Speed and eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion

Czuba, T.B., Rokers, B., Huk, A.C., & Cormack, L.K.

Psychology, Neurobiology, & Center for Perceptual Systems

The University of Texas at Austin

Copyright © 2010 by the American Physiological Society.
Abstract

Two binocular cues are thought to underlie the visual perception of 3D motion: a disparity-based cue, that relies on changes in disparity over time, and a velocity-based cue that relies on interocular velocity differences. The respective building blocks of these cues, instantaneous disparity and retinal motion, exhibit very distinct spatial and temporal signatures. Although these two cues are synchronous in naturally moving objects, disparity-based and velocity-based mechanisms can be dissociated experimentally. We therefore investigated how the relative contributions of these two cues change across a range of viewing conditions. We measured direction-discrimination sensitivity for motion though depth across a wide range of eccentricities and speeds for disparity-based stimuli, velocity-based stimuli, and “full cue” stimuli containing both changing disparities and interocular velocity differences. Surprisingly, the pattern of sensitivity for velocity-based stimuli was nearly identical to the full cue stimuli across the entire extent of the measured spatiotemporal surface, and both were clearly distinct from those for the disparity-based stimuli. These results suggest that for direction discrimination outside the fovea, 3D motion perception primarily relies on the velocity-based cue with little—if any— contribution from the disparity-based cue.

Keywords: Anticorrelation, Changing disparity, Interocular velocity difference, Motion in depth, Stereomotion
Introduction

Most research on motion perception has focused on two-dimensional (2D) frontoparallel motion (which is, of course, the easiest to generate on computer-driven displays). The neural computations that support the perception of motion through a more realistic 3D environment are considerably less well-understood. Here, we explore the relative contributions of two fundamental binocular cues to 3D motion (sometimes referred to as motion-in-depth or z-axis stereomotion): one based on changes in binocular disparity over time, and one based on interocular velocity differences. We investigated the conditions under which the visual system might preferentially rely upon one cue over the other to shed light upon how 3D motion is processed by the brain.

The first potential cue, changing disparity over time (CD), is classically assumed to be the pure, fundamental signal for binocular 3D motion perception (Cumming and Parker 1994; Gray and Regan 1996; Regan and Gray 2009). The CD cue can be computed simply by taking the time derivative of horizontal binocular disparity (Figure 1, upper). Although this cue is sufficient for the generation of 3D motion percepts, it is important to note that it is based on disparity signals, which are greatly degraded at far eccentricities and high temporal frequencies (e.g. Blakemore 1970; Julesz 1960; Norcia and Tyler 1984; Westheimer and Truong 1988). We therefore hypothesized that the CD cue might preferentially support 3D motion percepts for slowly-moving stimuli near fixation.

The second potential cue, the interocular velocity difference (IOVD), has been proposed to also contribute to 3D motion perception (Harris et al. 2008; but see Regan and Gray 2009). The IOVD cue can be computed by comparing monocular velocity
signals of the stimulus projections on each of the two retinas (Figure 1, lower). This computation exploits the fact that motion through depth projects different (and often opposite) motion upon the left and right eyes retinae. Although the IOVD cue has proven challenging to study in isolation, one might expect it to be perceptually robust across eccentricity and speed, given that it is based on monocular motion signals (e.g. McKee and Nakayama 1984; Wright 1987). We therefore hypothesized that the visual system may rely more heavily on the velocity-based IOVD cue to 3D motion for faster motions outside the fovea—the very conditions under which the disparity-based CD cue might be expected to be relatively ineffective.

In order to compare the relative contributions of the CD and IOVD cues, we assessed sensitivity with a common task and metric across all conditions. To do this, we asked observers to discriminate the direction of motion of a frontoparallel plane of dots moving towards or away through depth. We varied stimulus strength in a manner akin to prior manipulations of 2D motion coherence (i.e., the proportion of coherently-moving dots on the signal plane relative to noise dots moving incoherently through depth; Burr and Santoro 2001; Lankheet et al. 2000; Newsome and Paré 1988; Watamaniuk et al. 1989). FULL binocular cue displays contained signal plane dots moving towards or away from the observer in depth (and hence contained both the CD and IOVD cues); CD displays contained 1-frame lifetime signal plane dots to remove any structured monocular velocity information (Julesz 1971); IOVD displays contained binocularly-anticorrelated dots to greatly reduce the ability of the visual system to compute disparity (Harris and Rushton 2003; Rokers et al. 2008).
We measured direction discrimination thresholds for FULL, CD, and IOVD displays across a wide range of speeds and eccentricities. At low speeds, eccentricity degraded sensitivity similarly for all three display types. At faster speeds, however, the relative contributions of the two cues became much more distinct. For the FULL and IOVD stimuli, sensitivity showed a distinct bandpass speed tuning, with a peak at relatively fast speeds (~2.0°/sec monocular velocity). In contrast, sensitivity for the CD stimuli was clearly lowpass; in fact, direction discrimination for these stimuli became impossible at moderately high speeds, regardless of eccentricity.

These results imply that the visual system can compute 3D motion primarily from the IOVD cue across a majority of the visual field for a broad range of speeds, and may rely more strongly on the CD cue for direction discrimination at very slow speeds at or near fixation. Although our findings may at first seem to run counter to previous demonstrations that the CD cue is sufficient to explain 3D motion sensitivity (Cumming and Parker 1994; Gray and Regan 1996), our demonstration of the possible primacy of the IOVD cue bolsters a growing, recent literature demonstrating dissociable contributions of the IOVD cue (for a more detailed review, see Discussion, and also Harris et al. 2008; Regan and Gray 2009). Furthermore, this relative primacy of IOVDs may result from the fact that our stimulus parameters, discrimination task, and sensitivity metric are better thought of extending the methods used to study 2D / frontoparallel motion processing into the third spatial dimension— as opposed to being temporally-dynamic extensions of the methods used to probe disparity-based mechanisms. Finally, the methods introduced in this paper could easily be generalized for use in future neuroimaging and electrophysiological studies of interocular velocity.
differences and 3D motion perception.
Methods & Materials

Observers

Data were collected in three experienced psychophysical observers (three of the authors, males aged 26–44), all with good stereopsis, and normal or corrected-to-normal vision. Experiments were undertaken with the written consent of each observer, and all procedures were approved by the UT-Austin Institutional Review Board. A total of 156,240 trials were collected across the 3 observers.

General procedure

We measured the ability of observers to discriminate the direction of motion through depth (directly towards or away from the observer) for three different types of motion cue stimuli (FULL, IOVD, or CD; described more fully below), which contained different combinations of the two primary binocular cues to 3D motion. Performance for each motion cue type was measured in a fully crossed design manipulating stimulus speed and eccentricity. Across all conditions, observers viewed a frontoparallel plane of random dots moving towards or away from them through a 3D volume of noise dots. The signal plane started at a random depth within the volume and moved at a smooth, constant speed either towards or away from the observer, wrapping from front-to-back (or back-to-front) to complete one full cycle through the depth volume. The random starting location ensured that the starting, ending, or average disparity (or the time of the wrap) of the plane could not be used to do the task.

On each trial, observers viewed the display and reported the perceived direction of motion through depth (towards or away) with a left or right mouse click. The response
triggered the next stimulus presentation with a minimum delay of 200 ms between trials. Observers were instructed to report their percept of the smoothest motion through depth throughout the experiment, disregarding the perceived jump through depth that occurred when the signal plane wrapped. In all conditions, signal dots were relocated in the x-y plane upon z-axis wrapping to minimize apparent motion during this brief change in signal plane depth. No feedback was provided; observers could thus concentrate on the smooth motion and not train themselves (consciously or not) on other potential cues, such as the jump due to wrapping, or a utrocular identification combined with a monocular direction judgment.

For each combination of cue type, speed, and eccentricity, we measured the proportion of correct responses as a function of motion coherence, defined as the relative percentage of signal to noise dots. Psychophysical thresholds (84% correct) were then estimated for each condition from a fitted logistic function, and sensitivity was expressed as inverse threshold coherence \((\text{coh}^{-1})\). More details about the stimuli and experimental design are provided below.

General stimuli

Observers stereoscopically viewed moving random dot displays in which 80 dark (0.4 cd/m²) or light (129.7 cd/m²) binocularly paired dots were presented on a mid-gray (56.0 cd/m²) background (Figure 2, panel A). In each monocular half-image, half the dots were dark and half the dots were light. Individual dots subtended a visual angle of 15 arcmin (0.25°) and were anti-aliased to achieve subpixel position accuracy. Dot density and luminance values were not selected on a single fundamental principle, but
rather a balance of several factors. First and foremost, we wanted our stimuli to be comparable to many 2D frontoparallel motion studies that used motion coherence as a manipulation (i.e. number of dots on each display frame [80], stimulus area [126 deg²], and variable dot lifetimes, resulting in dot densities of 0.7–2.5 dots per degree and 3.7–11.5% dot coverage). We emphasize that many of our stimulus parameters may differ from prior studies of static disparity processing and stereomotion (e.g., lower density). Second, we chose parameters that were within hardware limitations (luminances that fit within the maximum linearized contrast range; dot numbers that did not overstep the available computational power to relocate each signal/noise dot on every display frame). Third, we chose parameters that allowed motion coherence manipulations to drive performance from chance to near perfect under all stimulus conditions.

Observers fixated a small central square (0.5°) with horizontal (black) and vertical (red) nonius lines. A single dot (bright, 0.25° diameter, 0 arcmin disparity) was placed in the fixation square to provide subjects with an object of fixation, and prevent fixation drift towards endpoints of the fixation square or nonius lines. To further aid in proper binocular alignment, four stationary dots (dark, 0.5° diameter, 10.6° eccentric, 0 arcmin disparity) were located beyond the stimulus on horizontal and vertical axes of each monocular half-image. We used a sparse set of reference dots to limit extraneous relative disparity cues at the outer edges of the display (Andrews et al. 2001) while still providing eccentric visual anchor points.

Manipulations of eccentricity and speed
To examine how 3D motion sensitivity varies across the visual field, stimuli were presented within three different eccentricity ranges: 3-7°, 7-11°, and 11-15° from fixation (Figure 2, panel B). The Near eccentricity stimulus consisted of a continuous annular region spanning 3-7° from fixation. From this annulus, four 90° annular segments were then shifted outward in oblique directions (45°, 135°, 225°, 315°) to Middle and Far eccentricities of 7-11° and 11-15°, respectively. Thus the number and density of signal dots were held constant across the eccentricity conditions. Stimulus disparities were constrained to a volume spanning ±0.6° of disparity (i.e., along the z-axis) from the plane of fixation. At our 70 cm viewing distance, this corresponded to a total (front to back) simulated depth interval of 16 cm. This z-axis depth, and hence the overall stimulus volume, remained constant across all conditions.

To examine how 3D motion sensitivity varies with stimulus speed, stimuli were presented at 5 different speeds (where we define “speed” as the monocular angular speed in each eye): 0.3, 0.6, 0.9, 1.8, and 2.7°/sec (Figure 2, panel C). Because we describe speed in °/sec per eye, and the monocular velocities were always opposite in the two eyes, one can simply multiply the monocular speeds times the number of eyes (2 in our case) to calculate the equivalent disparity change in °/sec.

Across all speeds, the total stimulus excursion through depth in a trial was always one full cycle through the stimulus volume (with a single “wrap” occurring on all trials, except those few in which the signal dots happened to begin at the very front or back of the volume). Stimulus presentations containing exactly one full cycle with a single depth wrap were chosen so that neither average depth (or disparity) over a trial, nor instantaneous depth (or disparity) at any point in the trial (e.g., starting or ending), could
be used to perform the task. Given this one-cycle constraint, the resulting stimulus
durations ranged from 2 s at the slowest speed to just over 0.2 s at the fastest speed
(corresponding to 120 and 13 video frames, respectively). The decision to fix the total
depth traveled (and not the overall duration of motion) was supported by three factors.
First, our main inferences are based on comparisons of sensitivity across FULL, CD,
and IOVD conditions, which effectively balances duration across the comparisons of
interest. Second, given that we observed peak sensitivity in the main experiment at
rather fast speeds (and hence, at short durations), we are confident that shorter
durations per se did not strongly impair performance. Third, exploratory manipulation of
duration at the medium and high speeds revealed only a very small effect of stimulus
duration which, in any event, was balanced across motion cue conditions in the main
experiment.

**Motion cue conditions: FULL, IOVD, and CD**

Three motion cue stimuli were employed: FULL, IOVD, and CD. All three stimulus
types contained a single plane of signal dots moving towards or away from the observer
through depth, in the presence of noise dots (described further below). Figure 3
schematizes the three motion cue stimulus types.

The FULL stimulus consisted of a moving random dot stereogram in which
binocularly-paired signal dots moved in opposite directions in the two eyes. The signal
dots thus contained both the IOVD and CD cues to 3D motion. Signal dots moved at
constant monocular speeds ranging from 0.3 to 2.7 degrees per second, corresponding
to 3D motion speeds of about 8 cm/sec to about 74 cm/sec at our 70 cm viewing
distance. Perceptually, the FULL stimulus resembled a fixed set of dots, not unlike a
group of flying insects in a fronto-parallel plane, moving directly towards or away from
the observer in synchrony. All dot pairs (signal and noise) were binocularly correlated
(i.e., were of the same contrast polarity across the eyes); noise dots are described in
detail later, and followed identical motion patterns across all three conditions.

The IOVD stimulus was identical to the FULL stimulus, except that all of the
dot pairs were binocularly anticorrelated: each dark dot in one eye was paired with a
corresponding bright dot in the other eye. Anticorrelation has been shown to disrupt
static disparity mechanisms (Cogan et al. 1995; Cumming et al. 1998; Neri et al. 1999),
while maintaining monocular velocity information (Harris and Rushton 2003), and
hence, retaining IOVDs in the presence of greatly-degraded disparity-based signals. In
addition, our group has previously used sparse, anticorrelated dot displays to isolate the
contribution of the IOVD cue (Rokers et al. 2009; Rokers et al. 2008). Perceptually, the
IOVD stimulus is phenomenologically rather interesting (although this is not the direct
subject of the current paper). Because static disparity information is greatly
compromised, one generally has the sensation of “something” moving towards or away
from the eyes, with neither a firm sense of a distinct plane of dots present in space, nor
a sense of the position-in-depth of these moving elements.

The CD stimulus was identical to the FULL stimulus, except that the signal dots
were randomly replotted in new x-y positions on the signal plane upon each screen
refresh (i.e. at 60 Hz). This reploting removed coherent monocular velocity information,
while preserving steadily-changing disparity information (Braddick 1974; Cumming and
Parker 1994; Julesz 1971). For the CD stimulus, the rate of disparity change matched
that of the \textit{FULL} and \textit{IOVD} stimuli. Perceptually, the \textit{CD} stimulus resembles a plane of TV snow moving through depth towards or away from the observer.

Manipulation of 3D motion coherence

3D motion coherence, defined as the ratio of signal dots to noise dots, was randomly varied on a trial-by-trial basis according to the method of constant stimuli. We determined direction discrimination thresholds in units of motion coherence from the resulting psychometric functions.

At the beginning of each trial, the number of signal dots (as determined by the coherence level pseudorandomly drawn for that trial) was selected (out of the 80 total dots). The remainder were designated as noise dots. The signal dots were randomly positioned on a single fronto-parallel plane moving towards or away from the observer. This signal plane began at a random position in depth, and moved throughout the entire depth range (wrapping when necessary), ending in the same position in depth as it began. Upon wrapping, each signal dot was assigned a new random \textit{x-y} position in addition to moving to the opposite end of the volume on the \textit{z}-axis. For a given speed, this implied that the signal dots followed a uniform distribution of lifetimes between 1 frame and the number of frames in a trial at that speed (i.e. 120 frames at the slowest speed, and 13 frames at the fastest speed). For any single trial, there were thus two signal dot lifetimes, one pre-wrap and one post-wrap, which summed to the total number of frames in that trial.

We designed our noise dots to satisfy multiple, somewhat competing demands: 1) remaining constant (statistically) across all conditions (cue type by speed by
268 eccentricity); 2) being capable of effectively masking the motion through depth of the
269 signal plane across all conditions (and thus allowing us to measure psychometric
270 functions spanning the full range of possible performance across all conditions); and 3)
271 effectively matching the spatiotemporal properties of the signal dots per se across all
272 conditions— that is, we did not want the signal dots themselves to “pop out” in any
273 condition due to either flashing on for one frame (CD condition), or persisting for
274 multiple frames that varied with speed (IOVD and FULL conditions).

To satisfy these constraints, we used variable-lifetime noise dots that
approximated random walks through depth. Their instantaneous (frame-to-frame)
velocity was variable along the z-axis, but was constrained to be less than or equal to
the signal dot velocity through depth (their x-y positions were fixed throughout each
lifetime). Each noise dot was assigned a random lifetime ranging from 1 to 12 frames
(16.7 to 200 ms) from an inverse squared distribution. Specifically, the probability of a
noise dot having a given lifetime, L (in frames) was proportional to $1/L^2$, with $L = 1, 2, ..., 12$. At the expiration of a noise dot’s lifetime, it was randomly repositioned within the
stimulus volume and assigned a new lifetime from this distribution (Our noise dots can
therefore be considered a hybrid of the random-position and random-walk same-
selection dot noise described in Scase et al. 1996). At any given time, then, the noise
consisted of a mixture of transient and persistent dots, like the CD signal dots on the
one hand and the IOVD and FULL signal dots on the other, but with the distribution
favoring the presence of short lifetime dots.

The distribution of noise dots included a higher proportion of shorter lifetimes,
because we reasoned that transient (flashing) elements are better at masking persistent
elements than vice versa. Noise composed of mostly transient elements would be expected to mask both transient and persistent signals (which is important, given that our CD condition contained transient signal due to the single-frame lifetimes of the signal dots, and our FULL and IOVD conditions contained more persistent signal due to the longer signal dot lifetimes). This argument can also be appreciated in the Fourier domain: transient noise elements will have broadband power in the temporal frequency domain, and thus would cover the spectral range of signal across all conditions; persistent noise elements would be better localized in the temporal frequency domain, and thus would not so broadly span the spectral range of signals of interest. Finally, pilot observations confirmed that noise from an inverse squared distribution yielded good subjective degradation of the signal plane’s direction of motion at high noise levels, provided good masking of the dots themselves across conditions (i.e., none of the signal dots popped out in any condition), and drove performance from an upper asymptote to chance levels for all motion cue types. We emphasize that these decisions allowed us to use noise dots that had the same distribution of motions and were subjected to the same manipulation of coherence (and hence, the same sensitivity metric) across all conditions; a crucial component that enabled direct comparison of sensitivities across cue conditions.

Experimental design

We measured observers’ ability to discriminate the direction of motion through depth (towards or away) across a range of 3D motion coherence levels (0, 3, 6, 12, 24 and 50% coherence) using the method of constant stimuli. We employed a fully-crossed
design containing all combinations of motion cue type (FULL, CD, IOVD), eccentricity (3-7°, 7-11°, and 11-15°), and speed (0.3, 0.6, 0.9, 1.8, and 2.7°/sec·eye). Within each run, we measured percent correct as a function of motion coherence for a single combination of these factors (resulting in a single estimate of the psychometric function).

Motion coherence was pseudo-randomized across trials within a run. Each run consisted of 40 trials per coherence level, resulting in 240 trials total. The order of runs was randomized. Each observer completed five runs of the 0.6, 0.9, and 2.7°/sec·eye speeds for each motion cue/eccentricity combination, and three runs for each of the 0.3, and 1.8°/sec·eye speeds. This resulted in either 720 or 1200 trials per observer per condition, and just under 45.4 kilotrials per observer for the main experiment. Two control experiments addressing position-in-depth and 2D motion discrimination (see Discussion) contributed an additional ~20.2 kilotrials across the same three observers.

Apparatus and display

To investigate 3D motion perception at large eccentricity with high temporal accuracy, stimuli were presented on a calibrated 42° LCD Display (Sharp LC-42D64U; 60Hz progressive scan, 1920 x 1080 pixel resolution) viewed through a mirror stereoscope with a concomitantly large field of view. The monitor was driven by Mac Pro computer and an NVIDIA GeForce 8800 GT video card.

Luminance calibrations were done at 10 locations across the display using an OptiCal photometer (CRS Ltd.). We verified that gamma correction tables for each location were the same across the entire luminance range, allowing all stimuli to be
accurately presented using a single linearizing gamma correction table. All 10 curves were nearly identical within a scale factor, demonstrating a high degree of spatial luminance homogeneity, nearly perfect contrast homogeneity, and the ability to implement good luminance linearization with this LCD display.

We achieved spatial luminance homogeneity by making internal display adjustments. Specifically, the duty-cycle of the LCD backlight was maximized by setting the “backlight” adjustment to the maximum level, while setting the “brightness” adjustment to the minimum level to maintain a comfortable luminance range. This provided the most homogenous display luminance, leaving at most a 10% residual luminance variation that was almost entirely constrained to the extreme edges of the display (where stimuli were not presented).

Display timing was verified using a fast photocell (Model 10AP, UDT Sensors Inc., Hawthorne, CA) and an oscilloscope. We used a splitter-cable so that we could measure the VBL signal directly while simultaneously measuring the instantaneous luminance on the monitor. Pixel updates were constant at 60 Hz and at a fixed phase relative to the VBL signal generated by the video card. The white-to-black transition was marginally faster than the black-to-white, with the later showing slight exponential characteristics. Nonetheless, the display easily followed repeating black-white and black-gray-white-gray cycles on a frame-by-frame basis at 60 Hz. Although the display was slow by modern CRT standards, it provided reliable timing of display updates, as well as luminance output that was easily linearized. Furthermore, all in-monitor enhancement modes (e.g. motion enhancement, dynamic contrast adjustment, etc.) were disabled, as they could yield undesirable display artifacts. In short, our
measurements suggested that our particular LCD was appropriate for use in our experiments; it remains to be seen whether similar results can be attained by similar adjustment and calibration using other LCDs.

Monocular half-images were presented separately on the left and right halves of the display, with a septum and various baffles positioned to assure that each half-image was only visible to the corresponding eye. Viewed through the 70 cm optical path length of the stereoscope, each monocular half-image subtended 30° of visual angle. This display arrangement was selected over traditional dual-display stereo or shutter goggle apparatus because it provided both perfect temporal synchronization between the two eyes and complete isolation of the monocular half-images. All stimuli were generated using the Psychophysics Toolbox (Brainard 1997) and MATLAB (2007a, The Mathworks Inc., Natick, MA).

Data analysis

For each condition (motion cue type x eccentricity x speed), we combined data across multiple runs for each subject, and fit a logistic psychometric function using the psignifit toolbox version 2.5.6 for Matlab (http://bootstrap-software.org/psignifit/). Threshold was defined as the 3D motion coherence yielding 84% accuracy. We bootstrapped confidence intervals (equivalent to ± 1 SEM) about these thresholds by resampling (with replacement) the binomial responses from each subject to create 500 repetitions of the experiment, fitting a psychometric function to each resampled experiment, and identifying the central 68% of the values. In instances where observers were unable to discriminate 3D motion direction we assigned a threshold coherence of
100% (pinning the thresholds at the maximum physically-realizable level was preferable to simply discarding those data, but our conclusions do not change if these are instead omitted). Across a total of 135 sensitivity estimates (3 observers, 5 speeds, 3 eccentricities, 3 cues) this occurred only 7 times, and was isolated to high speed CD stimulus conditions.iii

We applied a similar resampling approach when fitting the eccentricity and speed tuning curves (Figures 4 and 5). We plotted the median fit parameters (after checking that the median values were very similar to the means), because the medians had the advantage of being robust to the occasional extreme values that can arise in a small number of fits across the very large number of resampled datasets.
Results

Recall that, in all conditions, observers simply judged whether a plane of signal dots was moving towards or away from them. On each trial, stimuli were presented at one of six different motion coherence levels (0, 3, 6, 12, 24 and 50% coherence). Sensitivity (inverse 3D motion coherence threshold) was then estimated for each combination of eccentricity (Near, Middle, and Far), speed (0.3, 0.6, 0.9, 1.8, and $2.7^\circ$/sec·eye), and motion cue type ($FULL$, $CD$, and $IOVD$).

In each of the following sections, we first establish a baseline for 3D motion discrimination by describing the performance in the $FULL$ cue condition, which contained both disparity- and velocity-based 3D motion cues ($CD$ plus $IOVD$). We then compare the results to each isolated cue condition ($CD$, $IOVD$). In order to explore how sensitivity varied across the entire eccentricity-speed space, we address the results from three perspectives in the following three sections: the effects of eccentricity at different speeds, the effects of speed at different eccentricities, and finally, the full spatiotemporal (speed x eccentricity) sensitivity surface. (Unless otherwise noted, data points in the following figures represent mean sensitivity across all three observers).

Effects of eccentricity at different speeds

Here, we first describe the data as functions of eccentricity measured at different speeds. Figure 4 (left) shows the eccentricity effect on $FULL$ stimulus sensitivity across the range of 3D motion speeds. Increases in stimulus eccentricity caused a decrease in direction discrimination sensitivity in a speed dependent manner. The strength of this effect can be determined from the slope of linear fits at each stimulus speed. Sensitivity
to slower 3D motion (darker curves and symbols) was strongly diminished by increasing eccentricity (slope of $-0.83\ coh^{-1}\cdot\deg^{-1}$ at slowest speed, 0.3°/sec·eye), while sensitivity for faster 3D motions (lighter gray curves and symbols) was not strongly affected by eccentricity (slope of $0.02\ coh^{-1}\cdot\deg^{-1}$ at fastest speed, 2.7°/sec·eye). The eccentricity effect was smaller at the highest speed primarily due to a large improvement in sensitivity at the Far eccentricity; sensitivity at the Near eccentricity did not change as much. In summary, eccentricity reduced direction discrimination sensitivity in the FULL condition, but did so strongly for slower speeds, and less so for faster speeds.

Sensitivity to the IOVD stimulus (Figure 4, middle) showed a similar pattern of eccentricity dependence as the FULL stimulus. Sensitivity to IOVD motion was strongly affected by eccentricity at slower speeds (slope of $-0.60\ coh^{-1}\cdot\deg^{-1}$ at the slowest speed, 0.3°/sec·eye), and was less affected by eccentricity at higher speeds (slope of $-0.16\ coh^{-1}\cdot\deg^{-1}$ at the fastest speed, 2.7°/sec·eye). As with the FULL stimulus, the change in eccentricity function across speeds was more due to changes in sensitivity at Far eccentricity than at Near eccentricity. So, just as for the FULL condition, eccentricity effects for the IOVD condition were larger at slower speeds, and smaller at higher speeds.

Sensitivity to the CD stimulus (Figure 4, right) followed a strikingly different pattern than that seen from the FULL and IOVD stimuli. Larger eccentricity did yield poorer performance in general, and eccentricity effects were steepest at the slower speeds. At the higher speeds, however, performance was very poor regardless of eccentricity. At the highest speed, in fact, only one observer was able to reliably perform above chance. In other words, the lack of an eccentricity effect at high speeds is probably best thought
of not as a uniform sensitivity across eccentricities per se, but rather as an overall lack of sensitivity of the CD system to stimuli moving rapidly in depth.

In summary, the effects of eccentricity on FULL and IOVD sensitivity were quite similar, and showed similar dependencies on speed. In contrast, CD sensitivity followed a very different pattern of interactions between eccentricity and speed. The nature of this interaction becomes more clear in the next section.

Effects of speed at different eccentricities

It is perhaps more illuminating to consider the same sensitivity data as speed-tuning curves measured at different eccentricities, as shown in Figure 5 (note the log speed axis). For the FULL stimulus (Figure 5, left) similar bandpass speed tuning functions were evident at all stimulus eccentricities; peaking near the higher speeds measured. At Near eccentricity, sensitivity fell off sharply for speeds faster than 1.8°/sec·eye. Because of the bandpass appearance, we fit the data with Gaussian functions. The peak of the fitted Gaussian for the Near data was at 1.09°/sec·eye, with a full width at half the maximum height (FWHM) of 3.92°/sec·eye. At Middle and Far eccentricities, peak sensitivity shifted towards even higher stimulus speeds (Middle, 1.94°/sec·eye; Far, 2.27°/sec·eye). The stronger effects of eccentricity at low speeds had the effect of narrowing the bandpass speed tuning at Middle and Far eccentricities (FWHM, Middle, 3.23°/sec·eye; Far, 3.12°/sec·eye).

The pattern of IOVD sensitivity (Figure 5, middle) was again strikingly similar to the FULL stimulus. Sensitivity was bandpass for speed with a peak near the higher speeds. This bandpass tuning was evident at all eccentricities, and IOVD peak sensitivity also
followed a similar pattern as for the **FULL** stimulus. Peak speeds were 1.54, 2.06, 2.06°/sec·eye for Near, Middle, and Far eccentricities respectively, and likewise became more sharply speed-tuned with increasing eccentricity (FWHMs were 3.49, 3.23, and 2.91°/sec·eye, respectively). The similarity between the **FULL** and **IOVD** patterns of speed tuning is further supported by a point-by-point comparison of the two tuning functions; revealing that twelve of the fifteen points fall within 68% (± 1 SEM) confidence intervals of one another.

The **CD** condition demonstrated an altogether different pattern of sensitivity from those seen in the **FULL** or **IOVD** conditions (Figure 5, right). All of the speed tuning curves were clearly lowpass, falling off precipitously with increasing speed (i.e., only one observer was able to discriminate the highest speed **CD** stimuli at accuracies above chance). In fact, the difference in the tuning curves was pronounced enough that we couldn’t fit the **CD** data satisfactorily with Gaussians given reasonable parameter values, and we therefore fit them with straight lines. Fitted linear slopes were all strongly negative (in units of sensitivity, coh$^{-1}$ per °/sec: Near, −3.28; Middle, −2.09; Far, −1.10).

As with the eccentricity effects in the previous section, the slopes of the speed effects on the **CD** stimulus became less steep at far eccentricities simply because accuracy levels fell towards chance.

In summary, the analysis of speed tuning reveals that **IOVD**-based performance closely mirrored that of full-cue performance. **FULL** and **IOVD** speed tuning was bandpass, with a peak at relatively brisk 3D motion speeds. In contrast, **CD**-based performance exhibited dramatically different speed tuning that cannot account for most of the full-cue sensitivity. Instead, **CD** speed tuning was lowpass, and fell off steeply
near the speeds at which the IOVD and FULL conditions revealed maximal sensitivity. Although the stimuli and tasks were different, these speed tuning results are qualitatively consistent with the temporal frequency tuning results of Shioiri et al. (2008).

**Speed by eccentricity (SxE) sensitivity surface**

The preceding sections show that IOVD sensitivity is very similar to FULL sensitivity, and that CD sensitivity follows a rather different pattern, regardless of whether the data are viewed as slices of constant speed or constant eccentricity. The overall shape of the data for the three conditions can be appreciated more thoroughly in spatiotemporal sensitivity surface-contours that span both eccentricity and speed (Figure 6). The top row depicts the sensitivity surface for each motion cue condition; stimulus eccentricity by speed (SxE) on the x- and y-axes respectively, and direction discrimination sensitivity on the z-axis (height). The surface sensitivities are also projected down to contour maps on the z=0 plane. Bandpass speed sensitivity can be seen in both FULL and IOVD conditions, as can the weaker eccentricity effects closer to the best speed. In contrast, the CD condition shows a distinct pattern of roughly linear sensitivity falloff as speed and eccentricity increase.

The bottom row of surface-contour plots (Figure 6) show the differential sensitivity surfaces generated by subtracting the spatiotemporal sensitivity surfaces of each motion cue. Positive values are shown in the same warm color map as the original sensitivity surfaces, while negative values are shown in cool colors extending below the contour map. The left two panels of the bottom row of surface-contour plots show differential sensitivity surfaces generated by subtracting each isolated cue surface from
the FULL stimulus surface (i.e., FULL−IOVD, FULL−CD). The surface goes positive (above the contour map) where sensitivity is better in the FULL stimulus, and the surface goes negative (below the contour map) where sensitivity is better in an isolated cue condition (IOVD or CD). Bootstrapped p-values for the difference surfaces are shown in Table 1.

The FULL−IOVD surface shown in the lower left panel is nearly flat, indicating that the sensitivities are nearly identical at each combination of speed and eccentricity tested. In fact, only two of the fifteen points in the FULL−IOVD surface are significantly different from zero.

The FULL−CD surface shown in the lower middle panel, however, is distinctly not flat. For faster speeds, observers are much more sensitive to the FULL stimulus than the CD stimulus, indicating that the (lack of) CD sensitivity does not limit the observers' performance. For the slowest speeds, on the other hand, observers actually performed better in the CD condition than they did in the FULL condition, indicating that: (1) observers were unable to fully exploit the changing disparity information when interocular velocity differences were also present; and/or (2) there was richer changing-disparity information present in the CD stimulus than in the FULL stimulus (perhaps due to the faster temporal refresh rate of the signal dots in the CD condition). Regardless of which possibility is at work (both could be), this superiority of the CD condition for slow speeds demonstrates that the larger-scale dissimilarity between the FULL and CD conditions is not simply because the CD stimulus did not contain strong signal.

The lower right surface-contour plot shows the differential sensitivity surface generated by subtracting the CD sensitivity surface from the IOVD sensitivity surface
Central role for the velocity-based 3D motion cue

$(IOVD-CD)$. Obviously, this difference surface closely resembles the $FULL-CD$ surface, and it can also be thought of as a visualization of the relative utility of the two (isolated) cues in our experimental conditions.

The ability of each isolated cue to predict the $FULL$ sensitivity is shown in Figure 7. Individual observer sensitivities for each isolated cue condition ($y$-axis) are plotted as function of $FULL$ stimulus sensitivity ($x$-axis). Each data point, in other words, shows a pair of thresholds for a particular combination of observer, speed, and eccentricity. The left scatter plot shows a high level of correlation between the $IOVD$ and $FULL$ stimulus sensitivities ($r^2 = 0.75$). This strong correlation on an individual subject level suggests that $FULL$ sensitivity can be accurately predicted simply by measuring an individual’s corresponding $IOVD$ sensitivity in isolation. However, this relationship does not hold for the right scatter plot of $CD$ vs $FULL$ sensitivities ($r^2 = 0.05$). Thus, knowing an individual’s $CD$ sensitivity does not provide much information to predict how well the observer will be able to discriminate the direction of realistic (full-cue) motions through depth. Interestingly, Watanabe et al. (2008) found a very similar pattern of results when comparing their novel clinical test for motion through depth with a standard static stereo test (Titmus).
Discussion

Our experiments revealed distinctly different patterns of sensitivity to the changing disparity (CD) and interocular velocity difference (IOVD) cues for 3D motion direction discrimination. Sensitivity to the CD cue was highest at the shortest eccentricities and the slowest speeds. Increasing either speed or eccentricity had strong and independently deleterious effects on CD sensitivity. Sensitivity to the IOVD cue, on the other hand, was lowest at the nearest eccentricities and the slowest speeds. Increasing speed led to greater IOVD sensitivity and mitigated the effects of eccentricity. Overall, the pattern of IOVD sensitivity was nearly identical to the pattern of FULL sensitivity—sensitivity for stimuli containing both the CD and the IOVD cues—across the entire eccentricity-speed space. In contrast, the pattern of CD sensitivity across the eccentricity-speed space was markedly different than the FULL pattern. Although these patterns of relative sensitivity were not as straightforwardly dependent on speed and eccentricity as we initially hypothesized, they did reveal a surprisingly close correspondence between IOVD and FULL sensitivity across the majority of the wide spatiotemporal range we investigated. We therefore conclude that, at least outside the fovea, the human visual system can rely primarily on interocular velocity differences—not changing disparities—to discriminate the direction of 3D motion.

Distinguishing the contributions of the CD and IOVD cues

The differential patterns of sensitivity across the eccentricity-speed space provide clear evidence for a dissociation of the CD and IOVD cues. The disparity-based cue functions best at slow speeds and nearer eccentricities, while the velocity-based cue
exhibits bandpass sensitivity for higher speeds, with muted eccentricity effects. These
distinct patterns of sensitivity suggests that the CD cue may be useful for slow-moving
(para-)foveal 3D motions, and that the IOVD cue may be more useful for faster and
more peripheral 3D motions.

Our results also provide additional support to the notion that the IOVD cue can be
experimentally isolated. We addressed the effectiveness of anticorrelation in isolating
the IOVD cue by conducting two experimental controls: (1) to determine whether useful
position-in-depth information could be contributing to 3D motion sensitivities in the \textit{IOVD}
stimuli, and (2) to rule out the possibility that similarities in \textit{FULL} and \textit{IOVD} sensitivities
could result from simple monocular motion sensitivities.

Unlike the CD cue, the IOVD cue cannot be perfectly isolated in principle. Although
contrast anticorrelation is not a perfect form of isolation, we believe it is currently the
most effective method of removing useful disparity information from a stimulus (and thus
strongly biasing the stimuli in favor of IOVD mechanisms). There are two primary
concerns that are often raised about these anticorrelated stimuli. The first is that, in an
anticorrelated stimulus, there are many potential “false” matches of the same contrast
polarity (e.g. a given white dot in one eye could conceivably be matched with any other
white dot in the other eye’s image, even though the experimenter had specified that the
“corresponding” dot be of opposite contrast polarity). There several reasons why it is
unlikely that these potential unintended matches influenced the data in a meaningful
way. First, given the dot density of our stimuli, any unintended matches would have a
large disparity, usually both horizontal and vertical, and these would vary from match to
match at any given time (thus, a vertical vergence movement, for example, couldn’t
suddenly create a large number of plausible, predominantly horizontal disparities). As the effective signal of a binocular element falls off with the overall disparity, regardless of how “effective signal” is determined, the majority of these matches would not be a very effective stimulus (Blakemore 1970; Cormack et al. 1993; Prince et al. 2002; Stevenson et al. 1994; Stevenson et al. 1992). Second, at threshold values of coherence, the vast majority of these potential matches for a given signal dot in one eye would be with a noise dot in the other eye. The match would thus result in an additional binocular noise dot (or at least a much noisier signal dot). If performance in the IOVD conditions were based on these spurious disparity signals, then it would be quite poor indeed. Rather, we find that the data from the IOVD conditions tracks that from the FULL conditions, despite the huge difference in the quality of the disparity signals. Even under the most general assumptions about the presence of an unintended CD signal in our IOVD stimulus, our data are inconsistent with this explanation.

The second potential problem with anticorrelated stimuli is that each dot contains two vertical (on average) edge segments of opposite contrast polarity, so it is conceivable that, for example, the left edge of a white dot could be paired with the right edge of the corresponding black dot. However, these matches would be between regions of different overall (signed) contrast and local mean luminance, and it is known that unequal contrast between corresponding elements in the two eyes impairs stereopsis very dramatically—much more so than an overall contrast reduction (Cormack et al., 1991).

However unlikely, these concerns are valid in principle, so we addressed them by conducting a control experiment where observers performed coarse position-in-depth
judgments (2AFC discrimination of the signal dots as near or far relative to the plane of fixation), while viewing stimuli moving at 0.9°/sec-eye under each motion cue/eccentricity combination. The position-in-depth judgments were performed on stimuli nearly identical to those of the main experiment except that one eye’s image was flipped horizontally, so that the motion was in the same direction in the two eyes. This created moving (frontoparallel) stimuli with a fixed, random disparity offset in each trial while still maintaining the same monocular motions as in the main experiment.

When observers were asked to discriminate whether this plane of dots was near or far relative to the zero-disparity plane of fixation, performance with the anticorrelated (IOVD) stimulus was so poor that we could not measure psychometric functions and report thresholds. We therefore did the following analysis: For each eccentricity and observer, we noted the 84% correct threshold 3D motion coherence for the FULL stimuli. We then measured position-in-depth performance for each observer and eccentricity at this coherence for each stimulus type. So if, for example, there was rich disparity information extracted from the anticorrelated stimulus (due to an early rectifying nonlinearity, say), then performance in the disparity-based position-in-depth task should be close to 84% correct. If, on the other hand, the anticorrelated stimulus does indeed greatly reduce the available disparity information, performance should be much poorer than 84% correct.

Panel A of figure 8 is a plot of the mean performance across subjects in the position-in-depth task for each cue type as a function of eccentricity, and tested at the corresponding 3D motion threshold coherence. The subjects all behaved very similarly (error bars are ± 1 S.E.M. across the 3 subjects, and are smaller than the symbols for 4
of the 9 points). Two trends are clear. First, (static) depth judgments for the CD stimuli at the 3D motion coherence threshold are better than 84% correct, confirming the rich disparity signals present in these stimuli. Second, and more importantly for our purposes, performance for all observers at all eccentricities was at or near chance for the IOVD stimuli, even though the same stimulus coherences supported 84% correct performance on the 3D motion task. This indicates that, what ever the potential disparity information is in the anticorrelated stimuli, observers were unable to use it to do basic depth discriminations.

The near-absolute failure to accurately judge position-in-depth on IOVD stimuli implies that any disparity signals arising from the anticorrelated elements were not perceptually accessible for the purposes of performing a simple near-vs.-far task. These observations support the notion that direction-discrimination of our IOVD stimulus in the main experiment was based on the interocular comparisons of velocities—with effectively no contribution of residual disparity signals that might have been used to compute an additional CD signal. This control experiment replicates and expands our previous dissociation of percepts of motion through depth from percepts of position-in-depth using similar anticorrelated stimuli (Rokers et al. 2008). Although it could be argued that the computations of static and dynamic disparity mechanisms may be distinct processes, we believe that the term “changing disparity” should (and has been) defined as a signal that takes conventional supra-threshold static disparity signals as its input. Thus, if it can be demonstrated that a given stimulus configuration does not support depth judgments based upon static disparities, then it cannot support 3D motion judgements based on changing disparities. This disparity hierarchy appears to hold in
both classical and recent models of CD mechanisms (Cumming 1995; Peng and Shi 2010).

We ruled out another concern regarding the IOVD stimuli, which is that observers might have performed the direction-discrimination task on the basis of monocular direction discrimination instead of based on perceived 3D direction per se. If this were the case, the intrinsic monocular similarities of the FULL and IOVD stimuli could account for the similar sensitivities observed. Of course, this argument assumes that subjects were able to perform the task using concurrent utrocular identification and 2D direction discrimination under conditions of simultaneous stimulation in the two eyes, (which is rather unlikely, see Ono and Barbeito 1985; Porac and Coren 1986), and also correctly mapping the monocular motion to the 3D direction in the absence of feedback. Regardless, we addressed this concern empirically, by performing an additional experiment in which we measured observer’s frontoparallel (2D) direction discrimination sensitivity for FULL and IOVD stimuli, and compared them to their corresponding 3D motion sensitivities from the main experiment. This 2D direction-discrimination task was performed on identical stimuli as in the position-in-depth task (described above), except that observers were instructed to respond to the 2D direction of motion (leftward or rightward). The scatterplot of these data (figure 8, panel B) shows that sensitivity for 2D direction discrimination was several times higher than equivalent 3D motion sensitivities across all eccentricity and motion cue conditions. This replicates stereomotion suppression (Tyler 1971) and supports a larger body of research which has shown that IOVD performance cannot be explained on the basis of monocular stimulation (e.g.
Although prior work has shown that the CD cue can be experimentally isolated and is sufficient to yield percepts of 3D motion (Cumming and Parker 1994; Gray and Regan 1996; Julesz 1971; Norcia and Tyler 1984), our results suggest that the IOVD cue can be similarly studied in near-isolation by using binocularly anticorrelated elements outside the fovea, moving at moderately fast speeds, at relatively sparse densities (more akin to displays traditionally used to study frontoparallel motion than those used to study stereopsis). Our results also suggest that the IOVD cue is not only sufficient to yield percepts of 3D motion, but is also relied upon preferentially (relative to the CD cue) under many viewing conditions.

**Effects of speed on the CD and IOVD mechanisms**

The CD and IOVD mechanisms were affected very differently by manipulations of speed. CD sensitivity fell quickly with increased speed, exhibiting a lowpass sensitivity (or having a peak at or lower than the lowest speeds measured). IOVD sensitivity, on the other hand, had a clearly bandpass sensitivity peaking at a faster speed. FULL sensitivity also exhibited a bandpass speed tuning, one that was nearly identical to the IOVD pattern, but that contrasted sharply with that seen for the CD stimuli.

We described our stimulus motions in terms of retinal speed per eye (or, in the case of the CD stimulus, equivalent retinal speed per eye). For example, a speed of “1.8°/sec-eye” corresponds to rightward (or leftward) monocular motion in one eye at 1.8°/sec, and leftward (or rightward) monocular motion in the other eye at 1.8°/sec.
Such a speed, although relatively slow when viewed monocularly, yields a percept of relatively fast motion through depth towards or away from the observer. This speed was closest to the peak of the IOVD and FULL speed tuning curves. Likewise, at 0.3°/sec·eye, the qualitative percept was of very slow displacement over time, and observers reported that this condition did not yield a “direct” perception of motion through depth—rather, the phenomenology was of inferring motion from a change in position-in-depth over time. It is therefore noteworthy that this speed yielded the highest CD sensitivity.

The effects of speed suggest that the perception of 3D motion in natural stimuli (which inherently contain both binocular cues) appears to be supported by the IOVD cue more than the CD cue. That said, the CD cue appears well-suited to carry information for very slow 3D motions. One possibility is that the CD mechanism is not fundamentally a motion mechanism, but rather one optimized for objects that are (nearly) stationary. The (isolated) CD signal can drive vergence eye movements, but these are relatively sluggish—generally less than 1°/sec (Stevenson et al. 1994). We speculate that the CD mechanism may simply reflect the brain’s attempt at inferring the pattern and rate of change of signals from well-characterized “static” disparity detectors. In contrast, the IOVD mechanism appears to be more similar to other (2D) motion mechanisms, and is well-suited to quickly moving objects. The pattern of sensitivity we observed to the FULL cue displays further suggests that the human visual system is capable of relying on this IOVD mechanism. This is ecologically appropriate, given that sensitivity to objects moving quickly towards or away (even if they have not yet been
fixated) is likely a major element in successful interaction with a dynamic 3D environment.

Eccentricity effects on the CD and IOVD mechanisms

The patterns of sensitivity to the CD and IOVD cues across a wide eccentricity range complement the observed effects of speed. Sensitivity to FULL, CD, and IOVD stimuli decreased at larger eccentricities. When considered in isolation, the manipulation of eccentricity was actually less effective at distinguishing between the CD and IOVD cues. One might have expected that CD sensitivity would be particularly affected by eccentricity because the processing of static disparities is known to be much better in the fovea (e.g. Blakemore 1970; Tyler 1975). However, displays that simulate reasonable 3D motions subtend a disparity range that is at least an order of magnitude greater than the stereoacuity threshold at the eccentricities tested (Westheimer and Truong 1988). Thus, even the most eccentric stimuli (11°–15°) still contained range of disparities that could likely have supported CD sensitivity at higher eccentricities.

Instead, the effects of eccentricity are more informative when one considers them in conjunction with the manipulation of speed. IOVD sensitivity was relatively more robust to eccentricity for faster speeds of motion through depth. This improvement with speed is further consistent with the notion that our anticorrelated IOVD stimulus tapped a motion mechanism. For example, sensitivity to temporal frequency is if anything improved at far eccentricities (Rovamo and Raninen 1984; Wright 1987). Indeed, the eccentricity effects on IOVD sensitivity were lowest at fast speeds, perhaps because performance had achieved a maximal level. In contrast, eccentricity effects on CD
sensitivity were lowest at fast speeds as well, but in this case, the reason was because performance was approaching chance (instead of peak performance) levels.

    Taken together, the consideration of speed and eccentricity effects suggest that the CD mechanism is capable of supporting 3D motion direction discrimination for slow, (and particularly para-foveal) motions. Outside of this range, the IOVD mechanism appears much more capable of accounting for 3D direction discrimination when both cues are present. We interpret this as evidence for the relative primacy of the IOVD cue outside central vision. The visual system may exploit interocular velocity differences as a robust source of information for moderate- and fast-moving objects that one is not (yet) looking at. This suggests that classical motion detectors, typically studied in the domain of 2D processing, may also be utilized for perceiving 3D motion. The question as to how neural circuits implement the differencing operation upon eye-specific velocity signals remains an open question and a topic of ongoing work.

Relation to past work

Although our results demonstrate that the IOVD cue makes a significant contribution to 3D motion perception that is sometimes superior to the CD cue, it is important to recognize that this role for IOVDs may depend to some extent on the experimental conditions (Regan and Gray 2009). For example, we used a 2-alternative forced choice direction-discrimination task. We selected this task because it seemed most analogous to a particularly well-studied task in the 2D motion literature (i.e. Newsome and Paré 1988; Watamaniuk et al. 1995), but prior work has (understandably) investigated 3D motion perception using a variety of different tasks, including direction
estimation, speed discrimination, judging time to contact, and indicating whether motion
through depth is perceived (Brooks and Stone 2006; Harris and Dean 2003; Harris and
Watamaniuk 1995; Portfors-Yeomans and Regan 1996). Because each of these tasks
might require the observer to rely on and interpret 3D motion signals in different ways—
ways that we do not yet fully understand—it is difficult to generalize or compare results
across tasks. Indeed, it will be interesting to extend the approach described in this
paper to these other tasks in order to build a broader characterization of the relative
contributions of the CD and IOVD cues. Although many of these tasks tap important
perceptual capacities, we again emphasize that our conclusion that the velocity-based
cue plays a major role may reflect the fact that the task and stimuli we chose had strong
roots in literature on both the psychophysics and physiology of 2D motion processing
(e.g. Braddick 1974; Newsome and Paré 1988; Perrone and Thiele 2002). Given this
constraint, it also had an obvious real-world validity (judging whether something is
moving towards or away from your head).

Driven by the goal of maintaining consistency across FULL, CD, and IOVD
conditions, we decided to use a single signal plane, which was easily depicted in all cue
conditions, salient at high signal strengths, and allowed for straightforward
manipulations of both eccentricity and speed. However, it will be important to generalize
these results to other stimulus conditions in order to relate it to a larger body of prior
work. Some prior work has employed small stimuli relatively near fixation, various types
of spatial motion structure (i.e. sinusoidal oscillation, rotation, or oblique trajectories
through depth), and element densities ranging from a few percent to complete
coverage—all of which could affect the relative contributions of the two cues (e.g.
Andrews et al. 2001; Cumming and Parker 1994; Portfors-Yeomans and Regan 1996; Shioiri et al. 2008). Of course this dependence on specific stimulus factors is true whenever one studies a system that can use multiple sources of information. The key point here is that we have found a set of reasonable and simple conditions under which the IOVD cue makes a surprisingly strong contribution to the perception of 3D motion.

We used stimuli that moved across a wide range of constant speeds consistent with motions of real objects through depth: at our viewing distances, we simulated 3D motions of about 8 cm/sec to about 74 cm/sec (the latter corresponds to approximately one and three-quarters miles per hour; a reasonable walking speed for a human). Across this range, we observed large (approximately an order-of-magnitude) changes in overall sensitivity, as well as large relative changes between the CD and IOVD conditions. These changes in sensitivity suggest that the IOVD cue makes a major contribution to 3D motion perception at ecologically-important speeds.

An important prior study concluded that IOVDs did not contribute to 3D motion perception across a very wide range of temporal frequencies (Cumming and Parker 1994). There are several reasons that might explain why we arrived at a starkly complementary conclusion. First, we directly assessed the IOVD contributions using anticorrelated displays, instead of inferring them from the difference between FULL-cue and CD-only displays (although note that, under our experimental conditions, the latter method would still have revealed a large role for IOVDs, as shown in Figure 8). More importantly, we asked observers to perform a single interval direction-discrimination task on stimuli moving at a constant speed, which is quite different from the prior study’s use of a two interval signal-present vs. signal-absent task on stimuli that
oscillated sinusoidally through depth. It is possible that subjects could identify the
presence of the signal in this discrimination task by preferentially attending to the slower
parts of the sinusoidal oscillation at the extremes of the depth range. Such a strategy
could be supported almost exclusively by disparity-based mechanisms, and would
reveal little about the sensitivity of IOVD vs. CD mechanisms.

Despite some significant differences between our experiments and prior ones
that arrived at different conclusions, our finding of a central role for IOVDs does not
imply that our CD stimulus was somehow weak or at a particular disadvantage relative
to the other stimuli in our study. In fact, we found that sensitivity in the CD condition
actually exceeded that in the FULL (and IOVD) conditions at the slowest speeds,
demonstrating that the CD stimulus itself contained strong signals under the viewing
conditions that favored CD processing—thus, the relative inability of the CD condition to
account for FULL sensitivity at faster speeds almost certainly lies within the visual
system. The position-in-depth control experiment lends further support: for the same
signal and noise dots, performance on a position-in-depth task was nearly perfect for
the CD stimuli, but abysmal for the IOVD stimuli. Although not conclusive, this is
certainly evidence against the notion that CD signals were at a huge disadvantage due
to low level masking.

Moreover, the upper limit of speed sensitivity that we observed is not at odds with
some prior studies, most notably that of Cumming and Parker (1994). They collected
data at slower speeds, but the sensitivity in their temporally-correlated condition
(analogous to our FULL condition) clearly improves with temporal frequency for both
subjects shown, while the sensitivity in their dynamic condition (analogous to our CD
condition) suggests a roll-off around 2Hz. Although both of their subjects were, in fact, better overall in their “CD” condition at the temporal frequencies tested, the difference in sensitivity between the two conditions was also clearly diminishing rapidly with increasing temporal frequency: from their Figure 3, it is not at all unreasonable to suppose that the sensitivity in their “FULL” condition would begin to exceed that of the “CD” condition had higher temporal frequencies been tested.

In another study, Norcia and Tyler (1984) found that a CD-based depth percept was present up to 6 Hz, but they used a square-wave alternation in depth and noted that the percept changed from one of apparent motion (in depth) to one of pulsating semi-transparent depth planes as temporal frequency was increased. It is thus unclear what portion of their responses can be attributable to true 3D motion percepts, and what portion was due to a modulation of signal strength at different disparities. Moreover, as they themselves noted, their estimate of the temporal resolution of stereoscopic position change was higher than had been reported in previous work (Regan and Beverley 1973; Richards 1972). Overall, we find that the similarity of results to prior work (despite the differences with earlier studies’ initial conclusions) yields a rather coherent picture of the relative temporal sensitivity of the CD and IOVD cues.

More recent work has also provided evidence for characteristic dependencies of the IOVD cue on speed and eccentricity. Shioiri et al. (2008) reported greater sensitivity to higher temporal frequencies for uncorrelated dot displays (presumably mediated primarily by the IOVD mechanism) than for correlated cyclopean displays (processed exclusively by the CD mechanism). Although observers performed different tasks in the IOVD and CD conditions (single interval direction discrimination versus two interval
signal detection, respectively), and different motion characteristics were present in the
two conditions (rotation in depth versus oscillation in depth, respectively), the general
conclusions they arrived at are rather consistent with our speed tuning observations.
Likewise, Brooks & Mather (2000) reported evidence for an IOVD contribution to 3D
motion based upon manipulations of eccentricity. Reductions in perceived frontoparallel
speed at farther eccentricities mirrored reductions in perceived speed of 3D motion, but
were relatively independent of eccentricity effects on disparity-based judgments. Such a
result is consistent with the robust IOVD contributions across speeds that we observed
at middle and far eccentricities.
More generally, our results complement prior attempts to isolate IOVD
contributions using a variety of different approaches. Although stimuli containing only
CD information without any IOVDs can be straightforwardly generated using 1-frame dot
lifetimes, a stimulus containing only IOVDs without also containing any potential CD
information has not been developed (and may be impossible). Thus, prior work has
employed uncorrelated elements (Brooks 2002a; Shioiri et al. 2000), vertically-
unmatched strips of opposite motions (Shioiri et al. 2000), or monocular
adaptation (Brooks 2002b; Fernandez and Farell 2006). Although details of each of
these approaches require careful consideration (e.g., ruling out spurious disparities in
uncorrelated stimuli, assessing the effects of optical blur and neural spatial summation
in vertically-unmatched strip stimuli, and understanding the relationship between
monocular adaptation and subsequent dichoptic 3D processing), many of these studies
have included careful controls and have begun to form a coherent and compelling case
for the importance of IOVDs in 3D motion perception. The overall body of relevant work,
including ours, thus encompasses a wide range of tasks and stimuli. Despite this heterogeneity, there is broad agreement that IOVDs do make a distinct contribution to 3D motion perception. Furthermore, they point to the generalization that the CD mechanism is generally lowpass, even if estimates of the cut-off speed may vary slightly. Moreover, previous studies are generally consistent with the notion that the IOVD cue supports the perception of motion through depth at relatively high speeds—speeds that are beyond the upper limit of dynamic disparity processing. Our ability to compare FULL, IOVD, and CD sensitivities using a common stimulus geometry, task, and sensitivity metric allow us to further suggest that the IOVD cue not only contributes to 3D motion perception, but is in fact dominant in a variety of important conditions.

Implications for future work

At a practical level, our results demonstrate the feasibility of studying the IOVD cue in isolation, or at least in near-isolation. The use of anticorrelated displays provides a straightforward means for degrading disparity-based signals to reveal the role of the IOVD cue, while maintaining a simple stimulus geometry that supports direct comparison to other 3D motion displays. At a theoretical level, our results provide strong motivation to extend models of motion processing to consider the interocular comparison of monocular velocities. Canonical models of motion processing typically assume that later stages of motion processing operate on generic cyclopean representations (i.e. binocular properties are left unspecified), and thus the representation of eye-specific motions has not been considered (Perrone and Thiele 2002; Rust et al. 2006; Simoncelli and Heeger 1998). Because motion towards or away
the observer typically yields opposite directions of motion in roughly corresponding parts
of the two retinae, standard motion mechanisms that involve directional antagonism
(motion opponency) need to be modified to be specifically monocular (Tailby et al.
2010). Instead of subtracting these locally-opposite directions of motion (for a net result
of zero), the visual system must instead extract their signed difference as a cue to 3D
velocity.

Furthermore, our results also motivate extensions of models of binocular
processing to consider the contributions of monocular motions. Instead of being
depicted as an “impurity” relative to CD-only cyclopean stereomotion, our results
support a complementary perspective: that the IOVD cue be considered an integral part
of seeing motion in depth, and that, at least under a wide range of reasonable
experimental conditions, the CD cue makes a rather limited contribution. The perception
of 3D motion may thus better be thought of as a binocular form of motion processing,
rather than as a dynamic form of stereopsis.
FOOTNOTES

i A percept of motion through depth has been shown to arise from moving objects visible to one eye and camouflaged to the other (Brooks and Gillam, 2006b, 2007). This is a dynamic analog of the stereopsis without binocular correlation described by Kaye (1978). Although interesting, this is beyond the scope of the current paper.

ii Even though “FULL” is not an acronym, we spell it in all capital letters so it visually groups with the other two condition names.

iii Specifically, observer ACH was unable to discriminate CD stimulus direction at three highest speeds in the farthest eccentricity conditions (11° eccentric.; 0.9, 1.8, and 2.7°/sec·eye) and at the highest speed in the middle eccentricity condition (7° eccentric.; 2.7°/sec·eye). Observer LKC was unable to discriminate CD stimulus direction at any eccentricity at the highest speed (3, 7, 11° eccentric.; 2.7°/sec·eye).
ACKNOWLEDGMENTS

This work was supported by a National Science Foundation CAREER Award (BCS-0748413) to A.C.H., a National Science Foundation Grant (IIS-0917175) to L.K.C., and a Netherlands Organisation for Scientific Research grant to B.R. (2006/11353/ALW).
Bibliography


Figure 1. Binocular cues to 3D motion. The visual system could infer the direction of 3D motion (white and black spheres moving towards the eyes, in the cylinder at lower right) based on dynamic information from the corresponding 2D retinal projections (white and black dots, flat panel at upper left). The schematic depicts two known binocular cues to 3D motion. The changing disparity (CD) cue is derived from tracking changes in binocular disparity over time (black brackets and ellipsis). The interocular velocity difference (IOVD) cue is derived by comparing monocular velocity signals from corresponding regions in the two retinae (red brackets and arrows). Although both cues coexist in natural scenes, experimental stimuli can be used to study the cues together or in isolation (see Figure 3).

Figure 2. Schematic of stimulus and manipulations. (A) Scale version of the stimulus as presented to the right eye. For clarity, this shows only the lower right quadrant of the monocular stimulus. The fixation point and nonius lines are located in the upper lefthand corner, with bright and dark stimulus dots scattered across an annular region 3–7° eccentric. Under all conditions the grey background extended across the entire monitor. By design, a single frame from one monocular image was identical (statistically) across all three motion cue stimuli. (B) Depiction of stimulus eccentricities. At the smallest eccentricity, signal and noise dots were restricted to an annular stimulus volume 3–7° from fixation. The smallest eccentricity was divided into 4 equal area quadrants (only the left 2 are shown, but the full display was left-right symmetric) that were displaced further outward for the other two eccentricity conditions, yielding eccentricities of 3–7°, 7–11°, and 11–15°. (C) Oblique depiction of the
cyclopean 3D percept and speeds. A plane of signal dots moved through a cloud of
noise dots (signal plane outlined in red for clarity, individual dots that actually
constituted the plane not shown), and observers performed a 3D direction of motion
discrimination (towards vs. away). The signal plane moved at 1 of 5 different speeds
(colored arrows, in °/sec·eye), corresponding to motions through depth ranging from
about 8 cm/sec to about 72 cm/sec (given our viewing distance of 70 cm). Direction
discrimination sensitivity was measured at each of the 5 speeds (C) and 3 eccentricities
(B), for each of the 3 motion cue types (see Figure 3).

**Figure 3. Motion cue stimulus conditions.** Depiction of representative signal dots
from two example frames from each of the three stimulus conditions (from left to right:
FULL, IOVD, CD). The half images in each set can be free-fused (L, left eye view; R,
right eye view). In the FULL stimulus, corresponding dots moved in opposite directions
in the two eyes. Such a stimulus contains both changing disparities (the CD cue) and
inter-ocular velocity differences (the IOVD cue). Red circles and arrows were not
present in the actual stimulus, of course, but depict the respective motions of left and
right eye’s views of a single signal dot. Also note that actual dot densities were much
higher (see Figure 2A and Methods): many similar signal dots specified a plane moving
towards or away through depth. In the IOVD stimulus, corresponding dots also moved in
opposite directions (just as in the FULL stimulus), but the dots were binocularly
anticorrelated: a black dot in one eye was paired with a white dot in the other. This
greatly reduced the contribution of the disparity-based CD cue, but the IOVD signal was
preserved. In the CD stimulus, signal dots were randomly repositioned every frame, but
their disparities (dotted brackets) still specified a signal plane moving towards or away through depth. The 1-frame lifetimes of the dots removed any coherent monocular motions and hence eliminated the IOVD cue, while preserving the CD signal.

**Figure 4. 3D motion direction-discrimination sensitivity as function of eccentricity (and speed).** Direction-discrimination sensitivity (y-axis) as a function of eccentricity (x-axis), with speed as the grouping parameter (lighter shades of gray corresponding to faster speeds). Error bars represent 68% bootstrapped confidence intervals (i.e. standard errors). For both the FULL (left) and IOVD (middle) conditions, sensitivity decreased with increasing eccentricity. Moreover, sensitivity generally increased with increasing speed for these two conditions, particularly at higher eccentricities. Only for the very fastest speed did sensitivity begin to decrease (this effect was most pronounced at the smallest eccentricity). The pattern in the CD condition (right) was strikingly different: sensitivity did decrease with increasing eccentricity and more so for slowest speeds, but the overall order of the curves was reversed, with low speeds yielding much higher sensitivities than high speeds.

**Figure 5. 3D motion direction-discrimination sensitivity as function of speed (and eccentricity).** Direction-discrimination sensitivity (y-axis) as a function of speed (x-axis), with eccentricity as the grouping parameter (lighter shades of gray corresponding to larger eccentricities). The speed axis is logarithmic, error bars represent 68% bootstrapped confidence intervals. For both the FULL (left) and IOVD (middle) conditions, sensitivity shows distinct bandpass tuning, with peak sensitivity occurring
just before the highest speeds tested. In contrast, sensitivity in the CD condition (right) exhibited clear lowpass tuning, with maximal sensitivity for the speed closest to stationary. For all conditions, sensitivity generally increased with decreasing eccentricity (dark curves above light curves). Although at faster speeds sensitivity for FULL and IOVD conditions was generally higher than for the CD condition, note also that at the lowest speeds, sensitivity for the CD condition actually exceeded that of the FULL cue condition.

**Figure 6. 3D motion direction-discrimination sensitivity as a function of both speed and eccentricity.** Top row, sensitivity as a function of both speed and eccentricity for the three cue conditions. Each sensitivity surface represents combined subject data from 45,360 trials. Height of mesh (z-axis) indicates sensitivity, as a function of speed and eccentricity (x and y axes). Colored floor is a contour plot depicting the same sensitivity information. The FULL and IOVD surfaces are quite similar, whereas the CD surface is distinctly different (so much so that, from this perspective, the view is of the bottom of the mesh surface). Bottom row, differential sensitivity surfaces highlighting the similarity, or lack thereof, between the three conditions (z=0 plane has been raised to allow for negative values). The FULL–IOVD surface is nearly flat, indicating substantial similarity. In contrast, the other two surfaces (FULL–CD, IOVD–CD) show a systematic pattern of differences. At the slowest speed tested, CD sensitivity was higher than IOVD sensitivity and higher than the FULL sensitivity as well (cool colored mesh and surface).
Figure 7. Comparison of IOVD sensitivity and CD sensitivity to FULL sensitivity.

Scatterplots show IOVD sensitivity (A) and CD sensitivity (B) plotted on the y-axis against corresponding FULL sensitivity (x-axis). Each data point corresponds to the sensitivity of an individual subject for a particular eccentricity/speed/motion cue condition. Lighter gray symbols represent farther eccentricities. The dashed line shows unity. IOVD sensitivity generally matched FULL sensitivity ($r^2 = 0.75$), while CD sensitivity was relatively unrelated ($r^2 = 0.05$), illustrating FULL sensitivity is better predicted by IOVD sensitivity than CD sensitivity.

Figure 8. Isolation of IOVD cue. A. Position-in-depth performance (proportion correct for a near vs. far 2AFC discrimination) for the CD (open circles), IOVD (closed circles), and FULL stimuli (open squares). Each point represents the position-in-depth performance (averaged over 3 observers) as measured at the motion coherence corresponding to each observer’s 3D motion direction-discrimination threshold. CD position-in-depth performance was much better than IOVD at all eccentricities (and even better than that for the FULL stimulus in one case). Crucially, CD performance was always better than 84% correct (upper dashed line indicates the performance level for the 3D motion task at the coherences used), but IOVD performance was at or near chance (lower dashed line). B. Individual observer’s 2D motion direction-discrimination sensitivity for the FULL (open squares) and IOVD (closed circles) stimuli as a function of the corresponding 3D motion direction-discrimination sensitivity. All of the points fall above the dashed line (unity), indicating the much greater sensitivity to 2D frontoparallel motion than 3D motion. The dotted line (root-2 improvement 2D vs. 3D) suggests that
the greater sensitivity to 2D motion cannot be explained by simple within-direction binocular summation.
TABLE 1. *P*-values on 3D motion sensitivity difference surfaces

<table>
<thead>
<tr>
<th>Speed</th>
<th>Near</th>
<th>Middle</th>
<th>Far</th>
</tr>
</thead>
<tbody>
<tr>
<td>FULL–IOVD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3°/sec·eye</td>
<td>0.004*</td>
<td>0.416</td>
<td>0.278</td>
</tr>
<tr>
<td>0.6°/sec·eye</td>
<td>0.090</td>
<td>0.200</td>
<td>0.322</td>
</tr>
<tr>
<td>0.9°/sec·eye</td>
<td>0.476</td>
<td>0.002*</td>
<td>0.416</td>
</tr>
<tr>
<td>1.8°/sec·eye</td>
<td>0.436</td>
<td>0.462</td>
<td>0.462</td>
</tr>
<tr>
<td>2.7°/sec·eye</td>
<td>0.162</td>
<td>0.460</td>
<td>0.034</td>
</tr>
<tr>
<td>FULL–CD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3°/sec·eye</td>
<td>0.066</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
<tr>
<td>0.6°/sec·eye</td>
<td>0.012*</td>
<td>0.020*</td>
<td>0.006*</td>
</tr>
<tr>
<td>0.9°/sec·eye</td>
<td>0.000*</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
<tr>
<td>1.8°/sec·eye</td>
<td>0.000*</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
<tr>
<td>2.7°/sec·eye</td>
<td>0.000*</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
<tr>
<td>IVD–CD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3°/sec·eye</td>
<td>0.000*</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
<tr>
<td>0.6°/sec·eye</td>
<td>0.198</td>
<td>0.004*</td>
<td>0.002*</td>
</tr>
<tr>
<td>0.9°/sec·eye</td>
<td>0.000*</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
<tr>
<td>1.8°/sec·eye</td>
<td>0.000*</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
<tr>
<td>2.7°/sec·eye</td>
<td>0.000*</td>
<td>0.000*</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

* *p < 0.05
**FULL Stimulus**
- oppositely-moving dots binocularly correlated

**IOVD Stimulus**
- oppositely-moving dots binocularly anti-correlated

**CD Stimulus**
- randomly repositioned binocularly correlated
FULL Stimulus  IOVD Stimulus  CD Stimulus

Eccentricity (deg)  Eccentricity (deg)  Eccentricity (deg)

Sensitivity (coh⁻¹)

Speed
• 0.3
• 0.6
• 0.9
• 1.8
• 2.7
FULL Stimulus

IOVD Stimulus

CD Stimulus

Sensitivity (coh⁻¹)

Speed (°/sec·eye)