Bayesian models of binocular 3-D motion perception

Martin Lages

Department of Psychology, University of Glasgow, Glasgow, Scotland, UK

Psychophysical studies on three-dimensional (3-D) motion perception have shown that perceived trajectory angles of a small target traveling in depth are systematically biased. Here, predictions from Bayesian models, which extend existing models of motion-first and stereo-first processing, are investigated. These statistical models are based on stochastic representations of monocular velocity and binocular disparity input in a binocular viewing geometry. The assumption of noise in these inputs together with a plausible prior for 3-D motion leads to testable predictions of perceived trajectory angle and velocity. Results from two experiments are reported, suggesting that disparity rather than motion processing introduces perceptual bias.

Keywords: velocity difference, disparity change, motion in depth, uncertainty, stereo–motion

Introduction

Observers have to infer trajectory and velocity of three-dimensional (3-D) object motion from 2-D projections into the left and right eyes. Under natural viewing conditions, there are many depth cues to 3-D motion, but in a sparse environment, only binocular depth cues may be available. In the following psychophysical experiments, interocular velocity difference and disparity change are investigated because motion and disparity detectors feature early and prominently in the processing of visual information (Howard & Rogers, 2002).

This is reflected in current models for the perception of motion in depth. (1) The motion-first model postulates monocular motion processing followed by stereo processing (Lu & Sperling, 1995; Regan, Beverley, & Cynader, 1979). In this model, monocular motion is independently detected in the left and right eyes before motion in depth is established. (2) The stereo-first model assumes disparity encoding followed by binocular motion processing (Cumming & Parker, 1994). This model first extracts binocular disparities and then computes change of disparity over time. Note that tracking of spatial location is also required to recover a 3-D motion trajectory. (3) Finally, the stereo-motion model suggests joint encoding of binocular disparity and interocular delay (Carney, Paradiso, & Freeman, 1989; Morgan & Faehle, 2000; Qian & Andersen, 1997).

Psychophysical evidence based on detection and discrimination thresholds has been inconclusive, supporting interocular velocity difference (Brooks, 2002; Portfors-Yeomans & Regan, 1996; Shioiri, Saisho, & Yaguchi, 2000), changing disparity (Cumming & Parker, 1994; Tyler, 1971), or both (Brooks & Stone, 2004) as possible inputs to 3-D motion perception. However, detection and discrimination thresholds cannot reveal bias in 3-D motion perception, and therefore, accuracy rather than precision of observers’ perception was measured in recent psychophysical studies (Harris & Dean, 2003; Welchman, Tuck, & Harris, 2004). Using the method of adjustment, observers reported perceived trajectory angle of a previously seen stimulus moving in depth. The results indicate overestimation of trajectory angle for a range of trajectories. It was suggested that observers exploit cyclopean azimuth $\alpha$ by using the endpoint of motion relative to a fixation point straight ahead (Harris & Drga, 2005). Although appealing in its simplicity, this heuristic requires additional depth cues to solve the inverse problem and to recover 3-D motion.

A constantly moving target changes position in $x$ and $z$ over time. If the eyes remain verged and accommodated on a binocular fixation point straight ahead, then motion information is projected onto the retina of the left and right eyes as illustrated in Figure 1. The projection angles onto the retinas depend on the trajectory and velocity of the motion stimulus, as well as viewing distance $D$ and interpupilar distance IPD. The average of left and right angles approximates visual angle in cyclopean view, and the difference defines binocular horizontal disparity. If projection angles are interpreted as angular velocities, then their difference also defines interocular velocity difference. Although motion and disparity input share the same geometry and are mathematically equivalent, processing of these inputs may be subject to different noise, uncertainty, and bias.

In the following, the motion-first and stereo-first models are extended to uncertainty models in a Bayesian framework. Predictions of these models are derived for 3-D motion trajectories over the full range of $360$ deg. They are tested under experimental conditions that try to minimize the influence of monocular depth cues and response heuristics.
Bayesian models

Some promising Bayesian models have been developed in vision (cf. Knill & Richards, 1996). Most notably, Weiss, Simoncelli, and Adelson (2002) combined motion constraints of local motion detectors with a Gaussian prior for slow motion to predict perceived motion direction and velocity of luminance-defined objects in 2-D space. With this elegant approach, they could explain a variety of 2-D motion illusions.

Most objects in natural environments are stationary. If one assumes that objects move slowly on any trajectory in $x$–$z$ space, then a symmetric bivariate Gaussian probability distribution centered on the starting point of a target, as shown in Figure 2, provides a plausible prior for 3-D motion in $x$–$z$ space. Symmetric perspective projections of this world prior into the left and right eyes give rise to marginal Gaussian distributions defining velocity priors centered on zero velocity. Similarly, the difference of the marginal distributions in the left and right eyes defines a prior for disparity (change) centered on zero disparity. Thus, the same 3-D motion prior in the world results in Gaussian velocity and disparity priors depending on the processing of signals.

The binocular viewing geometry also suggests constraints for stimulus trajectory and velocity in both eyes. The intersection of visual constraint lines in $x$–$z$ space determines trajectory angle and velocity of a target moving in depth (see Appendix A). Under certainty, these constraint lines intersect at the same point, whether they are based on angular velocities or binocular disparities. If noise or uncertainty is introduced together with different priors for motion or disparity processing, then the intersection of constraints falls on different points, as illustrated in Figure 3.

There are several potential sources for uncertainty and noise in binocular motion processing. For example, small moving targets in a sparse 3-D environment offer limited motion and disparity input, thereby introducing a degree of uncertainty in the observer. Mini-saccades during fixation or early noise in the encoding system is another possible source (Hogervorst & Eagle, 1998).

First, assume that noise is present in the activation of monocular motion detectors optimally tuned to different velocities (see Appendix B). The representation of angular velocity in each eye is therefore not exact but subject to noise (Ascher & Grzywacz, 2000). The corresponding likelihood distributions for angular velocity in the left and right eyes are conveniently expressed as Gaussian distributions with equal variance centered on the true angular
Figure 3. Illustration of predicted trajectory and velocity for the (A) motion-first and (B) stereo-first Bayesian models in top view. The constraint lines and IOC (gray) shift as a consequence of biased velocity (blue) or disparity (red) processing (see Appendices B and C for details).

Figure 4. Predictions of trajectory angle and velocity in polar coordinates for the (A) motion-first and (B) stereo-first Bayesian models. Points denote model predictions for a target moving in three dimensions on trajectory angles of 10 to 350 deg, in steps of 20 deg. Uncertainty is modeled by the ratio of standard deviation for likelihood and prior, ranging from 0.1 to 1.9, in steps of 0.2. At a viewing distance of 1,140 mm, predictions are initially veridical, describing a circle of radius 33 mm, but with increasing uncertainty, they approximate a shrinking circle for the motion-first Bayesian model (blue) and a laterally oriented ellipse for the stereo-first Bayesian model (red; see text for explanation).
velocity of the stimulus in each eye. Each likelihood distribution is then combined with the velocity prior in each eye. Velocity priors favoring slow and smooth motion have been suggested in the context of 2-D motion (Ullman & Yuille, 1989; Weiss et al., 2002).

Alternatively, internal noise may be introduced by the activation of binocular disparity detectors tuned to different disparities (see Appendix C). The likelihood distribution for disparity (change) is conveniently expressed as a Gaussian distribution centered on the true disparity (change) of the stimulus. This representation of disparity is combined with the disparity prior favoring zero disparity. A similar disparity prior has been suggested in the context of sustained and transient stereo images (Read, 2002a, 2002b).

Following Bayes’s rule, likelihoods and priors are combined to establish a posterior distribution for each model and trajectory. Applying a simple decision rule to the posterior distributions provides a posteriori estimates of angular velocity and disparity. These estimates correspond to biased constraint lines, and their intersection point determines a predicted trajectory angle and radial distance in x–z space.

In Figure 4, predictions over the full range of stimulus trajectories are derived for the two alternative Bayesian models of binocular 3-D motion perception. Predictions are based on motion trajectories between 0 and 360 deg with a radius of 3.3 cm at a viewing distance of 114 cm and at an interpupilar distance of 6.5 cm. Uncertainty is modeled by the ratio of likelihood and prior standard deviation varying from 0.1 to 1.9 in steps of 0.2. Predictions are initially veridical, but with increasing uncertainty, they approximate a shrinking isotropic circle for the motion-first Bayesian model and an ellipse for the stereo-first Bayesian model. In Figure 4A, the unbiased prediction of trajectory angles or isotropic shape is the result of or multiplying left and right angular velocities by the same factor (see also Equation B5). In Figure 4B, the increasingly elliptical shape is the consequence of stronger bias for large disparities near the z-axis (0 and 180 deg) and weaker bias for small disparities near the x-axis (90 and 270 deg).

Figure 5. Illustration of adjustment methods. In separate blocks of trials, observers rotated a black line inside a circle or adjusted its length to indicate perceived trajectory and radial distance in top view.

Shown at 50% Michelson contrast using a ray-tracing technique to compensate for vergence and projection on a flat screen. Observers were seated in front of haploscopic mirrors with their head supported by a chin rest and a headrest. The experimental room remained dark with lights switched off during testing.

All observers had normal or corrected-to-normal visual acuity and were screened for stereo deficiencies. Before testing, observers attended a training block where they received feedback on trajectory angle and radial distance of each previously seen target motion. Observers who placed more than 25% of trajectory angles in the wrong quadrant were excluded. Informed written consent was obtained from naive observers before participation in experiments.

The method of adjustment was used to assess perceived trajectory angle and radial distance of the stimulus in x–z space after each stimulus presentation. As illustrated in Figure 5, observers saw a line probe inside a circle with a diameter of 88.8 mm, representing motion trajectory in top view. In four blocks of trials, they adjusted either orientation or length of the line probe. Observers could make coarse or fine adjustments by pressing corresponding keys on a keyboard. When observers adjusted line length, the probe was set to an orientation that corresponded to the correct trajectory angle. Once an adjustment was completed, observers press a key to confirm their setting and to continue with the next trial.

General methods

Tasks were programmed in MatLab using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and run on a Macintosh G4 computer with a 21-in. Sony GDM-F500R cathode-ray tube flat screen monitor. The monitor was calibrated for luminance using a Minolta photometer (Cambridge Research Systems). Stimuli were presented in a split-screen Wheatstone configuration at a viewing distance of 1,140 mm with a frame rate of 120 Hz. Stimuli were shown at 50% Michelson contrast using a ray-tracing technique to compensate for vergence and projection on a flat screen. Observers were seated in front of haploscopic mirrors with their head supported by a chin rest and a headrest. The experimental room remained dark with lights switched off during testing.

All observers had normal or corrected-to-normal visual acuity and were screened for stereo deficiencies. Before testing, observers attended a training block where they received feedback on trajectory angle and radial distance of each previously seen target motion. Observers who placed more than 25% of trajectory angles in the wrong quadrant were excluded. Informed written consent was obtained from naive observers before participation in experiments.

The method of adjustment was used to assess perceived trajectory angle and radial distance of the stimulus in x–z space after each stimulus presentation. As illustrated in Figure 5, observers saw a line probe inside a circle with a diameter of 88.8 mm, representing motion trajectory in top view. In four blocks of trials, they adjusted either orientation or length of the line probe. Observers could make coarse or fine adjustments by pressing corresponding keys on a keyboard. When observers adjusted line length, the probe was set to an orientation that corresponded to the correct trajectory angle. Once an adjustment was completed, observers press a key to confirm their setting and to continue with the next trial.
Figure 6. Results from four observers in Experiment 1. Polar plots of perceived trajectory angles and radial distances (black) and best fitting motion-first (blue) and stereo-first (red) Bayesian models. Black data points denote average adjustments to stimulus trajectories between 10 and 350 deg, in steps of 20 deg, as well as average radial distance; filled data points correspond to trajectories at 90 and 270 deg.
Ideally, observers would have made both adjustments after a single stimulus presentation, but pilot studies suggested that a dual task was too demanding for most naïve observers.

**Experiment 1: Variation of 3-D stimulus trajectory**

Using the method of adjustment, we measured perceived trajectory angle and velocity of a small target stimulus in separate blocks of trials. Motion-first and stereo-first Bayesian models were fitted to individual data sets, and results are compared.

**Methods**

**Observers**

Four unpaid undergraduate students from the University of Glasgow took part. They were naïve as to the aim of the experiment.

**Stimuli and procedure**

Stimuli were presented to the left and right eyes using the split-screen Wheatstone configuration. Participants viewed a single Gaussian dot subtending less than 4.4 arcmin, which was initially positioned 0.5 deg above a fixation cross. The fixation cross flanked by nonius lines above and below subtended 15.5 arcmin. The target dot moved 33.3 mm at 0.05 m/s on each of the 18 trajectories in x–z space for 666 ms. Note that monocular image velocity changed due to the projections of the animated stimulus in separate blocks of trials. Motion-first and stereo-first Bayesian models were fitted to individual data sets, and results are compared.

On each trial, observers verged on the fixation cross before they initiated motion of the target by key-press. Trajectory angle of motion ranged between 10 and 350 deg in steps of 20 deg. In a 2 × 2 block design, each observer judged stimulus trajectories either to the front or to the back and reported either perceived trajectory angle or radial distance. In each block, adjustments to nine different trajectories were repeated in 10 randomly intermixed trials.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Observer A</th>
<th>Observer B</th>
<th>Observer C</th>
<th>Observer D</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.M.</td>
<td>2.41</td>
<td>134.9</td>
<td>6.03*</td>
<td>1.44</td>
</tr>
<tr>
<td>K.K.</td>
<td>0.00</td>
<td>23.3</td>
<td>6.53*</td>
<td>1.00</td>
</tr>
<tr>
<td>L.B.</td>
<td>1.04</td>
<td>53.5</td>
<td>1.60</td>
<td>0.80</td>
</tr>
<tr>
<td>S.Y.</td>
<td>2.91</td>
<td>147.1</td>
<td>7.32*</td>
<td>1.51</td>
</tr>
</tbody>
</table>

Table 1. Results from four observers in Experiment 1. Parameter estimates of uncertainty and radius, as well as of goodness of fit for motion-first and stereo-first Bayesian models. Note. *p < .05.

Results and discussion

In Figure 6, individual data sets are depicted in polar plots where orientation of 0 deg indicates motion to the front; 90 deg, motion to the left; 180 deg, motion to the back; and 270 deg, motion to the right. Adjustments for each trajectory angle were averaged across 10 trials.

For each observer, average adjustments (black points) systematically deviate from veridical, that is, a circle with a 33.3-mm radius and trajectories from 10 to 350 deg in steps of 20 deg. Individual data averaged over 10 trials are noisy but appear more elliptical than circular. Trajectory angle was overestimated near the median plane and underestimated near the frontal plane, whereas radial distance was overestimated near the frontal plane and underestimated near the median plane.

Parametric fits to individual data are provided for the motion-first (blue) and the stereo-first (red) Bayesian model. The ratio of standard deviations for likelihood and prior was estimated (cf. Hürlimann, Kiper, & Carandini, 2002) together with radius for each model. Estimated ratios σv/σd for angular velocity and σr/σd for disparity are presented together with estimated radius r_v and r_d in Table 1. These two parameters were fitted in the maximum likelihood sense for each observer and model.

As shown in Table 1, the stereo-first model gives better fits, and parameter estimates assume more plausible values because the ratio of standard deviations varies around 1.0 and estimates of radius are close to the true value of 33.3 mm. For the motion-first model, the ratio of standard deviations indicates no uncertainty for Observer K.K. and radius is overestimated by a factor of 4 in Observers E.M. and S.Y.

Despite the wide range of trajectories in Experiment 1, there is concern that results were affected by response heuristics. The presentation of motion to the front and back in separate blocks of trials may have encouraged observers to judge trajectory by comparing angular velocities or cyclopean azimuth of the motion endpoint across trials (Harris & Drga, 2005). Monocular image velocity and cyclopean azimuth of motion endpoint are confounded with trajectory angle (Brooks & Stone, 2004). All observers assigned a crossed disparity to 90 deg trajectory, an uncrossed disparity to 270 deg trajectory, and relatively constant depths to motion trajectories in the same block. The assignment of stimulus motion to depth planes that correspond with the presentation of stimulus trajectories to the front and back in separate blocks could...
Figure 7. Results from four observers in Experiment 2. Polar plots for perceived trajectory angle and radial distance standardized to radius 1.0 (black) and best fitting motion-first (blue) and stereo-first (red) Bayesian models. Black data points denote average adjustments to stimulus trajectories between 0 and 350 deg in steps of 10 deg, as well as average standardized radial distance; filled data points correspond to cardinal trajectories.
be the result of sequential dependencies between trials (cf. Lages & Treisman, 1998).

### Experiment 2: Variation of 3-D stimulus trajectory and velocity

The same method and apparatus as in Experiment 1 were employed unless mentioned otherwise. In each block of trials, motion targets traveled on 36 trajectory angles at three different velocities in randomly intermixed trials. These changes in stimulus and design were introduced to eliminate response heuristics of observers who may rely on cues such as monocular velocity or endpoint of motion when adjusting trajectory angle and radial distance.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Radius $(r)$</th>
<th>Motion-first $\sigma_{d/\sigma}$</th>
<th>$r_d$</th>
<th>$\chi^2(2)$</th>
<th>Stereo-first $\sigma_{d/\sigma}$</th>
<th>$r_d$</th>
<th>$\chi^2(2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.G.</td>
<td>16.6</td>
<td>0.00</td>
<td>17.6</td>
<td>2.40</td>
<td>0.32</td>
<td>17.7</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
<td>0.00</td>
<td>20.6</td>
<td>6.63*</td>
<td>0.81</td>
<td>27.7</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>33.3</td>
<td>0.00</td>
<td>25.2</td>
<td>17.2*</td>
<td>1.01</td>
<td>32.9</td>
<td>5.79</td>
</tr>
<tr>
<td>M.L.</td>
<td>16.6</td>
<td>0.00</td>
<td>26.3</td>
<td>2.89</td>
<td>0.56</td>
<td>28.2</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
<td>0.00</td>
<td>30.8</td>
<td>4.07</td>
<td>0.49</td>
<td>34.3</td>
<td>3.09</td>
</tr>
<tr>
<td></td>
<td>33.3</td>
<td>0.00</td>
<td>34.4</td>
<td>9.24*</td>
<td>0.77</td>
<td>39.5</td>
<td>3.00</td>
</tr>
<tr>
<td>R.G.</td>
<td>16.6</td>
<td>1.83</td>
<td>113.5</td>
<td>1.40</td>
<td>0.48</td>
<td>28.8</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
<td>2.84</td>
<td>274.3</td>
<td>3.40</td>
<td>0.87</td>
<td>35.4</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>33.3</td>
<td>2.22</td>
<td>213.1</td>
<td>8.97*</td>
<td>1.00</td>
<td>43.2</td>
<td>1.45</td>
</tr>
<tr>
<td>S.S.</td>
<td>16.6</td>
<td>0.00</td>
<td>22.2</td>
<td>2.31</td>
<td>0.00</td>
<td>22.2</td>
<td>2.31</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
<td>0.00</td>
<td>36.1</td>
<td>1.45</td>
<td>0.58</td>
<td>39.3</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>33.3</td>
<td>0.82</td>
<td>64.5</td>
<td>1.74</td>
<td>0.64</td>
<td>44.5</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Table 3. Results from four observers and three stimulus velocities (radius in millimeters) in Experiment 2. Parameter estimates and goodness of fit for motion-first and stereo-first Bayesian models are shown. Note. *$p < .05$. 

### Methods

#### Observers

Four observers took part. A.G., R.G., and S.S. were naive as to the aim of the experiment, whereas M.L. is the author.

#### Stimuli and procedure

To increase binocular input and to reduce uncertainty, each observer viewed three Gaussian blobs presented above and below the fixation cross surrounded by a rectangular fusion lock at fixation depth. Each dot subtended less than 4.4 arcmin at 27.7, 38.8, and 50 arcmin above and below fixation. In randomly intermixed trials, the dots moved 16.6, 25, or 33.3 mm in $x$–$z$ space for 833 ms (0.02, 0.03, or 0.04 m/s) on 36 trajectories.

On each trial, the observer verged on the fixation cross before they initiated motion of the target dots by keypress. Trajectory angle ranged between 0 and 350 deg in steps of 10 deg. Each observer attended a total of 8 separate blocks (2 tasks x 4 repetitions), each comprising 108 trials (3 velocities x 36 trajectories).

#### Results and discussion

Individual adjustments of radial distance were first standardized to radius 1.0 and then averaged across 12 trials. As shown in Figure 7, all observers deviate systematically from veridical, that is, a circle with equidistant trajectories.
Figure 8. Results from four observers in Experiment 2. Polar plots for perceived trajectory angle and radial distance for stimulus velocities of 0.02 m/s or 16.6 mm (blue), 0.03 m/s or 25.0 mm (magenta), and 0.04 m/s or 33.3 mm (red) and best fitting stereo-first Bayesian models. Filled data points correspond to cardinal stimulus trajectories. With increasing stimulus velocity, estimates of radius and uncertainty increase, and model fits assume a more compressed elliptical shape (see Table 3 for details of both model fits and Supplements for separate plots of perceived vs. physical trajectory angle and radial distance).
and constant radius. Average adjustments (black points) appear elliptical rather than circular. Maximum likelihood parametric fits to individual data are provided for the motion-first (blue) and stereo-first (red) Bayesian models.

As in Experiment 1, standard deviation of the likelihood distribution for angular velocity or disparity together with radius served as free parameters. The two parameters were fitted, and results are shown for each observer and model in Table 2. The stereo-first Bayesian model gives better fits, and parameter values assume more plausible values for each observer, confirming that disparity processing determines perceived trajectory angle and velocity. However, radial adjustments for approaching and receding motions at 0 and 180 deg appear similar. Reduced variability of trajectory adjustments also suggests that opposite motion directions in the left and right eyes may have helped observers identify these trajectory angles.

In a more detailed analysis, adjustments of trajectory and radial distance were averaged across trials with the same stimulus velocity only. Two-parameter models were fitted to data of each observer and stimulus velocity, and results are shown in Table 3. Again, the stereo-first Bayesian model gives better fits, and parameter values assume more plausible values. As illustrated in Figure 8, fits of the stereo-first model for each stimulus velocity suggest that an increase in stimulus velocity systematically raises uncertainty. This trend appears in all four observers except for Observer M.L. in the 16.6-mm condition.

### General discussion

Rendering 3-D motion in a stereoscopic setup is difficult and can introduce various artifacts and cue conflicts. In the present experiments, constant size and blur of the target stimuli moving in depth could have influenced perceived depth (Watt, Akeley, Ernst, & Banks, 2005). On the other hand, cue conflict due to looming and accommodative cues are too small to account for the substantial and systematic bias found for small blurred targets that move up to ±3.3 cm in depth at a viewing distance of 114 cm. Using LEDs moving in real depth, Harris and Dean (2003) and Welchman et al. (2004) reported systematic overestimation of stimulus trajectories near the median plane, confirming the perceptual bias for trajectory angle.

In the two experiments reported here, angular velocities ranged from 0.03 to 2 deg/s and disparities ranged from 0.0 to 5.87 arcmin. It seems reasonable to assume that velocity-tuned and disparity-tuned channels have relatively constant bandwidth for this limited range of monocular velocities and binocular disparities. This was modeled by a single parameter estimate of the likelihood standard deviation for all trajectories. However, velocity discrimination thresholds tend to be lower for low velocities (e.g., McKee & Nakayama, 1984), and disparity discrimination thresholds tend to decrease close to zero disparity (e.g., Farell, Li, & McKee, 2004). Hence, it can be argued that uncertainty, modeled by the spread of the likelihoods, may be linearly related to monocular image velocities and binocular disparity for different trajectories.

Within a Bayesian framework, only the posterior is directly observable, whereas likelihoods and priors need to be estimated. Thus, just-noticeable differences in a standard discrimination task do not necessarily inform about the spread of likelihood distributions because reference and test stimuli are both exposed to a velocity or disparity prior (see however Stocker & Simoncelli, 2006). Nevertheless, if one assumes that standard deviations of likelihoods vary linearly with monocular image velocity and with binocular disparity, then predictions of the motion-first model are no longer isotropic but are compressed along the x-axis, whereas predictions of the stereo-first model are further compressed along the z-axis (see Supplements). When a linear relationship between input values and likelihoods is introduced, then the model predictions strengthen the claim that disparity processing is the source of the bias.

There is a third option for the encoding of 3-D motion that needs to be considered. A stereo–motion model usually refers to joint encoding of binocular disparity and interocular delay (Carney et al., 1989; Morgan & Fahle, 2000; Qian & Andersen, 1997). In neurophysiological studies, it was shown that a number of binocular complex cells in cats (Anzai, Ohzawa, & Freeman, 2001) and cells in V1 and MT of monkey (Pack, Born, & Livingstone, 2003) are tuned to interocular spatial–temporal shifts, but the significance of these findings has been questioned (Read & Cumming, 2005a).

In the present framework, it is difficult to test the possibility of joint velocity and disparity encoding, but a combined input of velocity and disparity is easily implemented in a Bayesian model with three parameters (see Appendix D). Uncertainty in velocity and disparity processing was combined, and both uncertainty parameters were estimated together with a parameter for perceived radius. The maximum likelihood fits of this stereo–motion model were almost identical to the stereo-first model fits with no significant improvement in goodness of fit for any of the individual data sets (results not shown).

### Conclusions

Bayesian model fits to perceived trajectory angle and radial distance promotes the idea that bias in 3-D motion
perception is introduced by disparity processing. This confirms previous findings in psychophysical studies that used different stimuli and methods (e.g., Cumming & Parker, 1994; Lages, Mamassian, & Graf, 2003). It is possible, however, that interocular velocity difference or optic flow contributes to 3-D motion perception, especially when stimuli are large and move on or near the median plane (Brooks & Stone, 2004).

In the stereo-first Bayesian model, disparity estimates are derived from the endpoint of stimulus motion rather than integrated over time. As a consequence, the stereo-first Bayesian model may be interpreted as (1) temporal integration of biased disparities or (2) biased temporal integration of disparity. The latter interpretation seems more plausible because uncertainty estimates increased systematically with stimulus velocity as reported in Table 3 of Experiment 2. If 3-D motion perception is based on velocity-tuned processing, a small change of stimulus velocity should have very little effect on uncertainty. Disparity-tuned processing, on the other hand, may well increase uncertainty levels for faster stimuli due to temporal limits of disparity integration (Read & Cumming, 2004).

In addition, using the same method and procedure as in Experiment 2, observers showed reduced bias in their adjustments of trajectory angle and radius when a static target stimulus appeared at the endpoint of each motion trajectory.

One of the main goals of visual processing is to segregate and identify objects in space and time. With increasing proximity or size of a moving object, monocular motion detectors signal a wider range of velocities. As a consequence, a system that processes motion input first needs to establish correspondence between different monocular motions before it can build a global 3-D motion percept. Computationally, it appears more parsimonious to solve the correspondence problem for disparity at successive points in time before deriving a 3-D motion percept. This argument also applies to joint motion and disparity encoding because the visual system would require a large number of detectors specifically tuned to all combinations of interocular spatial and temporal offsets to capture objects moving in three dimensions.

It is concluded that under the present experimental conditions, perceptual bias in 3-D motion trajectory and velocity is most likely the result of limited temporal integration for disparity change. This points to stereo-first or stereo-motion processing with negligible motion bias but rules out a motion-first mechanism that relies on interocular velocity difference only.

## Appendix

Details of viewing geometry and intersection of constraints (IOC) are provided in Appendix A. The three Bayesian models are described in Appendices B, C, and D.

### A. Viewing geometry and IOC

An approximation of trajectory angle \( \beta \) between an object’s direction of motion and the median plane is usually given by the following expression:

\[
\tan \beta = \frac{(\dot{a}_R + \dot{a}_L)}{2} \times \frac{\text{IPD}}{D} = \frac{\dot{a}}{\delta} \times \frac{\text{IPD}}{D} ,
\]

(A1)

where \( \dot{a}_L = d\alpha_L(t)/dt \) denotes the angular velocity in the left eye, \( \dot{a}_R = d\alpha_R(t)/dt \) is the angular velocity in the right eye, \( \dot{a} = d\alpha(t)/dt \) is the angular velocity of the cyclopean (fused) image, and \( \delta = d\delta(t)/dt \) describes the rate of disparity change. Both expressions on the right-hand side are mathematically equivalent. If the cyclopean point on the Vieth–Müller circle is approximated by IPD/2 between the eyes, then this introduces a computational error for shorter viewing distances.

In an alternative solution, the IOC was determined in Cartesian coordinates with the origin \( O \) located halfway on IPD as shown in Figure 1. Angles \( a_L \) and \( a_R \) measure deviation from the visual axes in the left and right eyes. The constraint lines in the left and right eyes intersect the \( x \)-axis at \(-\text{IPD}/2\) and \(+\text{IPD}/2\) with slopes \( \tan(a_L) \) and \( \tan(a_R) \), where angles \( a_L \) and \( a_R \) are measured counterclockwise and clockwise from the \( x \)-axis, respectively. They are adjusted for vergence angle \( \beta_0 \), so that \( a_L = 90 - \beta_0 + a_L \) and \( a_R = 90 - \beta_0 - a_R \).

The constraint lines for the left and right eyes

\[
z = \tan(a'_L) \times \left( \frac{\text{IPD}}{2} + x \right) ,
\]

(A2)

\[
z = \tan(a'_R) \times \left( \frac{\text{IPD}}{2} - x \right)
\]

are equated and solved for \( x \) and \( z \):

\[
x = \frac{\text{IPD}}{2} \times \frac{\sin(a'_R - a'_L)}{\sin(a'_R + a'_L)} \]

(A3)

\[
z = \text{IPD} \times \sin(a'_L) \times \left( \frac{\sin(a'_R)}{\sin(a'_R + a'_L)} \right)
\]

\( x \) is the horizontal distance to the cyclopean point, \( z \) is the vertical distance to the cyclopean point, and \( \alpha \) is the angular velocity of the cyclopean point. The constraint lines in the left and right eyes are mathematically equivalent. If the cyclopean point on the Vieth–Müller circle is approximated by IPD/2 between the eyes, then this introduces a computational error for shorter viewing distances.
From the coordinates of the intersection point, the clockwise trajectory angle $\beta$ is then estimated by

$$\beta = \arctan\left( \frac{x}{(z-D)} \right) \quad \text{for } z < D$$

$$\beta = \arctan\left( \frac{x}{(z-D)} \right) + 180 \quad \text{for } z \geq D$$

and radial distance $r$ by

$$r = \sqrt{x^2 + (z-D)^2}. \quad \text{(A5)}$$

This geometric solution requires knowledge about IPD and $D$. It is used to first determine constraint lines and IOCs and then estimate trajectories and radii for the Bayesian models.

**B. Motion-first Bayesian model**

Following Bayes’s rule, the likelihoods and priors of a scene $S$ and image $I$ are combined to produce a posterior for each model (e.g., Knill & Richards, 1996):

$$p(S|I) \propto p(I|S) \times p(S). \quad \text{(B1)}$$

In this framework, perceived angular velocity of motion-first processing is described as a product of likelihood and prior for the left and right eyes:

$$p(v_L|\beta) \propto p(\beta|v_L)p(v_L),$$

$$p(v_R|\beta) \propto p(\beta|v_R)p(v_R). \quad \text{(B2)}$$

The likelihood for the left eye is modeled as a Gaussian distribution of angular velocities centered on the true angular velocity with $\dot{a}_L(t)/dt$ abbreviated as $\dot{a}_L$. The standard deviation $\sigma_v$ of the likelihood distribution is left as a free parameter:

$$p(\beta|v_L; \sigma_v) = \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp\left[ \frac{-(v_L - \dot{a}_L)^2}{2\sigma_v^2} \right]. \quad \text{(B3)}$$

The likelihood for the right eye is modeled accordingly. The preference or prior for slow motion is modeled as a Gaussian distribution centered on zero velocity with fixed standard deviation $\sigma$:

$$p(v_R; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[ \frac{-(v_R)^2}{2\sigma^2} \right]. \quad \text{(B4)}$$

The product of the Gaussian likelihood distribution with prior $G(0,\sigma)$ gives the posterior distribution, which is the probability of each possible angular velocity taking into account both prior and trajectory. Through differentiation, the maximum a posteriori (MAP) estimates of angular velocity are found for the left and right eye:

$$\hat{a}_L = \frac{1/\sigma_v^2 \dot{a}_L}{1/\sigma_v^2 + 1/\sigma^2} = \frac{\dot{a}_L}{1 + (\sigma_v/\sigma)^2}$$

$$\hat{a}_R = \frac{1/\sigma_v^2 \dot{a}_R}{1/\sigma_v^2 + 1/\sigma^2} = \frac{\dot{a}_R}{1 + (\sigma_v/\sigma)^2}. \quad \text{(B5)}$$

For a given stimulus, the true trajectory $\beta$ and radius $r$ are inserted in Equations A4 and A5. These equations are solved for the coordinates of the IOC before Equation A3 is solved for true angular velocities in the left and right eyes. Bias in angular velocities is introduced according to Equation B5. These estimates are then inserted into Equation A3 to determine the new coordinates of the biased IOC. Inserting these coordinates in Equations A4 and A5 provides the predicted trajectory and radius for a given stimulus.

The ratio $\sigma_v/\sigma$ (cf. Hürlimann et al., 2002) in Equation B5 was estimated for all trajectories using the Nelder–Mead downhill simplex method (fmins in MatLab, MathWorks) with $\sigma$ constant. A parameter for radius was also included to compensate for overestimation of radial distance. Overestimation is typical for the representation of moving targets (Assad & Maunsell, 1995; Freyd & Finke, 1984).

**C. Stereo-first Bayesian model**

Similarly, the stereo-first Bayesian model describes perceived binocular disparity (change) as the product of likelihood and prior:

$$p(d|\beta) \propto p(\beta|d) \times p(d). \quad \text{(C1)}$$

The likelihood for binocular disparity (change) is modeled as a Gaussian distribution centered on the true
disparity (change) $\delta$ measured at the endpoint of stimulus motion. The standard deviation $\sigma_d$ of the distribution is left as a free parameter:

$$p(\beta|d; \sigma_d) = \frac{1}{\sqrt{2\pi\sigma_d^2}} \exp\left[\frac{-(d - \delta)^2}{2\sigma_d^2}\right].$$  \hfill (C2)

The preference or prior for small disparity (change) is modeled as a Gaussian distribution centered on zero disparity:

$$p(d; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(d - \delta)^2}{2\sigma^2}\right].$$  \hfill (C3)

The MAP estimate for disparity is then given by

$$\hat{\delta} = \frac{1/\sigma_d^2 \delta}{1/\sigma_d^2 + 1/\sigma^2} = \frac{\delta}{1 + (\sigma_d/\sigma)^2}. \hfill (C4)$$

Changing disparity information needs to be coupled with spatial location to recover 3-D motion. The cyclopean azimuth $\alpha$ is approximated by $(\alpha_L + \alpha_R)/2$, and disparity constraints are estimated relative to angle $\alpha$:

$$\hat{\alpha}_L = \alpha - \hat{\delta}/2, \quad \hat{\alpha}_R = \alpha + \hat{\delta}/2. \hfill (C5)$$

Predicted trajectories and velocities were calculated similar to the motion-first model: Bias in disparity was introduced according to Equation C4, and biased angles were found using Equation C5. The ratio $\sigma_d/\sigma$ was estimated using a constant $\sigma$ for all trajectory angles. As before, an additional parameter for radius was introduced to compensate for individual overestimation of radial distance.

Note that disparity estimates are not integrated over time. Instead, the endpoint of stimulus motion was used to derive trajectory angle and radius. Consequently, the stereo-first model and its prior can be interpreted as temporal integration of biased disparities or alternatively biased temporal integration of disparities.

D. Stereo-motion Bayesian model

If one estimates cyclopean azimuth $\hat{\alpha}$ by $(\hat{\alpha}_L + \hat{\alpha}_R)/2$ and insert the velocity estimates from Equation B5 into Equation C5, then velocity and disparity input are combined in a single Bayesian model with three parameters:

$$\hat{\alpha}_L = \hat{\alpha} - \hat{\delta}/2, \quad \hat{\alpha}_R = \hat{\alpha} + \hat{\delta}/2. \hfill (D1)$$

Maximum likelihood fits and parameters for this combined Bayesian model are not reported because model fits to each individual data set were almost identical to the stereo-first Bayesian model and because goodness of fit was only marginally better.

Acknowledgments

I would like to thank Pascal Mamassian, Larry Maloney, and two anonymous reviewers for helpful comments, Pamela Hunter and Katherine McArthur for collecting data. Parts of this research were supported by the Royal Society of London (UK) and presented at VSS 2005 and ECVP 2005.

Commercial relationships: none.
Corresponding author: Martin Lages.
Email: m.lages@psy.gla.ac.uk.
Address: 58 Hillhead Street, Glasgow, G12 8QB, UK.

References


