A BAYESIAN APPROACH TO THE APERTURE PROBLEM OF 3D MOTION PERCEPTION

Hongfang Wang, Suzanne Heron, James Moreland and Martin Lages

School of Psychology
University of Glasgow
58 Hillhead Street, Glasgow, G12 8QB, UK

ABSTRACT

We suggest a geometric-statistical approach that can be applied to the 3D aperture problem of motion perception. In simulations and psychophysical experiments we study perceived 3D motion direction in a binocular viewing geometry by systematically varying 3D orientation of a line stimulus moving behind a circular aperture. Although motion direction is inherently ambiguous perceived directions show systematic trends and a Bayesian model with a prior for small depth followed by slow motion in 3D gives reasonable fits to individual data. We conclude that the visual system tries to minimize velocity in 3D but that earlier disparity processing strongly influences perceived 3D motion direction. We discuss implications for the integration of disparity and motion cues in the human visual system.

Index Terms— binocular vision, motion, depth, stereomotion, inverse problem, correspondence problem.

1. INTRODUCTION

Perception may be understood as a form of statistical inference [10]. The Bayesian framework provides an optimal way of combining uncertain information extracted from images with prior assumptions about the nature of objects in the world. This approach has been successful when modeling human visual performance in a range of tasks. In 2D motion perception for example, perceived motion direction and speed can be inferred from noisy velocity constraints that are combined with a prior assumption for slow speed [24]. Similarly, in binocular depth perception, perceived depth can be derived from noisy disparity constraints that are combined with a prior for small depth [18].

Motion encoding is believed to be accomplished by orientation-specific (binocular) complex cells in V1 [8]. Similarly, disparity encoding is known to involve binocular complex cells in V1 tuned to positional and phase offsets [6]. Other early input may come from hypercomplex (end-stopped) cells. If an oriented line or edge moves inside a receptive field or local aperture then the endpoints of the stimulus are occluded and local velocity remains ambiguous. This is known as the aperture problem of local motion [1, 23].

Under binocular viewing, an oriented line or edge inside a circular aperture also has ambiguous 3D motion direction but surprisingly the 3D aperture problem has not been systematically addressed. Similar to the 2D aperture problem, the visual system needs to establish motion correspondence over time to solve the 3D aperture problem. In addition, it has to solve the stereo correspondence between input to the left and right eye. Thus, there are two possible sources of information to establish 3D motion: motion and (dynamic) disparity input.

Existing models of 3D motion perception [19, 5] are based on disparity-first or motion-first processing to predict trajectory and speed, typically for unambiguous features and objects moving on the horizontal plane. Here we investigate a Bayesian model that provides local 3D velocity estimates for oriented lines or edges moving inside an aperture. The generalization from a 2D to a 3D line stimulus is a significant departure from existing models because motion direction is ambiguous in 3D and the moving stimulus or contour can have spatial extent in depth.

First, we outline a general Bayesian approach before we report results from simulations and psychophysical experiments. We investigated perceived motion direction in the context of the 3D aperture problem and compare a Bayesian Motion with a two-stage Disparity-Motion model. Finally, we discuss implications of our results and draw conclusions about motion and disparity processing in the visual system.

2. BAYESIAN MODELS OF 3D MOTION PERCEPTION

The present approach extends existing probabilistic models [24, 11] to 3D motion and provides a velocity estimate for the 3D aperture problem [12]. The underlying geometric-statistical model is based on monocular constraints in a binocular viewing geometry. Likelihoods for velocity constraint planes are combined with a spherical 3D motion prior to estimate perceived velocity.

Funded by The Leverhulme Trust (F-00179B/G), EPSRC studentship to S.H., and Wellcome Biomedical Vacation Scholarship to J.M.

978-1-4799-1580-4/12/$31.00

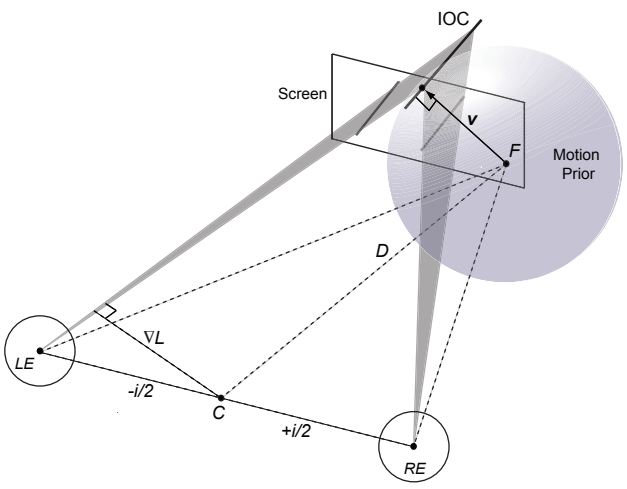

Fig. 1. Illustration of the Bayesian Model. Left and right eye velocity constraint planes (shaded triangles) and their intersection (IOC) are shown for an oriented line moving in 3D. A weak prior for slow motion in 3D is indicated by the sphere centered on the start point F which is also the fixation point. Without disparity bias and small noise in the constraints the 3D velocity estimate $\mathbf{v}$ approximates the vector normal of the IOC through F (arrow).

2.1. Viewing Geometry

For the sake of simplicity we assume a fixed viewing geometry in a cartesian coordinate system where positions on the image plane, nodal points of the eyes and fixation point are known. Since the visual system monitors and receives feedback on accommodation and vergence of the eyes in a binocular viewing geometry, it is reasonable to assume that the visual system computes not only local 2D intensity gradients which constrain 2D velocity in the image plane but extends constraints within a given binocular viewing geometry, effectively creating 3D velocity constraints. Assuming that local apertures are small and close to fixation these constraints may be approximated by planes cutting through 3D velocity space without perspective distortion (Fig. 1).

Geometrically the left and right eye velocity constraint can be expressed through a nodal point and the corresponding 2D motion gradient in the image plane. As a matter of convenience we use the fronto-parallel screen as the image plane for both eyes. The image plane is located at the fixation point $F = (0, 0, 0)$ which also serves as the origin of our co-ordinate system.

Two points on the (fronto-parallel) image plane

$$\begin{aligned} \mathbf{p}_1 &= [t_R, 0, 0]\,, \\ \mathbf{p}_2 &= [t_R - \cos(\theta_R), \sin(\theta_R), 0] \end{aligned} \tag{1}$$

define a constraint line (2D intensity gradients) for the right eye where t_R is the horizontal translation of the projected line and θ_R measures the line orientation from horizontal in the image plane. Similarly, t_L is the horizontal translation and line orientation θ_L for the left eye. We define interocular velocity difference (iovd) as $(t_L - t_R)$ and orientation disparity of the line as $(\theta_L - \theta_R)$ in the image plane. The nodal point of the right eye is given by

$$\mathbf{p}_0 = [+i/2, 0, -D] \tag{2}$$

where i denotes the internodal distance and D is the distance to the image plane and fixation point.

Then the 3D plane normal which describes the right constraint plane can be expressed as

$$\mathbf{n}_R = \frac{(\mathbf{p}_2 - \mathbf{p}_0) \times (\mathbf{p}_1 - \mathbf{p}_0)}{\| (\mathbf{p}_2 - \mathbf{p}_0) \times (\mathbf{p}_1 - \mathbf{p}_0) \|} = \frac{\nabla R}{\|\nabla R\|} \tag{3}$$

where $\times$ denotes the cross product and $\|.\|$ the norm. The three components of the normal may be understood as (first-order Taylor approximations of) intensity gradients for the left ∇L and and right eye ∇R, respectively. In a strict sense, intensity gradients are only defined on a surface. However, we can extend the 2D definition to 3D by setting 3D intensity gradients constant for all x, y, z that project to the same 2D image gradients. The resulting plane describe all possible 3D positions of the line. Alternatively, the constraint planes and their intersection (IOC) may be understood as the inferred line position in 3D when sampled over time. This implies that disparity rather than motion processing determines line positions in 3D before they are integrated by the 3D motion system (see Section 2.3) [5].

2.2. Bayesian Motion (BM) Model

The velocity constraint planes may be noisy due to microsaccades and motion encoding. If we make the simplifying assumption that spatial derivatives ∇L and ∇R of the constraint planes are precise but temporal derivatives L_t and R_t have additive noise

$$\begin{aligned} \tilde{L}_t &= L_t(x, y, z, t) + \eta(x, y, z, t), \\ \tilde{R}_t &= R_t(x, y, z, t) + \eta(x, y, z, t) \end{aligned} \tag{4}$$

where $\eta(x, y, z, t)$ has Gaussian density with zero mean and variance σ_v^2, or $\mathcal{N}(0, \sigma_v^2)$ for short, then the gradient constraint equation also holds in 3D for the left and right eye constraint planes. Given that the intensity of the line or edge does not change with position in 3D we can write

$$\mathbf{v}^T \nabla L + \tilde{L}_t \sim \mathcal{N}(0, \sigma_v^2),\ \mathbf{v}^T \nabla R + \tilde{R}_t \sim \mathcal{N}(0, \sigma_v^2) \tag{5}$$

If temporal noise is small and viewing distance large ($D >> i$) then adding noise to the temporal gradient approximates uncertainty of 3D line motion inside a local aperture. The

Gaussian noise simply blurs the 3D position of the constraint plane along the plane normal.

If 3D velocity $\mathbf{v}$ of the moving line is known then the probability or likelihood of observing the motion constraints for the left and right eye may be expressed as

$$p(\nabla L, \tilde{L}_t|\mathbf{v}) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left(-\left(\mathbf{v}^T\nabla L + \tilde{L}_t\right)^2/2\sigma_v^2\right),$$
$$p(\nabla R, \tilde{R}_t|\mathbf{v}) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left(-\left(\mathbf{v}^T\nabla R + \tilde{R}_t\right)^2/2\sigma_v^2\right) \tag{6}$$

In their influential paper on 2D motion illusions Weiss, Simoncelli and Adelson [24] suggested a 2D Gaussian motion prior for slow motion perception in v_x-v_y velocity space. Similarly, Lages [11] introduced a bivariate Gaussian prior to explain bias in perceived trajectory and speed of a target moving on the horizontal plane. If we assume that most features and objects in a scene are stationary or tend to move slowly on an arbitrary trajectory in 3D space then an isotropic 3D Gaussian provides a plausible world prior for binocular 3D motion perception.

$$p(\mathbf{v}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\mathbf{v}^T\mathbf{v}}{2\sigma^2}\right) \tag{7}$$

The posterior distribution is then expressed by combining left and right eye likelihood constraints with the 3D velocity prior using Bayes' rule.

$$p(\mathbf{v}|\nabla L, \tilde{L}_t; \nabla R, \tilde{R}_t) = \frac{p(\nabla L, \tilde{L}_t|\mathbf{v})p(\nabla R, \tilde{R}_t|\mathbf{v})\, p(\mathbf{v})}{\int p(\nabla L, \tilde{L}_t|\mathbf{v})p(\nabla R, \tilde{R}_t|\mathbf{v})\, p(\mathbf{v})d\mathbf{v}} \tag{8}$$

If we drop the denominator which is independent of $\mathbf{v}$ then

$$p(\mathbf{v}|\nabla L, \tilde{L}_t; \nabla R, \tilde{R}_t) \propto p(\nabla L, \tilde{L}_t|\mathbf{v})p(\nabla R, \tilde{R}_t|\mathbf{v})\, p(\mathbf{v}) \tag{9}$$

In order to find the most probable velocity or maximum a posteriori (MAP) estimate, we take the negative logarithm of the posterior, differentiate it with respect to $\mathbf{v}$ and set the derivative equal to zero.

$$\frac{d}{d\mathbf{v}}\left(-\log\left[p(\mathbf{v}|(\nabla L, \tilde{L}_t), (\nabla R, \tilde{R}_t)\right]\right) = 0$$
$$\frac{\mathbf{v}^T\nabla L\nabla L^T + \nabla L\,\tilde{L}_t}{\sigma_v^2} + \frac{\mathbf{v}^T\nabla R\nabla R^T + \nabla R\,\tilde{R}_t}{\sigma_v^2} + \frac{\mathbf{v}}{\sigma^2} = 0 \tag{10}$$

The logarithm of the posterior is quadratic in $\mathbf{v}$ so that the solution can be written in closed form

$$\hat{\mathbf{v}} = -\left(\mathbf{M} + \frac{1}{\sigma^2}\mathbf{I}\right)^{-1}\mathbf{b} \tag{11}$$

where $\mathbf{I}$ is the 3×3 identity matrix and

$$\mathbf{M} = \frac{1}{\sigma_v^2}\left(\nabla L\nabla L^T + \nabla R\nabla R^T\right),$$
$$\mathbf{b} = \frac{1}{\sigma_v^2}\left(\nabla L\,\tilde{L}_t + \nabla R\,\tilde{R}_t\right) \tag{12}$$

Thus, binocular motion constraints in a known viewing geometry together with a 3D velocity prior disambiguate the local 3D aperture problem and provide a unique velocity estimate.

2.2.1. Simulations

Line stimuli had a vertical (90 deg) or oblique (90 ± 45 deg from vertical) orientation in the image plane. Due to orientation disparity between the projected lines the stimulus line is perceived slanted in depth about the horizontal axis. Orientation disparity is defined as the difference between left and right eye line orientation on the image plane ranging from -6 deg to $+6$ deg. Viewing distance D was set to 55 cm and interocular distance i to 6.5 cm. The line stimulus on the screen always moved horizontally with 3.0 $\pm$0.23 deg/s and interocular velocity difference (iovd) of ± 0.46 deg/s. In the following we express estimated motion trajectories as azimuth and elevation angle in respect to the start point rather than the observer. Ground truth for motion direction of the line stimulus is then $+52.5$ deg azimuth and -52.5 deg azimuth at 0 deg elevation. Note that these angles should be perceived by an observer as horizontal motion to the front and left, and to the back and left, respectively.

We varied the noise in the two likelihoods σ_v^2 while keeping the noise of the prior $\sigma^2 = 1.0$ constant. The noise in the likelihoods therefore reflect the observer's uncertainty about the position of the left and right velocity constraint plane over time.

The predictions for an oblique line stimulus slanted in depth about the horizontal axis are illustrated in Fig. 2. MAP velocity estimates approximate the vector normal (VN) solution if likelihoods are virtually noise-free. MAP estimates are only slightly biased compared to the vector normal if the likelihoods contain very little noise. As noise in the likelihoods increases velocity estimates shorten and move away from the IOC line towards the normal of the right or left constraint plane whichever is closer to the start point.

2.3. Bayesian Disparity Motion (BDM) Model

If disparity rather than motion encoding provides the initial input to 3D motion perception then the perceived position of the constraint planes may be biased. It is well known that observers underestimate slant about the x-axis. This may be modelled by estimation bias in an over-determined equation system [9] or by an explicit disparity prior in a probabilistic approach [11, 18].

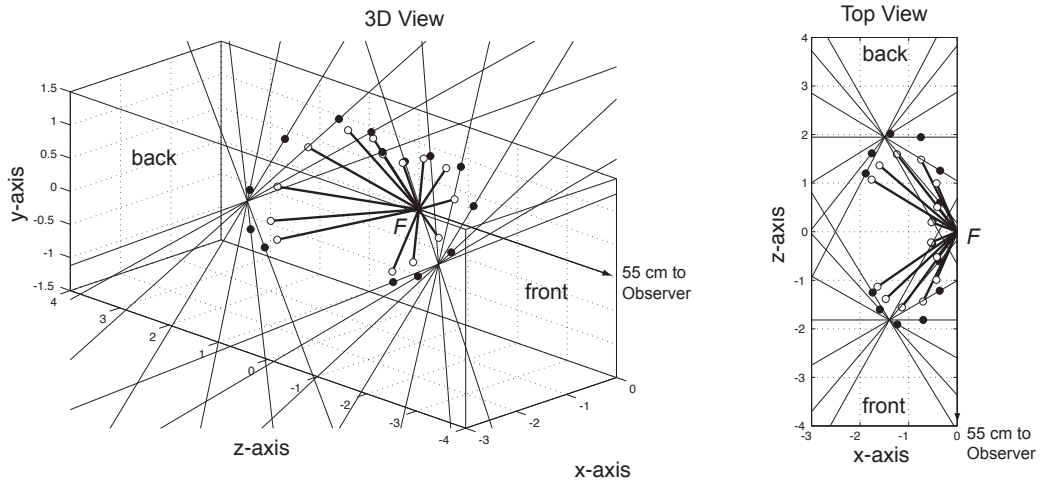

Fig. 2. Simulation results for oblique line stimuli slanted in depth by -6 deg to $+6$ deg in steps of 2 deg and moving to the front left or back left (iovd ± 0.46 deg/s). Bayesian velocity estimates (MAP) with noise ratio $\sigma_v^2 : \sigma^2 = 0.001 : 1$ are shown as thick lines with open circles attached. Endpoints of vector normals (VN) are indicated by filled circles superimposed on thin IOC lines for each orientation disparity.

In an extension of the BM model we introduce disparity computation in the first processing stage. As a consequence points defining the line on the image plane may not be perceived veridically but may be biased due to (orientation) disparity processing.

If the likelihood for orientation disparity $d = (\theta_L - \theta_R)$ is a Gaussian centered on the true disparity δ with standard deviation σ_d

$$p(\delta|d) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left(\frac{-(d-\delta)^2}{2\sigma_d^2}\right) \tag{13}$$

and the Gaussian prior is centered on zero orientation disparity with fixed $\sigma = 1.0$

$$p(d) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-d^2}{2\sigma^2}\right) \tag{14}$$

then we can approximate the posterior by

$$p(d|\delta)) \propto p(\delta|d)p(d) \tag{15}$$

The MAP estimate for orientation disparity is used to adjust the orientation of the left and right line projection $\hat{\theta}_L = (\theta_L + \theta_R)/2 - \hat{d}/2$ and $\hat{\theta}_L = (\theta_L + \theta_R)/2 + \hat{d}/2$ on the image plane. (Alternatively, if we apply disparity estimates to the average of the x-components of the two points $\mathbf{p}_1$ and $\mathbf{p}_2$ defining the line projection on the image plane, then this affects both perceived line position and orientation of the constraint planes.)

The disparity-adjusted lines on the image plane together with the nodal points also define intersecting 3D velocity constraint planes with adjusted gradients $\nabla\tilde{L}$ and $\nabla\tilde{R}$. The estimation of 3D velocity is then achieved as in the BM Model by combining the adjusted likelihoods with the 3D motion prior to derive an estimate of perceived 3D velocity.

3. PSYCHOPHYSICAL RESULTS

Perceived 3D motion direction was measured in repeated trials without time limit. During each trial observers were instructed to continuously fixate a hairline cross at the centre of a black circular aperture with diameter 4.83 deg. An oriented line stimulus of the same mid-gray as the surround appeared and oscillated from left to right and back inside the aperture with a horizontal velocity of $3.0\pm\ 0.23$ deg/s on the left and right screen. The moving line was anti-aliased and blended into the gray surround so that no line endpoints were visible. The fixation cross was positioned 55 cm in depth and the midpoint of the depth probe was placed 2.5 cm behind the fixation point. Motion direction of the line stimulus was measured from the start point rather than observer and was expressed as azimuth angle from the fronto-parallel x-axis and elevation angle from horizontal. Horizontal motion served as ground truth with ± 52.5 deg azimuth from the fronto-parallel and 0 deg elevation from horizontal.

The matching task was programmed in MatLab (MathWorks, Natick, MA) using the Psychophysics Toolbox exten-

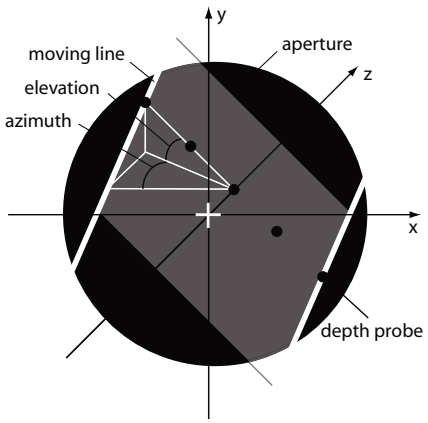

Fig. 3. Illustration of the stimulus display and the matching task. In each trial an oriented line (white) moved back and forth up to inflection points behind a circular aperture on an (invisible) plane (gray) oriented in depth. The observer's task was to fixate the cross at the centre and to adjust orientation and horizontal disparity of a string of five dots until they matched the perceived direction of the moving line.

sions [4, 15] and run on a Macintosh G4 computer with two 21 inch Sony GDM-F500R cathode-ray tube flat screen monitors in a Wheatstone configuration. The monitors were calibrated for luminance using a Minolta photometer (Cambridge Research Systems). Images were presented stereoscopically through haploscopic mirrors at a viewing distance of 55 cm and a frame rate of 120 Hz. Stimuli were shown at 50% Michelson contrast. Observers were comfortably seated in front of the mirrors with their head supported by a chin- and head-rest. The experimental room remained dark with lights switched off during testing. After observing repeated oscillations of the line the observer pressed the spacebar to see a string of five equally-spaced dots inside the aperture. The observer made online adjustments of orientation and horizontal disparity of the probe by pressing corresponding keys on a keyboard. When observers adjusted orientation the string of dots rotated around their midpoint and when they adjusted horizontal disparity the string of probe dots changed position in depth with the dot at the center anchored at a constant depth. Once the observer was confident that the adjusted probe matched perceived motion direction of the line stimulus they pressed a key to confirm their orientation and disparity setting and continued with the next trial. Matching was repeated four times in blocks of 28 randomized trials with two interocular velocity difference (iovd ± 0.46 deg/s), two line orientations ($-45, +45$ deg) and six orientation disparities (ranging between -6 deg and $+6$ deg in steps of 2 deg). Note that the depth probe itself was not perceived veridically and perceived bias in disparity of the five dots was measured in a separate task where the observer indicated perceived depth on a protractor aligned with the fontoparallel plane.

Observers were three University students and one staff member. All observers had normal or corrected-to-normal visual acuity and were screened for stereo deficiencies (Random dot E test). Before experimental testing observers attended several training blocks. Observer M.L. and S.H. were authors. Informed written consent was obtained from naïve observers T.N. and S.W. before participation in experiments and experimental procedures were approved by the Faculty Ethics Committee at Glasgow University according to the Declaration of Helsinki.

3.1. Experiment 1: Perceived Motion Direction of Oblique Line

The moving line stimulus inside the aperture had an orientation of 90 ± 45 deg from vertical. Orientation disparity between the line stimulus projected into the left and right eye varied between ± 6 deg in steps of 2 deg. Adjustments of the depth probe was used to establish perceived motion direction in terms of azimuth and elevation angle from the start point in the centre of the aperture. It was tested whether perceived motion direction of oblique stimulus lines were systematically affected by orientation disparity [12].

3.2. Results of Experiment 1

First, we compared individual data settings to the horizontal stimulus direction and found systematic departures from ground truth in all conditions. Next, we compared the data to predictions of the geometric VN model and found little agreement. Although individual data show the characteristic curvature of the VN predictions azimuth and elevation angles are very much reduced. Fig. 4 illustrates the results from two observers. All observers perceived line motion direction as a function of orientation disparity. Line stimuli are perceived as moving from the start point at an elevation and azimuth angle that reflects the shortest distance to the IOC but only if lines are perceived with reduced orientation disparity and iovd.

We fitted the BM and BDM model with one and two free parameters, respectively. The first parameter reflects an influence of temporal noise whereas the second parameter describes noise or uncertainty in (orientation) disparity processing. Bayesian parameter estimates and model selection are summarized in Table 1. The parameter estimates suggest a strong influence of a disparity prior centered on zero orientation disparity ($\sigma_d : \sigma > 1.0$) and a weak influence of a motion prior centered on zero velocity ($\sigma_v : \sigma < 0.1$) in all four observers. Note that a BM (BDM) with small temporal noise σ_v (and small spatial noise σ_d) approximates the geometric VN strategy. The estimates indicate a strong influence of the zero disparity prior but relatively little temporal noise for motion integration. Again, this suggests a vector normal (VN) strategy for moving line stimuli with a strong bias for zero (orientation) disparity.

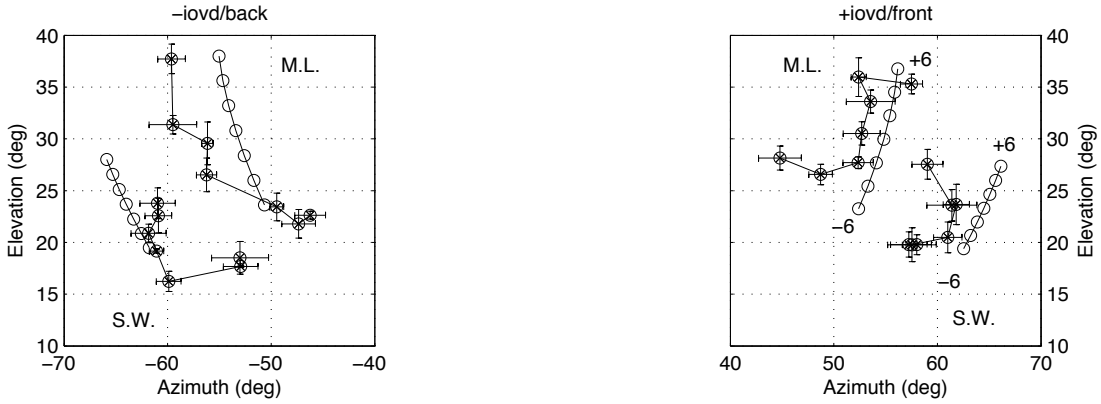

Fig. 4. Left (-iovd/back) and right plot (+iovd/front) shows results from 3D motion direction adjustments for a moving oblique (45 deg) line with orientation disparity between -6 deg to $+6$ deg in steps of 2 deg. Data points are averaged across four repeated trials are shown as black circles with $\pm 1SE$ error bars. Superimposed are the best-fitting estimates of the Bayesian Disparity Motion Model.

Table 1. Bayesian noise estimates (with $\sigma = 1$) and Bayesian model selection for 3D motion direction of oblique line

Obs	**BDM**			**BM**		**Model**
mot	σ_v/σ	σ_d/σ	$\chi^2(11)$	σ_v/σ	$\chi^2(12)$	BF
M.L.	0.056	2.59	8.26	0.088	832	377
T.N.	0.050	2.95	15.4	0.086	766	187
S.H.	0.047	2.76	8.71	0.085	851	366
S.W.	0.031	3.66	12.5	0.083	924	276
sta	σ_v/σ	σ_d/σ	$\chi^2(11)$	σ_v/σ	$\chi^2(12)$	BF
M.L.	0.041	1.67	25.0	0.080	855	128
T.N.	0.051	1.71	20.3	0.083	780	144
S.H.	0.037	3.54	7.51	0.083	945	471
S.W.	0.042	3.13	16.2	0.084	901	208

In a control experiment we investigated the effect of disparity on static stimulus lines. Observers adjusted perceived surface orientation as defined by three parallel lines slanted in depth. The three lines were samples from the line motion display. They correspond to the line at the start position and the two inflection points at frame 1, 60 and 180 during movement. The task for the observer was to indicate the 3D orientation of the surface described by the three parallel lines.

Perceived orientation of the surface corresponds to perceived motion direction of the moving line (not shown). However, compared to the motion condition noise in disparity processing was reduced in two observers.

The only cue that is available in both the static as well as motion experiment is (orientation) disparity between lines inside the aperture. The present data therefore suggest that disparity processing influenced position and orientation of the line before 3D motion integration occurred.

3.3. Experiment 2: Perceived Motion Direction of Vertical Line

A vertical line stimulus (90 deg) slanted in depth about the horizontal x-axis moved behind the circular aperture. Again we varied orientation disparity (-6 deg to $+6$ deg in steps of 2 deg) and two interocular velocity differences (iovd ± 0.46 deg/s).

First, we compared the individual data to geometric predictions of the VN strategy [12] and found little agreement. Again, individual data sets showed reduced azimuth and elevation compared to VN predictions. Next, we fitted the Bayesian Motion (BM) model with one and a Bayesian Disparity-Motion (BDM) model with two free parameters. The first parameter describes the influence of temporal noise whereas the second parameter captures uncertainty in (orientation) disparity. Parameter estimates and Bayesian model selection using an approximation of the Bayes Factor (BF) are summarized in Table 2. As before, the estimates indicate a strong influence of the zero disparity prior but relatively little noise from temporal processing. This suggests a VN strategy with a strong bias for reduced (orientation) disparity.

In a control condition copies of the line stimulus from three time points were displayed simultaneously suggesting a slanted and titled surface. Observers were instructed to adjust the probe so that the string of dots indicated the orientation of the surface. Similar to the motion condition and the previous experiment these adjustments reflect a bias in (orientation) disparity processing with a preference for the shortest distance.

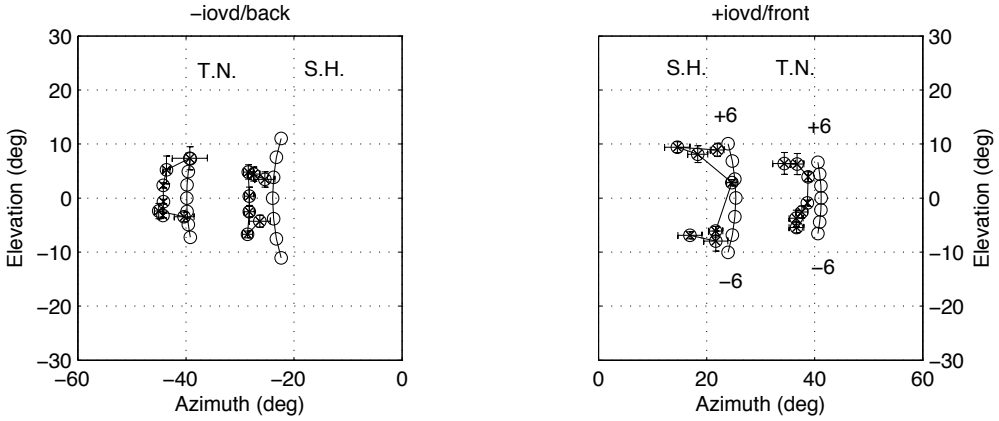

Fig. 5. Left (-iovd/back) and right plot (+iovd/front) shows results from 3D motion direction adjustments for a moving vertical (90 deg) line with orientation disparity varied between ± 6 deg in steps of 2 deg. Data points are averaged across four repeated trials are shown as circles with $\pm 1SE$ error bars. Note that for better illustration all data for SH are shifted by ± 15 deg in azimuth. Superimposed are the best-fitting estimates of the Bayesian Disparity-Motion model.

Table 2. Bayesian noise ratio estimates with $\sigma = 1.0$ and Bayesian model selection for 3D motion direction of vertical line

Obs	**BDM**			**BM**		**Model**
mot.	σ_v/σ	σ_d/σ	$\chi^2(11)$	σ_v/σ	$\chi^2(12)$	BF
M.L.	0.061	1.98	10.8	0.062	358	124
T.N.	0.070	0.79	27.8	0.070	600	80.7
S.H.	0.064	1.44	30.5	0.064	419	51.3
S.W.	0.062	1.00	40.1	0.062	521	49.6
stat.	σ_v/σ	σ_d/σ	$\chi^2(11)$	σ_v/σ	$\chi^2(12)$	BF
M.L.	0.061	0.92	28.8	0.055	910	116
T.N.	0.056	0.19	29.5	0.060	575	74.6
S.H.	0.065	1.08	26.8	0.066	708	98.9

4. DISCUSSION

For both experiments we compared the one-parameter Bayesian Motion (BM) with the two-parameter Bayesian Disparity-Motion (BDM) model approximating the Bayes Factor (BF) by Bayesian Information Content (BIC). For all cases the BF was larger than 40 indicating strong evidence in favour of the BDM model. Assuming that the two model variants are equally plausible the posterior probability in favour of the BDM model accumulated over observers was at least .99 in each condition of the two experiments. This also constitutes very strong evidence in favor of the BDM model [17]. Thus, a BDM model with a strong bias for zero orientation disparity provides a far better account of perceived motion direction than a BM model without disparity bias.

What enables the human visual system to instantaneously perceive 3D motion and to infer direction and speed of a moving object? It seems likely that the visual system exploits many cues in concert to make this difficult inference as reliable and veridical as possible. The diverse set of effective local and global cues has been documented in a number of psychophysical studies [22, 3] and points at late (hMT+/V5+) rather than early integration within the visual processing hierarchy.

It is tempting to assume that gradients, velocity constraints, and motion prior are the result of motion encoding and processing only. However, the stereo correspondence between differently oriented lines in the left and right eye indicates that some contribution from disparity processing is necessary to achieve 3D motion perception. Moreover, we have shown that bias in disparity processing even influences perceived 3D motion direction of a line stimulus. Similarly, perceived orientation of a surface defined by static lines was affected by orientation disparity in the stimulus. Taken together this suggests that motion and disparity information is not integrated early and jointly to estimate 3D velocity but that early disparity processing affects motion constraints and 3D velocity estimation.

Incidentally, combining disparity or depth information with velocity constraints at a later stage provides a flexible scheme that can exploit intermediate depth processing such as relative disparity in V2 [2, 21] and orientation disparity in V4 [7]. Furthermore, it seems possible that velocity constraints may be processed in the ventral stream and (dynamic) binocular disparity and other depth cues in the dorsal stream [16]. It is also neuroanatomically and neurophysiologically plausible that integration of motion and disparity occurs late in subre-

gions of human MT/V5 [14, 20] if not in areas beyond human MT/MST [13].

So far we have only considered translation in 3D space but a moving line or edge may also change orientation over time. It is immediately clear that modeling *velocity* constraints by intersecting planes creates difficulties when encoding rotational stimulus movement. However, if line rotation is captured by disparity processing followed by 3D motion integration then it is possible to represent rotational 3D line motion by updating and integrating position over time.

In summary, our basic geometric-statistical approach extends existing models of 3D motion perception under uncertainty. It provides testable predictions in the context of the 3D aperture problem and captures aspects of human visual perception of local 3D motion direction and 3D surface orientation.

5. REFERENCES

[1] E.H. Adelson and J.A. Movshon. Phenomenal coherence of moving visual patterns. *Nature*, 300:523–525, 1982.

[2] J.S. Bakin, K. Nakayama, and C.D. Gilbert. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*, 20:8188–8198, 2000.

[3] M.F. Bradshaw and B.G. Cumming. The direction of retinal motion facilitates binocular stereopsis. *Proc Royal Soc London B*, 264:1421–1427, 1997.

[4] D.H. Brainard. The psychophysics toolbox. *Spatial Vision*, 10:433 – 436, 1997.

[5] B.G. Cumming and A.J. Parker. Binocular mechanisms for detecting motion-in-depth. *Vision Research*, 34:483–495, 1994.

[6] DeAngelis D.C. and Newsome W.T. Perceptual "readout" of conjoined direction and disparity maps in extrastriate area mt. *PLoS Biology*, 2:0349, 2004.

[7] D.A Hinkle and C.E. Connor. Three-dimensional orientation tuning in macaque area v4. *Nat Neurosci*, 5(7):665–670, Jul 2002.

[8] D.H. Hubel and T.N. Wiesel. Receptive fields , binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 195:215–243, 1968.

[9] H. Ji and C. Fermüller. Noise causes slant underestimation in stereo and motion. *Vision Research*, 46:3105–3120, 2006.

[10] D.C. Knill, D. Kersten, and A.L. Yuille. A bayesian formulation of visual perception. In D.C. Knill and W. Richards, editors, *Perception as Bayesian Inference*. Cambridge University Press, Cambridge, UK, 1996.

[11] M. Lages. Bayesian models of binocular 3-d motion perception. *Journal of Vision*, 6(4):508–522, 2006.

[12] M. Lages and S. Heron. On the inverse problem of binocular 3d motion perception. *PLoS Computational Biology*, 6(11):e1000999, 2010.

[13] L.T. Likova and C.W. Tyler. Stereomotion processing in the human occipital cortex. *Neuroimage*, 38(2):293–305, Nov 2007.

[14] N. Majaj, M. Carandini, and J.A. Movshon. Motion integration by neurons in macaque mt is local not global. *J Neurosci*, 27:366–370, 2007.

[15] D.G. Pelli. The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4):437 – 442, 1997.

[16] C.R. Ponce, S.G. Lomber, and R.T Born. Integrating motion and depth via parallel pathways. *Nature Neuroscience*, 11:216–223, 2008.

[17] A.E. Raftery. Bayes factors and BIC. *Sociological Methods and Research*, 27:411–427, 1999.

[18] J.C.A. Read. A Bayesian approach to the stereo correspondence problem. *Neural Computation*, 14(6):1371–92, 2002.

[19] D. Regan. Binocular correlates of the direction of motion in depth. *Vision Research*, 33(16):2359–2360, 1993.

[20] B. Rokers, L.K. Cormack, and A.C. Huk. Disparity- and velocity-based signals for three-dimensional motion perception in human mt+. *Nat Neurosci*, 12:1050–1055, 2009.

[21] O.M. Thomas, B.G. Cumming, and A.J. Parker. A specialization for relative disparity in v2. *Nature Neuroscience*, 5(5):472–478, 2002.

[22] van Ee R. and B.L. Anderson. Motion direction, speed and orientation in binocular matching. *Nature*, 410:690–694, 2001.

[23] H. Wallach. Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, 20:325–380, 1935.

[24] Y. Weiss, E.P. Simoncelli, and E.H. Adelson. Motion illusions as optimal percepts. *Nature Neuroscience*, 5:598–604, 2002.