TPM®

TESTING PROBABILISTIC MODELS OF BINOCULAR 3D MOTION PERCEPTION

MARTIN LAGES
UNIVERSITY OF GLASGOW

Geometric constraints for the perception of three-dimensional (3D) binocular motion are discussed in a probabilistic framework. Two alternative Bayesian models of binocular integration are put forward to explain perceptual bias under uncertainty. The models exploit biologically plausible constraints of local motion and disparity processing in a binocular viewing geometry. Results from computer simulations and model selection support the idea that disparity processing rather than motion processing introduces perceptual bias in 3D motion. This suggests that the binocular motion system combines motion and disparity constraints relatively late when establishing a 3D motion percept.

Key words: Visual perception; Motion; Disparity; Bayesian model; Bayesian Information Content.

Correspondence concerning this article should be addressed to Martin Lages, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, UK. Email: Martin.Lages@glasgow.ac.uk

The perceptual inference of the three-dimensional (3D) external world from two-dimensional (2D) retinal input is a fundamental problem (Berkeley, 1709/1975; Gregory, 1980; von Helmholtz, 1910/1962) that the visual system has to solve through neural computation (Pizlo, 2001; Poggio,Torre, & Koch, 1985). This is true for static scenes as well as for dynamic events. For dynamic events the inverse problem implies that the visual system estimates motion in 3D space from local encoding and spatio-temporal processing.

Velocity in 3D space is described by motion direction and speed. Motion direction is conveniently expressed in terms of spherical coordinates or as a vector in a 3D Cartesian coordinate system. Estimating local motion vectors is highly desirable for a visual system because local estimates of direction and speed in a vector field provide the basis for the perception of 3D object motion. This information is essential for segmenting objects from the background, for identifying and interpreting objects, as well as for planning and executing actions in a dynamic environment.

If a single moving point, corner, or other unique feature serves as binocular input then intersection of constraint lines or triangulation in a binocular viewing geometry provides a straightforward and unique geometrical solution to the inverse problem. If, however, the moving stimulus has spatial extent, such as an oriented line or contour inside a circular aperture or receptive field then local motion direction of corresponding receptive fields in the left and right eye remains ambiguous, and additional constraints are needed to solve the inverse problem in 3D (Lages & Heron, 2010).

The inverse optics and the aperture problem are well-known problems in computational vision, especially in the context of stereo processing (Mayhew & Longuet-Higgins, 1982; Poggio et al., 1985), structure from motion (Koenderink & van Doorn, 1991), and optic flow (Hildreth, 1984). Gradient constraint and related methods (Johnston, McOwan, & Benton, 1999) are based

on image intensities and belong to the most widely used techniques of optic-flow computation. They can be divided into local area-based (Lucas & Kanade, 1981) and into more global optic flow methods (Horn & Schunck, 1981). Both techniques typically employ brightness constancy and smoothness constraints in the image to estimate velocity in an over-determined equation system. It is important to note that optical flow only provides a constraint in the direction of the image gradient, the normal component of the optical flow. As a consequence some form of regularization or smoothing is needed. Various algorithms have been developed implementing error minimization and regularization for 3D stereo-motion detection (e.g., Bruhn, Weickert, & Schnörr, 2005; Min & Sohn, 2006; Scharr & Küsters, 2002; Spies, Jähne, & Barron, 2002). These algorithms effectively extend processing principles of 2D optical flow to 3D scene flow (Carceroni & Kutulakos, 2002; Vedula, Baker, Rander, Collins, & Kanade, 2005).

However, these computational studies of 3D motion are usually concerned with fast and efficient encoding. Here we are less concerned with the efficiency or robustness of a particular algorithm and implementation. Instead we want to understand local and binocular constraints in order to explain characteristics of human 3D motion perception such as perceptual bias under uncertainty.

The entire perceptual process may be understood as a form of statistical inference (Knill, Kersten, & Yuille, 1996) and motion perception has been modeled as an inferential process for 2D object motion (Weiss & Fleet, 2001; Weiss, Simoncelli, & Adelson, 2002) and 3D surfaces (Ji & Fermüller, 2006). Models of binocular 3D motion perception however are typically deterministic and predict only azimuth angle or change in depth (Regan & Gray, 2009). In the following we discuss probabilistic models of 3D motion perception that are based on velocity constraints and can explain perceptual bias under uncertainty as well as motion perception under ambiguity.

The purpose of this paper is to promote Bayesian models of binocular 3D motion perception because they provide local estimates of 3D velocity under uncertainty. We will focus on local constraints of motion perception using basic features such as single dots moving in 3D space because these features are regarded as primitives of local encoding in the early stages of visual processing.

For the sake of simplicity we exclude eye, head and body movements of the observer and consider only passively observed, local motion. Smooth motion pursuit of the eyes and self-motion of the observer during object motion throw up related questions (Harris, 2006; Miles, 1998; Rushton & Warren, 2005) but are beyond the scope of this paper.

BACKGROUND

Like many other predators in the animal kingdom humans have two eyes that are set a short distance apart so that an extensive region of the world is seen simultaneously by both eyes from slightly different points of view. Vision in this region of binocular overlap has a special quality that has intrigued artists, philosophers, and scientists.

Under natural viewing conditions the human visual system seems to effortlessly establish a 3D motion percept from local inputs to the left and right eye. The instantaneous integration of binocular input is essential for object recognition, navigation, action planning and execution. It appears obvious that many depth cues help to establish 3D motion perception under natural viewing conditions but local motion and disparity input features prominently in the early processing stages of the visual system (Howard & Rogers, 2002).

Any biologically plausible solution to the inverse 3D motion problem needs to rely on binocular sampling of local spatio-temporal information (Beverley & Regan, 1973, 1974, 1975). There are at least three known cell types in primary visual cortex V1 that may be involved in local encoding of 3D motion: simple and complex motion detecting cells (DeAngelis, Ohzawa, & Freeman, 1993; Hubel & Wiesel, 1962, 1968; Maunsell & van Essen, 1983), binocular disparity detecting cells (Hubel & Wiesel, 1970; Ohzawa, DeAngelis, & Freeman, 1990), and joint motion and disparity detecting cells (Anzai, Ohzawa, & Freeman, 2001; Bradley, Qian, & Andersen, 1995; DeAngelis & Newsome, 1999).

Therefore, it may be not surprising that three approaches to binocular 3D motion perception emerged in the literature: (i) interocular velocity difference (IOVD) is based on monocular motion detectors, (ii) changing disparity over time (CDOT) monitors output of binocular disparity detectors, and (iii) joint encoding of motion and disparity (JEMD) relies on binocular motion detectors also tuned to disparity.

(i) The *motion-first* model postulates monocular motion processing followed by stereo processing (Lu & Sperling, 1995; Regan & Beverley, 1973; Regan, Beverly, Cynader, & Lennie, 1979). In this model monocular motion is independently detected in the left and right eye before IOVD, typically along the horizontal, establishes motion in depth.

(ii) The stereo-first model assumes disparity encoding followed by binocular motion processing (Cumming & Parker, 1994; Peng & Shi, 2010). This model first extracts binocular disparities and then computes CDOT. Note that tracking of spatial position is also required to recover a 3D motion trajectory.

(iii) Finally, the stereo-motion model suggests JEMD or binocular disparity and interocular delay (Carney, Paradiso, & Freeman, 1989; Morgan & Fahle, 2000; Qian, 1994; Qian & Andersen, 1997). In neurophysiological studies it was shown that a number of binocular complex cells in cats (Anzai et al., 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, 2005b).

These three approaches have generated an impressive body of results but psychophysical experiments have been inconclusive and the nature of 3D motion processing remains an unresolved issue (Harris, Nefs, & Grafton, 2008; Regan & Gray, 2009).

In psychophysical studies vision researchers have tried to isolate motion and disparity input by creating different motion stimuli. These stimuli are rendered in stereoscopic view and typically consist of many random dots in so-called random dot kinematograms (RDKs) that give rise to the perception of a moving surface, defined by motion, disparity or both. However, psychophysical evidence based on detection and discrimination thresholds using these stimuli has been inconclusive, supporting interocular velocity difference (Brooks, 2002; Fernandez & Farell, 2005; Portfors-Yeomans & Regan, 1996; Rokers, Cormack, & Huk, 2008; Shioiri, Saisho, & Yaguchi, 2000), changing disparity (Cumming & Parker, 1994; Tyler, 1971) or both (Brooks & Stone, 2004; Lages, Mamassian, & Graf, 2003; Rokers, Cormack, & Huk, 2009) as possible inputs to 3D motion perception.

Another limitation of random-dot stimuli is that random dots moving in depth may invoke intermediate and higher processing stages similar to structure from motion and global object motion. A surface defined by dots can invoke mid-level surface and even high-level object processing and therefore may not reflect characteristics of local motion encoding alone. Although the

involvement of higher-level processing has always been an issue in psychophysical studies it is of particular concern when researchers are relating behavioral or neuroscientific measures of surface and object motion to characteristics of early motion processing.

In addition, detection, and discrimination thresholds for RDKs often do not reveal non-veridical or biased 3D motion perception. Accuracy rather than precision of observers' perceptual performance needs to be measured to investigate and establish characteristics of motion and disparity processing in psychophysical studies (Harris & Dean, 2003; Rushton & Duke, 2007; Welchman, Tuck & Harris, 2004). Using the method of adjustment observers reported the perceived trajectory angle of a previously seen stimulus moving in depth. The results indicated overestimation of trajectory angle for a range of trajectories approaching the observer. To explain this perceptual bias it was suggested that observers exploit the cyclopean average by using the angle relative to the direction straight ahead along the line of sight (Harris & Drga, 2005; Harris & Rushton, 2003). Although appealing in its simplicity this strategy only constrains perceived motion direction in *x-z* but does not solve the inverse problem of binocular 3D motion.

PERCEPTUAL BIAS UNDER UNCERTAINTY

The binocular viewing geometry imposes obvious constraints for stimulus trajectory and velocity. For a moving dot for example the intersection of constraint lines in *x-z* space determines trajectory angle and speed of the target moving in depth as illustrated in Figure 1. If we assume that the eyes remain accommodated and verged at angle β_0 on a fixation point *F* straight ahead then motion information is projected onto the retina of the left and right eye as illustrated in Figure 1. The projection angles onto the retinae depend on the azimuth β and speed r of the motion stimulus, as well as viewing distance D and interpupillary distance i. The average of the left and right projection angle approximates the visual angle α in cyclopean view through point *C* and the difference defines a binocular horizontal disparity δ. If the projection angles are interpreted as angular velocities then their difference also describes interocular velocity difference.

Motion and disparity constraint lines intersect at the same point in space, regardless of whether they are based on the computation of angular velocities or binocular disparities. Although binocular motion and changing disparity input share the same geometry and are mathematically equivalent, different neural encoding and processing of these inputs may be subject to noise resulting in characteristic perceptual bias. If, however, noise or uncertainty is introduced together with a motion or disparity prior then the intersection of constraints lines generates different predictions. We will exploit these characteristics when modeling data from a psychophysical experiment.

Baynesian Model of Notion Perception

Some promising Bayesian models have been developed in vision (Knill & Richards, 1996). Following Bayes's Rule the likelihoods and priors of a scene *S* and image *I* are combined to produce a non-normalized posterior

$$p(S|I) \propto p(I|S)p(S) \tag{1}$$

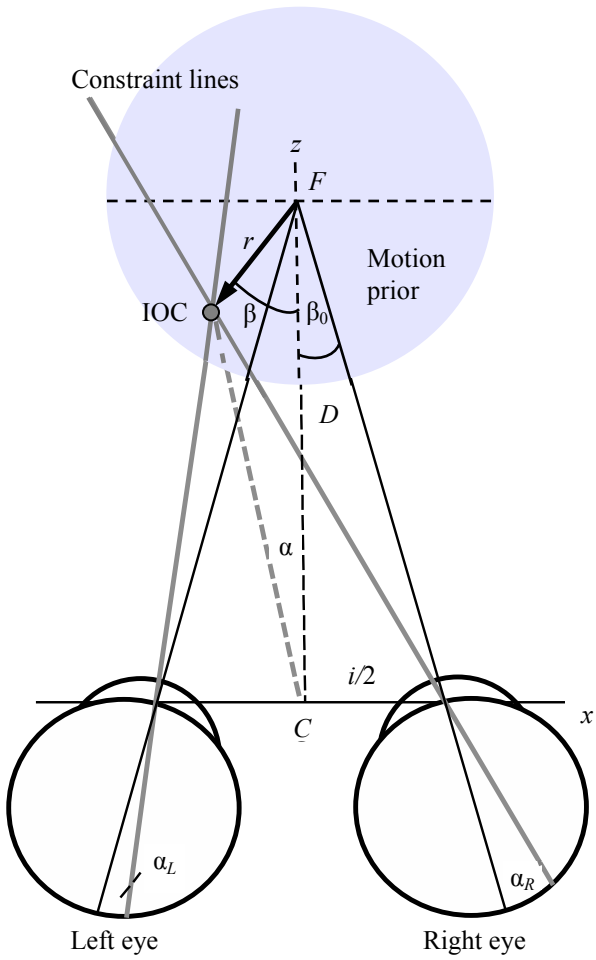

FIGURE 1
Binocular viewing geometry in top view.
If the two eyes are verged on a fixation point at viewing distance D with angle β_0 then projections of a moving target (arrow) with angle α_L in the left eye and α_R in the right eye constrain motion of the target in x-z space. The intersection of constraints (IOC) determines stimulus trajectory β and radius r. The Gaussian motion prior is indicated by the gray disc around the fixation point which also features as the start point of motion.

Various quantities given in the images, such as motion and disparity, can be used to infer aspects of a scene. Weiss et al. (2002), for example, 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 moving in 2D space. With this elegant approach they could explain a range of 2D motion illusions.

Most objects in natural scenes are stationary. If we assume that objects tend to move slowly on an arbitrary trajectory in x-z space then a bivariate Gaussian probability distribution centered on the starting point of a stimulus provides a plausible prior for 3D motion perception in x-z space. Symmetric perspective projections of this world prior into the left and right eye give rise to marginal Gaussian distributions defining motion priors centered on zero velocity. Similarly, the difference of the marginal distributions in the left and right eye defines a prior for dis-

parity (change) centered on zero disparity. Thus, the same 3D motion prior in the world results in a Gaussian velocity and disparity prior on the retinae.

There are several potential sources for uncertainty and noise in binocular motion processing. For example, local moving targets in a sparse 3D environment offer limited motion and disparity input and other depth cues thereby introducing different degrees of uncertainty in the observer. Mini-saccades during fixation, or early noise in the encoding system are possible sources of uncertainty (Hogervorst & Eagle, 1998). In the following we extend the motion-first (IOVD) and stereo-first (CDOT) processing models to probabilistic models by adding Gaussian noise to the input and postulating a plausible prior for each processing scheme.

Bayesian Motion-First Model (BIOVD)

First assume that noise is present in the activation of monocular motion detectors optimally tuned to velocities in the left and right eye. 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 eye are conveniently expressed as Gaussian distributions with equal variance centered on the true angular velocity of the stimulus in each eye. Each likelihood distribution is then combined with the motion prior. Motion priors favoring slow motion have been suggested in the context of 2D motion (Ascher & Grzywacz, 2000; Ullman & Yuille, 1989; Weiss et al., 2002).

In this framework perceived angular velocity of motion-first processing may be described as a product of likelihood and prior for the left and right eye

$$p(v_L|\beta) \propto p(\beta|v_L)p(v_L); \quad p(v_R|\beta) \propto p(\beta|v_R)p(v_R) \tag{2}$$

using the same prior $p(v_L) = p(v_R) = p(v)$.

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

$$p(\beta|v_L;\sigma_v) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp\left[\frac{-(v_L-\dot{\alpha}_L)^2}{2\sigma_v^2}\right] \tag{3}$$

The likelihood for the right eye is modeled accordingly. The preference or prior for slow motion is described by a Gaussian distribution with unknown but constant standard deviation σ centered on zero velocity

$$p(v;\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[\frac{-(v)^2}{2\sigma^2}\right] \tag{4}$$

The product of the Gaussian likelihood distribution with such a conjugate prior $N(0,\sigma)$ defines a posterior distribution, that is the probability of each possible angular velocity taking into account both prior and likelihood of the trajectory. Through differentiation the *maximum a posteriori* (MAP) estimates of angular velocity are found for the left eye and right eye, respectively.

$$\hat{\alpha}_L = \frac{1/\sigma_v^2\,\dot{\alpha}_L}{1/\sigma_v^2 + 1/\sigma^2} = \frac{\dot{\alpha}_L}{1+(\sigma_v/\sigma)^2}; \quad \hat{\alpha}_R = \frac{1/\sigma_v^2\,\dot{\alpha}_R}{1/\sigma_v^2 + 1/\sigma^2} = \frac{\dot{\alpha}_R}{1+(\sigma_v/\sigma)^2} \tag{5}$$

The noise ratio between likelihood and prior σ_v/σ is the only free parameter in this model (Hürlimann, Kiper, & Carandini, 2002).

Bayesian Stereo-First Model (BCDOT)

Alternatively, internal noise may be introduced by the activation of binocular disparity detectors tuned to different disparities. The likelihood distribution for disparity (change) is also conveniently expressed as a Gaussian distribution centered on the true disparity (change) of the stimulus. The disparity likelihood is then 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).

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

$$p(d\,|\,\beta) \propto p(\beta\,|\,d)p(d) \tag{6}$$

The likelihood for binocular disparity (change) is modeled as a Gaussian distribution centered on the true disparity δ (or disparity change) measured at the endpoint of stimulus motion. The standard deviation σ_d of the distribution is left as a free parameter

$$p(\beta\,|\,d;\sigma_d) = \frac{1}{\sqrt{2\pi}\sigma_d}\exp\left[\frac{-(d-\delta)^2}{2\sigma_d^2}\right] \tag{7}$$

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

$$p(d;\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[\frac{-(d)^2}{2\sigma^2}\right] \tag{8}$$

The MAP estimate for disparity is given by

$$\hat{\delta} = \frac{1/\sigma_d^2\,\delta}{1/\sigma_d^2 + 1/\sigma^2} = \frac{\delta}{1+(\sigma_d/\sigma)^2} \tag{9}$$

Changing disparity information needs to be coupled with spatial position to recover 3D motion. The cyclopean azimuth $\dot{\alpha}$ is approximated by $(\dot{\alpha}_L + \dot{\alpha}_R)/2$ and disparity constraints are therefore given relative to angle $\dot{\alpha}$

$$\hat{\alpha}_L = \dot{\alpha} - \hat{\delta}/2,\quad \hat{\alpha}_R = \dot{\alpha} + \hat{\delta}/2 \tag{10}$$

Bayesian Stereo-Motion Model (BJEMD)

In the present framework velocity and disparity input can be combined in a Bayesian model with different noise ratios for motion and disparity processing. Uncertainty in velocity and disparity processing are combined and both uncertainty parameters are estimated together.

If we estimate cyclopean azimuth $\hat{\alpha}$ by $(\hat{\alpha}_L + \hat{\alpha}_R)/2$ and insert the velocity estimates from Equation (5) into Equation (10) then velocity and disparity input can be combined in a single Bayesian model

$$\hat{\alpha}_L = \hat{\alpha} - \hat{\delta}/2, \quad \hat{\alpha}_R = \hat{\alpha} + \hat{\delta}/2 \tag{11}$$

with estimates $\left(\hat{\alpha}_L, \hat{\alpha}_R, \hat{\delta}\right)$ based on noise parameters for velocity σ_v and disparity σ_d processing.

Following Bayes' rule, likelihoods and priors are combined to establish a posterior distribution for each model and trajectory. Applying a simple decision rule, such as the *maximum a posteriori* rule, provides a posteriori estimates of angular velocity and disparity.

The estimates describe biased constraint lines and their intersection determines an azimuth angle and radial distance (speed) in *x*-*z*.

Simulation of Perceived Velocity under Uncertainty

In this Bayesian framework uncertainty is modeled by the ratio of standard deviations between a Gaussian likelihood and prior. If uncertainty is negligible model predictions of azimuth angle are veridical but with increasing uncertainty model predictions approximate a shrinking circle for the motion-first and a compressed ellipse for the stereo-first Bayesian model. In Figure 2 predictions over the full range of 360° of stimulus azimuths are plotted for the two Bayesian models.

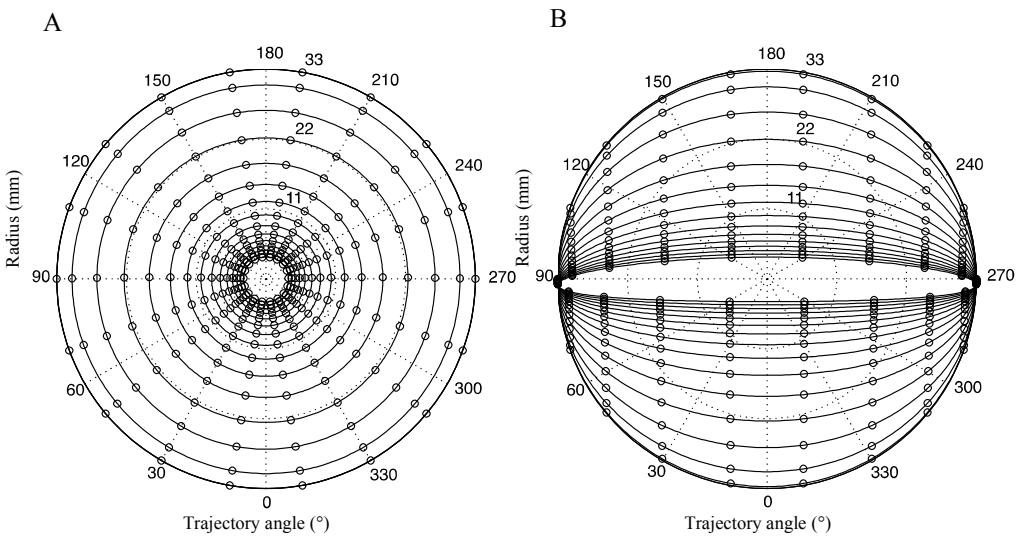

FIGURE 2
Simulation results for (A) motion-first Bayesian IOVD and (B) stereo-first Bayesian CDOT (and Bayesian JEMD).
Plots show model predictions of trajectory angle and velocity in polar co-ordinates for azimuth angles of 10° to 350° in steps of 20° at a viewing distance of 114 cm.
Uncertainty is modelled by the ratio of likelihood and prior ranging from 0.1 to 3.0 in steps of 0.2.

The unbiased prediction of azimuth angles, its circular shape, is the result of multiplying left and right angular velocity by the same factor. The increasingly flat elliptical shape is the consequence of a stronger bias for larger disparities near the *z*-axis or azimuths of 0° and 180° and weaker bias for smaller disparities near the *x*-axis or azimuths of 90° and 270°.

Perception of Horizontal Trajectories under Uncertainty

Under natural viewing conditions there are many monocular cues to 3D motion but in a sparse environment only binocular motion and disparity cues may be available. In a psychophysical experiment Lages (2006) investigated perceived bias of motion trajectories of small target dots and used the Bayesian models of interocular velocity difference (BIOVD) and disparity change (BCDOT) as well as joined encoding (BJEMD) to explain the results. Here we reanalyze the data of Experiment 2 (Figure 3) and perform additional model selection (Table 1).

Stimuli were presented to the left and right eye using a split-screen Wheatstone configuration. Each of our four observers was tested for stereo deficits before the experiment commenced (Heron & Lages, 2012). In the experiment they viewed three anti-aliased dots presented above and below a fixation cross, surrounded by a rectangular fusion lock, at a viewing distance of 114 cm. 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 on the horizontal *x-z* plane for 833 ms (0.02, 0.03, and 0.04 m/s) on 36 different trajectories.

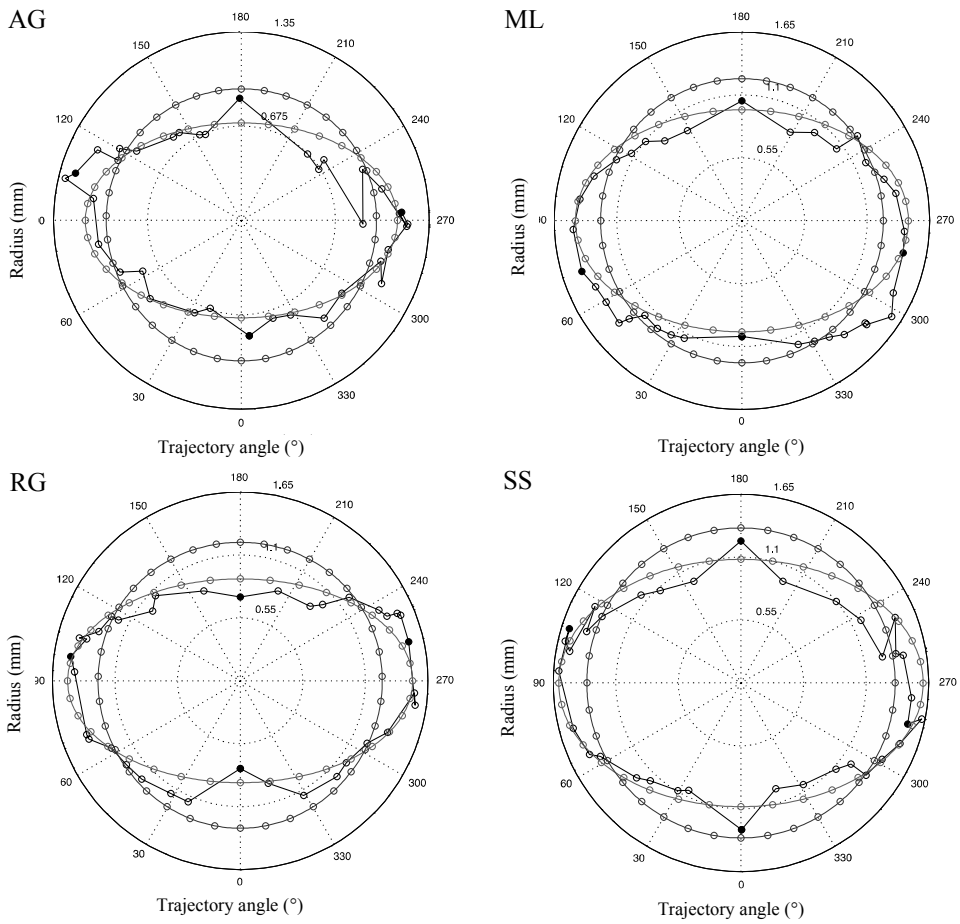

FIGURE 3
Illustration of empirical results from four observers (adapted from Lages 2006, Figure 7).
Polar plots for perceived azimuth and radial distance (speed) averaged across three stimulus velocities and best model fits of the motion-first (circle; BIOVD) and stereo-first Bayesian model (ellipse; BCDOT).
Filled data points correspond to cardinal stimulus trajectories at 0°, 90°, 180°, and 270°.

TABLE 1
Results from four observers and three stimulus speeds (radius in mm) from Lages (2006, Experiment 2)

Observers	Radius	Motion-first model				Stereo-first model			Model selection		
	r	σ_v:σ	r_v	$\chi^2(34)$	σ_d:σ	r_d	$\chi^2(33)$	$-2\log(\Lambda)$	BF_{BIC}	P_{BIC}
AG	16.6	0.00	17.6	2.40	0.32	17.7	2.29	0.04	2.22	0.69
	25.0	0.00	20.6	6.63	0.81	27.7	2.12	0.99	3.57	0.78
	33.3	0.00	25.2	17.20	1.01	32.9	5.79	0.95	3.49	0.78
ML	16.6	0.00	26.3	2.89	0.56	28.2	1.92	0.36	2.60	0.72
	25.0	0.00	30.8	4.07	0.49	34.3	3.09	0.24	2.45	0.71
	33.3	0.00	34.4	9.24	0.77	39.5	3.00	0.98	3.55	0.78
RG	16.6	1.83	113.0	1.40	0.48	28.8	1.32	0.05	2.23	0.69
	25.0	2.84	274.0	3.40	0.87	35.4	0.58	1.54	4.69	0.82
	33.3	2.22	213.0	8.97	1.00	43.2	1.45	1.58	4.80	0.83
SS	16.6	0.00	22.2	2.31	0.00	22.2	2.31	0.00	2.18	0.69
	25.0	0.00	36.1	1.45	0.58	39.3	1.14	0.21	2.42	0.71
	33.3	0.82	64.5	1.74	0.64	44.5	1.35	0.22	2.43	0.71

Note. Parameter estimates and goodness-of-fit are reported for motion-first and stereo-first Bayesian model as well as model selection based on log-likelihood ratio test $-2\log(\Lambda)$ and BIC (Bayesian Information Content) approximation of Bayes factor BF_{BIC} with the corresponding predictive probability.

On each trial the observer verged on the fixation cross before they initiated motion of target dots by key-press. Azimuth angle of the target ranged between 0° and 350° in steps of 10°. Each observer attended a total of eight separate blocks (2 Tasks × 4 Repetitions) each comprising 108 trials (3 Velocities × 36 Trajectories). In each block of trials observers judged either motion azimuth or radial distance. Adjustments to 36 trajectories and three velocities were repeated four times in randomly intermixed trials. The individual data shown in Figure 3 are averaged across stimulus speeds.

In maximum likelihood (ML) fits individual adjustments of azimuth and radial distance were averaged across trials with the same stimulus velocity. Bayesian models with one free parameter (noise ratio between likelihood and prior determines perceived speed) for the motion-first model and two free parameters (perceived speed and noise ratio between likelihood and prior) for the stereo first model were fitted to data of each observer and three stimulus speeds and results are summarized in Figure 3 and in Table 1. The stereo-first Bayesian model gives slightly better fits and parameter values assume more plausible values than the motion-first model.

In the following ML fits and parameters for the stereo-motion Bayesian model (BJEMD) are not reported because results were almost identical to the stereo-first Bayesian model with no significant improvement for any of the individual data sets. As shown in Table 1 individual fits of the stereo-first Bayesian model for each stimulus velocity suggest that an increase in stimulus velocity systematically raised uncertainty in the noise ratio. This trend appears in all observers except for Observer ML in the 16.6 mm condition.

A conventional log-likelihood test ($-2\log(\Lambda)$ in Table 1), approximated by a χ^2-distribution, did not indicate statistically significant differences between BCDOT and BIOVD model fits. We

approximated the Bayes factor (BF) through Bayesian Information Content (BIC). The BIC for model M_i is defined as

$$\mathrm{BIC}(M_i) = -2\log(L_i) + k_i \log(n) \tag{12}$$

where L_i and k_i is the maximum likelihood and number of free parameters of model M_i, respectively, and n is the number of observations (Raftery, 1995).

Assuming that the prior odds for model M_0 and M_i are the same, the Bayes factor is the ratio of the prior predictive probabilities. Prior predictive probabilities can be obtained as $P(D|M_i) = \exp[-\mathrm{BIC}(M_i)/2]$. Comparing two competing models a BIC approximation of the BF is then given by

$$\mathrm{BF}_{01} = \frac{P(D|M_0)}{P(D|M_1)} \propto \exp(\Delta\mathrm{BIC}_{10}/2) \tag{13}$$

where $\Delta\mathrm{BIC}_{10} = \mathrm{BIC}(M_1) - \mathrm{BIC}(M_0)$. Bayes factors of $\mathrm{BF}_{\mathrm{BIC}} > 1.0$ in Table 1 indicate odds in favour of the stereo-first (BCDOT) model when compared to the motion-first (BIOVD) model.

Bayes factors of $\mathrm{BF}_{\mathrm{BIC}} > 1.0$ in Table 1 indicate odds in favour of the stereo-first (BCDOT) model when compared to the motion-first (BIOVD) model. Bayesian Information Content as an approximation of the Bayes factor suggests weak to positive evidence in favor of the stereo-first model (Raftery, 1995). The $\mathrm{BF}_{\mathrm{BIC}}$ and the posterior predictive probabilities $\mathrm{P}_{\mathrm{BIC}}$ indicate that given the empirical data the stereo-first model (BCDOT) is on average three times more probable than the motion-first model (BIOVD).

Rendering 3D motion in a stereoscopic set-up is difficult and can introduce various artifacts and cue conflicts. In this experiment constant size and blur of the target stimuli moving in depth may have influenced perceived depth (Watt, Akeley, Ernst, & Banks, 2005). On the other hand, cue conflicts due to looming and accommodative cues are probably too small to account for the substantial and systematic bias found for small blurred targets that move maximally ±3.3 cm in depth at a viewing distance of 114 cm. Using LEDs moving in depth Harris and Dean (2003), as well as Welchman et al. (2004) reported systematic overestimation of perceived azimuths near the fronto-parallel plane, confirming that perceptual bias also exists for real-world stimuli at various trajectory angles.

Model fits for perceived azimuth angle and speed (radial distance) and Bayesian model selection promotes the idea that bias in 3D 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 et al., 2003). It is possible, however, that interocular velocity difference or optical flow contributes to 3D motion perception, especially when stimuli define surfaces that move on a trajectory near the observer's line of sight (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 (i) temporal integration of biased disparities or (ii) biased temporal integration of disparity. The latter interpretation appears more plausible since uncertainty estimates increased systematically with stimulus velocity as reported in Table 1 and Figure 3.

If 3D motion perception is based on velocity-tuned processing the relatively small change of stimulus velocity in our experiment should have very little effect on uncertainty. Disparity-tuned processing on the other hand may increase uncertainty levels for faster stimuli due to the

temporal limits of disparity integration (Read & Cumming, 2005b; Tyler, 1971) in a transient stereo-system (Edwards & Schor, 1999).

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 local motion detectors signal a wider range of velocities. As a consequence a system that processes motion input first needs to establish correspondence between rather different monocular motions before it can build a percept of 3D object motion. Computationally it appears more parsimonious to solve the stereo correspondence problem before deriving a 3D motion percept. This argument also applies to joint encoding of motion and disparity (JEMD) because early encoding of true 3D motion would require a large number of detectors specifically tuned to all combinations of spatial frequency, orientation, and interocular spatio-temporal offsets to capture all possible local 3D motions.

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

General Discussion

Motion-first (IOVD) and stereo-first (CDOT) are extreme models because they only consider horizontal motion or disparity input, respectively. IOVD excludes contributions from binocular disparity processing but requires early stereo correspondence. It does not solve the inverse problem for local 3D line motion and is confined to horizontal 3D motion in the *x-z* plane.

CDOT on the other hand excludes contributions from motion processing and therefore has problems to establish motion correspondence and direction. Without further assumptions it is confined to motion in depth along the line of sight.

If motion-only or disparity-only input determines 3D motion perception as suggested by IOVD and CDOT then processing of any other input needs to be silenced or disengaged. This would require mutual inhibition or top-down interference rather than integrated motion and disparity processing in the visual system. Instead, we suggest that the visual system takes advantage of motion and disparity input (Bradshaw & Cumming, 1997; Lages & Heron, 2008) as well as additional cues before it combines them in an optimal fashion. This would favor parallel over serial processing and late integration over early joint encoding (Lages, Dolia, & Graf, 2007). The inverse problem of local 3D motion is ill-posed for JEMD processing and requires a population read-out at a later stage to approximate 3D motion. It is unclear how this can be achieved without additional input that disambiguating local motion.

Combining disparity or depth information with velocity constraints at a later stage solves the inverse problem of local 3D motion and provides a flexible scheme that can exploit intermediate depth processing such as relative and orientation disparity in V2 and V4 (Hinkle & Connor, 2002; Thomas, Cumming, & Parker, 2002). Velocity constraints may be processed in the ventral stream and binocular disparity together with other depth cues in the dorsal stream (Ponce, Lomber, & Born, 2008). It seems anatomically and neurophysiologically plausible that integration of motion and disparity occurs relatively late in subregions of human V5/MT (DeAngelis & Newsome, 2004;

Majaj, Carandini, & Movshon, 2007; Orban, 2008; Rokers et al., 2009) if not in areas beyond V5/MT (Ban, Preston, Meeson, & Welchman, 2012; Likova & Tyler, 2007).

What enables the 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 in psychophysical studies (Bradshaw & Cumming, 1997; van Ee & Anderson, 2001) already points at late integration within the visual processing hierarchy (Marr, 1982).

More specifically, we suggest that binocular 3D motion perception is based on parallel streams of motion and disparity processing. Thereby motion processing captures coarse spatio-temporal constraints in the scene whereas disparity processing provides a fine and frequently updated depth map that helps to disambiguate motion direction and to maintain a detailed spatial representation of the scene. Late integration of local motion and disparity constraints in combination with other cues solves the inverse problem of local 3D motion and allows the visual system to remain flexible when binding and segmenting local inputs from different processing stages into a global 3D motion percept. Parallel processing and late integration may explain why, compared to 2D motion perception, 3D motion perception shows reduced spatio-temporal tuning characteristics (Lages et al., 2003; Tyler, 1971) and why motion perception can retain relatively fine spatial detail at slow speeds. The combination of local motion constraints with a more global dynamic depth map from higher-order features may even explain the perception of different types of non-linear motion, such as non-rigid and 2nd order motion.

The notion of parallel pathways feeding functionally different aspects of motion perception into a later stage is not new and has been advanced in the context of 2D speed perception (Braddick, 1974, 1980), 2D pattern motion (Adelson & Movshon, 1982; Weiss et al., 2002; Yo & Wilson, 1992), eye movements (Masson & Castet, 2002; Rashbass & Westheimer, 1961), and the processing of higher order motion (Ledgeway & Smith, 1994; Lu & Sperling, 2001). Surprisingly however, it has not sufficiently been addressed in the context of binocular 3D motion perception (Lu & Sperling, 2001; Regan et al., 1979).

When considering the ill-posed inverse problem and the underdetermined characteristics of local binocular motion constraints, then parallel processing and late integration of motion and disparity as well as other cues appears promising. Solving the inverse problem for local 3D motion adds a functional significant aspect to the notion of parallel and modular streams of dynamic disparity and motion processing. It will require considerable efforts to unravel the entire process but geometric-probabilistic models can achieve motion and disparity integration under uncertainty and ambiguity (Lages, 2013; Wang, Heron, Moreland, & Lages, 2012; Wang & Lages, 2011).

Conclusions

Bayesian models of binocular motion perception for dots or local features moving on a horizontal depth plane (Lages, 2006) can be extended to 3D motion perception of lines or edges, predicting perceived azimuth and elevation under ambiguity (Heron & Lages, 2009; Lages & Heron, 2010). Using generalized Bayesian motion models (Lages, 2013; Wang et al., 2012; Wang & Lages, 2011), noisy velocity constraint planes define velocity likelihoods that, combined with a 3D motion prior, can explain perceptual bias under uncertainty and motion perception un-

der ambiguity. This suggests that the visual system integrates velocity constraints with feature tracking from disparity processing to arrive at velocity estimates of moving features and objects.

How exactly the visual system establishes binocular 3D motion perception from image-based local motion and disparity input remains a difficult and unresolved issue. We hope, however, that the present line of research improves understanding of local constraints in binocular 3D motion perception. The results should be of interest to researchers in psychology, neuroscience, and computer vision and may inform development of new 3D technologies in applied areas.

Our results also suggest that a geometric-statistical approach as exemplified here provides a powerful framework to model binocular 3D motion perception. However, more empirical data are needed to evaluate perceived 3D trajectories and systematic distortions. More specifically, in order to validate the Bayesian approach it would be important to verify the motion prior for slow motion through real-world measurements of scene flow as well as experimental data from discrimination tasks adapted to binocular 3D motion. Note that due to environmental constraints the motion prior may not be isotropic and Gaussian. There is empirical evidence that the motion prior in 2D velocity space has heavier tails than a Gaussian (Stocker & Simoncelli, 2006). Also, if experience shapes the motion prior it may not only reflect slow motion but also horizontal motion along a ground plane and possibly downward motion aligned with the pull of gravity. As a consequence, the motion prior may not be entirely isotropic. On the other hand, any of these effects is likely to depend on multisensory integration and top-down processing, mainly affecting global object rather than local motion perception (Graf, Adams & Lages, 2004; Lages, Jenkins, & Hillis, 2008).

References

Adelson, E. H., & Movshon J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523-525. doi:10.1038/300523a0

Anzai, A., Ohzawa, I., & Freeman, R. D. (2001). Joint encoding of motion and depth by visual cortical neurons: Neural basis of the Pulfrich effect. *Nature Neuroscience*, *4*, 513-518. doi:10.1038/87462

Ascher, D., & Grzywacz, N. N. (2000). A Bayesian model for the measurement of visual velocity. *Vision Research*, *40*, 3427-3434. doi:10.1016/S0042-6989(00)00176-0

Ban, H., Preston, T. J., Meeson, A., & Welchman, A. E. (2012). The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature Neuroscience*, *15*, 636-643. doi:10.1038/nn.3046

Berkeley, G. (1975). *Philosophical works including the works on vision* (M. R. Ayers, Ed.). London, UK: Dent. (Original work published 1709).

Beverley, K. I., & Regan, D. (1973). Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space, *The Journal of Physiology*, *235*, 17-29.

Beverley, K. I., Regan, D. (1974). Temporal integration of disparity information in stereoscopic perception. *Experimental Brain Research*, *19*, 228-232. doi:10.1007/BF00238537

Beverley, K. I., Regan, D. (1975). The relation between discrimination and sensitivity in the perception of motion in depth. *The Journal of Physiology*, *249*, 387-398.

Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research*, *14*, 519-527. doi:10.1016/0042-6989(74)90041-8

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of Royal Society B*, *290*, 137-151. doi:10.1098/rstb.1980.0087

Bradley, D. C., Qian, N., & Andersen, R. A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature*, *373*, 609-611. doi:10.1038/373609a0

Bradshaw, M. F., & Cumming, B. G. (1997). The direction of retinal motion facilitates binocular stereopsis. *Proceedings of the Royal Society of London B*, *264*, 1421-1427. doi:10.1098/rspb.1997.0198

Brooks, K. R. (2002). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision*, *2*, 218-231. doi:10.1167/2.3.2

Brooks, K. R. & Stone, L. S. (2004). Stereomotion speed perception: Contributions from both changing disparity and interocular velocity difference over a range of relative disparities. *Journal of Vision*, *4*, 1061-1079. doi:10.1167/4.12.6

Bruhn, A., Weickert, J., & Schnörr, C. (2005). Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, *61*, 211-231. doi:10.1023/B:VISI.0000045324.43199.43

Carceroni, R. L., & Kutulakos, K. N. (2002). Multi-view scene capture by surfel sampling: From video streams to non-rigid 3D motion, shape and reflectance. *International Journal of Computer Vision*, *49*, 175-214. doi:10.1023/A:1020145606604

Carney, T., Paradiso, M. A., & Freeman, R. D. (1989). A physiological correlate of the Pulfrich effect in cortical neurons of the cat. *Vision Research*, *29*, 155-165. doi:10.1016/0042-6989(89)90121-1

Cumming, B. G., & Parker, A. J. (1994). Binocular mechanisms for detecting motion in depth. *Vision Research*, *34*, 483-495. doi:10.1016/0042-6989(94)90162-7

DeAngelis, G. C., & Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *The Journal of Neuroscience*, *19*, 1398-1415.

DeAngelis, G. C., & Newsome, W. T. (2004). Perceptual "read-out" of conjoined direction and disparity maps in extrastriate area MT. *PLoS Biology*, *2*, e0394. doi:10.1371/journal.pbio.0020077

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *Journal of Neurophysiology*, *69*, 1091-1117.

Edwards, M., & Schor, C. M. (1999). Depth aliasing by the transient stereo-system. *Vision Research*, *39*, 4333-4340. doi:10.1016/S0042-6989(99)00149-2

Fernandez, J. M., & Farell, B. (2005). Seeing motion-in-depth using inter-ocular velocity differences. *Vision Research*, *45*, 2786-2798. doi:10.1016/j.visres.2005.05.021

Graf, E. W., Adams, W. J., & Lages, M. (2004). Prior monocular information can bias motion perception. *Journal of Vision*, *4*, 427-433. doi:10.1167/4.6.2

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of Royal Society B*, *290*, 181-197. doi:10.1098/rstb.1980.0090

Harris, J. M. (2006). The interaction of eye movements and retinal signals during the perception of 3-D motion direction. *Journal of Vision*, *6*, 777-790.

Harris, J. M., & Dean, P. J. A. (2003). Accuracy and precision of binocular 3-D motion perception. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 869-881. doi:10.1037/ 0096-1523.29.5.869

Harris, J. M., & Drga, V. F. (2005). Using visual direction in three-dimensional motion perception. *Nature Neuroscience*, *8*, 229-233. doi:10.1038/nn1389

Harris, J. M., Nefs, H. T., & Grafton, C. E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, *21*, 531-547. doi:10.1163/156856808786451462

Harris, J. M., & Rushton, S. K. (2003). Poor visibility of motion-in-depth is due to early motion averaging. *Vision Research*, *43*, 385-392. doi:10.1016/S0042-6989(02)00570-9

Heron, S., & Lages, M. (2012). Screening and sampling in studies of binocular vision. *Vision Research*, *62*, 228-234. doi:10.1016/j.visres.2012.04.012

Hildreth, E. C. (1984). The computation of the velocity field. *Proceedings of the Royal Society of London B*, *221*, 189-220. doi:10.1098/rspb.1984.0030

Hinkle, D. A., & Connor, C. E. (2002). Three-dimensional orientation tuning in macaque area V4. *Nature Neuroscience*, *5*, 665-670. doi:10.1038/nn875

Hogervorst, M. A., & Eagle, R. A. (1998). Biases in three-dimensional structure-from-motion arise from noise in the early visual system. *Proceedings of the Royal Society: Biological Sciences*, *265*, 1587-1593. doi:10.1098/rspb.1998.0476

Horn, B. K. P., & Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, *17*, 185-203. doi:10.1016/0004-3702(81)90024-2

Howard, I. P., & Rogers, B. J. (2002). *Seeing in depth (Vol. 2). Depth Perception*. Ontario, Canada: I. Porteous.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*, 106-154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*, 215-243.

Hubel, D. H., & Wiesel, T. N. (1970). Stereoscopic vision in macaque monkey. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature*, *225*, 41-42. doi:10.1038/225041a0

Hürlimann, F., Kiper, D. C., & Carandini, M. (2002). Testing the Bayesian model of perceived speed. *Vision Research*, *42*, 2253-2257. doi:10.1016/S0042-6989(02)00119-0

Ji, H., & Fermüller, C. (2006). Noise causes slant underestimation in stereo and motion. *Vision Research*, *46*, 3105-3120. doi:10.1016/j.visres.2006.04.010

Johnston, A., McOwan, P. W., & Benton, C. (1999). Robust velocity computation from a biologically motivated model of motion perception. *Proceedings of the Royal Society of London B*, *266*, 509-518. doi:10.1098/rspb.1999.0666

Knill, D. C., Kersten, D., & Yuille, A. L. (1996). Introduction: A Bayesian formulation of visual perception. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian Inference* (pp. 1-21). Cambridge, UK: Cambridge University Press.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press.

Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, *8*, 377-385. doi:10.1364/JOSAA.8.000377

Lages, M. (2006). Bayesian models of binocular 3-D motion perception. *Journal of Vision*, *6*, 508-522. doi:10.1167/6.4.14

Lages, M. (2013). Straight or curved? From deterministic to probabilistic models of binocular 3D motion perception. *Frontiers in Behavioral Neuroscience*, *7*, 1-3. doi:10.3389/fnbeh.2013.00079

Lages, M, Dolia, A., & Graf, E. W. (2007). Dichoptic motion perception limited to depth of fixation? *Vision Research*, *47*, 244-252. doi:10.1016/j.visres.2006.10.001

Lages, M., & Heron, S. (2008). Motion and disparity processing informs Bayesian 3D motion estimation. *Proceedings of the National Academy of Science of the United States of America*, *105*, E117. doi:10.1073/pnas.0809829105

Lages, M., & Heron, S. (2010). On the inverse problem of local binocular 3D motion perception. *PLoS Computational Biology*, *6*, e1000999. doi:10.1371/journal.pcbi.1000999

Lages, M., Jenkins, R., & Hillis, J. M. (2008). Anticipation of gravity alters perception of average speed [Abstract]. *Perception*, *37*(*Suppl.*), 28.

Lages, M., Mamassian, P., & Graf, E. W. (2003). Spatial and temporal tuning of motion in depth. *Vision Research*, *43*, 2861-2873. doi:10.1016/j.visres.2003.08.006

Ledgeway, T., & Smith, A. T. (1994). Evidence for separate motion-detecting mechanisms for first-order and 2nd-order motion in human vision. *Vision Research*, *34*, 2727-2740. doi:10.1016/0042-6989 (94)90229-1

Likova, L. T., & Tyler, C. W. (2007). Stereomotion processing in the human occipital cortex. *Neuroimage*, *38*, 293-305. doi:10.1016/j.neuroimage.2007.06.039

Lu, Z.-L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*, 2697-2722. doi:10.1016/0042-6989 (95)00025-U

Lu, Z.-L., & Sperling, G. (2001). Three systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A*, *18*, 2331-2370. doi:10.1364/JOSAA.19.002142

Lucas, B. D., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In P. J. Hayes (Ed.), *Proceedings of the 7th International Joint Conference on Artificial Intelligence* (pp. 674-679). Vancouver, BC, Canada: William Kaufmann Inc.

Majaj, N., Carandini, M., & Movshon, J. A. (2007). Motion integration by neurons in macaque MT is local not global. *The Journal of Neuroscience*, *27*, 366-370. doi:10.1523/JNEUROSCI.3183-06.2007

Marr, D. (1982). *Vision*. New York, NY: W. H. Freeman & Co.

Masson, G. S., & Castet, E. (2002). Parallel motion processing for the initiation of short-latency ocular following in humans. *The Journal of Neuroscience*, *22*, 5149-5163.

Maunsell, J. H., & van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey: I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, *49*, 1127-1147.

Mayhew, J. E. W., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature*, *297*, 376-378. doi:10.1038/297376a0

Miles, F. A. (1998). The neural processing of 3-D visual information: Evidence from eye movements. *European Journal of Neuroscience*, *10*, 811-822. doi:10.1046/j.1460-9568.1998.00112.x

Min, D., & Sohn, K. (2006). Edge-preserving simultaneous joint motion-disparity estimation. *Proceedings of the 18th International Conference on Pattern Recognition* (ICPR 2006), *2*, 74-77. doi:10.1109/ICPR.2006.470

Morgan, M. J., & Fahle, M. (2000). Motion-stereo mechanisms sensitive to interocular phase. *Vision Research*, *40*, 1667-1675. doi:10.1016/S0042-6989(00)00016-X

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, *249*, 1037-1041. doi:10.1126/ science.2396096

Orban, G. A. (2008). Higher order visual processing in macaque extrastriate cortex. *Physiological Review*, *88*, 59-89. doi:10.1152/physrev.00008.2007

Pack, C. C., Born, R. T., & Livingstone, M. S. (2003). Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, *37*, 525-535. doi:10.1016/S0896-6273(02)01187-X

Peng, Q., & Shi, B. E. (2010). The changing disparity energy model. *Vision Research*, *50*, 181-192. doi:10.1016/j.visres.2009.11.012

Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision Research*, *41*, 3145-3161. doi:10.1016/ S0042-6989(01)00173-0

Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature*, *317*, 314-319. doi:10.1038/317314a0

Ponce, C. R., Lomber, S. G., & Born, R. T. (2008). Integrating motion and depth via parallel pathways. *Nature Neuroscience*, *11*, 216-223. doi:10.1038/nn2039

Portfors-Yeomans, C. V., & Regan, D. (1996). Cyclopean discrimination thresholds for the direction and speed of motion in depth. *Vision Research*, *36*, 3265-3279. doi:10.1016/0042-6989(96)00065-X

Qian, N. (1994). Computing stereo disparity and motion with known binocular cell properties. *Neural Computation*, *6*, 390-404. doi:10.1162/neco.1994.6.3.390

Qian, N., & Andersen, R. A. (1997). A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Research*, *37*, 1683-1698. doi:10.1016/S0042-6989(96)00164-2

Raftery, A. E. (1995). Baysian model selection in social research. *Sociological Methodology*, *25*, 111-163.

Rashbass, C., & Westheimer, G. (1961). Disjunctive eye movements. *The Journal of Physiology*, *159*, 339-360.

Read, J. C. A. (2002a). A Bayesian approach to the stereo correspondence problem. *Neural Computation*, *14*, 1371-1392. doi:10.1162/089976602753712981

Read, J. C. A. (2002b). A Bayesian model of stereopsis depth and motion direction discrimination. *Biological Cybernetics*, *86*, 117-136. doi:10.1007/s004220100280

Read, J. C. A., & Cumming, B. G. (2005a). Effect of interocular delay on disparity-selective V1 neurons: Relationship to stereoacuity and the Pulfrich effect. *Journal of Neurophysiology*, *94*, 1541-1553. doi:10.1152/jn.01177.2004

Read, J. C. A., & Cumming, B. G. (2005b). The stroboscopic Pulfrich effect is not evidence for the joint encoding of motion and depth. *Journal of Vision*, *5*, 417-434. doi:10.1167/5.5.3

Regan, D., & Beverley, K. I. (1973). Some dynamic features of depth perception, *Vision Research*, *13*, 2369-2379. doi:10.1016/0042-6989(73)90236-8

Regan, D., Beverley, K. I., Cynader, M., & Lennie, P. (1979). Stereoscopic subsystems for position in depth and for motion in depth. *Proceedings of the Royal Society of London B*, *42*, 485-501. doi:10.1098/rspb.1979.0042

Regan, D., & Gray, R. (2009). Binocular processing of motion; some unresolved problems. *Spatial Vision*, *22*, 1-43.

Rokers, B., Cormack, L. K., & Huk, A. C. (2008). Strong percepts of motion through depth without strong percepts of position in depth. *Journal of Vision*, *8*, 1-10. doi:10.1167/8.4.6

Rokers, B., Cormack, L. K., & Huk, A. C. (2009). Disparity- and velocity-based signals for three-dimensional motion perception in human MT+. *Nature Neuroscience*, *12*, 1050-1055. doi:10.1038/nn.2343

Rushton, S. K., & Duke, P. A. (2007). The use of direction and distance information in the perception of approach trajectory. *Vision Research*, *47*, 899-912. doi:10.1016/j.visres.2006.11.019

Rushton, S. K., & Warren, P. A. (2005). Moving observers, relative retinal motion, and the detection of object movement. *Current Biology*, *15*, R542. doi:10.1016/j.cub.2005.07.020

Scharr, H., & Küsters, R. (2002). A linear model for simultaneous estimation of 3D motion and depth. *Proceedings of the IEEE Workshop on Motion and Video Computing* (pp. 220-225). Orlando, FL: IEEE. doi:10.1109/MOTION. 2002.1182240

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, *40*, 2565-2572. doi:10.1016/S0042-6989(00)00130-9

Spies, H., Jähne, B. J., & Barron, J. L. (2002). Range flow estimation. *Computer Vision Image Understanding*, *85*, 209-231. doi:10.1006/cviu.2002.0970

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, *9*, 578-585. doi:10.1038/nn1669

Thomas, O. M., Cumming, B. G., & Parker A. J. (2002). A specialization for relative disparity in V2. *Nature Neuroscience*, *5*, 472-478. doi:10.1038/nn837

Tyler, C. W. (1971). Stereoscopic depth movement: Two eyes less sensitive than one. *Science*, *174*, 958-961. doi:10.1126/science.174.4012.958

Ullman, S., & Yuille, A. (1989). Rigidity and smoothness of motion. In S. Ullman & W. Richards (Eds.), *Image understanding* (pp.163-183). Norwood, NJ: Ablex Publishing Corporation.

van Ee, R., & Anderson, B. L. (2001). Motion direction, speed and orientation in binocular matching. *Nature*, *410*, 690-694. doi:10.1038/35070569

Vedula, S., Baker, S., Rander, P., Collins, R., & Kanade, T. (2005). Three-dimensional scene-flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*, 475-480. doi:10.1109/TPAMI. 2005.63

von Helmholtz, H. (1962). *Treatise on physiological optics*. In J. P. Southall (Ed. & Trans.), *Helmholtz's Treatise on physiological optics* (Vol. 1, pp. 312-313). New York, NY: Dover. (Original work published 1910).

Wang, H., Heron, S., Moreland, J., & Lages, M. (2012). A Bayesian approach to the aperture problem of 3D motion perception. *Proceedings of the International Conference on 3D Imaging (IC3D)* (pp. 1-8). Liege, Belgium. doi:10.1109/IC3D.2012.6615116

Wang, H., & Lages, M. (2011, May). *A biologically-inspired Bayesian model of 3D velocity estimation*. Paper presented at the AVA Meeting, Cardiff, UK.

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision*, *5*, 834-862. doi:10.1167/5.10.7

Weiss, Y., & Fleet, D. J. (2001). Velocity likelihoods in biological and machine vision. In R. P. N. Rao, B. Olshausen, & M. S. Lewicki (Eds.), *Probabilistic models of the brain: Perception and neural function* (pp. 81-100). Cambridge, MA: MIT Press.

Weiss, Y., Simoncelli, E. P. & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*, 598-604. doi:10.1038/nn858

Welchman, A. E., Lam, J. M., & Bülthoff, H. H. (2008). Bayesian motion estimation accounts for a surprising bias in 3D vision. *Proceedings of the National Academy of Science of the United States of America*, *105*, 12087-12092.

Welchman, A. E., Tuck, V. L., & Harris, J. M. (2004). Human observers are biased in judging the angular approach of a projectile. *Vision Research*, *44*, 2027-2042. doi:10.1016/j.visres.2004.03.014

Yo, C., & Wilson, H. R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research*, *32*, 135-147.