimized in 3D space a percept is predicted that is equivalent to a circular disk that rotates on a surface without slipping. Based on that result it may be argued that rather than minimizing perceived motion, observers adjusted the stimulus so that the percept does not slip. However, the experimental data show that observers do not select a percept that is compatible with

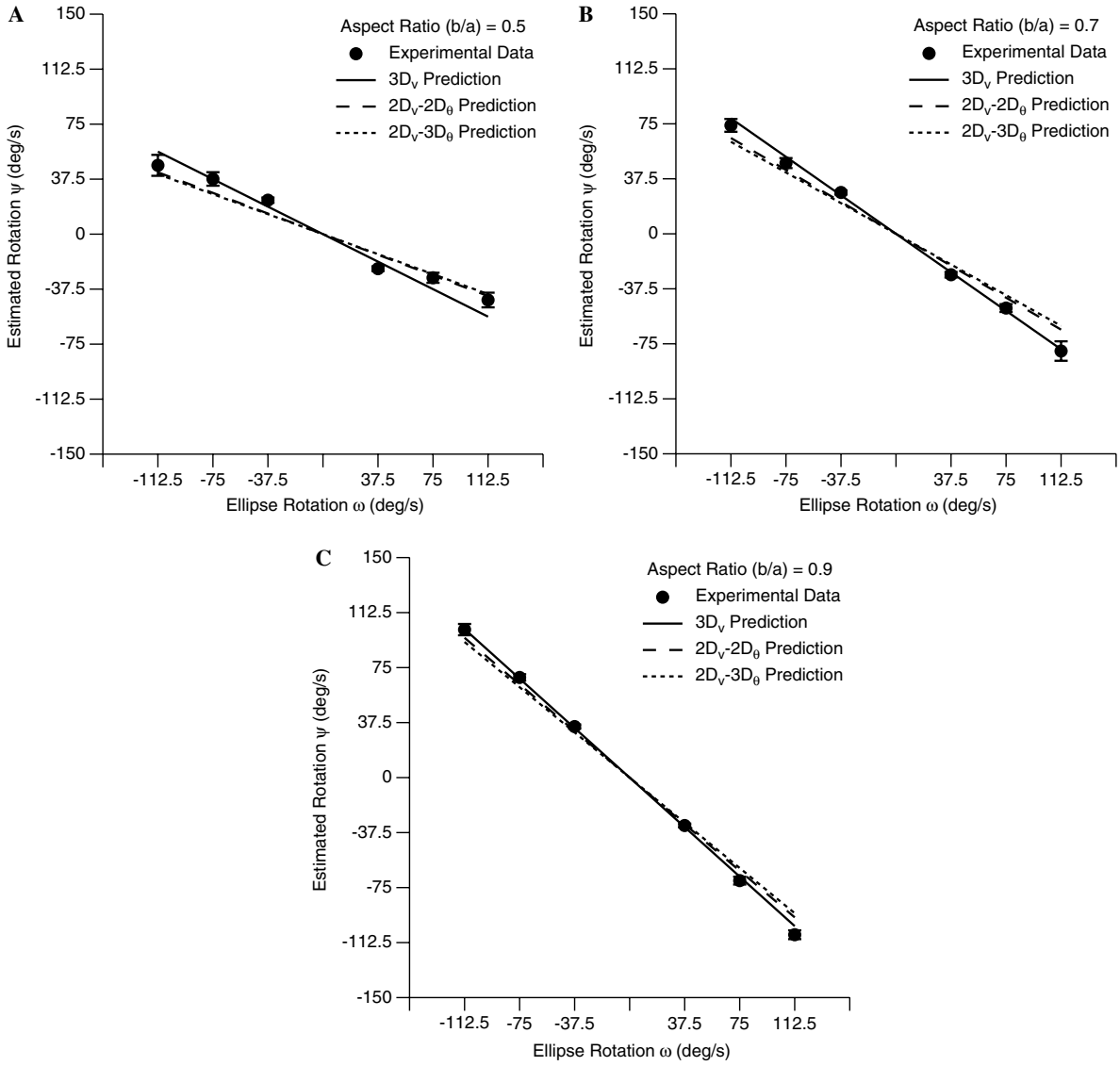

Fig. 5. Empirical results and theoretical predictions for Experiment 2. A dashed line circle that was rotated with angular velocity ω around the z axis was viewed from the side, i.e., in the Oyz plane. Observers perceived a circular disk tilted in depth. Observers were asked to match the perceived motion around the surface normal of the disk to a stimulus that had a solid line contour, but was otherwise identical. The actual rotation ψ around the surface normal (black dots) was revealed when observers matched the motion of the dashed line stimulus to the motion of the corresponding solid line stimulus. The theoretical data (dashed and solid lines) show the predicted rotation ψ around the surface normal if the visual system minimized the motion of a 3D rigid percept. Theoretical predictions closely matched the data. Experimental data and predictions were a function of rotation speed in the image plane ω and aspect ratio (b/a) of the ellipse. Aspect ratio of the ellipse in the image plane equaled (A) 0.5, (B) 0.7, and (C) 0.9. Error bars depict standard error of the mean ($n = 28$).

the $3D_v$ scenario. The percept compatible with the experimental data appears to show some slipping.

We hypothesized that the visual system uses prior assumptions of rigidity and motion slowness to generate the 3D percepts. Based on these assumptions we predicted that the perceived rotation ψ around the surface normal of a 3D disk should be a function of the speed of rotation ω and the aspect ratio b/a of the stimulus in the image plane. However, the experimental data were bound between the $2D_v$ and $3D_v$ predictions rather than matched by any of the theoretical predictions.

The fact that the experimental data were bound by the predictions may have been due to a number of factors. Neurophysiological (Maunsell & Van Essen, 1983) and behavioral studies (Harris, McKee, & Watamaniuk, 1998) have shown that the visual system is less sensitive to motion in depth compared to motion in the image plane. This may be due to the different reliability of these cues. The combination of both visual cues and prior assumptions based on their reliability has accounted for some percepts reported by observers (Landy, Maloney, Johnston, & Young, 1995; Mamassian & Landy, 2001). If motion priors in the image plane and in depth were similarly combined for stereokinetic stimuli, such a combination would predict a perceived rotation ψ that falls in between the $2D_v$ and $3D_v$ predictions.

It could have been that observers underestimated the slant of the percept. Previous research has found that observers underestimate the perceived slant of stationary stimuli (Epstein, 1977; Gibson, 1950). However when observers were asked to report the perceived slant of stationary stimuli considerable between observer variability has been reported (Mamassian & Kersten, 1996; Todd, Koenderink, van Doorn, & Kappers, 1996). Our results show little variability between observers, although we cannot exclude the possibility that slant may have been underestimated equally by all observers in our study. Previous results have been based on stationary stimuli that contained varying curvature across their surface and may therefore have produced larger between subject variability, but it seems hard to reconcile these results with our data.

The stereokinetic stimuli were compatible with the percept of any number of rigid 3D elliptical disks tilted in depth. Although observers were instructed that their percept should be a tilted rigid circular disk, we did not verify what percept observers actually achieved. However, when all points on the disk are perceived to be equidistant from its center, motion along the contour can be characterized by the rotation ψ around the surface normal of the disk. If the stimulus was perceived as the projection of an elliptical disk the rotation ψ around the surface normal would not be uniform along the stimulus contour. The high within and between subject agreement on the perceived contour rotation ψ in our experiments suggests that the observers did perceive the stimulus as the projection of a circular disk rather than an elliptical disk.

Previous work has suggested that 2D percepts minimize both motion slowness and smoothness. For simplicity we provided derivations based on motion slowness alone in the main body of the paper. Although the predictions based on the combination of prior assumptions of motion slowness and smoothness were mostly identical, the $2D_v$–$2D_\theta$ scenario predictions were somewhat different. We provided these predictions in Appendix A. When taking both slowness and smoothness into account the $2D_v$–$2D_\theta$ scenario predictions are closest to the data in Experiment 1 and Experiment 2. Although more experimentation is required, this would suggest that the 3D percept is generated based on motion assumptions that are applied purely in the image plane.

Slowness and smoothness assumptions have been used to account for the 2D percepts generated by rotating ellipse stimuli (Hildreth, 1984; Weiss & Adelson, 1998). The most recent accounts rely on the varying intensity of an image over time, without reference to shapes (Weiss et al., 2002). This approach may be applied to account for the 3D precepts generated by rotating ellipse stimuli as well and may address the question of whether the slowness and smoothness assumptions alone, without a rigidity assumption, can account for the 3D percepts generated by the stereokinetic stimuli. However shape-based assumptions, such as the rigidity assumption, are not sufficient to account for these 3D percepts. The rigidity assumption does not distinguish between different rigid interpretations. A shape assumption like the compactness assumption would predict that an ellipse in the image plane is perceived as a circle tilted in depth. Similarly, other shape-based assumptions, like compactness, the assumption that stimuli that are 'compact' in the image plane are perceived as 'compact' in space, are not sufficient to determine the perceived motion of stereokinetic stimuli. Assuming a circular disk tilted in depth constrains but does not determine the motion of the 3D percept.

Until now the mechanism underlying the stereokinetic effect has been unknown. Based on the evidence presented here, we believe that the prior assumptions of motion slowness and motion smoothness account for both the 2D and 3D percepts generated by a rotating ellipse stimulus. We predicted the motion of the 3D percept in analytical terms and compared the predictions to the 3D percept reported by observers. If observers were not using prior assumptions of shape and motion there is no reason to expect that they would have similar percepts. Prior assumptions of motion may help account for the percepts generated by other stereokinetic stimuli and structure from motion stimuli in general.

Acknowledgments

We thank Hongjing Lu, Denis Schluppeck, Ben Thompson, Pascal Mamassian, James Thomas, Georgia Scheele, and two anonymous reviewers for their valuable comments and suggestions.

Appendix A

Models of perceptual organization using prior assumptions of motion have generally posited constraints on both motion slowness and smoothness (Grzywacz & Yuille, 1991). In this Appendix, we derive the predicted 3D percept generated by the 2D stereokinetic stimulus when we incorporate both constraints. To derive the predicted rotation ψ around the surface normal of a 3D circular disk we minimize

$$F(\psi, t) = \sum_{n=0}^{2} c_n \int \left| \frac{\partial^n \vec{v}}{\partial s^n} \right|^2 \mathrm{d}s, \tag{32}$$

where

$$c_n = \frac{\lambda^{2n}}{n! 2^n}. \tag{33}$$

$F(\psi)$ is a weighted linear sum of the 0th to 2nd order derivatives of velocity integrated along the stimulus contour. The 0th order derivative corresponds to motion slowness. The 1st and 2nd order derivatives correspond to first and second order smoothness, respectively. The only free parameter in the derivation of $\psi_{\min}$ is λ, which determines the relative weight of the slowness and smoothness components in $F(\psi, t)$.

The derivations below are based on Eqs. (3)–(5) that specify the velocity of a 3D circular disk. Since the rigidity of the circular disk stipulates that ψ is constant with respect to θ, we substitute θ for $\phi + \psi t$ for notational convenience.

A.1. Experiment 1

A.1.1. Case i: Minimizing 3D motion along the 3D contour

When we assume that the visual system integrates 3D motion along the 3D contour, then

$$\left|\frac{\partial^n \vec{v}}{\partial s^n}\right|^2 = \left(\frac{\partial^n v_x}{\partial s^n}\right)^2 + \left(\frac{\partial^n v_y}{\partial s^n}\right)^2 + \left(\frac{\partial^n v_z}{\partial s^n}\right)^2 \tag{34}$$

and

$$\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2} = a\mathrm{d}\theta. \tag{35}$$

Using

$$\frac{\mathrm{d}^n \sin\theta}{\mathrm{d}\theta^n} = \sin\left(\theta + \frac{n\pi}{2}\right) \tag{36}$$

and, similarly,

$$\frac{\mathrm{d}^n \cos\theta}{\mathrm{d}\theta^n} = \cos\left(\theta + \frac{n\pi}{2}\right) \tag{37}$$

it follows

$$\left|\frac{\partial^n \vec{v}}{\partial s^n}\right|^2 = \frac{(\psi a + \omega b)^2}{a^{2n}} \sin^2\left(\theta + \frac{n\pi}{2}\right) + \frac{(\omega a + \psi b)^2}{a^{2n}} \times \cos^2\left(\theta + \frac{n\pi}{2}\right) + \frac{\psi^2(a^2 - b^2)}{a^{2n}} \cos^2\left(\theta + \frac{n\pi}{2}\right). \tag{38}$$

Then

$$F_{3D-3D}(\psi) = \sum_{n=0}^{2} \frac{c_n}{a^{2n}} \int_0^{2\pi} \left(a^2\psi^2 + 2ab\omega\psi + \omega^2 \times \left(a^2 \sin^2\left(\theta + \frac{n\pi}{2}\right) + b^2 \cos^2\left(\theta + \frac{n\pi}{2}\right)\right)\right) a\mathrm{d}\theta. \tag{39}$$

Solving the integral gives

$$F_{3D-3D}(\psi) = \sum_{n=0}^{2} \frac{\pi c_n}{a^{2n-1}} (2a^2\psi^2 + 4ab\omega\psi + \omega^2(a^2 + b)^2). \tag{40}$$

So that after we set $\frac{\mathrm{d}F_{3D-3D}(\psi)}{\mathrm{d}\psi} = 0$ and given that ψ is independent of θ, we find

$$\psi_{\min} = -\frac{b}{a}\omega \tag{41}$$

for all nth order derivatives. Since all nth order derivatives provided the same solution, the prediction is independent of the parameter λ. Recall that this parameter determines the relative weight of the slow and smooth motion components.

A.1.2. Case ii: Minimizing 2D motion along the 3D contour

Rather than minimizing the perceived 3D velocity, the visual system may only minimize motion components in the image plane. We called this the $2D_v$–$3D_\theta$ scenario. To find the predicted contour rotation $\psi_{\min}$ that minimizes motion in the image plane we use

$$\left|\frac{\partial^n \vec{v}}{\partial s^n}\right|^2 = \left(\frac{\partial^n v_x}{\partial s^n}\right)^2 + \left(\frac{\partial^n v_y}{\partial s^n}\right)^2 \tag{42}$$

while all other equations are identical to those used in Case i above. It follows:

$$F_{2D-3D}(\psi) = \sum_{n=0}^{2} \frac{c_n}{a^{2n}} \int_0^{2\pi} a\mathrm{d}\theta\big(\psi^2(a^2 \sin^2\vartheta + b^2\cos^2\vartheta) + 2ab\omega\psi + \omega^2(a^2\cos^2\vartheta + b^2\sin^2\vartheta)\big), \tag{43}$$

where $\vartheta = \theta + \frac{n\pi}{2}$. When solving the integral

$$F_{2D-3D}(\psi) = \sum_{n=0}^{2} \frac{\pi c_n}{a^{2n-1}} \big(\psi^2(a^2 + b^2) + 4ab\omega\psi + \omega^2(a^2 + b^2)\big). \tag{44}$$

So that after we set $\frac{\mathrm{d}F_{2D-3D}(\psi)}{\mathrm{d}\psi} = 0$ it follows

$$\psi_{\min} = -\frac{2ab}{a^2 + b^2}\omega \tag{45}$$

for all nth order derivatives. Like in the $3D_v$–$3D_\theta$ scenario, the result in the $2D_v$–$3D_\theta$ scenario is independent of the parameter λ.

A.1.3. Case iii: Minimizing 3D velocity along the 2D contour

Thus far we have presented analytical solutions of the two scenarios that integrate motion across the stimulus contour in 3D space. The visual system may integrate motion across the stimulus contour in the image plane instead. Under the $3D_v$–$2D_\theta$ scenario we minimized the 3D velocity along the 2D contour. The derivation is similar to the $3D_v$–$3D_\theta$ scenario in Case i above. However, when integrating along the stimulus contour in the image plane, we use

$$\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2} = \mathrm{d}\theta\sqrt{a^2 \sin^2\theta + b^2\cos^2\theta}. \tag{46}$$

We derived predictions assuming that motion is integrated in the image plane numerically since we were unable to obtain analytical solutions. The rotation around the surface normal that minimized motion flow appeared identical to

$$\psi_{\min} = -\frac{b}{a}\omega \tag{47}$$

for all nth order derivatives. This solution appears to be identical to the $\psi_{\min}$ predicted in the $3D_v$–$3D_\theta$ scenario.

A.1.4. Case iv: Minimizing 2D velocity along the 2D contour

In the $2D_v$–$2D_\theta$ scenario 2D velocity was integrated along the 2D elliptical contour. The numerical analysis proceeded based on the $2D_v$–$3D_\theta$ scenario derivation in Case ii with the exception that Eq. (23) was used to

integrate across the image plane. We obtained a numerical solution for $\psi_{\min}$ that lies in between the $2D_v$–$3D_\theta$ and $3D_v$ predictions, thus

$$-\frac{2ab}{a^2+b^2}\omega < \psi_{\min} < -\frac{b}{a}\omega. \tag{48}$$

Numerical analysis revealed that the $2D_v$–$2D_\theta$ scenario predicted a $\psi_{\min}$ that appeared linear in ω. Thus, in all four scenarios discussed, a linear relationship between image plane rotation ω and rotation around the surface normal $\psi_{\min}$ appeared to be predicted.

Under the $2D_v$–$2D_\theta$ scenario, predictions were not identical for all nth order derivatives. As a result, the $2D_v$–$2D_\theta$ predictions were a function of λ. When λ approached ∞ the highest order smoothness term dominated the prediction and when λ approached 0 the slowness term dominated so that the prediction quickly approached the slowness only prediction. For all results presented λ was set to 1. Similar results were obtained for a large range of λ. The predicted rotation ψ appeared to remain a function of aspect ratio b/a and a linear function of rotation ω in the image plane.

Fig. 6A shows the theoretical predictions and the rotation reported by observers in Experiment 1. We estimated the perceived rotation ψ as a function of the rotation ω using linear regression for each aspect ratio b/a. Error bars show the standard error of the estimates.

Inclusion of a motion smoothness constraint slightly altered the predicted rotation ψ around the surface normal. In the $3D_v$ scenarios the predictions appeared to be identical to the prediction incorporating a motion slowness constraint only. In the $2D_v$–$3D_\theta$ scenario the predicted rotation ψ appeared to be numerically equal to the prediction obtained in the slowness only derivation. However, in the $2D_v$–$2D_\theta$ scenario the predicted ψ varied slightly from the prediction derived in the slowness only derivation, although the relationship between the variables of rotation around the surface normal ψ, rotation in the image plane ω and aspect ratio of the ellipse b/a remained the same.

A.2. Experiment 2

In Experiment 2 the viewing direction was along the x-axis, so that the Oyz plane became the image plane. Naturally, the prediction based on the $3D_v$–$3D_\theta$ scenario was independent of the viewing direction. However, in the $3D_v$–$2D_\theta$ scenario we integrate along the stimulus contour in the Oyz plane, so that

$$\begin{aligned} ds &= \sqrt{(dy)^2+(dz)^2} \\ &= d\theta\sqrt{(-a\sin\theta\sin\omega t + b\cos\theta\cos\omega t)^2 + (a^2-b^2)\cos^2\theta}. \end{aligned} \tag{49}$$

In this case, and the cases that follow, we derived the solutions numerically. In these cases motion flow was a function of time, so that we integrated over both time t and the stimulus contour s. Motion flow F was defined as

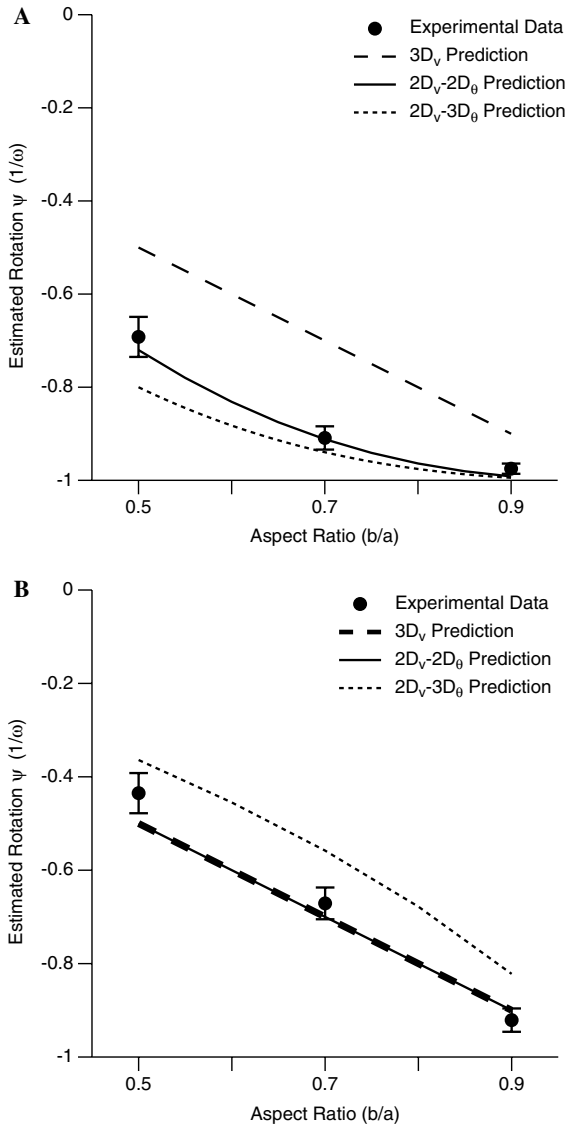

Fig. 6. Estimated and predicted rotation ψ around the surface normal of the 3D percept when minimizing both motion slowness and smoothness. (A) Results based on Experiment 1. (B) Results based on Experiment 2. In (B) the $3D_v$ and $2D_v$–$2D_\theta$ predictions overlap. Error bars show standard error of the mean. When taking motion smoothness into account the $2D_v$–$2D_\theta$ prediction is closest to the experimental data.

$$F(\psi,t) = \sum_{n=0}^{2} c_n \int\int \left|\frac{\partial^n \vec{v}}{\partial s^n}\right|^2 ds\,dt. \tag{50}$$

Numerical analysis suggested that the ψ that minimized motion flow in the $3D_v$–$2D_\theta$ scenario was the same as that in the $3D_v$–$3D_\theta$ scenario regardless of λ.

In the $2D_v$–$3D_\theta$ and $2D_v$–$3D_\theta$ scenarios the motion flow in the image plane is defined as

$$\left|\frac{\partial^n \vec{v}}{\partial s^n}\right|^2 = \left(\frac{\partial^n v_y}{\partial s^n}\right)^2 + \left(\frac{\partial^n v_z}{\partial s^n}\right)^2. \tag{51}$$

In the $2D_v$–$3D_\theta$ scenario we used $ds = a d\theta$ to integrate along the stimulus contour in 3D space. Numerical analysis suggested that the ψ that minimized motion flow was a linear function of ω and was independent of λ.

Finally in the $2D_v$–$2D_\theta$ scenario we used Eq. (49) to integrate along the stimulus contour in the image plane. Numerical analysis suggested that ψ was a linear function of ω, but the predicted ψ was a function of λ.

Fig. 6B shows the theoretical predictions and experimental results for Experiment 2 when taking motion smoothness into account. We estimated the perceived rotation ψ as a function of the rotation ω using linear regression for each aspect ratio b/a. Error bars show the standard error of the estimates. Like in Experiment 1, the $2D_v$–$2D_\theta$ prediction is in close accordance with the experimental data. These results suggest that the visual system minimizes the perceived motion slowness and smoothness of the 3D percept in the image plane.

References

Cavanagh, P., & Anstis, S. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Research, 31*(12), 2109–2148.

Chichilnisky, E. J., Heeger, D., & Wandell, B. A. (1993). Functional segregation of color and motion perception examined in motion nulling. *Vision Research, 33*(15), 2113–2125.

Epstein, W. (1977). *Stability and constancy in visual perception: Mechanisms and processes*. New York, NY: John Wiley & Sons.

Freeman, W. T. (1994). The generic viewpoint assumption in a framework for visual perception. *Nature, 368*(6471), 542–545.

Gibson, J. J. (1950). The perception of visual surfaces. *American Journal of Psychology, 63*(3), 367–384.

Gregory, R. L. (1985). Movement nulling: For heterochromatic photometry and isolating channels for 'real' and 'apparent' motion. *Perception, 14*(2), 193–196.

Grzywacz, N. M., & Yuille, A. L. (1991). Theories for the visual perception of local velocity and coherent motion. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing*. Cambridge, MA: MIT Press.

Harris, J. M., McKee, S. P., & Watamaniuk, S. N. (1998). Visual search for motion-in-depth: Stereomotion does not 'pop out' from disparity noise. *Nature Neuroscience, 1*(2), 165–168.

Hildreth, E. C. (1984). *The measurement of visual motion*. Cambridge, MA: MIT Press.

Jansson, G., & Johansson, G. (1973). Visual perception of bending motion. *Perception, 2*, 321–326.

Kersten, D., Bülthoff, H. H., Schwartz, B. L., & Kurtz, K. J. (1991). Interaction between transparency and structure from motion. *Neural Computation, 4*(4), 573–589.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35*(3), 389–412.

Mamassian, P., & Kersten, D. (1996). Illumination, shading and the perception of local orientation. *Vision Research, 36*(15), 2351–2367.

Mamassian, P., & Landy, M. S. (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research, 38*(18), 2817–2832.

Mamassian, P., & Landy, M. S. (2001). Interaction of visual prior constraints. *Vision Research, 41*(20), 2653–2668.

Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology, 49*(5), 1148–1167.

Musatti, C. (1924). Sui fenomeni stereocinetici. *Archivio Italiano di Psicologia, 3*, 105–120.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science, 257*(5075), 1357–1363.

Rock, I. (1983). *The logic of perception*. Cambridge, Massachusetts: The MIT Press.

Shiffrar, M., & Pavel, M. (1991). Percepts of rigid motion within and across apertures. *Journal of Experimental Psychology: Human Perception and Performance, 17*(3), 749–761.

Todd, J. T., Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology-Human Perception and Performance, 22*(3), 695–706.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 203*(1153), 405–426.

Ullman, S. (1984). Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception, 13*(3), 255–274.

von Helmholtz, H. (1924). *Helmholtz's Treatise on Physiological Optics. Translated from the Third German Edition. Vol. 3*. The Optical Society of America.

Wallach, H., Weisz, A., & Adams, P. A. (1956). Circles and derived figures in rotation. *American Journal of Psychology, 69*, 48–59.

Weiss, Y., & Adelson, E. H. (1998). Slow and Smooth: A Bayesian theory for the combination of local motion signals in human vision. In: A.I. Memo No. 1624: M.I.T.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5*(6), 598–604.

Yang, Z., & Purves, D. (2003). A statistical explanation of visual space. *Nature Neuroscience, 6*(6), 632–640.

Yuille, A. L., & Grzywacz, N. M. (1988). A computational theory for the perception of coherent visual motion. *Nature, 333*(6168), 71–74.