

Depth estimation from retinal disparity requires eye and head orientation signals

Gunnar Blohm

Centre for Vision Research, York University,
Toronto, Ontario,
Canadian Institutes of Health Research, Canada,
Université Catholique de Louvain,
Louvain-la-Neuve, Belgium, &
Centre for Neuroscience Studies, Queen's University,
Kingston, Ontario, Canada



Aarlenne Z. Khan

Centre for Vision Research, York University,
Toronto, Ontario, Canada,
Group for Action and Perception,
Canadian Institutes of Health Research, Canada, &
The Smith-Kettlewell Eye Research Institute,
San Francisco, CA, USA



Lei Ren

Centre for Vision Research and Departments of Psychology,
Biology, and Kinesiology and Health Sciences,
York University, Toronto, Ontario, Canada, &
Group for Action and Perception,
Canadian Institutes of Health Research, Canada



Kai M. Schreiber

Center for Molecular and Behavioral Neuroscience,
Rutgers University Newark, Newark, NJ, USA



J. Douglas Crawford

Centre for Vision Research and Departments of Psychology,
Biology, and Kinesiology and Health Sciences,
York University, Toronto, Ontario, Canada, &
Group for Action and Perception,
Canadian Institutes of Health Research, Canada



To reach for an object, one needs to know its egocentric distance (absolute depth). It remains an unresolved issue which signals are required by the brain to calculate this absolute depth information. We devised a geometric model of binocular 3D eye orientation and investigated the signals necessary to uniquely determine the depth of a non-foveated object accounting for naturalistic variations of eye and head orientations. Our model shows that, in the presence of noisy internal estimates of the ocular vergence angle, horizontal and vertical retinal disparities alone are insufficient to calculate the unique depth of a point-like target. Instead the brain must account for the 3D orientations of the eye and head. We tested the model in a behavioral experiment that involved reaches to targets in depth. Our analysis showed that a target with the same retinal disparity produced different estimates of reach depth that varied consistently with different eye and head orientations. The experimental results showed that subjects accurately account for this extraretinal information when they reach. In summary, when estimating the distance of point-like targets, all available signals about the object's location as well as body configuration are combined to provide accurate information about the object's distance.

Keywords: binocular vision, computational modeling, depth, eye movements, spatial vision

Citation: Blohm, G., Khan, A. Z., Ren, L., Schreiber, K. M., & Crawford, J. D. (2008). Depth estimation from retinal disparity requires eye and head orientation signals. *Journal of Vision*, 8(16):3, 1–23, <http://journalofvision.org/8/16/3/>, doi:10.1167/8.16.3.

Introduction

To accurately plan a goal-directed action, such as a reach, the visual system first reconstructs the 3D location of the goal relative to gaze; or put another way, a 3D retinocentric representation relative to the fovea (Batista, Buneo, Snyder, & Andersen, 1999; Battaglia-Mayer, Caminiti, Lacquaniti, & Zago, 2003; Crawford, Medendorp, & Marotta, 2004; Snyder, 2000). This representation must account for all 3 dimensions of space; horizontal and vertical directions relative to the fovea, as well as the distance of the target from the eye. Some believe that the direction component of binocular images are merged into a cyclopean representation (Ding & Sperling, 2006; Khokhotva, Ono, & Mapp, 2005; Ono, Mapp, & Howard, 2002). This cyclopean representation of direction can be viewed as the mean of two vectors emanating from corresponding points relative to the fovea on the left and right retinas.

The third dimension, depth (that is, the egocentric distance of the object from the cyclopean eye), has to be extracted from retinal disparity: the difference between the images obtained from the right and left eyes. However, it is still controversial how the brain computes depth, in particular how it uses retinal disparity information to decode absolute target distance (DeAngelis, Cumming, & Newsome, 1998; Palanca & DeAngelis, 2003; Tsutsui, Taira, & Sakata, 2005; Uka & DeAngelis, 2002). Theoretical studies have suggested that horizontal and vertical disparities are sufficient to compute depth (without the need for any additional signals; Bishop, 1989; Mayhew & Longuet-Higgins, 1982) for at least five distinct targets (Horn, 1990). The brain also relies on the ocular vergence angle (Collewijn & Erkelens, 1990; Mon-Williams, Tresilian, & Roberts, 2000; Richard & Miller, 1969; Ritter, 1977; Viguier, Clement, & Trotter, 2001) and horizontal version (Backus, Banks, van Ee, & Crowell, 1999; Gonzalez & Perez, 1998; Mueller, 1826; Vieth, 1818). Additional potential cues for depth perception are retinal blur (Mather, 1997; O'Shea, Govan, & Sekuler, 1997), ocular accommodation (Mon-Williams & Tresilian, 1999, 2000), and object features (shading, texture, perspective, etc.; Gonzalez & Perez, 1998; Johnston, 1991; Johnston, Cumming, & Parker, 1993; O'Shea, Blackburn, & Ono, 1994).

A limitation of most previous investigations is that they focused on the depth problem for targets that were foveated and/or that were located in the cyclopean straight-ahead line. Although we often interact with foveally viewed objects, we are also capable of reaching toward targets that are viewed in the retinal periphery (Blohm & Crawford, 2007; Henriques & Crawford, 2000; Van Pelt & Medendorp, 2008). Similarly, it has been shown the oculomotor system can program accurate versional and vergence movements toward peripherally glimpsed targets (Collewijn, Erkelens, & Steinman, 1997;

Qing & Kapoula, 2004). However, it is still an open question how absolute depth is calculated for non-foveated targets in the visual periphery.

In this paper, we test the hypothesis that depth estimation is influenced by the 3D eye-in-head orientation, which in turn is affected by head orientation with respect to gravity. A recent study examined the role of eye orientation in depth perception (Erkelens & van Ee, 1998). However, for visually guided action it is also necessary to consider the influence of *combined* eye and head movements, because head orientation influences 3D eye-in-head orientation (Crawford & Vilis, 1991) and thus the geometry of retinal projection (Misslisch, Tweed, & Hess, 2001). With the head upright, 3D eye rotations are behaviorally constrained to two dimensions, confining the eye rotation axis to a plane in space known as Listing's plane (Haslwanter, 1995; Hepp, 1990; Tweed, 1997a). Vergence causes the Listing's planes of the two eyes rotate outward like saloon doors as a function of vergence angle (Mok, Ro, Cadera, Crawford, & Vilis, 1992; Van Rijn & Van den Berg, 1993). The additional influence of head orientation relative to gravity comes—for static head orientations—from the static vestibulo-ocular reflex (sVOR; Bockisch & Haslwanter, 2001; Haslwanter, Straumann, Hess, & Henn, 1992). The sVOR is responsible for the ocular counter-roll during head roll (head rotations around the anterior–posterior axis) and also causes Listing's plane to tilt forward or backward as a function of head pitch angle. Therefore, 3D eye orientation for a specific cyclopean gaze direction depends not only on vergence but also on head pitch and roll angles. In particular, modulations of Listing's plane change the torsional state of both eyes and thus alter the location onto which a visual stimulus is projected (Schreiber, Crawford, Fetter, & Tweed, 2001). It is well established that these various modulations of binocular eye position alter retinal disparity and binocular correspondence (Schreiber, Tweed, & Schor, 2006; Tweed, 1997b) but it is not presently clear to what degree these various states are accounted for in calculating absolute depth.

Here, we show that the visual system has to account for not just retinal disparity and vergence but also the complete 3D geometry of the (cyclopean) eye and head in order to uniquely compute the absolute distance of a single point-like object from the eyes. Using a theoretical model, we show that a point-like target placed at different depths can produce the *same* binocular retinal position (2D retinal position and horizontal and vertical retinal disparities) when paired with certain combinations of 3D eye and head orientations. We further show that when the vergence angle is noisy or inaccurate (Brenner & van Damme, 1998; Collewijn & Erkelens, 1990; Foley, 1980; Harwerth, Smith, & Siderov, 1995; Viguier et al., 2001) the visual system cannot solve for depth without accounting for 3D eye and head orientations. We then validate this prediction experimentally by means of a reaching

task. By corollary, we show that in real-world situations the visual system has the capacity to use extraretinal copies of 3D eye and head orientations to decode the depth of a target from binocular retinal signals.

Methods

Theoretical model

Using a geometrical model, we attempted to determine which retinal and/or extraretinal signals are needed to decode the egocentric distance of a target (relative to the location of the cyclopean eye) in the peripheral visual field viewed under different eye-in-head and head relative to gravity orientations. To answer this question, we designed our model to solve the inverse problem, with the goal of finding different 3D eye–head-vergence orientations for which the binocular retinal target projection rays (constant binocular retinal input) had an intersection. If these intersections can be found, then the brain must be using extraretinal signals about 3D eye and head rotations to uniquely decode target distance from a given retinal input. Briefly, we fixed the retinal stimulation points for both eyes, projected the retinal target positions out into viewing space (Figure 1—black solid lines from target onto retina), and searched for different eye and head orientations for which these projection rays intersect; we systematically changed head pitch and roll angles for each 3D eye-in-head fixation position (specifying horizontal and vertical versions as well as horizontal vergence). This is described in detail in Appendix A.

To do so, it was necessary to compute the geometry of the retinal target projection lines (Appendix A). Therefore, we calculated the 3D orientation of both eyes (Haslwanter, 1995; Hepp, 1990; Mok et al., 1992; Tweed,

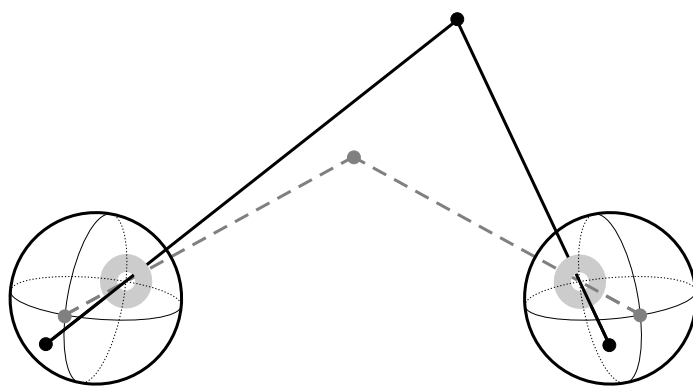


Figure 1. Projection geometry. The right and left eyes' fixation lines (gray dotted lines) and the target projection lines (black solid) are shown. Due to Listing's law and the static VOR, the torsional state of the eyes changes for different eye and head orientations, modifying the retinal location of the target projections.

1997a; Van Rijn & Van den Berg, 1993) as a function of fixation position in 3D space and accounting for changes to this 3D eye-in-head orientation due to head roll and pitch angles through the static VOR (Bockisch & Haslwanter, 2001; Haslwanter et al., 1992; see Appendix B). We then chose a cyclopean retinal position (we supposed—without any restrictions on the generality of the results—that the cyclopean eye is located in the center of the interocular axis) and a given retinal disparity (Fick coordinates) and projected those target rays out into 3D space.

To simplify mathematical expressions, we performed all computations in a cyclopean-eye-centered, head-fixed reference frame. This reference frame is attached to the head and thus all results are relative to the head orientation. We also assumed that the fixation target and the reach target are single point-like objects without any physical extent.

Most previous studies investigating depth estimation have done so while subjects fixated on the target and with the head and eye in straight-ahead orientation. These two configurations are not only unreflective of realistic conditions but they limit the solutions possible, thereby providing a false impression of simplicity. Because we included targets that were not viewed foveally and we considered all possible eye and head configurations, this increased the complexity of the problem. Figure 2A shows a typical pattern of retinal disparities resulting from peripherally viewing iso-distant targets separated by 10° horizontally and vertically and placed at a radial distance of 50 cm from the cyclopean eye while the eye and head were directed straight-ahead. The bars show the retinal disparity (horizontal/vertical disparity is represented by the horizontal/vertical components of the bars) associated with each target (dots) in cyclopean angular retinal coordinates (expressed in Fick coordinates; for visibility, the length of the bars is doubled). The four panels in Figure 2 show how the retinal disparity pattern changes for different eye and head orientations (horizontal: H; vertical: V) as compared to primary orientation (panel A, PP), i.e., straight-ahead gaze and head upright. Importantly, we also show how the retinal disparity pattern changes with head movements, e.g., when using Donder's strategy (see Appendix B), specifying that for a given gaze angle, the head always has the same unique 3D orientation.

Those changes of the retinal projection pattern across eye–head orientations are mainly due to ocular torsion. A torsional eye movement is a rotation of the eye around the line of gaze. As lined out in Appendix B, we use the quaternion description of 3D eye position, which leads to a definition of torsion as being each individual eye's angular-vector component in depth. Cyclovergence is then the virtual cyclopean eye's torsion, whereas cyclovergence is the difference between right and left eye torsions. The amount of ocular torsion varies differently for each eye with different eye and head orientations (due to Listing's law and sVOR), and therefore, depending on the

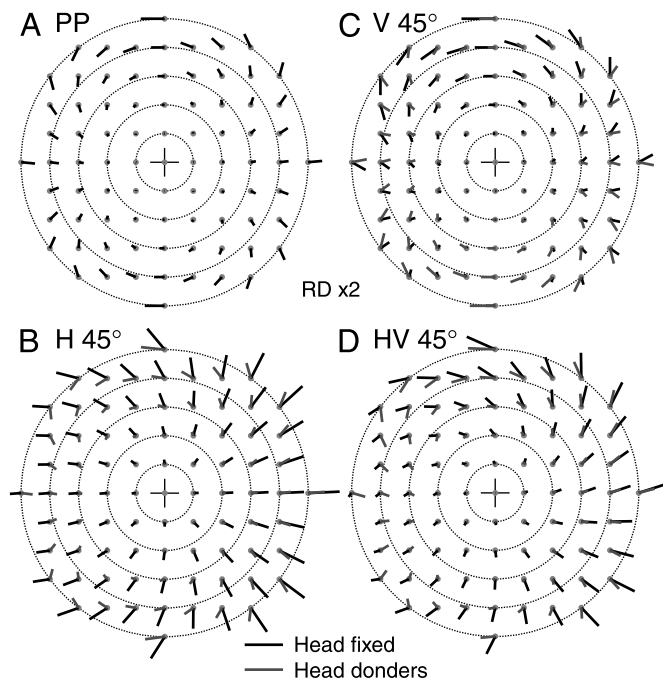


Figure 2. Retinal disparities for different eye and head orientations. Gray dots correspond to different cyclopean-eye-fixed targets (in 10° horizontal and vertical intervals arranged on a hemisphere at 50-cm distance) and the bars attached to them correspond to the disparity of the right and left eye's retinal images. The bars show the direction and amplitude (length) of the retinal disparity associated with the cyclopean retinal target positions to which the bars are attached. Target and fixation distance from the cyclopean eye was always 50 cm. Dotted circles are 10° intervals of retinal eccentricity. The central cross indicates the fixation position and fovea. (A) The target projection pattern and associated retinal disparity was computed for eye and head in primary position (PP), i.e., straight-ahead gaze and upright head orientation. Different eye-only gaze orientations (head fixed, black disparity lines) and combined eye-head gaze orientations (Donder's law, gray lines) influence the retinal disparity pattern. This is shown in panel B for a 45° horizontal gaze shift. Since for primary position (panel A) both retinal disparities are identical, only the black disparity lines are visible. Panel C shows a 45° vertical gaze shift and panel D illustrates the effect of a 45° oblique gaze shift on the retinal disparity pattern (representation in Fick coordinates).

eye-head configuration, the same peripheral target will project onto different parts of the retina for the two eyes.

The range of behaviorally plausible values of right and left eye torsions as described by our model is shown in Figure 3 (gray shaded area) and the typical influence of different eye and head orientation changes on these torsional combinations is also shown as individual lines. The ranges were very large; for example, for a 0 degree left eye torsional orientation, the torsional range of the right eye was about 16 degrees (8 deg clockwise to 8 deg counterclockwise), which is in accordance with previous

findings (Goonetilleke, Mezey, Burgess, & Curthoys, 2008). In this example, different combinations of horizontal and vertical versions, vergence, and head pitch and roll angles used in our model made it possible to change the right eye torsion angle over its whole physiological range while the left eye torsion angle remained constant.

Knowing that the static VOR modifies Listing's law only for head roll and pitch relative to gravity (not for yaw; Bockisch & Haslwanter, 2001; Haslwanter et al., 1992), we varied only these two parameters when looking for a possible solution, i.e., when retinal target rays intersected. For each fixation target (i.e., each 3D gaze convergence point in head-fixed, eye-centered coordinates) that had a possible solution we calculated the relative (with respect to fixation distance) and absolute target distance from the cyclopean eye (see Appendix B).

The particular mathematical procedure of ray intersections that we used in our model does not necessarily reflect the way the brain reconstructs target depth. This is because in many situations, there is no such intersection. Despite this, the visual system interprets the retinal input and attributes a depth estimate to it. This is the case in the induced effect (Liu, Berends, & Schor, 2005; Ogle, 1938) or when viewing stereograms from an incorrect viewing

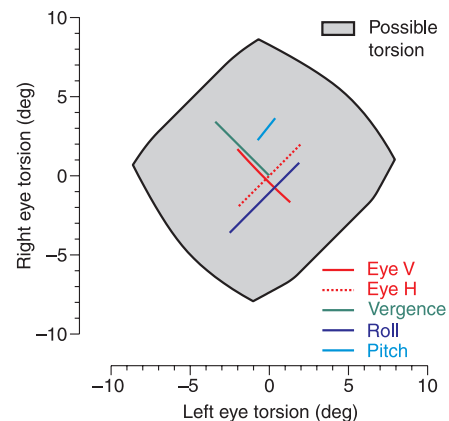


Figure 3. Relationship between the torsional orientations of the right and left eyes in the model. The gray area delimits all possible combinations of right and left eye torsions for different eye and head orientations and different vergence angles used in the model. The different colored lines show how horizontal (dotted red) and vertical (solid red) eye movements, vergence (solid green), and head roll (solid dark blue) and pitch (solid light blue) in isolation affect the combination of right and left eye torsions. For example, if only vertical vergence changes and all other variables remain constant, the right and left eye torsional values evolve along the solid red line. Note that horizontal version only changes torsion for non-zero vergence (cyclovergence) because of the saloon-door-like rotation of Listing's plane with vergence (Mok et al., 1992). Therefore, different combinations of values for all variables allow reaching all possible left and right eye torsional combinations described by the gray area. These combinations are non-linear interactions; effects do not add up linearly.

position (Girshick & Banks, 2005). The fact that even in normal stereovision not all (and actually only very specific) retinal target projection rays intersect has led to a theory of depth perception that might use the point of shortest distance between the retinal target rays. The ensemble of these points is called the empirical or extended horopter (Schreiber et al., 2006).

Experimental procedures and data analysis

We tested the model in a behavioral task (Figure 4) designed to test whether identical binocular retinal position resulted in different estimations of target distance with different eye and head orientations. We created identical retinal stimuli across different eye and head orientations by inducing a visual afterimage at a reference position and then asking subjects to reorient their eyes and head before making a depth judgment about the retinal afterimage object. The use of a visual afterimage allowed us to specifically test the inverse problem, that is, the afterimage was fixed on both retinas, which provided us with retinal target projection rays leading to the exact same retinal input regardless of movement of the eye or head. If information about the 3D eye and head orientations is indeed used to compute target depth, the subjects should point to different depths depending on the eye and head angles. The specific eye and head orientations used in the depth estimation experiment were based on predictions from our model that used each individual subject's eye movement parameters (Listing's law and

sVOR). Reaching or pointing tasks have previously been used to address absolute depth estimation (Blohm & Crawford, 2007; e.g., Carey, Dijkerman, & Milner, 1998; Knill, 2005; Mon-Williams & Tresilian, 1999; Viguier et al., 2001) and have been reported to be more accurate than verbal or other judgments (Viguier et al., 2001).

Subjects

We recruited 5 subjects (3 male, 4 naive), aged between 21 and 30 years with no known visual, oculomotor, or other neurological disorders. Subjects provided written informed consent for their participation in this study, pre-approved by the York University Human Participants Review Subcommittee.

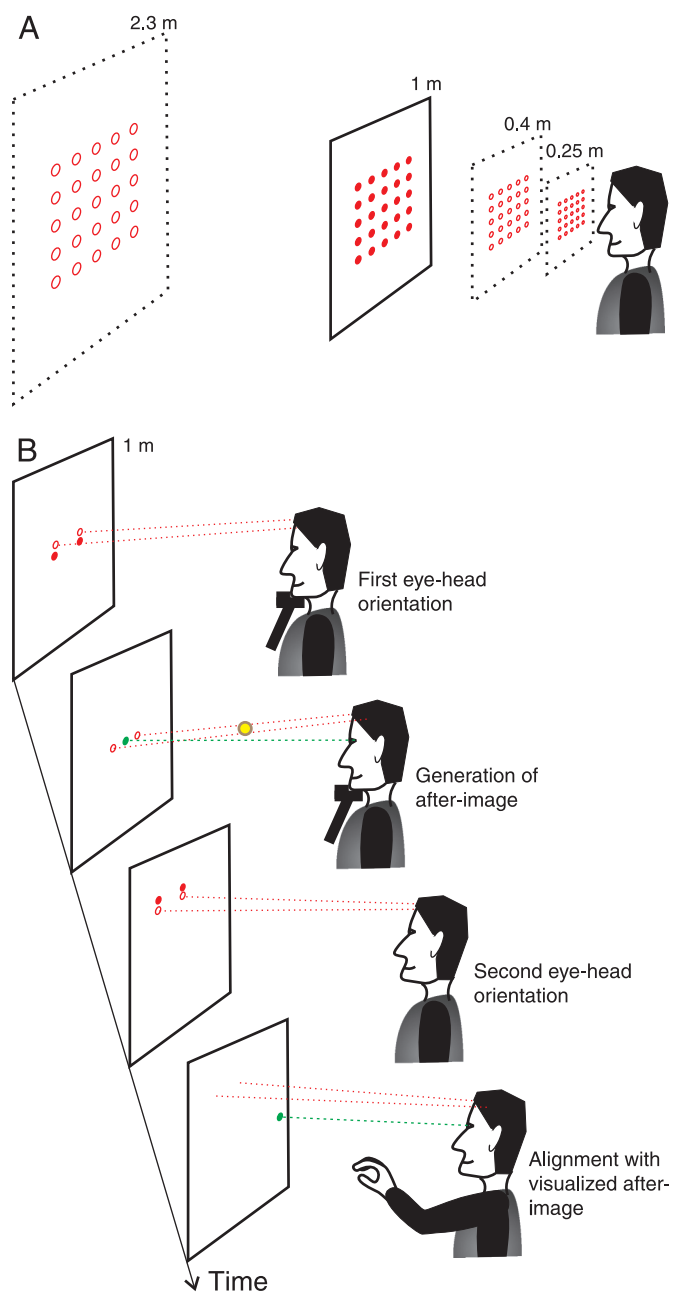


Figure 4. Experimental set-up and paradigm. (A) Set-up used in the first recording session to identify each subject's Listing's law and sVOR parameters. Targets were presented on screens at 4 different distances. (B) Paradigm used in the second recording session to test the model predictions. First, subjects were asked to bring their eyes and head into primary position by using a chin rest and aligning the head orientation feedback markers with the required head orientation placeholders (1st panel; red dotted lines represent the head alignment). Then subjects fixated (2nd panel; green dotted lines represent gaze alignment) 1-m distant straight-ahead target while an ultra-bright target (yellow filled circle) was flashed to generate a retinal afterimage (at 5° obliquely upright at 50-cm distance). Next (3rd panel), subjects lifted their head off the chin rest and oriented their head into another orientation, where the model predicted that the retinal target lines should intersect in space (with a different estimation of depth). Finally (4th panel), subjects fixated a new fixation position while maintaining the head in the required orientation (head orientation targets were switched off). During this time, they visualized the afterimage and aligned their fingertip with the perceived afterimage object in 3D space. All procedures were performed in complete darkness and at the time of the fingertip alignment only the fixation spot was visible (never the hand). The approximately 1-m fixation distance ensured that subjects never accidentally hit the screen.

Apparatus

Visual targets projected onto the screen were red laser spots that were oriented by means of mirror galvanometers (GSI Lumonics, Billerica, LA) controlled by an onboard real-time microprocessor. 3D orientation of both eyes was recoded using combination search coils (Skalar Medical BV, Delft, The Netherlands). Subjects sat in complete darkness in the center of a custom-built 2-m coil frame in a chair with their head restrained by a bite bar, which could be rotated independently around the roll and pitch axes. Head and hand positions and orientations were recorded using an Optotrak motion analysis system (200 Hz; NDI, Waterloo, Ontario, Canada) and we also measured cyclopean eye position by placing a marker on the upper nose, in between both eyes. Cyclopean eye and head position and orientation signals were used to provide online feedback at 50 Hz about current head orientation in the second part of the experiment (see below).

Procedures and data analysis: Experimental session 1

The experiment was divided into 2 sessions performed on 2 consecutive days. In the first session we evaluated each individual subject's Listing's law and sVOR parameters. These values were then used in our theoretical model to compute the experimental test set for each subject for the depth pointing experiment in the second session.

In the first recording session (Figure 4A), subjects were asked to fixate a series of targets that were projected onto a fronto-parallel screen, while the head was constrained at different roll and pitch angles. We used 4 different screen distances, 2.3 m, 1 m, 0.4 m, and 0.25 m. For each head orientation, there were 25 fixation spots that were arranged in a square and presented in a random order for 1 s each. For the 2.3-m and 1-m screen distances, the targets were 12.5° horizontally and vertically apart and thus spread over a ±25° square. For the two near distances, the targets were 15° apart and thus spread over a ±30° square. The center of the square was always aligned with vertical head orientation (relative to the head's median plane), i.e., it moved up and down for different head pitch angles. We first presented targets at the most distant screen where we changed head roll angle from 30° counterclockwise (CCW) to 30° clockwise (CW) in 15° steps (pitch was held constant at 0°). Then we consecutively presented targets on the 3 nearer screens. For each of the nearer screen distances, we changed head pitch orientation from 30° up to 30° down in 15° steps (and roll was kept constant at 0°). All 25 fixation targets were presented for each head orientation condition and subjects were instructed to fixate the red laser target at the different positions on the screen. During this experiment, we recorded 3D eye orientation of both eyes as well as head orientation. To ensure that the coils did not slip during the head orientation procedures, we recorded throughout all head rotations and reviewed the signals offline.

We used the 3D eye orientation recordings to extract each eye's ocular torsion for each horizontal/vertical version, head roll, head pitch, and vergence angle. Next, we used a simple non-linear least-squares fit (Gauss–Newton) to identify each subject's Listing's law and static VOR parameters from the 3D eye and head orientation measurements. This was done by using the following relationship for ocular torsion (T):

$$T = \alpha_0 + c_{OCR} \cdot \beta_R + \sin(c_P \cdot \alpha_P) \cdot E_V + \sin(\delta \cdot v) \cdot E_H. \quad (1)$$

In Equation 1, α_0 is the tilt angle of Listing's plane for upright head orientations, c_P is the gain for the gravity modulation of this tilt related to the pitch angle α_P , c_{OCR} is the gain for the static ocular counter-roll of the head roll angle β_R , δ is the gain for the rotation of Listing's plane due to vergence v , and E_H and E_V were horizontal and vertical versions, respectively. The results of this analysis are summarized in Table 1.

After obtaining each subject's parameters, we used these parameters to perform simulations to find eye and head orientations for which the retinal target projection rays had an intersection point in 3D space. This was done in three steps:

1. we calculated subjects' theoretical binocular projections of a 50-cm distant target that was presented 5 deg up and to the right (on the 45 deg oblique axis) with respect to straight-ahead fixation at 1-m fixation distance, as this would be the experimental condition in the following depth estimation experiment,
2. we searched for possible solutions as described in the Methods section, and
3. we chose a subset of solutions that yielded different depth estimations to perform the experiment.

When doing so, we took care that for each subject there was no consistent correlation between depth and vergence angle. For each subject, we chose 3 solutions, i.e., 3 sets of eye–head configurations that would lead to different depth estimations.

Procedures and data analysis: Experimental session 2

In the second experimental session (Figure 4B), we used the simulated eye–head configurations to perform the

Subject	OCR gain c_{OCR}	Pitch gain c_P	Vergence gain δ	Pitch offset α_0	IOD
GB	0.0815	0.1407	0.5518	0.8312°	6.4 cm
GS	0.1566	0.0741	0.3309	0.1911°	6.3 cm
JC	0.0270	0.0307	0.2917	0.2536°	6.1 cm
KR	0.0362	0.1150	0.2622	1.2554°	6.5 cm
LO	0.1482	0.0380	0.2269	4.9700°	6.7 cm

Table 1. Identified Listing's law and static VOR parameters for each subject. IOD = interocular distance.

actual test determining whether different eye and head orientations lead to different estimates of target depth. (It is believed that the identified Listing's law and sVOR gains are approximately constant over short time scales (Schor, Maxwell, McCandless, & Graf, 2002) and therefore performing the second session on a consecutive day would not invalidate our simulation results.) Subjects sat in complete darkness during which we presented a series of experimental blocks. Each block was composed of the 3 selected eye–head configurations, presented in random order. Targets were projected onto a 1-m distant tangential screen. Each trial began with an initial head orientation period that lasted 5 s (Figure 4B, 1st panel). Subjects used a chin rest to ensure identical head location and were required to orient the head into a straight-ahead orientation. The combination of chin rest and head orientation procedure ensured that the head was in the exact same position and orientation at the beginning of each trial. For the head orientation, we presented 4 laser points on the screen: two indicated the desired head orientation (the center of mass indicated the pitch and yaw orientation and the tilt of the two points indicated desired roll) and two points indicated current head position and orientation in real time (see below). Subjects were instructed to match the four dots by moving their head.

After the initial head orientation period, a central fixation spot was visible for 5.5 s and subjects were instructed to maintain fixation on it (Figure 4B, 2nd panel). 2 s after the onset of the fixation spot, we flashed an ultra-bright LED for 3 s (10 Hz, 60 kcd, ~1 W) that was placed at 5 deg up and to the right of the straight-ahead gaze orientation and at a distance of 50 cm. This LED was used to induce a retinal afterimage. After the central fixation spot disappeared, there was another head orientation period (similar to the first one) lasting for 5 s (Figure 4B, 3rd panel). Once subjects had reoriented their head, they were instructed not to move their head until the end of the trial. Then the 4 head orientation targets were extinguished and subjects were required to refixate a new fixation target on the screen for another 20 s (Figure 4B, 4th panel). The set of eye and head orientations used here that would produce ray intersections given the points of stimulation on the two retinas was found from the experimental knowledge of the observer's binocular eye movements (recording session 1). While fixating the new fixation target, subjects were asked to visualize the afterimage of the flash and to align their fingertip in darkness with the perceived spatial location of the virtual object created by the afterimage. During this period of time, the only light in the experimental room was the dim laser fixation spot. Subjects did not see their hand and performed the alignment task in otherwise complete darkness. The ability to accurately align finger position in depth in complete darkness has previously been reported (e.g., Blohm & Crawford, 2007; Carey et al., 1998; Knill, 2005; Mon-Williams & Tresilian, 1999; Viguier et al., 2001). The fact that the tangential screen used to project the visual stimuli

was at 1-m distance and that the afterimage was created by a ultra-bright LED located at 50 cm from the subject ensured perceived virtual afterimage object locations around this 50-cm LED depth and prevented subjects from accidentally hitting the screen. Subjects typically needed 1–2 blocks of three trials to get used to the experiment and to be able to visualize the afterimage of the LED. We then recorded another 5 blocks of test trials.

For the head orientation periods, it was necessary to provide subjects with online head orientation feedback. This was obtained by placing 3 infrared markers (IREDs, Optotrak, NDI, Waterloo, Ontario, Canada) onto the subjects' head and sample their 3D spatial position at 50 Hz. From these 3 IREds, we computed the quaternion of head orientation. A fourth IRED was placed in between the eyes and provided real-time cyclopean eye location in space. Eye translation in space and head rotation were then used to adjust the position of the displayed targets on the screen in real time in order to ensure the correct eccentricity. At the same time, the finger, eye, and head IRED positions in space were recorded at 200 Hz and stored for offline analysis.

We verified that head orientation was accurate, marked the final finger positions, and calculated the difference vector between the cyclopean eye IRED position and the finger position to compute observed (=reported) target depth as the length of this difference vector.

Results

We will first describe model results reproducing previous findings when all variables in the brain are exactly known and then analyze the model for the case of noisy estimates of ocular vergence. We will then describe the test of the model's predictions in a behavioral experiment.

Our model describes the exact geometry of 3D binocular eye orientation for different head and vergence angles, which allows us to analyze the projection of a target in space onto both retinas. This projection geometry of the target onto the back of the eyes changes with eye and head orientations (Figure 2) due to different combinations of right and left eye torsions (Figure 3). The fact that the head contribution to a gaze orientation influences the retinal projection suggests that binocular retinal stimulation alone might not be sufficient to find a unique solution, as has been suggested previously (Bishop, 1989; Mayhew & Longuet-Higgins, 1982), i.e., if a solution exists, not only one but multiple depths might be decoded.

Ideal model geometry for depth estimation

In order to find different eye–head–vergence geometries leading to intersection of the retinal target projection rays in space in our model, we systematically changed

horizontal and vertical versions as well as the vergence angle of the eyes while keeping the binocular retinal target positions constant. Each such combination of horizontal/vertical version and vergence corresponds to a single particular fixation position (the gaze convergence point) in eye-centered, head-fixed 3D space. We then systematically varied head roll (keeping head pitch constant at 0°) for each fixation position in order to search for a head roll angle that made the two retinal target projection rays

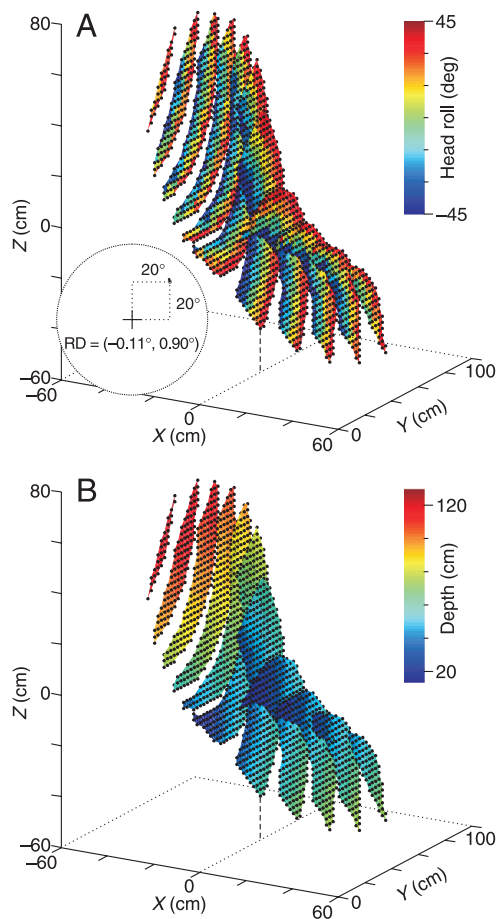


Figure 5. Fixation positions that lead to the intersection of the retinal target rays in space when only head pitch was kept constant (head roll changed). Black dots show different cyclopean eye-centered, head-fixed gaze fixation positions (X : lateral, Y : forward, Z : vertical axes relative to the head), i.e., fixation positions are plotted relative to the head, for which the target lines intersect. Note that all these possibilities arise from exactly the same binocular retinal stimulation. To enhance visibility, we show slices through the 3D volume of possible fixation positions. The slices are separated by 10 cm on the X -axis. (A) The color code of the slices indicates the head roll angle that allowed the target lines to intersect in space. The inset specifies the retinal target position and disparity values used for this simulation in the same representation as Figure 2. (B) Same plot as in panel A but now the distance of the intersection points in space is color coded.

intersect in space. Figure 5 shows the result of these computations. Each black dot corresponds to one particular fixation position in eye-centered, head-fixed space (X : lateral; Y : forward; Z : vertical axes relative to the head) for which a solution exists, i.e., a head roll angle could be found that made the retinal target projection rays intersect (black dots are identical for panels A and B).

We found an entire 3D volume of fixation positions that provided such a solution. If the eyes are fixating one of these black dots in Figure 5, then the two retinal target projection rays intersect for a given head roll angle. (For means of visibility, we only show slices of the complete 3D volume of fixation positions providing a solution.) The presence of such a large volume of solutions already indicates that binocular retinal information alone (horizontal and vertical retinal positions and disparities) is not mathematically sufficient to obtain a unique estimate of target depth, as this has been previously suggested (Mayhew & Longuet-Higgins, 1982). Indeed, this volume showed that the same retinal disparity pattern could be produced by targets at different distances in 3D head-centered space. So the question is how can the brain disambiguate these different depths? We will show that both the relative depth and the absolute depth from the cyclopean eye are not uniquely determined by the binocular retinal input and depend on 3D eye and head orientations. Again, remind that we only consider a single point-like target here where no whole-field disparity pattern can be computed. This is quite different from natural viewing conditions of complex visual stimuli where the visual system can scale and correct disparities from the disparity pattern alone (Backus et al., 1999; Rogers & Bradshaw, 1995).

From the model, we calculated head roll angle and the absolute target distance for each solution. This is represented by the color code in Figure 5. In Figure 5, a color corresponding to the head roll angle (panel A) and absolute target distance (panel B) is associated with each fixation position (black dot) for which the retinal target projection rays intersected. The range of absolute target distances was quite large (roughly between 20 cm and 120 cm). This means, for example, that targets more than 1 m apart in radial distance can produce identical binocular retinal stimulations for different fixation positions. Given this ambiguity of retinal stimulation, we asked which extraretinal signals are geometrically necessary to obtain a correct estimate for a particular viewing situation.

In Figure 6A we plotted relative target depth (the difference between absolute target depth as shown by the color code in Figure 5B and fixation distance) as a function of fixation distance. Each gray dot in Figure 6A corresponds to one fixation position (black dot) in Figure 5. From our model, the binocular retinal input was not sufficient to uniquely estimate target depth of a point target. If binocular retinal stimulation had been sufficient, we would have expected a unique relationship between relative (and absolute) distance and fixation distance, which was not the case (Figure 6A).

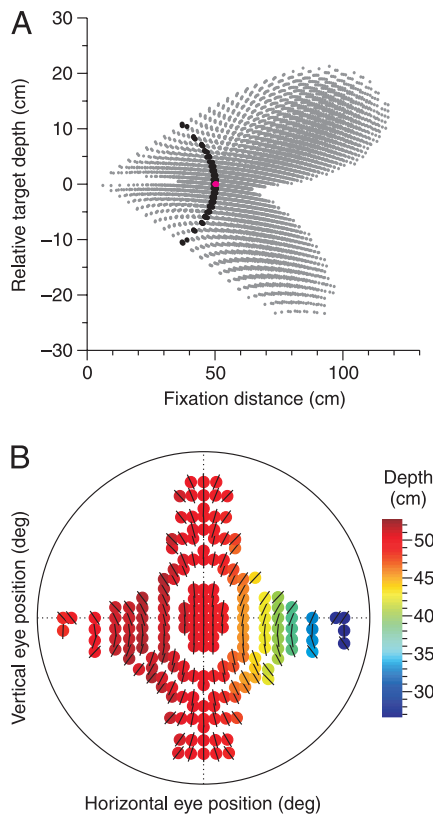


Figure 6. Information needed to estimate depth from binocular retinal stimulation for constant pitch. Panel A shows the non-unique relationship between relative target depth and the fixation distance. Gray dots correspond to all the possible solutions from the volume depicted in Figure 5, i.e., gray dots here correspond to all black dots in Figure 5. Black dots here represent a subset of solutions with constant vergence angle (vergence = 6.9° , corresponding to straight-ahead fixation at 50 cm). The magenta dot shows a further subset of solutions where horizontal vergence was zero (0°). (B) Solutions for the same retinal stimulation and vergence as a function of horizontal and vertical vergences (back circle: $\pm 50^\circ$ range). Color codes the target depth and the orientation of the small black lines indicates the associated head roll orientation.

Since it is well known that the ocular vergence angle influences depth perception, we analyzed the effect of providing this information. The black dots in Figure 6A show all the solutions that had a constant vergence angle of 6.9° (corresponding to fixating a straight-ahead 50-cm distant position). Although reducing the range of possible depths, there was still a 20-cm possible range of target depths for the particular vergence angle in our example (black dots in Figure 6A). It was only when adding horizontal version (magenta dot corresponding to 0° horizontal version) that depth could be uniquely estimated.

To show the influence of eye orientation on absolute depth estimation from binocular vision when the vergence

angle was known, we expanded the constant-vergence results (black dots in Figure 6A) and plotted them as a function of horizontal and vertical versions in Figure 6B. Each colored dot thus represents one isovergence cyclopean eye fixation position with a solution. The color associated with each dot indicates the absolute distance of the intersection point for this particular eye orientation, which had a range of approximately 25 cm. The horizontal color gradient indicated that knowledge of horizontal version was crucial to obtain an accurate estimation of absolute target depth. It is worth mentioning that absolute target depth is not symmetric (with respect to straight-ahead) for right and left eye orientations. Because the cyclopean retinal target position we used for these simulations (see insert in Figure 5A) was 20° horizontal (to the right) and 20° vertical (up), the maximal absolute target depth was obtained when the eyes were directed horizontally 20° to the left (due to the Vieth–Mueller circle).

We have only analyzed solutions related to changes of the head roll angle. Changing head pitch angle provided qualitatively similar results. This is shown in Appendix C.

Estimation of depth with uncertain vergence

We have shown that absolute depth of a peripherally viewed target could be accurately estimated if the brain had access to the true vergence and version via extraretinal signals. However, this may be an unrealistic assumption. Therefore, we will now analyze the consequence of uncertainty of the extraretinal ocular vergence signal on the estimation of target depth.

Evidence for a noisy internal representation of vergence comes from absolute distance judgment studies (Brenner & van Damme, 1998; Collewijn & Erkelens, 1990; Foley, 1980; Harwerth et al., 1995; Rogers & Bradshaw, 1995; Viguier et al., 2001). When subjects were required to judge the absolute distance of a previously seen, foveated point-like target in darkness (by cursor alignment with the memorized depth), they typically underestimated depth (Viguier et al., 2001). The main result of this experiment is reproduced in Figure 7 where the perceived depth was plotted as a function of the actual target depth. Since targets were foveated, depth estimation had to rely on the vergence angle alone. Therefore, the variability in depth estimation (error bars, *SD*) provides insight into vergence variability. For example, for the 80-cm fixation distance, the variability of depth estimation was 30 cm, which corresponds to more than 2° of vergence change.

If the vergence angle was not correct, one expects errors in depth estimation. This is mainly due to the fact that we need to know fixation distance to obtain absolute target distance. Therefore, the question arises whether other extraretinal signals could be used to compensate for this and allow a correct estimation of depth despite uncertain vergence information. We assumed that a similar amount

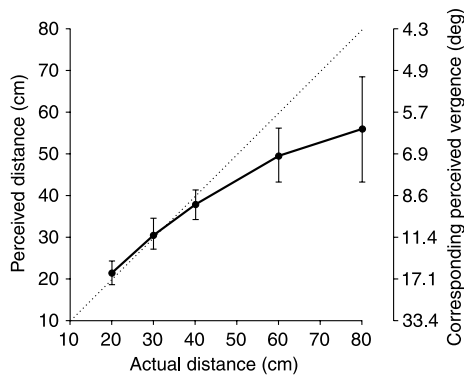


Figure 7. The internal representation of the vergence angle is noisy. The data show the perceived distance of a foveated, point-like target in complete darkness as a function of the real object distance. Subjects had to align a cursor in depth with a previously viewed target. Since the target was foveated, subjects could only rely on the ocular vergence signal to estimate depth. The large variability (*SD*) in the data provides evidence for a noisy internal representation of vergence. Data adapted from Viguier et al. (2001).

of noise in extraretinal eye and head orientation signals would have less of an effect on depth estimation (Figure 6). Table 2 tests this assumption within our model simulations. To do so, we calculated the range of depths obtained when allowing horizontal and vertical versions, as well as head pitch and roll angles to be within a 10 deg interval under certain conditions. Consistent with previous findings (Erkelens & van Ee, 1998), this analysis shows that even relatively large uncertainties about eye version and head orientation only have a very limited influence on depth estimation, i.e., the effect was 1–2 orders of magnitude smaller than for noise in the vergence signal (see below).

We used our model to investigate the role of extraretinal eye and head orientation signals for depth estimation with uncertain vergence in a similar way as in the previous section. We systematically varied horizontal and vertical versions, vergence (fixation depth), and—in this case—head roll angle and tried to find a head pitch orientation for which for a given constant binocular retinal stimulation the two retinal target projection rays intersect in space to provide a solution. We did this for a subset of eye orientations, shifting in 15° intervals between -45° and 45° both horizontally and vertically.

The results of these computations are shown in Figure 8. Gray dots in Figure 8A show all absolute depths of the intersections of the retinal target projection rays (that is, when there was a solution for the inverse problem as a function of fixation distance). This plot was similar to Figure 6A, but we now plot absolute instead of relative depth. The subset of solutions shown as black dots (Figure 8A) shows all possibilities when vergence was uncertain between 3° and 5° (Tresilian, Mon-Williams, & Kelly, 1999; Viguier et al., 2001), an uncertainty magni-

tude previously used (Erkelens & van Ee, 1998). Uncertain vergence resulted in a large range of possible absolute target depths. The dotted rectangular area including the black dots for which vergence was confined between 3° and 5° (Figure 8A) was magnified in Figure 8B (now light gray dots).

To investigate if additional signals could reduce the number of possible solutions, we gradually added other extraretinal information. First, we added horizontal version, which reduced the number of possibilities to the dark gray dots in Figure 8B (horizontal version = -30°), but still left large variations in estimated absolute depth. Adding vertical version (15° , black dots, Figure 8B) further reduced the number of possible solutions for the depth estimate. This was the same when adding head pitch angle (0° , cyan dots, Figure 8B). To obtain a unique estimate of depth, however, we also needed to add the head roll angle (5° , magenta dot, Figure 8B). Using head roll in this final step allowed us to uniquely infer target depth. This was because, all other variables being specified, there now was a unique relationship between head roll, vergence, and target depth (Figure 8C). Therefore, in the case where all other variables are held constant the brain could theoretically rely on head roll in order to estimate target depth, because head roll and vergence were uniquely related.

We quantified the amplitude of depth uncertainty as a function of the amplitude of vergence uncertainty. Figure 9 shows the results from this analysis for different vergence uncertainties centered on a vergence angle of 5° both for the relative (panel A) and absolute depths (panel B). When only retinal information was available (dashed line, Figures 9A and 9B), depth estimation was very poor and decreased in accuracy with increasing vergence uncertainty. As expected, the increase of absolute depth uncertainty (Figure 9B) was much steeper than the increase of relative depth uncertainty (Figure 9A) because absolute depth includes vergence (specifying fixation distance).

Specifying horizontal version and head roll (solid lines, Figures 9A and 9B) reduced the error range for estimated

Signal	Uncertainty	Depth range (cm)	Conditions
Horizontal version	10 deg	2.50	Vergence = 5 deg Vertical version = cst.
Vertical version	10 deg	0.18	Vergence = 5 deg Horizontal version = cst.
Head roll	10 deg	0.23	Horizontal/vertical version = cst.
Head pitch	10 deg	1.48	Horizontal/vertical version = cst.

Table 2. Influence of uncertainty in eye and head orientation signals on distance estimation.

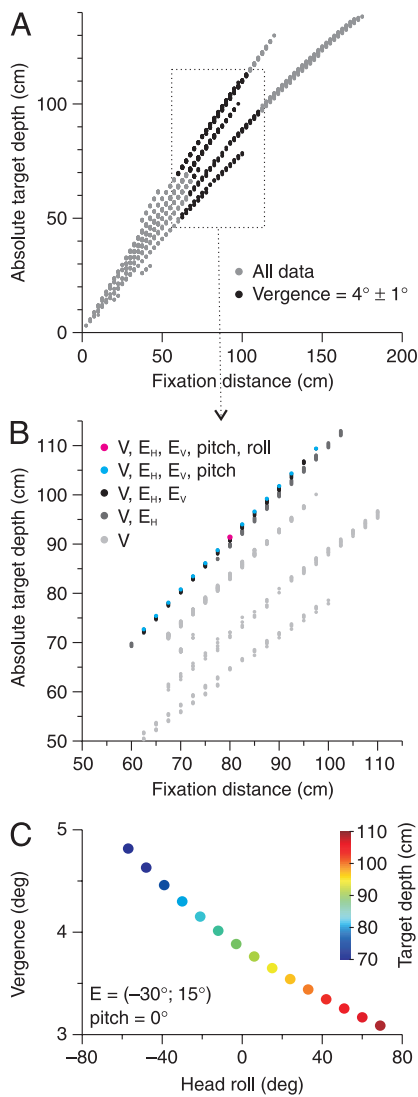


Figure 8. Signals needed for depth estimation with noisy vergence. (A) Absolute target depth is represented as a function of fixation distance for simulations where both head pitch and roll varied (gray dots). Black dots represent a subset for which vergence was chosen to lie between 3° and 5°. The dotted box indicates the magnification in (B). (B) Light gray dots are the same as the black dots in panel A. When adding horizontal vergence, the number of possibilities of the retinal target rays intersection points in space is reduced (dark gray dots). Adding vertical vergence (black dots), head pitch (cyan dots), and head roll (magenta dots) information gradually reduces the solution to uniqueness. (C) Explicit link between head orientation and vergence. Multiple solutions still arise when horizontal and vertical vergences as well as head pitch was held constant in addition to using constant binocular retinal stimuli (same as for all other simulations). The relationship between vergence, head roll, and target depth is represented. In this case, knowing head roll allows a unique estimation of depth if the vergence angle is not well known.

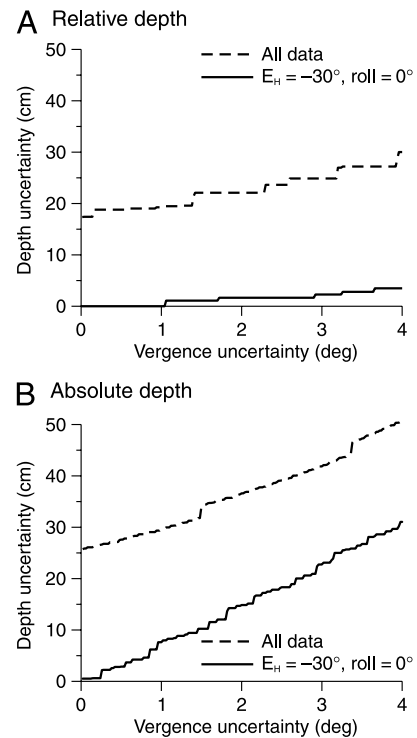


Figure 9. Depth uncertainty changes with vergence uncertainty. Example of a range of depth error observed in simulation results as a function of the vergence range (here, centered on 6.9°) for all data (dashed line) and for a subset of data with fixed horizontal vergence and fixed head roll angle (solid line). The remaining uncertainty has to be resolved using head pitch and vertical vergence. (A) Relative depth uncertainty. (B) Absolute depth uncertainty.

depths, but this error increased with increasing uncertainty of the vergence angle. This was because vertical version as well as head pitch could still vary. The relative depth error induced by vergence uncertainty when horizontal version and head roll were known was negligible. Therefore, relative depth could be reasonably accurately estimated using a subset of all available eye and head orientation signals when vergence was noisy. However, the absolute depth estimate suffered greatly from the vergence uncertainty, even if horizontal version and head roll angle were known. Therefore we argue that the brain has to use all available extraretinal eye and head orientation signals in order to obtain an accurate estimate of absolute depth from binocular visual signals if vergence is uncertain.

Validating the depth estimation model in a behavioral test

Our model simulations predicted that the brain should ideally use all eye and head orientation information (although those might be noisy too) in order to accurately infer depth from retinal stimulation. In order to test this prediction, we set out to perform an

experiment, where subjects were asked to align their fingertip in complete darkness with the perceived 3D location of a target for different eye and head orientations. We designed the experiment using retinal afterimage targets and computed individual test sets with different eye–head configurations that predicted different absolute distances for the target (see [Methods](#) section for more details). Subjects fixated the fixation spot on the screen but were otherwise in complete darkness, i.e., they could never see their hand or fingers. Therefore, vergence uncertainty only affected the distance estimate of the afterimage object, but not the estimate of finger distance. As a consequence, any difference in the depth judgment of the virtual afterimage must result from the fact that vergence is unreliable and that under these conditions different eye and head orientation signals were used to interpret the retinal input.

[Figure 10](#) shows the results from the experiment for all subjects where the observed afterimage target depth (i.e.,

the measured distance of the finger alignment with the afterimage target from the center of the interocular line) was plotted as a function of the predicted depth as computed by our geometrical model. In a hypothetical case where subjects relied on retinal stimulation alone, observed depth should not change across different eye–head orientations, because the binocular retinal stimulation remained identical across all trials. In contrast, if subjects used all available extraretinal signals to estimate depth from the binocular retinal information, we would expect observed depth to match the theoretically predicted depth. As can be observed, all subjects modulated their reported depth in the direction predicted by our model. The slope was significantly (*t*-test, $p < 0.01$) different from 0 for all subjects and varied between 0.6 and 3.3 (mean slope = 1.37, [Figure 11A](#)). Although the individual values seem to be far from the ideal value of 1 (see [Discussion](#) section), subjects did use extraretinal signals to modulate their depth estimate. Most subjects also showed a global

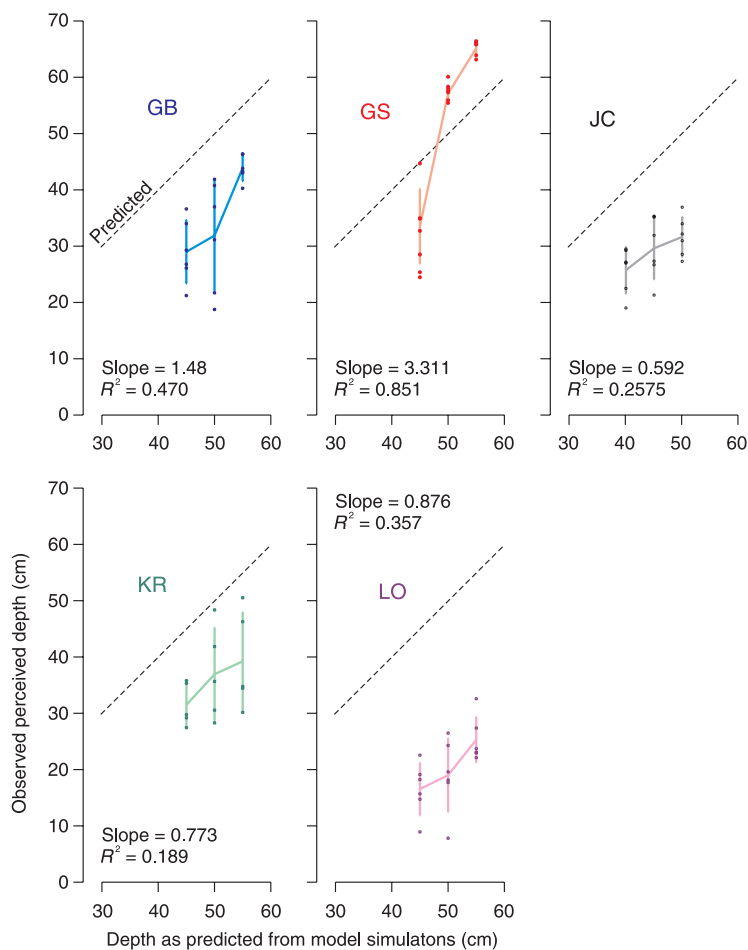


Figure 10. Experimental results. Observed pointing depth was plotted as a function of the theoretical depth predicted by our model when using the individual subject's parameters of Listing's law and sVOR. A slope of 0 would indicate that subjects did not make use of extraretinal signals to infer depth from retinal stimulation, whereas a slope of 1 (dotted line) indicates that subjects used full extraretinal information to estimate depth from vision. The 5 panels show data from each subject separately. The slope of the regression and the R^2 values are given.

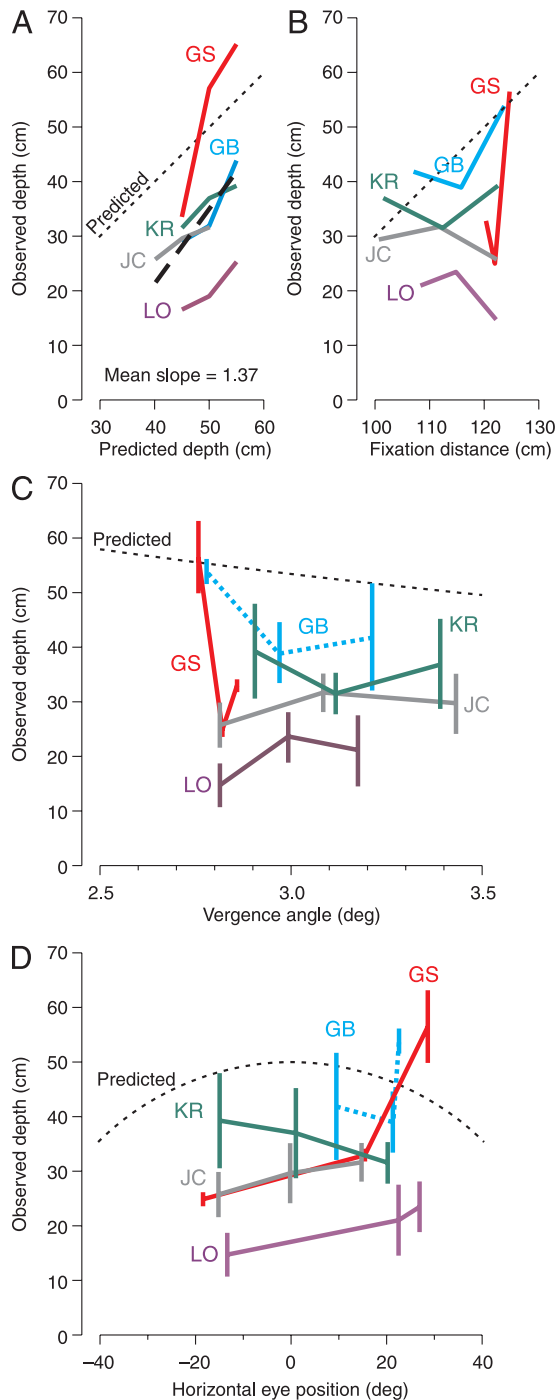


Figure 11. Evaluation of experimental results against alternative hypotheses. (A) Observed depth as a function of predicted depth (same as Figure 10) across all subjects. (B) Observed depth is not reliably predicted by fixation distance alone. If this were the case, we would expect a slope of 1 for all subjects (predicted, dotted line). (C) Similar to fixation distance, the vergence angle is not a good predictor for the observed depth variations either. Colored lines indicate the mean (\pm SD) for each subject. (D) Horizontal vergence is not a good predictor of the observed depth modulations in pointing. The pointing pattern does not follow the predicted (dotted) curve.

underestimation of depth (Mon-Williams & Tresilian, 1999; Tresilian et al., 1999; Van Pelt & Medendorp, 2008; Viguier et al., 2001).

Since it is well known that horizontal vergence plays a role in depth estimation, the observed modulation of reported depth could simply be due to the slight differences in fixation distance/vergence angle between test points (although we took care to compile a test set in which the vergence angles were not correlated with the predicted depth). In addition, differences in horizontal version could also account for the observed differences in reach depth (because for isovergence, the distance estimate also depends on horizontal version, as described by the Vieth–Muller circle). To rule out these possibilities, we re-plotted the observed depth reported by the subjects as a function of vergence angle/fixation distance and horizontal version in Figure 11. Figure 11A summarizes the mean results from Figure 10 and also shows the average results across all subjects (black dashed line). Figure 11B shows there was no consistent modulation of the reported depth with fixation distance. This is confirmed in Figure 11C where subjects' performance was plotted as a function of vergence angle (the dotted line shows the predictions of the model if observed depth only relied on vergence angle). In Figure 11D, we show that horizontal version was also not a good predictor for the observed modulations in reported depth.

To summarize our experimental validation, we asked subjects to judge the absolute depth of a virtual afterimage under different eye–head configurations. Subjects modulated the reported target depth as predicted by our model. We ruled out that this could simply be due to changes in vergence and/or horizontal version, which demonstrates that subjects made use of extraretinal eye and head orientation signals in order to infer depth from binocular retinal stimulation.

Discussion

We have shown theoretically that retinal disparity alone does not provide a unique estimation of absolute target depth of peri-foveal targets for different natural eye–head orientations. In order to accurately estimate absolute target distance, the brain must account for the complete 3D eye–head geometry. It could rely on vergence and horizontal version alone if the vergence signal was perfectly known. However, if vergence was uncertain, vertical version, head pitch, and head roll are also needed in order to uniquely estimate absolute depth from retinal disparity. We validated the findings of this theoretical evaluation through a behavioral experiment, which showed that different 3D eye–head configurations are taken into account when estimating target depth.

Noisy extraretinal signals

Our model showed that the brain could in principle interpret the same binocular retinal stimulation as resulting from targets located at different egocentric distances in space by taking extraretinal signals into account. To see if this is valid for human vision, we asked subjects to point to a retinotopically invariant afterimage of a point under different eye and head orientations and found that the estimated absolute distance was indeed modulated by eye and head orientations, as predicted by the model. Overall, we found that subjects varied their distance estimate in a fashion that was qualitatively similar to our model predictions. These different depth estimates cannot be explained by differences in vergence or horizontal version. However, the observed depth estimates differed quantitatively from those predicted by the model, i.e., for some subjects gain values were different from the predicted gain of 1. There are a number of possible reasons for this variation. First, it is well known that depth judgments—in particular for targets in isolation without any other visual cues or environment—have been shown to be biased away from the subject (Gogel, 1972); however most studies show a bias toward the observer (Howard & Rogers, 1995; Mon-Williams & Tresilian, 1999; Tresilian et al., 1999; Viguier et al., 2001), which is consistent with most of our subjects' behavior. Second, these variations may be due to changes in the parameters for Listing's law and the static VOR used to compute the experimental test set. Despite the fact that these parameters are thought to be relatively stable over a short period of time (Schor et al., 2002), they might have slightly changed, which would change projection geometry. Finally, it is possible that the extraretinal signals used to estimate egocentric depth from binocular visual information are not only noisy but might also be biased. Such a bias could also induce the observed changes in the gain values.

In terms of noisy signals, in the current investigation we focused on the uncertainty of the vergence signal (Brenner & van Damme, 1998; Collewijn & Erkelens, 1990; Foley, 1980; Harwerth et al., 1995; Viguier et al., 2001) as opposed to other extraretinal signals such as horizontal version. The underlying reason for this becomes clear in our analysis (Figures 6–9 and Table 2); the effect of small changes in the estimated vergence angle on the estimated absolute distance was typically much larger than errors due to variability in eye or head orientation signals. Since all extraretinal signals are noisy, we suggest that the brain uses all available signals in order to estimate depth, including the noisy vergence. A statistically optimal combination of all those signals would ensure the best possible outcome for depth reconstruction.

In the case of noise or biases in the internal representation of eye and head orientations as well as the vergence angle, the brain faces an additional problem that we have not considered in our analysis. In order to compute the

solutions to the inverse problem, the projection lines were required to intersect in space. However, this may not be the normal case for the brain, e.g., if the extraretinal signals were erroneous such as in the induced effect (Liu et al., 2005; Ogle, 1938) or for incorrect viewing positions with stereograms (Girshick & Banks, 2005). Since these signals are indeed prone to noise, we might even expect this to be the norm. How would the brain then deal with this? We suggest that the brain simply uses the point at which the two retinal target projection rays are closest to each other in 3D space as the presumed intersection point. This would be similar to what has been proposed as the extended horopter (Schreiber et al., 2006) addressing the retinal correspondence problem. We expect these presumed retinal target projection ray intersection points to provide results that are consistent with our other findings, but this problem remains to be investigated.

In a related modeling study, Erkelens and van Ee (1998) have analyzed how extraretinal eye orientation signals could be used in addition to binocular visual input in order to solve depth perception based on computing head-centric disparity. However, they focused on eye orientation signals and did not consider head-position induced changes to the ocular orientation due to VOR. More importantly, in their study they were interested in perceptual phenomena related to full-field stimuli, a condition in which the brain could—at least in theory—rely solely on retinal information (Horn, 1990). Our approach was very different; we only used a single point-like target in order to investigate, from a purely geometric point of view, the signals the brain should theoretically use to solve the absolute depth estimation problem. Therefore, whereas Erkelens and van Ee (1998) were mainly interested in how depth percepts (i.e., relative depths) were differentially influenced by different extraretinal signals, we concentrated on absolute distance computations for action. Although theoretically, both the action and perception systems could rely on the same computations when estimating depth, differences between the ventral (perception) and dorsal (action) visual streams have been reported in the past (Goodale, 2001; Goodale & Westwood, 2004) and thus different signals and/or computations could be used in the perception and action streams.

Signals for depth estimation from retinal disparity

Relative vs. absolute depth

The distance of a peripherally viewed target with respect to the fixation distance (=relative depth) is mainly estimated from the retinal disparity of the target (Foley, 1980; Mayhew & Longuet-Higgins, 1982; Mon-Williams & Tresilian, 1999; Ohzawa, 1998; Ritter, 1977). In addition, horizontal version also plays an important role in interpreting relative depth from retinal disparity

(Figure 9A, dashed line). On the other hand, vergence information adds relatively little information and reduces the possible range of relative depths by only a few centimeters (Figure 9A, solid line). Of course, this makes sense since relative depth only requires information about the position of the object relative to fixation and not about the fixation position itself. This relative distance information would be sufficient to judge whether a peripherally viewed object was closer or farther from the fixation target. However, in order to interact with the target, e.g., to make an eye movement to it or to reach out for it, information about the absolute distance from the observer is needed. Ideally, this can be obtained from retinal disparity (and the associated mean retinal position), by adding horizontal version as well as precise knowledge of the vergence angle (Figure 9B). However, as suggested before, the vergence angle might be prone to a bias or uncertainty (Brenner & van Damme, 1998; Collewijn & Erkelens, 1990; Foley, 1980; Harwerth et al., 1995; Viguier et al., 2001). Therefore, we argue that additional eye and head orientation signals have to be used in order to uniquely estimate absolute target distance from retinal disparity (Figure 8). Ocular accommodation could also be used as a cue to viewing distance (Mon-Williams & Tresilian, 1999, 2000). In both our study and Viguier et al. (2001) the vergence uncertainty might have resulted from a combination of vergence and accommodation that we cannot tease apart. However, this does not take away from the generalizability of our findings. If viewing distance was estimated from both vergence and accommodation, then the used extraretinal signal would simply reflect that.

Solving for ocular torsion

Theoretically, the problem the brain faces is to determine the exact 3D orientation of both eyes and in particular the exact torsional angle of both eyes independently. With the knowledge of the correct 3D binocular ocular geometry, the intersection point in depth of the retinal target projection rays could be computed. Yet, evidence that the brain has access to ocular torsion information is lacking. Indeed, judgments of vertical slant perception for different cyclovergence states indicate that the brain has no knowledge of the eyes' torsional angles (Banks, Hooge, & Backus, 2001), whereas we are not aware of convincing evidence against extraretinal signal associated with cyclovergence (if enough retinal information is available, cyclovergence could be estimated from retinal shear disparity (Banks et al., 2001)). Therefore, the brain must estimate ocular torsion from other extraretinal signals (Schreiber et al., 2001; van Ee & van Dam, 2003).

An alternative to estimating binocular ocular torsion is to use the sensory parameters that modulate torsion along with an internal model to reconstruct the torsional angles of the eyes. An internal model of binocular Listing's law (also called L2) and its modification with the static VOR would implement the necessary requirements. Thus, in

order to accurately estimate absolute distance, the brain would have to make use of horizontal and vertical versions and head pitch and roll in an internal model of binocular ocular torsion.

There are open questions that remain in terms of the signals the brain theoretically requires in order to estimate absolute depth from retinal disparity. One problem concerns the measured variability of the Listing's plane, i.e., its "thickness". This "thickness" for constant (horizontal and vertical) eye and head (pitch and roll) orientations results from variability in the torsional states of the eyes. How does the brain take this variability into account when our model suggests that exact torsional information is needed to estimate absolute depth? Two possible answers come to mind. First, the brain might not know about the exact torsional state and estimates depth based on the available sensory information. If that was the case, then we would expect that part of the observed variability in reach depth might be due to this misestimation of the torsional angles from the internal model. Second, part of this apparent natural variability of torsion might be due to cyclodisparity-induced ocular torsion, i.e., cyclovergence (Hooge & van den Berg, 2000; Howard, Sun, & Shen, 1994; van Rijn, van der Steen, & Collewijn, 1992). In that case, the brain could use cyclodisparity information (if available) in order to estimate its contribution to ocular torsion within the internal model.

In our model, we only address the question of estimating absolute depth for a single point-like target and without any other visual information. In more natural environments, where multiple targets or full-field views are available, it has been suggested that the visual system can estimate the relative 3D orientation of both eyes purely from vision (Horn, 1990). However, we have shown experimentally, that in addition to vergence (Rogers & Bradshaw, 1995) and version (Backus et al., 1999), the brain makes also use of other extraretinal signals specifying the complete 3D orientation of both eyes (see above). It is thus likely that redundant retinal and extraretinal signals are combined in everyday life.

Hypothetical neural substrates

The visual system calculates an egocentric 3D representation of target location within the dorsal stream responsible to guiding motor actions toward the target. Whereas this is a relatively simple task for the angular direction of the target, we have shown that absolute distance estimation is more complex than previously thought.

Potential areas involved in estimating absolute target depth should not only have information about retinal disparities but should also be modulated by vergence as well as extraretinal eye and head orientation signals. The brain could either integrate all these signals together to

directly compute the absolute distance from the visual input or, alternatively, proceed in two steps, first computing the depth relative to the fixation distance and then computing absolute distance. In both cases, we do not believe that the brain performs the ray intersection computations explicitly but uses simple distributed computing mechanisms as has been previously suggested (Pouget & Sejnowski, 1994).

It is well known that the early visual system contains neurons that are specifically selective for absolute retinal disparities (e.g., Cumming & DeAngelis, 2001; Neri, 2005; Parker & Cumming, 2001). A retinocentric representation including the angular direction and the target's radial distance is believed to exist at the level of the posterior parietal cortex (Genovesio & Ferraina, 2004; Gnadt & Mays, 1995; Hasebe et al., 1999). Thus, the visual system must build up this retinocentric representation of 3D egocentric target location somewhere between the striate cortex and the posterior parietal cortex.

Modulation of disparity responses with fixation distance (vergence) has been observed in V1, V2 (Rosenbluth & Allman, 2002) as well as in area MST (Roy, Komatsu, & Wurtz, 1992). Eye orientation and vergence modulation is also largely present in the dorsal stream (DeSouza, Dukelow, & Vilis, 2002; Gnadt & Mays, 1995), but far less is known about head orientation signals in dorsal stream areas, besides the posterior parietal cortex (Brotchie, Andersen, Snyder, & Goodman, 1995). An area of particular interest for the mechanisms we propose here seems to be area MT/V5, which projects to the posterior parietal cortex (Blatt, Andersen, & Stoner, 1990) and has been shown to be crucial for absolute distance perception (Cumming & DeAngelis, 2001; DeAngelis et al., 1998; Krug, Cumming, & Parker, 2004). Activity of neurons in area MT is also modulated by eye orientation (DeSouza et al., 2002), but to our knowledge, the influence of vergence or head orientation signals on the neural activity in this area is unknown. Probing for those extraretinal signals in area MT would be an important next step in the search for the underlying neural substrates of 3D egocentric depth reconstruction.

Conclusion

The internal reconstruction of 3D target position, i.e., an egocentric measure of object location with respect to the line of sight, is necessary for successful interaction with the environment. The present study provides evidence that retinal disparity alone is insufficient to accurately estimate the absolute distance of an object from the observer. Instead, the brain must account for various extraretinal eye and head orientation signals in order to reconstruct the 3D orientation of both eyes in space. We suggest that all available signals (including a noisy estimate of vergence)

are used and combined in a statistically optimal fashion in order to reduce the overall uncertainty in absolute depth estimation.

Appendix A

The 3D geometry of fixation and target lines

If the eyes fixate a target $\vec{F} = (F_X, F_Y, F_Z)$ (expressed in cyclopean-eye-centered, head-fixed coordinates), then the fixation lines L^G of the right and left eyes can be written as simple geometrical expressions:

$$L_R^G : s \cdot \vec{g}_R + \vec{g}_{R,0} = 0 \quad (\text{A1})$$

$$L_L^G : u \cdot \vec{g}_L + \vec{g}_{L,0} = 0.$$

In Equation A1, s and u are parameters, \vec{g}_R and \vec{g}_L are the unit gaze direction vectors of the right and left eyes, respectively, and $\vec{g}_{R,0}$ and $\vec{g}_{L,0}$ are the locations of the right and left eyes in cyclopean-eye-centered, head-fixed coordinates, i.e.,

$$\vec{g}_{R,L} = \frac{(\vec{F} - \vec{g}_{R,L,0})}{\|\vec{F} - \vec{g}_{R,L,0}\|}, \quad (\text{A2})$$

$$\vec{g}_{R,0} = -\vec{g}_{L,0} = \begin{pmatrix} 3 \text{ cm} \\ 0 \\ 0 \end{pmatrix}. \quad (\text{A3})$$

On the other hand let us consider a given cyclopean retinal target position \vec{p}_T and retinal disparity d_T (both arbitrarily expressed in eye-centered, eye-fixed coordinates (also called retinal coordinates) using the Fick convention, i.e., $\vec{p}_T = \begin{pmatrix} \vartheta_T \\ \phi_T \end{pmatrix}$ and $\vec{d}_T = \begin{pmatrix} \Delta\vartheta_T \\ \Delta\phi_T \end{pmatrix}$, where ϑ is the horizontal and ϕ is the vertical component) associated with a potential reach target at the cyclopean-eye-centered, eye-fixed location \vec{T} . Then, the lines L^T defined by the projection of the target onto both retinas can be written as

$$L_R^T : v \cdot \vec{t}_R + \vec{g}_{R,0} = 0 \quad (\text{A4})$$

$$L_L^T : w \cdot \vec{t}_L + \vec{g}_{L,0} = 0.$$

The unit target directions vectors \vec{t}_i (i stands for R or L) are calculated as the cyclopean-eye-centered, head-fixed

direction vector of the target ($\vec{t}_{i,PP}$, standing for Primary Position):

$$\vec{t}_{i,PP} = \begin{pmatrix} \sin \vartheta_i \cdot \cos \varphi_i \\ \cos \vartheta_i \cdot \cos \varphi_i \\ \sin \varphi_i \end{pmatrix}. \quad (\text{A5})$$

The target directions (ϑ_i , φ_i) are computed from the retinal target position (\vec{p}_T) and the retinal disparity (\vec{d}_T) associated with the reach target and written as

$$\begin{aligned} \vartheta_i &= \vartheta_T \pm \frac{\Delta \vartheta_T}{2} \\ \varphi_i &= \varphi_T \pm \frac{\Delta \varphi_T}{2}. \end{aligned} \quad (\text{A6})$$

The eye-centered, eye-fixed target direction vector $\vec{t}_{i,PP}$ then has to be translated and rotated into the right and left eye orientations. This has to account for the 3D eye-in-head orientation and will be developed in the next section. Importantly, the two target rays L_R^T and L_L^T must intersect. This means that their spatial distance D_T must be zero (see next section). If the two target rays L_R^T and L_L^T intersect, then the target position $\vec{T} = \vec{g}_{R,0} + v \cdot \vec{t}_R$ in head-fixed, eye-centered space can be found using Equation A4, where the parameter v is written as

$$v = \frac{(\vec{\sigma} \times \vec{t}_L) \cdot (\vec{t}_R \times \vec{t}_L)}{\|\vec{t}_R \times \vec{t}_L\|^2}, \quad \text{with } \vec{\sigma} = \vec{g}_{L,0} - \vec{g}_{R,0}. \quad (\text{A7})$$

Appendix B

3D eye orientation constraints on the projection geometry

In order to rotate the eye-centered, eye-fixed target direction vector $\vec{t}_{i,PP}$ into a head-fixed representation, we need to account for the binocular version of Listing's law. Listing's law (Hepp, 1990) constrains the three-dimensional eye-in-head rotation vectors to a two-dimensional plane, called Listing's plane. In addition, Listing's law is also modulated by pitch and roll head orientations and this is known as the gravity pitch of Listing's plane and the ocular counter-roll, respectively.

We will use Hamilton's dual quaternion algebra (Clifford, 1873) to describe the translation and rotation

of the eye-centered, eye-fixed target direction vector $\vec{t}_{i,PP}$ to account for the right and left eye's 3D position and orientation. The advantage of using dual quaternions over any other formalism is that it allows us to describe eye rotation independently of rotation sequences. In addition, dual quaternions provide a simple way of calculating the skew distance between two lines in 3D space. Finally, dual quaternions provide certain mathematical and numerical gains over possible alternatives (Aspragathos & Dimitros, 1998). However, the use of dual quaternions is an arbitrary choice and any other formalism would give the same results.

A dual quaternion operator \hat{Q} can be written as the sum of two quaternion operators Q and Q_0 , of which one is multiplied by a duality operator ε , i.e., $\hat{Q} = Q + \varepsilon Q_0$, where Q describes the rotational component and Q_0 implements the translation. A dual quaternion can also be represented as an eight-dimensional vector $\hat{Q} = \begin{bmatrix} Q \\ Q_0 \end{bmatrix}$. For a rotation θ around the axis \vec{r} applied in \vec{a} and a translation d along \vec{r} , the dual quaternion components are

$$\begin{aligned} Q &= \begin{bmatrix} \cos \frac{\theta}{2} \\ \vec{r} \cdot \sin \frac{\theta}{2} \end{bmatrix} \quad \text{and} \\ Q_0 &= \begin{bmatrix} -\frac{d}{2} \cdot \sin \frac{\theta}{2} \\ \frac{d}{2} \cdot \vec{r} \cdot \cos \frac{\theta}{2} + (\vec{r} \times \vec{a}) \cdot \sin \frac{\theta}{2} \end{bmatrix}. \end{aligned} \quad (\text{B1})$$

A chain of translations and rotations can be expressed in the dual quaternion product $\hat{Q} = \prod \hat{Q}_i$. Dual quaternion multiplication has the following i property (quaternion multiplication applies to the individual elements):

$$\begin{aligned} \hat{A}\hat{B} &= (A + \varepsilon A_0)(B + \varepsilon B_0) \\ &= AB + \varepsilon(A_0B + AB_0). \end{aligned} \quad (\text{B2})$$

Quaternion algebra has been described elsewhere (Brand, 1947; Hamilton, 1899; Haslwanter, 1995; Martin, 1983; Tait, 1890). The 3D eye-centered, eye-fixed target line in space with directional vector $\vec{t}_{i,PP}$ and passing through the center of the eyes $\vec{g}_{i,0}$ can be represented by the dual quaternion line $\hat{L}_{i,0} = [0 \ \vec{t}_{i,PP} \ 0 \ \vec{g}_{i,0} \times \vec{t}_{i,PP}]^T$. Using the appropriate dual quaternion $\hat{Q}_{eh,i}$ representing the 3D eye-in-head rotation, we can then rotate the lines $\hat{L}_{i,0}$ according to gaze direction and obtain the eye-centered, head-fixed target lines $\hat{L}_i = \hat{Q}_{eh,i} \hat{L}_{i,0} \hat{Q}_{eh,i}^C$, where $\hat{Q}^C = Q^C + Q_0^C$ is the dual quaternion component conjugate and Q^C is the quaternion conjugate.

The eye-in-head rotation quaternion $\hat{Q}_{eh,i}$ is composed by the binocular Listing's law (also called L2) quaternion $Q_{L2,i}$ and the ocular counter-roll operator Q_{OCR} , i.e., $\hat{Q}_{eh,i} = [Q_{L2,i} Q_{OCR}]$. The calculation of the Listing's law quaternion was presented in details elsewhere (Tweed, 1997a). In brief, the primary orientation of Listing's plane Q_{LP} accounted for the static vestibulo-ocular reflex (sVOR) by inducing a so-called gravity pitch (α_P) of the normal vector defining the Listing's plane (Q_{LP}). To compute the Listing's plane in the binocular extension ($Q_{LP2,i}$), we accounted for the ocular vergence angle ν ($\cos \nu = \vec{g}_R \cdot \vec{g}_L$). Vergence then rotates the Listing's planes out "like saloon doors" (rotation $Q_{V,i}$). We hypothesize that this rotation is performed in head-fixed coordinates, but this is not known to date. This allowed us to compute the rotation quaternion $Q_{L2,i}$ that brings the eyes from the primary position ($Q_{PP,i}$) into the appropriate Listing's plane as

$$Q_{L2,i} = Q_{L,i} Q_{PP,i}, \quad (\text{B3})$$

$$Q_{L,i} = \sqrt{-\begin{bmatrix} 0 \\ \vec{g}_i \end{bmatrix}} Q_{LP2,i}, \quad (\text{B4})$$

$$Q_{PP,i} = \sqrt{Q_{LP2,i}^{-1} [0 \ 0 \ 1 \ 0]^T}, \quad (\text{B5})$$

$$Q_{LP2,i} = Q_{LP} Q_{V,i}, \quad (\text{B6})$$

$$Q_{LP} = \begin{bmatrix} 0 \\ 0 \\ \cos \alpha \\ -\sin \alpha \end{bmatrix}, \quad \text{with } \alpha = \alpha_0 + c_P \cdot \alpha_P, \quad (\text{B7})$$

$$Q_{V,i} = \begin{bmatrix} \cos \frac{\delta_i \cdot \nu}{2} \\ 0 \\ 0 \\ \sin \frac{\delta_i \cdot \nu}{2} \end{bmatrix}, \quad (\text{B8})$$

$$Q_{OCR} = \begin{bmatrix} \cos(-c_{OCR} \cdot \beta_R) \\ 0 \\ \sin(-c_{OCR} \cdot \beta_R) \\ 0 \end{bmatrix}. \quad (\text{B9})$$

We used a tilt angle of Listing's plane for the head-upright orientation that was $\alpha_0 = 5^\circ$ and the gain for the gravity modulation of this tilt related to the pitch angle α_P was $c_P = 0.05$ (Bockisch & Haslwanter, 2001; Haslwanter et al., 1992). The gain for the static ocular counter-roll of the head roll angle β_R was $c_{OCR} = 0.05$ (Bockisch & Haslwanter, 2001; Haslwanter et al., 1992). The gain δ_i for the rotation of Listing's plane due to vergence was 1/4 and $\text{sign}(\delta_i) = +1$ for the left eye and -1 for the right eye. However, different values ranging between 1/6 and 1/2 have been reported in the literature (Mok et al., 1992; Tweed, 1997b; Van Gisbergen & Minken, 1994; Van Rijn & Van den Berg, 1993).

The skew distance \hat{D} between the two dual quaternion target lines in eye-centered, head-fixed coordinates can be computed as $\hat{D} = \hat{L}_R \hat{L}_L^{-1}$. The first component of \hat{D} provides the angular distance θ between the two lines, i.e., $\theta = 2 \cdot \cos^{-1} \hat{D}(1)$, whereas the closest metric distance d of these lines can be computed from the fifth component of \hat{D} , i.e., $d = -2 \cdot \frac{\hat{D}(5)}{\sin(\theta/2)}$. This distance must be zero if both rays intersect, i.e., if the set of parameters used has a solution.

To investigate how a given cyclopean eye-centered, cyclopean eye-fixed target at position \vec{p}_C projects onto the left and right retinas, we computed the (monocular) dual quaternion of cyclopean eye rotation $\hat{Q}_{eh,C}$ to position the target into a cyclopean-eye-centered, head-fixed reference frame. Then we projected this target onto both retinas to calculate the individual right and left eye retinal positions as well as the retinal disparity associated with the cyclopean retinal position. The point \vec{p}_C represented by the dual quaternion $\hat{P}_C = [1 \ 0 \ 0 \ 0 \ 0 \ \vec{p}_C]^T$ can then be transformed with $\hat{Q}_{eh,C}$ into the cyclopean-eye-centered, head-fixed reference frame, i.e., $\hat{P}_H = \hat{Q}_{eh,C} \hat{P}_C \hat{Q}_{eh,C}^{DC}$, where the dual quaternion double conjugate is $\hat{Q}_{eh,C}^{DC} = Q_{eh,C}^C - \varepsilon Q_0^C$ with the quaternion conjugate $Q_{eh,C}^C$. Then the projection of the cyclopean-eye-centered, head-fixed target \hat{P}_H onto both retinal writes $\hat{P}_{E,i} = \hat{Q}_{eh,i}^{DC} \hat{P}_H \hat{Q}_{eh,i}$. To extract the translational part from a dual quaternion, one can use the following expression:

$$\begin{bmatrix} p_0 \\ \vec{p} \end{bmatrix} = H(Q)^T Q_0, \quad (\text{B10})$$

where $H(Q)$ is the negative Hamiltonian of the quaternion Q defined as

$$H(Q) = \begin{bmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & q_3 & -q_2 \\ q_2 & -q_3 & q_0 & q_1 \\ q_3 & q_2 & -q_1 & q_0 \end{bmatrix}. \quad (\text{B11})$$

Appendix C

Changes in head pitch angle

In this section we analyze how changes in head pitch angle affect the inverse problem. Since both head roll and head pitch modify ocular torsion in a similar way (blue lines in Figure 3), a torsional change due to a alteration in head pitch angle can be compensated by a head roll movement. This is illustrated in Figure C1A where we show the combinations of head roll and pitch angles that produce the same amount of ocular torsion for different vertical version. Vertical version influenced this relationship because head pitch tilts Listing's plane forward and this tilt produced larger torsional angles for larger vertical version. Therefore we expected that the same binocular retinal stimulation could also result from different horizontal/vertical version-vergence combinations when only head pitch angle changed and head roll was kept at 0° .

To show this we performed similar model simulations for head pitch than for roll. However, now we imposed an additional constraint, namely that head orientation obeys Donders' law, i.e., the head contribution to a given gaze shift was specified. With this additional constraint, we fixed only horizontal version and searched all vertical version angles (and thus pitch angles, due to Donders' law) for a solution. The result was a 2D hyperbolic surface in 3D space (not shown). All fixation positions on this surface had an associated head pitch angle for which the retinal target projection lines intersected. Thus, even when reducing the degrees of freedom by constraining the head movements, purely visual information was not sufficient to provide unique target depth. This is shown in Figure C1B. As in Figure 6A, the solutions associated with all possible fixation positions are plotted as gray dots. The knowledge of vergence (black dots, 6.9° corresponding to fixating a straight-ahead 50-cm distant position) in addition to binocular retinal input was not sufficient to infer target distance. As for head roll, an accurate estimate needed additional horizontal version information (0° horizontal

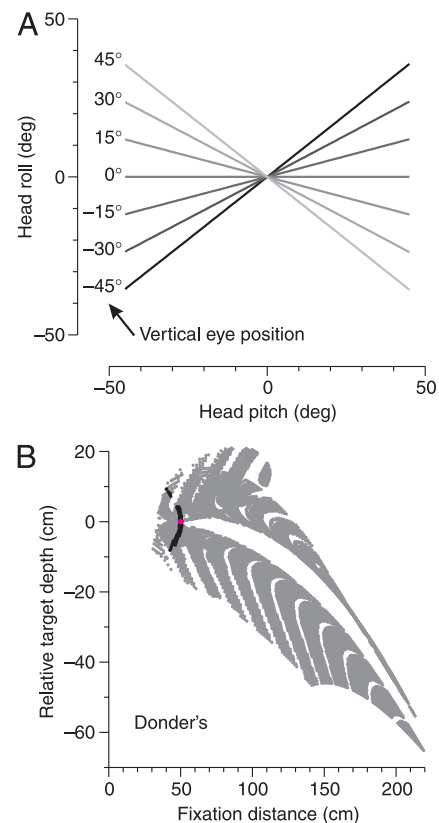


Figure C1. Depth estimation for constant roll. (A) Relationship between head pitch and roll angles that produce the same torsional state of the eyes (cyclovergence and cyclovergence) as a function of vertical vergence. (B) Possible combinations of relative target distance and fixation distance when head pitch changes and head roll is fixed. The degrees of freedom of the problem was further reduced by the Donders' strategy that specifies the head pitch contribution to a given gaze shift. This did not reduce the complexity. As in Figure 6, black dots represent a subset of solutions with constant vergence angle (vergence = 6.9° , corresponding to straight-ahead fixation at 50 cm). The magenta dot shows a further subset of solutions where horizontal vergence was zero (0°).

version subset is shown in magenta in Figure C1B). To summarize, the brain required horizontal version as well as vergence in addition to binocular retinal input (2D retinal position and horizontal and vertical disparities) in order to uniquely estimate the relative and absolute distance of a peripherally viewed target if all signals are precise and accurate.

Acknowledgments

We thank Dr. Xiaogang Yan and Saihong Sun for their kind help in the development of the hardware and software required for the experiment. We are also thankful

to Dr. Laurence R. Harris who suggested using a virtual afterimage object in the experiment. This work was supported by the Canadian Institutes of Health Research (CIHR). GB was supported by a Marie Curie International fellowship within the 6th European Community Framework Program and CIHR (Canada). AZK holds a CIHR (Canada) postdoctoral fellowship. JDC holds a Canada Research Chair.

Commercial relationships: none.

Corresponding author: Dr. J. Douglas Crawford.

Email: jdc@yorku.ca.

Address: Centre for Vision Research, York University, 4700 Keele Street, Toronto, Ontario M5P 2L3, Canada.

References

- Aspragathos, N. A., & Dimitros, J. K. (1998). A comparative study of three methods for robot kinematics. *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, 28, 135–145. [[PubMed](#)]
- Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, 39, 1143–1170. [[PubMed](#)]
- Banks, M. S., Hooge, I. T., & Backus, B. T. (2001). Perceiving slant about a horizontal axis from stereopsis. *Journal of Vision*, 1(2):1, 55–79, <http://journalofvision.org/1/2/1/>, doi:10.1167/1.2.1. [[PubMed](#)] [[Article](#)]
- Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates. *Science*, 285, 257–260. [[PubMed](#)]
- Battaglia-Mayer, A., Caminiti, R., Lacquaniti, F., & Zago, M. (2003). Multiple levels of representation of reaching in the parieto-frontal network. *Cerebral Cortex*, 13, 1009–1022. [[PubMed](#)] [[Article](#)]
- Bishop, P. O. (1989). Vertical disparity, egocentric distance and stereoscopic depth constancy: A new interpretation. *Proceedings of the Royal Society of London B: Biological Sciences*, 237, 445–469. [[PubMed](#)]
- Blatt, G. J., Andersen, R. A., & Stoner, G. R. (1990). Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *Journal of Comparative Neurology*, 299, 421–445. [[PubMed](#)]
- Blohm, G., & Crawford, J. D. (2007). Computations for geometrically accurate visually guided reaching in 3-D space. *Journal of Vision*, 7(5):4, 1–22, <http://journalofvision.org/7/5/4/>, doi:10.1167/7.5.4. [[PubMed](#)] [[Article](#)]
- Bockisch, C. J., & Haslwanter, T. (2001). Three-dimensional eye position during static roll and pitch in humans. *Vision Research*, 41, 2127–2137. [[PubMed](#)]
- Brand, L. (1947). *Vector and tensor analysis* (9th ed.). London: John Wiley and Sons.
- Brenner, E., & van Damme, W. J. (1998). Judging distance from ocular convergence. *Vision Research*, 38, 493–498. [[PubMed](#)]
- Brotchie, P. R., Andersen, R. A., Snyder, L. H., & Goodman, S. J. (1995). Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature*, 375, 232–235. [[PubMed](#)]
- Carey, D. P., Dijkerman, H. C., & Milner, A. D. (1998). Perception and action in depth. *Consciousness and Cognition*, 7, 438–453. [[PubMed](#)]
- Clifford, W. (1873). Preliminary sketch of bi-quaternions. *Proceedings of the London Mathematical Society*, 4, 381–395.
- Collewijn, H., & Erkelens, C. J. (1990). Binocular eye movements and the perception of depth. *Reviews of Oculomotor Research*, 4, 213–261. [[PubMed](#)]
- Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1997). Trajectories of the human binocular fixation point during conjugate and non-conjugate gaze-shifts. *Vision Research*, 37, 1049–1069. [[PubMed](#)]
- Crawford, J. D., Medendorp, W. P., & Marotta, J. J. (2004). Spatial transformations for eye–hand coordination. *Journal of Neurophysiology*, 92, 10–19. [[PubMed](#)] [[Article](#)]
- Crawford, J. D., & Vilis, T. (1991). Axes of eye rotation and Listing’s law during rotations of the head. *Journal of Neurophysiology*, 65, 407–423. [[PubMed](#)]
- Cumming, B. G., & DeAngelis, G. C. (2001). The physiology of stereopsis. *Annual Review of Neuroscience*, 24, 203–238. [[PubMed](#)]
- DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, 394, 677–680. [[PubMed](#)]
- DeSouza, J. F., Dukelow, S. P., & Vilis, T. (2002). Eye position signals modulate early dorsal and ventral visual areas. *Cerebral Cortex*, 12, 991–997. [[PubMed](#)] [[Article](#)]
- Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 1141–1146. [[PubMed](#)] [[Article](#)]
- Erkelens, C. J., & van Ee, R. (1998). A computational model of depth perception based on headcentric disparity. *Vision Research*, 38, 2999–3018. [[PubMed](#)]
- Foley, J. M. (1980). Binocular distance perception. *Psychological Review*, 87, 411–434. [[PubMed](#)]

- Genovesio, A., & Ferraina, S. (2004). Integration of retinal disparity and fixation-distance related signals toward an egocentric coding of distance in the posterior parietal cortex of primates. *Journal of Neurophysiology*, *91*, 2670–2684. [PubMed] [Article]
- Girshick, A. R., & Banks, M. S. (2005). Do people compensate for incorrect viewing position when looking at stereograms? *Perception*, *34*, ECVF Abstract Supplement.
- Gnadt, J. W., & Mays, L. E. (1995). Neurons in monkey parietal area LIP are tuned for eye-movement parameters in three-dimensional space. *Journal of Neurophysiology*, *73*, 280–297. [PubMed]
- Gogel, W. C. (1972). Scalar perceptions with binocular cues of distance. *American Journal of Psychology*, *85*, 477–497. [PubMed]
- Gonzalez, F., & Perez, R. (1998). Neural mechanisms underlying stereoscopic vision. *Progress in Neurobiology*, *55*, 191–224. [PubMed]
- Goodale, M. A. (2001). Different spaces and different times for perception and action. *Progress in Brain Research*, *134*, 313–331. [PubMed]
- Goodale, M. A., & Westwood, D. A. (2004). An evolving view of duplex vision: Separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology*, *14*, 203–211. [PubMed]
- Goonetilleke, S. C., Mezey, L. E., Burgess, A. M., & Curthoys, I. S. (2008). On the relation between ocular torsion and visual perception of line orientation. *Vision Research*, *48*, 1488–1496. [PubMed]
- Hamilton, W. R. (1899). *Elements of quaternions*. Cambridge, UK: Cambridge University Press.
- Harwerth, R. S., Smith, E. L., III, & Siderov, J. (1995). Behavioral studies of local stereopsis and disparity vergence in monkeys. *Vision Research*, *35*, 1755–1770. [PubMed]
- Hasebe, H., Oyamada, H., Kinomura, S., Kawashima, R., Ouchi, Y., Nobezawa, S., et al. (1999). Human cortical areas activated in relation to vergence eye movements—A PET study. *Neuroimage*, *10*, 200–208. [PubMed]
- Haslwanter, T. (1995). Mathematics of three-dimensional eye rotations. *Vision Research*, *35*, 1727–1739. [PubMed]
- Haslwanter, T., Straumann, D., Hess, B. J., & Henn, V. (1992). Static roll and pitch in the monkey: Shift and rotation of Listing's plane. *Vision Research*, *32*, 1341–1348. [PubMed]
- Henriques, D. Y., & Crawford, J. D. (2000). Direction-dependent distortions of retinocentric space in the visuomotor transformation for pointing. *Experimental Brain Research*, *132*, 179–194. [PubMed]
- Hepp, K. (1990). On Listing's law. *Communications in Mathematical Physics*, *132*, 285–292.
- Hooge, I. T., & van den Berg, A. V. (2000). Visually evoked cyclovergence and extended Listing's law. *Journal of Neurophysiology*, *83*, 2757–2775. [PubMed] [Article]
- Horn, B. K. P. (1990). Relative orientation. *International Journal of Computer Vision*, *4*, 59–78.
- Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford, UK: Oxford University Press.
- Howard, I. P., Sun, L., & Shen, X. (1994). Cyclovergence and cyclovergence: The effects of the area and position of the visual display. *Experimental Brain Research*, *100*, 509–514. [PubMed]
- Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, *31*, 1351–1360. [PubMed]
- Johnston, E. B., Cumming, B. G., & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. *Vision Research*, *33*, 813–826. [PubMed]
- Khokhotva, M., Ono, H., & Mapp, A. P. (2005). The cyclopean eye is relevant for predicting visual direction. *Vision Research*, *45*, 2339–2345. [PubMed]
- Knill, D. C. (2005). Reaching for visual cues to depth: The brain combines depth cues differently for motor control and perception. *Journal of Vision*, *5*(2):2, 103–115, <http://journalofvision.org/5/2/2/>, doi:10.1167/5.2.2. [PubMed] [Article]
- Krug, K., Cumming, B. G., & Parker, A. J. (2004). Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. *Journal of Neurophysiology*, *92*, 1586–1596. [PubMed] [Article]
- Liu, B., Berends, E. M., & Schor, C. M. (2005). Adaptation to the induced effect stimulus normalizes surface slant perception and recalibrates eye position signals for azimuth. *Journal of Vision*, *5*(10):5, 808–822, <http://journalofvision.org/5/10/5/>, doi:10.1167/5.10.5. [PubMed] [Article]
- Martin, R. R. (1983). Rotation by quaternions. *Mathematical Spectrum*, *17*, 42–48.
- Mather, G. (1997). The use of image blur as a depth cue. *Perception*, *26*, 1147–1158. [PubMed]
- Mayhew, J. E., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature*, *297*, 376–378. [PubMed]
- Misslisch, H., Tweed, D., & Hess, B. J. (2001). Stereopsis outweighs gravity in the control of the eyes. *Journal of Neuroscience*, *21*, RC126. [PubMed] [Article]
- Mok, D., Ro, A., Cadera, W., Crawford, J. D., & Vilis, T. (1992). Rotation of Listing's plane during vergence. *Vision Research*, *32*, 2055–2064. [PubMed]

- Mon-Williams, M., & Tresilian, J. R. (1999). Some recent studies on the extraretinal contribution to distance perception. *Perception*, *28*, 167–181. [PubMed]
- Mon-Williams, M., & Tresilian, J. R. (2000). Ordinal depth information from accommodation? *Ergonomics*, *43*, 391–404. [PubMed]
- Mon-Williams, M., Tresilian, J. R., & Roberts, A. (2000). Vergence provides veridical depth perception from horizontal retinal image disparities. *Experimental Brain Research*, *133*, 407–413. [PubMed]
- Mueller, J. (1826). *Zur vergleichenden physiologie des gesichtssinnes des menschen und der thierte*. Leipzig, Germany: Cnobloch.
- Neri, P. (2005). A stereoscopic look at visual cortex. *Journal of Neurophysiology*, *93*, 1823–1826. [PubMed] [Article]
- Ogle, K. N. (1938). Induced size effect: I. A new phenomenon in binocular space perception associated with the relative sizes of the images of the two eyes. *Archives of Ophthalmology*, *20*, 604–623.
- Ohzawa, I. (1998). Mechanisms of stereoscopic vision: The disparity energy model. *Current Opinion in Neurobiology*, *8*, 509–515. [PubMed]
- Ono, H., Mapp, A. P., & Howard, I. P. (2002). The cyclopean eye in vision: The new and old data continue to hit you right between the eyes. *Vision Research*, *42*, 1307–1324. [PubMed]
- O’Shea, R. P., Blackburn, S. G., & Ono, H. (1994). Contrast as a depth cue. *Vision Research*, *34*, 1595–1604. [PubMed]
- O’Shea, R. P., Govan, D. G., & Sekuler, R. (1997). Blur and contrast as pictorial depth cues. *Perception*, *26*, 599–612. [PubMed]
- Palanca, B. J., & DeAngelis, G. C. (2003). Macaque middle temporal neurons signal depth in the absence of motion. *Journal of Neuroscience*, *23*, 7647–7658. [PubMed] [Article]
- Parker, A. J., & Cumming, B. G. (2001). Cortical mechanisms of binocular stereoscopic vision. *Progress in Brain Research*, *134*, 205–216. [PubMed]
- Pouget, A., & Sejnowski, T. J. (1994). A neural model of the cortical representation of egocentric distance. *Cerebral Cortex*, *4*, 314–329. [PubMed]
- Qing, Y., & Kapoula, Z. (2004). Saccade-vergence dynamics and interaction in children and in adults. *Experimental Brain Research*, *156*, 212–223. [PubMed]
- Richard, W., & Miller, J. F. (1969). Convergence as a cue to depth. *Perception & Psychophysics*, *5*, 317–320.
- Ritter, M. (1977). Effect of disparity and viewing distance on perceived depth. *Perception & Psychophysics*, *22*, 400–407.
- Rogers, B. J., & Bradshaw, M. F. (1995). Disparity scaling and the perception of frontoparallel surfaces. *Perception*, *24*, 155–179. [PubMed]
- Rosenbluth, D., & Allman, J. M. (2002). The effect of gaze angle and fixation distance on the responses of neurons in V1, V2, and V4. *Neuron*, *33*, 143–149. [PubMed] [Article]
- Roy, J. P., Komatsu, H., & Wurtz, R. H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience*, *12*, 2478–2492. [PubMed] [Article]
- Schor, C. M., Maxwell, J. S., McCandless, J., & Graf, E. (2002). Adaptive control of vergence in humans. *Annals of the New York Academy of Sciences*, *956*, 297–305. [PubMed]
- Schreiber, K., Crawford, J. D., Fetter, M., & Tweed, D. (2001). The motor side of depth vision. *Nature*, *410*, 819–822. [PubMed]
- Schreiber, K. M., Tweed, D. B., & Schor, C. M. (2006). The extended horopter: Quantifying retinal correspondence across changes of 3D eye position. *Journal of Vision*, *6*(1):6, 64–74, <http://journalofvision.org/6/1/6/>, doi:10.1167/6.1.6. [PubMed] [Article]
- Snyder, L. H. (2000). Coordinate transformations for eye and arm movements in the brain. *Current Opinion in Neurobiology*, *10*, 747–754. [PubMed]
- Tait, P. G. (1890). *An elementary treatise on quaternions* (3rd ed.). Cambridge, UK: Cambridge University Press.
- Tresilian, J. R., Mon-Williams, M., & Kelly, B. M. (1999). Increasing confidence in vergence as a cue to distance. *Proceedings of the Royal Society B: Biological Sciences*, *266*, 39–44. [PubMed] [Article]
- Tsutsui, K., Taira, M., & Sakata, H. (2005). Neural mechanisms of three-dimensional vision. *Neuroscience Research*, *51*, 221–229. [PubMed]
- Tweed, D. (1997a). Three-dimensional model of the human eye–head saccadic system. *Journal of Neurophysiology*, *77*, 654–666. [PubMed] [Article]
- Tweed, D. (1997b). Visual-motor optimization in binocular control. *Vision Research*, *37*, 1939–1951. [PubMed]
- Uka, T., & DeAngelis, G. C. (2002). Binocular vision: An orientation to disparity coding. *Current Biology*, *12*, R764–R766. [PubMed] [Article]
- van Ee, R., & van Dam, L. C. (2003). The influence of cyclovergence on unconstrained stereoscopic matching. *Vision Research*, *43*, 307–319. [PubMed]
- Van Gisbergen, J. A. M., & Minken, A. W. H. (1994). Conjugate and disconjugate contributions to bifoveal fixations studied from a 3-D perspective. In J. M. Delgado-Garcia, E. Godaux, & P. P. Vidal (Eds.),

Information processing underlying gaze control.
Oxford, UK: Pergamon Press.

- Van Pelt, S., & Medendorp, W. P. (2008). Updating target distance across eye movements in depth. *Journal of Neurophysiology*, *99*, 2281–2290. [[PubMed](#)]
- Van Rijn, L. J., & Van den Berg, A. V. (1993). Binocular eye orientation during fixations: Listing's law extended to include eye vergence. *Vision Research*, *33*, 691–708. [[PubMed](#)]
- van Rijn, L. J., van der Steen, J., & Collewijn, H. (1992). Visually induced cyclovergence and cyclovergence. *Vision Research*, *32*, 1875–1883. [[PubMed](#)]
- Vieth, G. U. A. (1818). Ueber die Richtung der Augen. *Annalen der Physik*, *58*, 233–253.
- Viguiet, A., Clement, G., & Trotter, Y. (2001). Distance perception within near visual space. *Perception*, *30*, 115–124. [[PubMed](#)]