Eye-specific pattern-motion signals support the perception of three-dimensional motion

Sung Jun Joo

Department of Psychology, Pusan National University, Busan, Republic of Korea ✉

Devon A. Greer

Center for Perceptual Systems, The University of Texas at Austin, Austin, TX, USA

Lawrence K. Cormack

Center for Perceptual Systems, The University of Texas at Austin, Austin, TX, USA
Department of Psychology, The University of Texas at Austin, Austin, TX, USA

Alexander C. Huk

Center for Perceptual Systems, The University of Texas at Austin, Austin, TX, USA
Department of Psychology, The University of Texas at Austin, Austin, TX, USA
Department of Neuroscience, The University of Texas at Austin, Austin, TX, USA

An object moving through three-dimensional (3D) space typically yields different patterns of velocities in each eye. For an interocular velocity difference cue to be used, some instances of real 3D motion in the environment (e.g., when a moving object is partially occluded) would require an interocular velocity difference computation that operates on motion signals that are not only monocular (or eye specific) but also depend on each eye's two-dimensional (2D) direction being estimated over regions larger than the size of V1 receptive fields (i.e., global pattern motion). We investigated this possibility using 3D motion aftereffects (MAEs) with stimuli comprising many small, drifting Gabor elements. Conventional frontoparallel (2D) MAEs were local—highly sensitive to the test elements being in the same locations as the adaptor (Experiment 1). In contrast, 3D MAEs were robust to the test elements being in different retinal locations than the adaptor, indicating that 3D motion processing involves relatively global spatial pooling of motion signals (Experiment 2). The 3D MAEs were strong even when the local elements were in unmatched locations across the two eyes during adaptation, as well as when the adapting stimulus elements were randomly oriented, and specified global motion via the intersection of constraints (Experiment 3). These results bolster the notion of eye-specific computation of 2D pattern motion (involving global pooling of local, eye-specific motion signals) for the purpose of computing 3D motion, and highlight the idea that classically "late" computations such as pattern motion can be done in a manner that retains information about the eye of origin.

Introduction

The perception and neural processing of frontoparallel two-dimensional (2D) motion have been studied extensively. In the classical primate motion-processing circuit, directionally selective neurons in striate cortex (V1) respond to the one-dimensional (1D) velocity component orthogonal to their preferred orientation. The disambiguated 2D direction of complex objects and patterns is then encoded explicitly by later stages of motion processing, most notably in area MT (Adelson & Movshon, 1982; Rodman & Albright, 1989; Stoner & Albright, 1992). V1 neurons are also known to exhibit high degrees of selectivity for other stimulus features, including spatial and temporal frequency, and have a gradual contrast response function. MT neurons, on the other hand, are notably insensitive to many visual features (Sclar, Maunsell, & Lennie, 1990), and their responses saturate more quickly with respect to contrast.

It is widely accepted that the initial extraction of 1D component motion occurs within the relatively local spatial scale of V1 receptive fields. Although MT receptive fields are considerably larger (Felleman & Kaas, 1984), the later encoding of pattern motion does not as cleanly map on to the spatially large and generally coarse tuning properties of MT neurons. Indeed, both psychophysical and physiological studies have suggested that 2D pattern-motion encoding in MT probably reflects computations on the afferent signals from earlier visual areas. For example, the contrast and spatial frequency of individual component gratings composing pattern motion affects coherent pattern-motion perception (Adelson & Movshon, 1982). The pattern-direction selectivity of MT neurons also becomes weaker when component gratings are placed in different spatial locations within the receptive field of an MT neuron (Majaj, Carandini, & Movshon, 2007) or when component gratings are presented to each eye (Tailby, Majaj, & Movshon, 2010). Furthermore, a cascade model based on the information flow from V1 to MT captures the pattern-direction selectivity of MT neurons using simple pooling and input–output mappings (Rust, Mante, Simoncelli, & Movshon, 2006).

Compared with this rather thorough understanding of 2D motion processing for a single input stream, our understanding of the perception and processing of three-dimensional (3D) motion coming through a binocular input stream is quite nascent (Cormack, Czuba, Knöll, & Huk, 2017). Conventionally, 3D motion has often been considered as cyclopean motion, in which motion integration occurs at the point of binocular combination and disparity extraction (Julesz, 1960, 1971; Carney & Shadlen, 1993; Shadlen & Carney, 1986). These models suggest that 3D motion is extracted at the level of the binocular disparity computation, early in the visual-processing pathway (prior to area MT), and that this 3D motion information is fed into standard stages of later motion processing (Patterson, 1999). A key piece of evidence for this is that 3D motion can be seen in dynamic random-element stereograms, which contain no coherent monocular motion signals (Norcia & Tyler, 1984; Tyler & Julesz, 1978). Others, however, have argued that the motion seen in such stimuli is actually the result of high-level feature tracking and thus outside the scope of the canonical motion pathway (Lu & Sperling, 1995). Regardless, the spatial scale of binocular integration is still "local" under this view, especially for binocular combination, which requires binocular correspondence (i.e., stimulation of corresponding local retinal regions within the upper disparity limit).

There is, however, a growing body of work inconsistent with this local-disparity-based explanation of how 3D motion processing should work, and which instead indicates that at least some eye-specific motion signals are exploited at a relatively late stage. Recent findings have shown the existence of 3D direction selectivity involving either disparity-based or velocity-based cues in the extrastriate area MT (Joo et al., 2016; Sanada & DeAngelis, 2014). Disparity-based and velocity-based 3D motion computation seem to be processed separately in area MT (Joo et al., 2016). Critically, eye-specific velocity signals—even without conventional binocular correspondence—may be available after the canonical point of binocular combination in primary visual cortex for a velocity-based computation of 3D direction in extrastriate cortex (Rokers, Czuba, Cormack, & Huk, 2011).

In the present study, we sought to more directly characterize how eye-specific velocity signals are integrated between the eyes and across space in supporting perceptual sensitivity to 3D direction, and to test whether 3D motion processing depends on mechanisms that are either different from those for processing frontoparallel 2D motion, or perhaps integrate local 2D motion signals in a way that is unique to eye-specific processing for 3D motion computations. First, based on previous findings showing strong 3D motion selectivity in area MT with relatively large receptive-field size compared to V1 (Sanada & DeAngelis, 2014; Rokers, Cormack, & Huk, 2009), we expected that 3D motion processing would be spatially global—showing robust direction-selective adaptation effects regardless of the precise spatial match between the adaptor and test stimulus. Such global processing is consistent with the larger receptive fields in extrastriate visual areas and contrasts with more spatially local 2D motion processing which correspond to the smaller receptive fields in earlier visual areas (Hedges et al., 2011; Kohn & Movshon, 2003). We used a psychophysical motion-adaptation paradigm using visual arrays consisting of small oriented Gabor patches (Amano, Edwards, Badcock, & Nishida, 2009; Hisakata, Hayashi, & Murakami, 2016; Lee & Lu, 2010; Rokers et al., 2011; Scarfe & Johnston, 2011), in which we measured motion aftereffects (MAEs). We manipulated the spatial congruency between adapting and test stimuli and tested whether there were strong 3D MAEs when test stimuli were presented at unadapted locations (Snowden & Milne, 1997).

Furthermore, to test whether eye-specific velocity signals are combined for 3D direction selectivity at later stages of motion processing, we then used "pseudo-plaid" stimuli, comprising small Gabor patches having random orientation with velocities specified by the intersection of constraints (Adelson & Movshon, 1982) to yield consistent global pattern motion. If eye-specific velocity signals are combined at the site of binocular combination (i.e., V1), 3D direction information would be lost because 3D motion based on the comparison of spatially local velocities between two eyes would not yield a consistent or coherent 3D direction. In contrast,

if eye-specific velocity signals are present in stages at or after the estimation of pattern motion, estimates of 3D direction could still be recovered by comparing more spatially global estimates of eye-specific pattern motion. Together, the results from these applications of a psychophysical motion-adaptation paradigm point to the existence of global eye-specific motion signals that are indeed exploited in the binocular computation of a 3D motion signal at a relatively late stage—quite likely past V1, the canonical site of binocular combination for static stereopsis.

Methods

Observers

Two authors (SJJ and DAG) and two observers who were unaware of our aims participated in the study. All had normal or corrected-to-normal vision and were experienced psychophysical observers. Written consent was obtained from all observers. All procedures were approved by the University of Texas at Austin's institutional review board and in adherence to the Declaration of Helsinki. All observers were recruited within the university's community and all data were collected on the University of Texas at Austin campus.

Apparatus

Experiments were programmed in MATLAB (MathWorks, Natick, MA) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). They were run on a Quad-Core Intel (Intel Corporation, Santa Clara, CA) Mac Pro computer (Apple, Inc., Cupertino, CA) with an ATI Radeon HD 5870 graphics card (Advanced Micro Devices, Inc., Sunnyvale, CA). Stimuli were displayed on a Sharp Aquos HDMI monitor (Sharp Corporation, Sakai-ku, Sakai, Japan; 1,920 × 1,080 resolution at 60 Hz) with a viewing distance of 73 cm. At this viewing distance, one pixel subtended slightly less than 1 arcmin. The luminance output of the display was linearized using standard gamma-correction procedures. A mirror stereoscope was used to combine left- and right-eye half images. This apparatus has been described in more detail previously (Czuba, Rokers, Huk, & Cormack, 2010).

Stimuli

A static annotated example of the stimulus is shown in Figure 1A. All stimuli were presented within an annular aperture subtending 1°–9° eccentricity. Ob-

servers were instructed to fixate the center of the display, which was marked by a black-and-white bull's-eye with gray horizontal and vertical nonius lines. The stimulus aperture was surrounded by a static texture of 100 dark (0.4 cd/m^2) and 100 light (129.7 cd/m^2) dots presented in a circular band (subtending 10°–12° eccentricity) to assist in maintaining a stable vergence posture.

Within the annular aperture in each eye, 60 (Experiments 1 and 2) or 20 (Experiment 3) drifting Gabor patches were placed in random spatial locations (Figure 1A). Each Gabor had a Gaussian envelope with a standard deviation (*SD*) of 0.15°, which made the diameter of each Gabor about 0.9° at ±3 *SD*. Gabors had a Michelson contrast of 35%, spatial frequency of 2 c/°, and drift speed of 0.5°/s. The shortest distance between Gabors (edge to edge) was approximately either 1° (Experiments 1 and 2) or 2° (Experiment 3). Each Gabor's central location was jittered by 0.1° to reduce any perception of grouping between nearby elements. The starting phase of each Gabor element was randomized in one eye, and the corresponding element in the other (if present; Experiments 1 and 2) was placed in antiphase to make the direction of disparity ambiguous. The effective disparity was 1.48°, which is outside the conventional range of Panum's fusional area (Mitchell, 1966; Panum, 1858).

During adaptation in the motion-through-depth (3D motion) experiments (Experiment 2 and 3), left and right eyes viewed opposite motion directions (simultaneously) to produce simulated motion toward or away from the observer. In Experiment 2, the stimulus elements were in corresponding locations in the two eyes but drifting in opposing directions. The large red arrows (Figure 1B) in each eye show these global, opposing motions consistent with motion through depth.

The motion strength (coherence) was defined as the percentage of Gabor patches (signal) that were drifting in the same (Experiment 1; 2D adaptation) or opposite (Experiments 2 and 3; 3D adaptation) direction between eyes among all the Gabor patches. The remainder of the Gabors (noise) flickered in counterphase at the same temporal frequency as the signal elements. Negative motion strength represents leftward/away (2D case/3D case) motion and positive motion strength indicates rightward/toward (2D case/ 3D case) motion. Zero coherence is the case where all gratings flicker. Using these motion strengths, we measured each individual's psychometric function for motion-direction discrimination. Before adaptation, one would predict that the point of subjective equality (PSE) would be around 0% coherence (all elements flicker, and hence there is no net motion direction). Motion adaptation shifts the psychometric function of motion-direction discrimination: After adaptation to
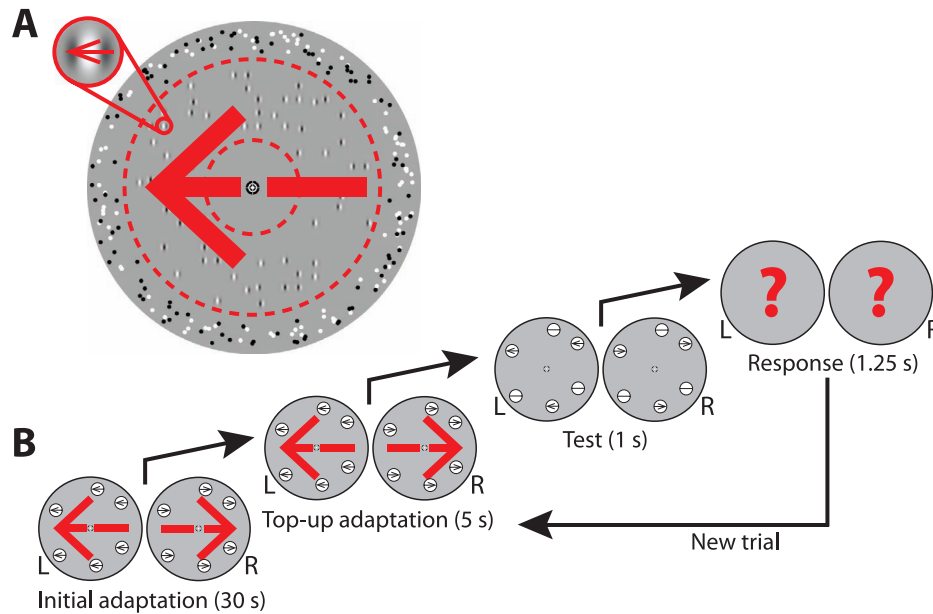
Figure 1. Stimulus schematic and adaptation procedure. (A) Example stimulus for a single eye. Large red arrow shows the global motion rendered by the drifting Gabor patches. Dashed red circles represent the boundaries of the annular aperture. Zoomed red circle shows one of the Gabor elements to illustrate orientation and drifting direction. (B) Adaptation procedure (organization of an example experimental run). Initial adaptation was 30 s, followed by a repeating trial sequence of top-up adaptation, blank, test, and response window. Adaptation periods (both initial and top-up) used 100% coherence. Test stimuli varied in coherence. Black outlined circles schematize Gabor element locations (not drawn to scale). Black arrows represent motions, and plain lines represent counterphase flickering noise.

one direction, the PSE would be shifted toward the adapted direction, suggesting that one needs more motion energy toward the adapted direction to perceive net zero motion. Measuring MAEs using this procedure is conceptually identical to the well-established technique of manipulating coherence in random-dot motion stimuli (Alais & Blake, 1999; Czuba, Rokers, Guillet, Huk, & Cormack, 2011; Lankheet & Verstraten, 1995; Sohn & Lee, 2009). In fact, the two techniques only differ in whether the noise is produced by random directions of dot motion or by carrier flicker. Our use of sparse Gabor elements has some advantages over using random-dot motion stimuli to measure 3D MAEs. First, individual Gabor elements carry signal or noise by local phase modulations, avoiding the ambiguities of dot matches and cross-matches between different signal and noise dots. Furthermore, because of the peripheral locations of the Gabor elements and the known broad spatial pooling of motion mechanisms, individual element signal and noise are not easily separable (Castet, Keeble, & Verstraten, 2002).

Task and procedure

Figure 1B shows the timeline of an experimental trial and run. Each run began with an initial 30-s adaptation period (Figure 1B, far left) in which the observer passively viewed all Gabor elements drifting at 100% coherence (i.e., the maximum motion strength). Depending upon the experiment, the Gabor elements viewed in the left eye drifted in the same direction (Experiment 1; 2D motion) or in the opposite direction (Experiments 2 and 3; 3D motion) as the right eye. Observers viewed the left and right images simultaneously by each eye. Key details for each condition are specified in the Results section for each experiment.

Once the initial adaptation period was complete, the MAE trial loop began (Figure 1B). Every trial began with a 4-s top-up adaptation followed by a blank period (1 s). Observers next viewed a test display drifting at a probe coherence value selected by the QUEST algorithm (see below; 750-ms stimulus duration, followed by a 250-ms blank). The observer was then prompted to indicate the direction of target motion (either left/right or toward/away) in a 1.25-s response window with a mouse click. Note that observers were never asked to judge coherence (nor attempt to null the motion), they simply attempted to judge either toward versus away or left versus right on each trial. The limited response window was used to prevent the adaptation state from being affected by long response times. In rare instances, if a response was not made during the response window, then the trial,

			Adapting Stimuli			Test Stimuli		
Experiment #	Condition		Direction	Correlated	Pseudoplaid	Direction	Correlated	= Adapter
1		A	X	Yes	No	X	Yes	Yes
1		B		Yes	No		Yes	No
2		A	Z	Yes	No	Z	Yes	Yes
2		B			No			No
3		A	Z	No	No	Z	Yes	No
3		B	Z		Yes	Z		No

Table 1. Experiment design with condition organization. *Notes*: Blue cells indicate an important comparison across conditions within an experiment. Yellow cells indicate an important comparison across successive experiments. Direction: X represents frontoparallel (2D) motion and Z indicates 3D motion; Correlated: stimulus elements were interocularly correlated—in the same location in each eye; Pseudoplaid: element orientations were random, and phase-drift velocities determined by the intersection of constraints to produce a unique global motion direction; = Adapter: test stimulus elements were in the same locations as the adapting stimulus elements.

including a new top-up adaptation period, was repeated.

Before participating in each of the actual experiments, all observers completed two to eight full-length practice sessions to stabilize performance. They then completed a minimum of two sessions (see below) of each adapting direction for each condition in all experiments.

Two interleaved adaptive staircases (QUEST; Watson & Pelli, 1983) were run during each session. Each staircase consisted of 25 trials total, and the initial coherence for each staircase was set to 50% leftward/rightward (Experiment 1; 2D motion) and 50% toward/away (Experiments 2 and 3; 3D motion). After 25 trials the two staircases converged within the estimated threshold standard deviation calculated by QUEST.

There were three key manipulations done to the stimuli across the experimental conditions: whether or not the local stimulus elements in the test stimuli occurred in the same locations as those in the adapting stimulus (Experiments 1 and 2; 2D and 3D motion, respectively); whether or not the adapting stimuli were pseudoplaid stimuli, having random orientation with velocities specified by the intersection of constraints (Adelson & Movshon, 1982) to yield consistent global motion (Experiment 3); and whether or not the stimulus elements were in the same retinal location across the two eyes (Experiment 2 vs. Experiment 3). The resulting taxonomy of experimental conditions is summarized in Table 1.

Experiment 1 thus tested whether a strong 2D MAE could be obtained from global motion, when the individual moving elements in the test stimulus were in different locations than they were in the adapting stimulus, or whether the 2D MAE was strictly local, requiring adapting and test elements in the same retinal locations. Previous studies using multielement Gabor arrays similar to our stimuli have found mixed results. Scarfe and Johnston (2011) found no 2D MAEs when test stimuli were displayed in the unadapted locations, suggesting local processing. However, Lee and Lu (2012) showed some 2D MAEs in the unadapted locations, although observers reported no MAEs more frequently. Thus, our Experiment 1 was important to characterize the spatial pooling of motion signals in 2D motion processing given our particular stimuli and methods.

Experiment 2 was identical to Experiment 1 in all respects except that the motion of the stimulus elements was in the opposite direction in the two eyes, yielding 3D motion consistent with an approaching or receding object.

In Experiment 3, the local motion elements were in different locations not only between adaptor and test but also in the two eyes within both adaptor and test. Further, in the second condition (Pseudoplaid), the orientation of each Gabor element in the adapting stimulus was drawn randomly from a uniform distribution between 20° and 70° from the vertical, and each Gabor element was constrained to have a single velocity based on the intersection of constraints. The stimuli in this experiment represent our attempt to absolutely ensure that, were a motion aftereffect to be observed, it could only have arisen from motion signals that were computed both globally and separately for the two eyes.

Data analysis

We modified the QUEST module in PsychToolbox to use the cumulative Gaussian function instead of a
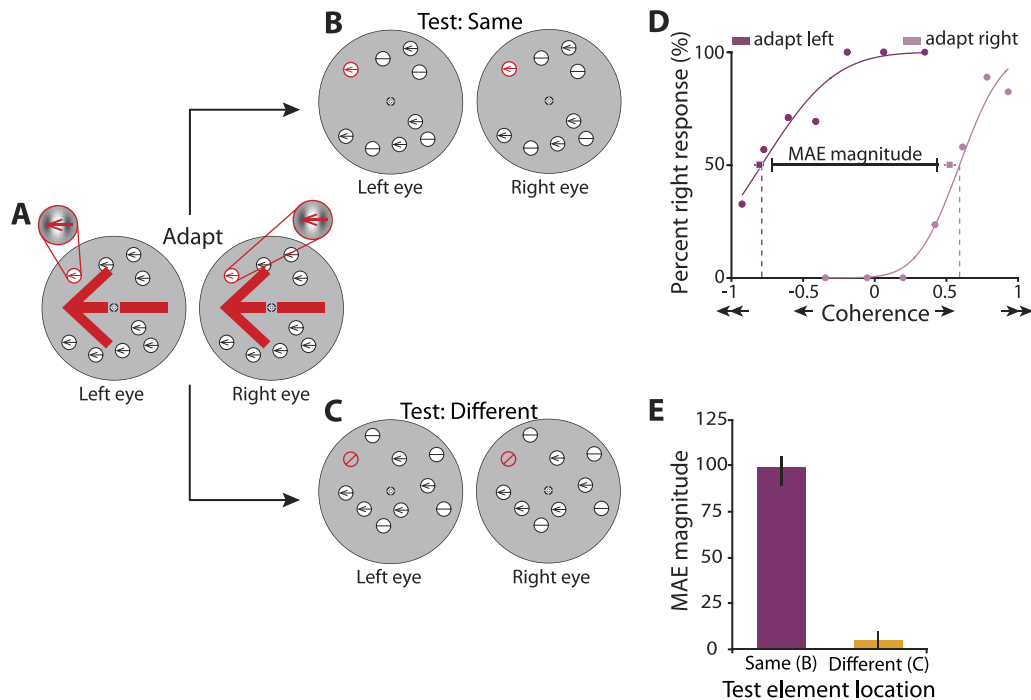
Figure 2. 2D motion aftereffects (MAEs) reflect local motion processing. (A) During adaptation, observers viewed 100% coherent motion in the same direction in both eyes. Small arrows inside the circles show the local motion element directions, and the large red arrow represents the global motion integrated across the elements. In this figure and the next, an example element is spotlighted by the red circles to illustrate, for each condition, the relationship of the local elements both across the two eyes and between the adaptation and test periods. (B) Experiment 1A—local motion test: Elements of the test stimuli were in the same location as the adapting stimuli. (C) Experiment 1B—global motion test: Elements of the test stimuli were all locally unadapted locations; that is, elements of the test stimuli were constrained to locations that had not been occupied by any element of the adapting stimuli. (D) Example psychometric functions and estimated points of subjective equality of one observer after adaptation. MAE magnitude was defined by the absolute value of difference in points of subjective equality between adapting directions. The dotted lines are the points at which psychometric functions cross 50% on the y-axis. (E) The averaged MAE magnitudes in both 2D adaptation conditions. We observed very strong MAEs in Experiment 1A, whereas MAEs for Experiment 1B were very weak, indicating that global motion does not support the 2D MAE. Error bars represent 95% bootstrapped confidence intervals.

Weibull function as the underlying psychometric function to measure the PSE, the coherence at which the direction of the stimulus motion is ambiguous (presumably because the signal in the stimulus is canceled by the effects of adaptation). The last estimates of threshold from the QUEST procedure were used to define the PSE estimates. The MAE magnitude was defined as the difference in PSEs between the opposite-direction adapting conditions ($MAE_{right} - MAE_{left}$ for 2D motion; $MAE_{toward} - MAE_{away}$ for 3D motion).

We also ensured that the PSE estimates from the QUEST staircases were similar to the point at which psychometric functions cross 50% on the y-axis (Figure 2D) by fitting the staircase data with a cumulative Gaussian function. Because the QUEST procedure samples the signal strength adaptively by definition, we binned the exact coherence values that were presented in a session into nine bins from −100% (left for Experiment 1; away for Experiments

2 and 3) to 100% (right for Experiment 1; toward for Experiments 2 and 3). We then calculated the proportion of rightward (for 2D) or toward (for 3D) responses in each bin. We used maximum-likelihood estimation to fit a Gaussian cumulative distribution function to the data. The results were similar whether we used the last estimates of threshold from the QUEST procedure or PSEs from the best-fitting psychometric function.

Results

Experiment 1: Local versus global adaptation in 2D frontoparallel motion processing

In the 2D motion-adaptation condition, observers adapted to and were tested with 2D motion stimuli,
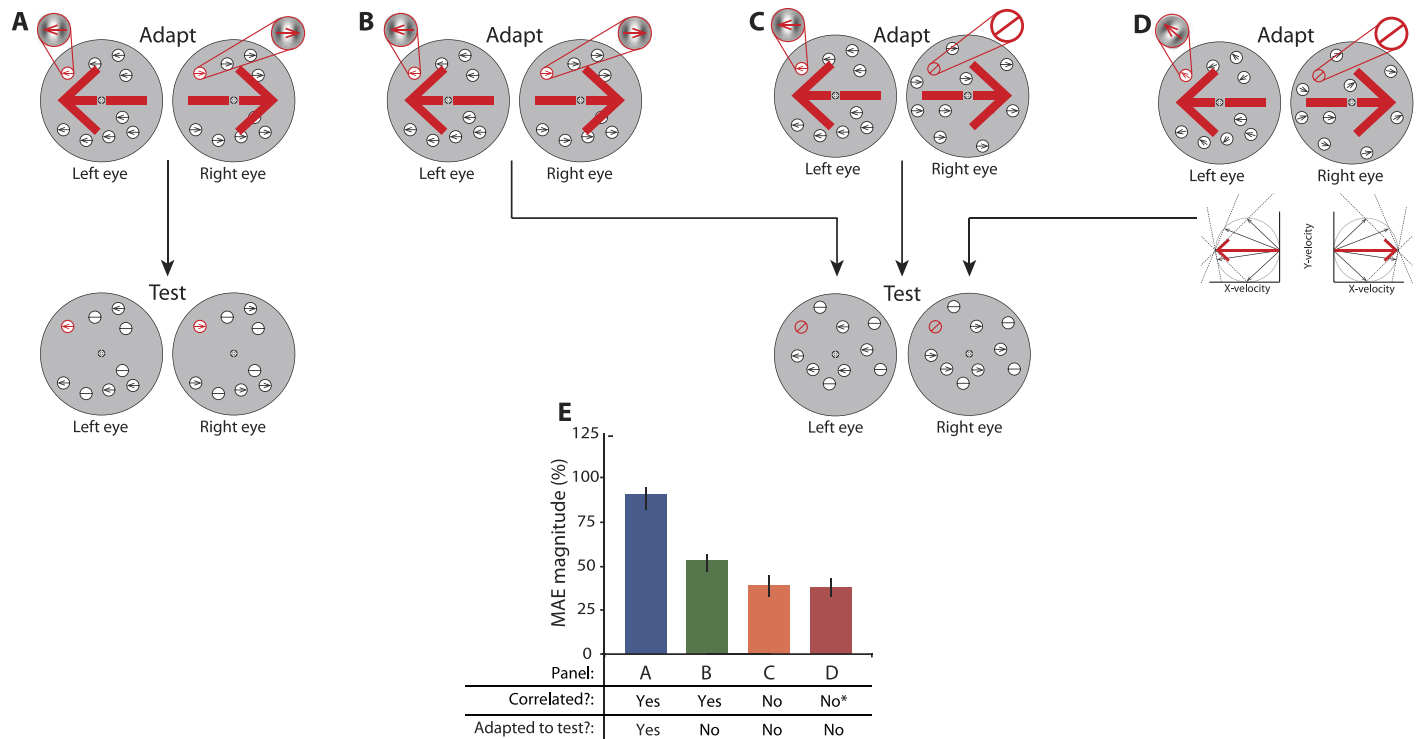
Figure 3. 3D motion aftereffects reveal global 3D motion integration and existence of eye-specific velocity information after binocular combination. (A–B) Experiment 2: Adaptor Gabor elements in each eye match binocularly. (A) Experiment 2A (3D Local): Elements of the test stimuli were in the same location as the adapting stimuli. (B) Experiment 2B (3D Global): Elements of the test stimuli were all at locally unadapted locations. (C–D) Experiment 3: Adaptor Gabor elements were constrained to fall on noncorresponding retinal locations in the two eyes. (C) Experiment 3A: All Gabor elements in both eyes had the same orientation and motion, and thus all of the local motion velocities were the same as the global motion velocity. (D) Experiment 3B (Pseudoplaid): The orientation of each Gabor element was drawn randomly from a uniform distribution between 20° and 70° from vertical, and each Gabor element was constrained to have a single velocity based on intersection of constraints. The inset shows an example of constructing a single velocity using four velocity components. (E) The averaged magnitudes of motion aftereffects for each of the 3D adaptation conditions. The error bars are 95% bootstrapped confidence intervals. * = pseudoplaid condition.

both comprised small drifting Gabor patches (Figure 2A). Test stimuli were presented at either the adapted or the unadapted locations (Figure 2B and 2C).

Figure 2D shows the PSE estimates and psychometric functions, after adaptation to each direction (left and right), for one example observer. The PSE estimates from the QUEST staircases were similar to the point at which psychometric functions crossed 50% on the y-axis. The magnitudes of 2D MAEs were dependent on the spatial congruency between adapting and test stimuli: There were strong 2D MAEs when the test Gabor elements were situated in the same location as the adapting Gabor elements (Figure 2E; the Local condition), while there were very weak, if any, 2D MAEs when the same test Gabor elements were placed in locations different from the adapting Gabor elements (Figure 2E; the Global condition). These results demonstrate that 2D motion adaptation is inherently local, consistent with previous single-neuron recordings (Kohn & Movshon, 2003).

Experiment 2: Local versus global adaptation in 3D motion-through-depth processing

Next, we measured 3D MAEs using the same drifting Gabor patches as in the 2D MAE experiment presented in the same retinal locations with opposite velocities between the left and right eye. Test stimuli were the same as adapting stimuli, but they were presented in either adapted (Figure 3A; the Local condition) or unadapted locations (Figure 3B; the Global condition).

When test Gabor elements were placed in adapted locations (Experiment 2A), there were strong 3D MAEs—confirming that our stimuli yielded 3D direction-selective adaptation (Figure 3E; blue bar). Crucially, even when test Gabor elements were positioned in unadapted locations (Experiment 2B), there were robust 3D MAEs, though smaller compared to when adapting and test Gabor elements shared the same retinal locations (Figure 3E; green

bar). These results suggest that 2D motion processing is spatially local, while 3D motion processing is more spatially global.

Experiment 3: 3D motion mechanisms adapt to globally integrated motion signals

In Experiment 3, we assessed whether eye-specific velocity information is combined at the site of binocular combination (i.e., V1) or at later stages of motion processing to give rise to 3D direction selectivity. We used two adapting stimuli consisting of Gabor patches with opposite directions and nonoverlapping retinal locations in each eye. Each Gabor patch in one eye could drift either in the same direction (Experiment 3A; Figure 3C) or at a variety of constrained velocities to create a pattern motion from the component velocities (pseudoplaids) that were compatible with the left/right direction (Experiment 3B; Figure 3D). The test stimuli remained the same across the 3D MAE experiments.

We reasoned that if 3D direction selectivity arises at the level of binocular combination and is carried to downstream motion-processing mechanisms, we would not observe 3D MAEs using these adapting stimuli, because there were no matching local motion signals between the eyes during adaptation. However, any measurable 3D MAEs would suggest that eye-specific global motion information remains available, and is used, to compute 3D motion at later stages of motion processing.

We did in fact find robust 3D MAEs for the adapting stimuli in which there was no binocular correspondence (Figure 3E; orange bar). Further, we also found similar 3D MAE magnitudes by adaptation to the dichoptic pseudoplaid stimulus (Figure 3E; red bar), suggesting that eye-specific 2D pattern-motion information is preserved at later stages of visual processing to compute 3D motion. Supplementary Movies S1 through S3 show example trials (4-s top-up adaptation followed by 1-s test with zero motion coherence) in our experiments. Observers should perceive the opposite direction of adapting direction (rightward in Supplementary Movie S1A and Supplementary Movie S1B, and toward in Supplementary Movie S2A, Supplementary Movie S2B, Supplementary Movie S3A, and Supplementary Movie S3B) after several repetitions.

Discussion

The preceding set of adaptation experiments strengthens the notion that monocular motion signals exist at a relatively late stage of visual processing.

Further, our results suggest that these signals are constructed by integrating local motion signals over a fairly large area before comparing across the eyes to yield a 3D motion signal. The results add additional support for the growing case to reconsider the nature of binocularity in the visual cortices; the common assumption that eye-of-origin information is wholly lost as signals leave V1 is inconsistent with the way that eye-specific velocities are exploited in 3D motion processing.

Our observations of local 2D direction-selective adaptation are consistent with findings in previous single-unit recording studies. Monkey MT neurons show location-specific 2D direction-selective adaptation effects within the receptive field (Kohn & Movshon, 2003). Although a hallmark of MT motion processing is the emergence of 2D pattern-motion direction sensitivity (as distinct from simpler 1D component-motion sensitivity), this encoding is primarily local—that is, at the scale of receptive fields smaller than area MT, perhaps V1 or V2 (Hedges et al., 2011; Majaj et al., 2007).

In contrast, 3D direction-selective adaptation effects revealed more global integration. Adaptation effects were robust to a lack of local spatial congruency between adapting and test stimuli, implying neural encoding within large, monocular receptive fields. Moreover, 3D adaptation was robust to a lack of strict binocular correspondence between the eyes, indicating that the large spatial integration almost certainly took place before the binocular construction of the 3D signal. Our results cannot be explained by the standard model (i.e., cyclopean view) that eye-specific signals are effectively merged into a single stream at the point of binocular combination within V1.

We have shown that monocular adaptation cannot account for 3D MAEs (Czuba et al., 2011; Rokers et al., 2009). Czuba et al. (2011) showed that monocular MAEs are virtually identical in magnitude to binocularly viewed frontoparallel MAEs (quantified as 19% and 18% motion coherence, respectively). Furthermore, if 3D MAEs simply reflect monocular MAEs and their binocular interaction, 3D MAE magnitude would be explained by an appeal to monocular MAEs that result from adaptation to one direction in one eye and the other direction in the other. However, we found that monocular MAEs after adaptation to 3D motion were very small (9% motion coherence). Thus, monocular adaptation cannot come close to accounting for 3D MAEs (44% motion coherence). In the current study, we showed in Experiment 1 that there were virtually no 2D MAEs when test stimuli were presented in unadapted locations. If the 3D MAE in Experiment 2 had been due to monocular (2D) MAEs, we would not have observed robust 3D MAEs in Experiment 2B, in which test stimuli were displayed in unadapted

locations. Contrary to the notion of 2D MAEs being required to create 3D MAEs, we found large 3D MAEs in spatial-mismatch conditions that do not yield 2D MAEs. In concert with a similar quantitative dissociation in fMRI work (Rokers et al., 2009), 3D MAEs are most simply explained by an additional stage of adaptation selective for 3D direction.

It is known that ocular dominance in area MT is relatively weak but of course is not perfectly balanced in every neuron (DeAngelis & Newsome, 1999; Maunsell & van Essen, 1983). Theoretical work, however, has shown that an unmixing algorithm can successfully retrieve the left and right images from binocularly mixed signals despite loss of explicit eye-of-origin information (Lehky, 2011), providing a possible explanation for interactions between binocular neurons at multiple visual-processing sites after the convergence of signals from the two eyes in V1 (Blake & Logothetis, 2002; Leopold & Logothetis, 1996). It is theoretically plausible that left and right velocity information can also be unmixed for 3D motion computations at later stages of visual processing after binocular combination. In fact, we have recently shown that small differences in monocular sensitivity can, in fact, be leveraged at the population level to give rise to 3D directional discrimination, so it is possible that 3D information that requires eye of origin can be computed statistically (Bonnen, Czuba, Kohn, Cormack, & Huk, 2017) without requiring strictly monocular responsivity such as is present in the optic nerves and lateral geniculate nucleus of primates.

Pattern-direction selectivity of MT neurons seems to disappear when two component gratings are presented in different spatial locations or presented dichoptically within the receptive field (Majaj et al., 2007; Tailby et al., 2010). Given these electrophysiological findings, it is possible that 3D direction selectivity in the dichoptic pseudoplaid condition might arise in different visual areas compared to other conditions. However, note that our stimuli are different in many key features (size, speed, density, and eccentricity) compared to those used in previous electrophysiological studies. Furthermore, classical single-neuron recordings were conducted using anesthetized monkeys, leaving a possibility of pattern-direction selectivity in V1 in awake behaving animals (Pack & Born, 2001; but see also Movshon, Albright, Stoner, Majaj, & Smith, 2003). The use of the dichoptic pseudoplaid stimuli, for which higher areas than area MT might be involved to resolve the local ambiguity, could be a useful means of studying the locus of 3D direction selectivity (Rider, Nishida, & Johnston, 2016).

It is important to distinguish between pattern motion (plaid stimuli) and global pattern motion (pseudoplaid stimuli). Plaid stimuli consist of two gratings with different orientations superimposed at the same location. There are some controversies about the site of pattern selectivity, whether MT or V1 (Movshon et al., 2003; Pack & Born, 2001; van Kemenade, Seymour, Christophel, Rothkirch, & Sterzer, 2014). However, this does not weaken our conclusions regarding spatially-global 3D computations beyond V1. Our pseudoplaid stimuli comprise many small gratings with different orientations and velocities specified by the intersection of constraints. Although we do not know the exact site of global pattern-motion selectivity (Majaj et al., 2007), it is certain that global pattern motion cannot be computed at the level of V1.

In summary, our findings show that key aspects of 3D direction selectivity reflect eye-specific velocity information at later stages of processing (i.e., beyond V1) within the canonical motion pathway. Whether such eye-specific velocity information involves truly monocular signals, a meaningful exploitation of imperfect ocular balance (i.e., ocular dominance), or some other computation will require additional investigation (Adams & Horton, 2009; Huk, 2012; Lehky, 2011; Schwarzkopf, Schindler, & Rees, 2010). Single-neuron recordings using the sorts of stimuli used in our psychophysical studies may be particularly effective in this dissection.

Acknowledgments

Corresponding author: Sung Jun Joo.
Email: sungjun@pusan.ac.kr.
Address: Department of Psychology, Pusan National University, Busan, Republic of Korea.

References

Adams, D. L., & Horton, J. C. (2009). Ocular dominance columns: Enigmas and challenges. *The Neuroscientist*, *15*(1), 62–77.

Adelson, E. H., & Movshon, J. A. (1982, December 9). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525.

Alais, D., & Blake, R. (1999). Neural strength of visual attention gauged by motion adaptation. *Nature Neuroscience*, *2*, 1015–1018.

Amano, K., Edwards, M., Badcock, D. R., & Nishida,

S. Y. (2009). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *Journal of Vision*, *9*(3):4, 1–25, https://doi.org/10.1167/9.3.4. [PubMed] [Article]

Barbur, J. L., Watson, J. D., Frackowiak, R. S., & Zeki, S. (1993). Conscious visual perception without VI. *Brain*, *116*(6), 1293–1302.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13–21.

Bonnen, K., Czuba, T., Kohn, A., Cormack, L., & Huk, A. (2017). Encoding and decoding in neural populations with non-Gaussian tuning: The example of 3D motion tuning in MT. *Journal of Vision*, *17*(10): 409, https://doi.org/10.1167/17.10.409. [Abstract]

Brainard, D. H (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Carney, T., & Shadlen, M. N. (1993). Dichoptic activation of the early motion system. *Vision Research*, *33*(14), 1977–1995.

Castet, E., Keeble, D. R. T., & Verstraten, F. A. J. (2002). Nulling the motion aftereffect with dynamic random-dot stimuli: Limitations and implications. *Journal of Vision*, *2*(4):3, 302–311, https://doi.org/10.1167/2.4.3. [PubMed] [Article]

Cormack, L. K., Czuba, T. B., Knöll, J., & Huk, A. C. (2017). Binocular mechanisms of 3D motion processing. *Annual Review of Vision Science*, *3*, 297–318.

Czuba, T. B., Rokers, B., Guillet, K., Huk, A. C., & Cormack, L. K. (2011). Three-dimensional motion aftereffects reveal distinct direction-selective mechanisms for binocular processing of motion through depth. *Journal of Vision*, *11*(10):18, 1–18, https://doi.org/10.1167/11.10.18. [PubMed] [Article]

Czuba, T. B., Rokers, B., Huk, A. C., & Cormack, L. K. (2010). Speed and eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion. *Journal of Neurophysiology*, *104*(5), 2886–2899.

DeAngelis, G. C., & Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *The Journal of Neuroscience*, *19*(4), 1398–1415.

Felleman, D. J., & Kaas, J. H. (1984). Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *Journal of Neurophysiology*, *52*(3), 488–513.

Hedges, J. H., Gartshteyn, Y., Kohn, A., Rust, N. C., Shadlen, M. N., Newsome, W. T., & Movshon, J.

A. (2011). Dissociation of neuronal and psychophysical responses to local and global motion. *Current Biology*, *21*(23), 2023–2028.

Hisakata, R., Hayashi, D., & Murakami, I. (2016). Motion-induced position shift in stereoscopic and dichoptic viewing. *Journal of Vision*, *16*(13):3, 1–13, https://doi.org/10.1167/16.13.3. [PubMed] [Article]

Huk, A. C. (2012). Multiplexing in the primate motion pathway. *Vision Research*, *62*, 173–180.

Joo, S. J., Czuba, T. B., Cormack, L. K., & Huk, A. C. (2016). Separate perceptual and neural processing of velocity- and disparity-based 3D motion signals. *The Journal of Neuroscience*, *36*(42), 10791–10802.

Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell Labs Technical Journal*, *39*(5), 1125–1162.

Julesz, B. (1971). *Foundations of cyclopean perception*. Oxford, UK: University of Chicago Press.

Kohn, A., & Movshon, J. A. (2003). Neuronal adaptation to visual motion in area MT of the macaque. *Neuron*, *39*(4), 681–691.

Lankheet, M. J., & Verstraten, F. A. (1995). Attentional modulation of adaptation to two-component transparent motion. *Vision Research*, *35*, 1401–1412.

Lee, A. L. F., & Lu, H. (2010). A comparison of global motion perception using a multiple-aperture stimulus. *Journal of Vision*, *10*(4):9, 1–16, https://doi.org/10.1167/10.4.9. [PubMed] [Article]

Lee, A. L. F., & Lu, H. (2012). Two forms of aftereffects induced by transparent motion reveal multilevel adaptation. *Journal of Vision*, *12*(4):3, 1–13, https://doi.org/10.1167/12.4.3. [PubMed] [Article]

Lehky, S. R. (2011). Unmixing binocular signals. *Frontiers in Human Neuroscience*, *5*, 78.

Leopold, D. A., & Logothetis, N. K. (1996, February 8). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*(6565), 549–553.

Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.

Majaj, N. J., Carandini, M., & Movshon, J. A. (2007). Motion integration by neurons in macaque MT is local, not global. *The Journal of Neuroscience*, *27*(2), 366–370.

Maunsell, J. H., & van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of Neuroscience*, *3*(12), 2563–2586.

Mitchell, D. E. (1966). A review of the concept of Panum's fusional areas. *Optometry & Vision Science*, 43(6), 387–401.

Movshon, J. A., Albright, T. D., Stoner, G. R., Majaj, N. J., & Smith, M. A. (2003). Cortical responses to visual motion in alert and anesthetized monkeys. *Nature Neuroscience*, 6(1), 3.

Norcia, A. M., & Tyler, C. W. (1984). Temporal frequency limits for stereoscopic apparent motion processes. *Vision Research*, 24(5), 395–401.

Pack, C. C., & Born, R. T. (2001, February 22). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, 409(6823), 1040–1042.

Panum, P. L. (1858). *Physiologische Untersuchungen über das Sehen mit zwei Augen*. Keil, Germany: Schwerssche Buchhandlung.

Patterson, R. (1999). Stereoscopic (cyclopean) motion sensing. *Vision Research*, 39(20), 3329–3345.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.

Rider, A. T., Nishida, S. Y., & Johnston, A. (2016). Multiple-stage ambiguity in motion perception reveals global computation of local motion directions. *Journal of Vision*, 16(15):7, https://doi.org/10.1167/16.15.7. [PubMed] [Article]

Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, 75(1), 53–64.

Rokers, B., Cormack, L. K., & Huk, A. C. (2009). Disparity- and velocity- based signals for 3D motion perception in human MT+. *Nature Neuroscience*, 12(8), 1050–1055.

Rokers, B., Czuba, T. B., Cormack, L. K., & Huk, A. C. (2011). Motion processing with two eyes in three dimensions. *Journal of Vision*, 11(2):10, https://doi.org/10.1167/11.2.10. [PubMed] [Article]

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, 9(11), 1421–1431.

Sanada, T. M., & DeAngelis, G. C. (2014). Neural representation of motion-in-depth in area MT. *The Journal of Neuroscience*, 34(47), 15508–15521.

Scarfe, P., & Johnston, A. (2011). Global motion coherence can influence the representation of ambiguous local motion. *Journal of Vision*, 11(12):6, 1–11, https://doi.org/10.1167/11.12.6. [PubMed] [Article]

Schwarzkopf, D. S., Schindler, A., & Rees, G. (2010). Knowing with which eye we see: Utrocular discrimination and eye-specific signals in human visual cortex. *PLoS One*, 5(10), e13775.

Sclar, G., Maunsell, J. H. R., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, 30(1), 1–10.

Shadlen, M., & Carney, T. (1986, April 4). Mechanisms of human motion perception revealed by a new cyclopean illusion. *Science*, 232(4746), 95–98.

Snowden, R. J., & Milne, A. B. (1997). Phantom motion aftereffects—Evidence of detectors for the analysis of optic flow. *Current Biology*, 7, 717–722.

Sohn, W., & Lee, S.-H. (2009). Asymmetric interaction between motion and stereopsis revealed by concurrent adaptation. *Journal of Vision*, 9(6):10, 1–15, https://doi.org/10.1167/9.6.10. [PubMed] [Article]

Stoner, G. R., & Albright, T. D. (1992, July 30). Neural correlates of perceptual motion coherence. *Nature*, 358(6385), 412–414.

Tailby, C., Majaj, N. J., & Movshon, J. A. (2010). Binocular integration of pattern motion signals by MT neurons and by human observers. *The Journal of Neuroscience*, 30(21), 7344–7349.

Tyler, C. W., & Julesz, B. (1978). Binocular cross-correlation in time and space. *Vision Research*, 18(1), 101–105.

van Kemenade, B. M., Seymour, K., Christophel, T. B., Rothkirch, M., & Sterzer, P. (2014). Decoding pattern motion information in V1. *Cortex*, 57, 177–187.

Watson, A. B., & Pelli, D. G. (1983) QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2), 113–120.

Supplementary Material

Supplementary Movie S1A. An example trial in Experiment 1A (2D adaptation). The stimulus elements are interocularly correlated. The test stimulus elements are in the same locations as the adapting stimulus elements. The movie shows 4-s adaptation to 2D motion (rightward) followed by 1-s test stimuli with zero coherence. There was a 1-s blank period between adapting and test stimuli. The red circle is for demonstration only and indicates whether the stimuli are interocularly correlated and the test stimulus elements are in the same locations as the adapting stimulus elements.

Supplementary Movie S1B. An example trial in Experiment 1B (2D adaptation). The stimulus elements

are interocularly correlated. The test stimulus elements are in different locations from the adapting stimulus elements.

Supplementary Movie S2A. An example trial in Experiment 2A (3D adaptation). The stimulus elements are interocularly correlated. The test stimulus elements are in the same locations as the adapting stimulus elements.

Supplementary Movie S2B. An example trial in Experiment 2B (3D adaptation). The stimulus elements are interocularly correlated. The test stimulus elements are in different locations from the adapting stimulus elements.

Supplementary Movie S3A. An example trial in Experiment 3A (3D adaptation). The stimulus elements are not interocularly correlated. The test stimulus elements are in different locations from the adapting stimulus elements.

Supplementary Movie S3B. An example trial in Experiment 3B (3D adaptation; the Pseudoplaid condition). The stimulus elements are not interocularly correlated. The test stimulus elements are in different locations from the adapting stimulus elements.