

Conjunctions between Motion and Disparity Are Encoded with the Same Spatial Resolution As Disparity Alone

Fredrik Allenmark and Jenny C. A. Read

Institute of Neuroscience, Newcastle University, Newcastle upon Tyne NE2 4HH, United Kingdom

Neurons in cortical area MT respond well to transparent streaming motion in distinct depth planes, such as caused by observer self-motion, but do not contain subregions excited by opposite directions of motion. We therefore predicted that spatial resolution for transparent motion/disparity conjunctions would be limited by the size of MT receptive fields, just as spatial resolution for disparity is limited by the much smaller receptive fields found in primary visual cortex, V1. We measured this using a novel “joint motion/disparity grating,” on which human observers detected motion/disparity conjunctions in transparent random-dot patterns containing dots streaming in opposite directions on two depth planes. Surprisingly, observers showed the same spatial resolution for these as for pure disparity gratings. We estimate the limiting receptive field diameter at 11 arcmin, similar to V1 and much smaller than MT. Higher internal noise for detecting joint motion/disparity produces a slightly lower high-frequency cutoff of 2.5 cycles per degree (cpd) versus 3.3 cpd for disparity. This suggests that information on motion/disparity conjunctions is available in the population activity of V1 and that this information can be decoded for perception even when it is invisible to neurons in MT.

Introduction

Human spatial resolution for disparity-defined depth is much worse than for luminance information. This is believed to be because spatial resolution for disparity is limited by the overall sizes of receptive fields (RFs) in primary visual cortex, whereas spatial resolution for luminance is limited by the size of their ON/OFF subregions (Banks et al., 2004; Nienborg et al., 2004; Filippini and Banks, 2009). Thus, information about the fine detail of disparity, potentially available within the photoreceptor activations, is lost at an information bottleneck in V1. Further information regarding disparity is lost in bottlenecks between V1 and subsequent levels of cortical processing. For example, V1 neurons encode absolute disparity with much greater sensitivity than humans can perceive (Cumming and Parker, 1999; Prince et al., 2000), and sensitivity can be reduced further by changes to the stimulus that do not alter the information available in V1 (McKee et al., 2004). Similarly, binocular fusion causes us to lose sensitivity to monocular position (McKee and Harrad, 1993). In this paper, we examined whether a similar information bottleneck affects the spatial resolution with which humans can detect conjunctions between horizontal motion and disparity.

A minority of V1 neurons are tuned both to disparity and to direction of motion (Grunewald and Skoumbourdis, 2004; Read and Cumming, 2005b). Many direction-selective V1 neurons

project to cortical area MT (Movshon and Newsome, 1996), where most neurons are tuned to both disparity and direction of motion (Bradley et al., 1995; DeAngelis and Uka, 2003). MT is widely believed to be critical for the perception of motion (Newsome and Paré, 1988; Salzman et al., 1992; Britten et al., 1996; Majaj et al., 2007). However, MT receptive fields are ~10 times larger than in V1 (Gattass and Gross, 1981). We wondered therefore whether V1 information about conjunctions between disparity/motion might be lost perceptually.

Of course, receptive fields in all extrastriate areas are larger than those in V1, and yet as we have seen, the spatial resolution of disparity perception is as good as the V1 representation. This is because neurons in many extrastriate areas respond best to variations in disparity across their receptive fields (Janssen et al., 1999; Sakata et al., 1999; von der Heydt et al., 2000; Nguyenkim and DeAngelis, 2003; Bredfeldt and Cumming, 2006). In contrast, no studies have reported neurons that prefer variations in joint motion/disparity information—for example, that respond best to stimuli containing leftward motion/far disparity in one receptive field subregion and to rightward motion/near disparity in another. Rather, MT neurons seem to prefer similar conjunctions of motion and disparity all across their large receptive fields. We predicted that this would be reflected in a low spatial resolution for motion/disparity conjunctions.

To test this prediction, we introduced a “joint motion/disparity grating,” a random-dot pattern in which the pairing between horizontal motion and disparity alternated as a function of vertical position. That is, in alternate horizontal strips, near dots moved left while far dots moved right, or near dots moved right while far dots moved left (Fig. 1A). This is different from either a pure disparity grating built from moving dots (Fig. 1B) or a pure motion grating built from two depth planes (Fig. 1C), both of which we also used for comparison. In each case, we asked sub-

Received July 8, 2011; revised July 25, 2012; accepted July 26, 2012.

Author contributions: F.A. and J.C.A.R. designed research; F.A. performed research; F.A. analyzed data; F.A. and J.C.A.R. wrote the paper.

This work was supported by a Newcastle University Institute of Neuroscience PhD studentship (F.A.) and by Royal Society University Research Fellowship UF041260 and Medical Research Council New Investigator Award 80154 (J.C.A.R.).

Correspondence should be addressed to Fredrik Allenmark at the above address. E-mail: fredrik.allenmark@ncl.ac.uk.

DOI:10.1523/JNEUROSCI.3495-11.2012

Copyright © 2012 the authors 0270-6474/12/3214331-13\$15.00/0

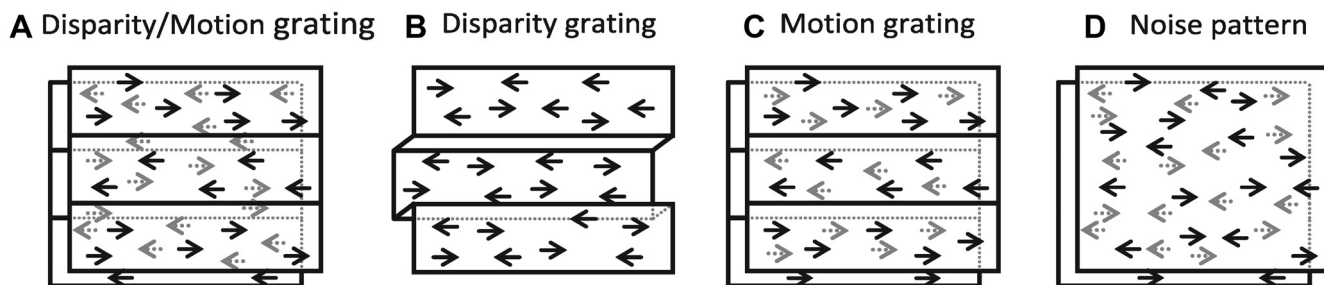


Figure 1. Sketches of the different types of stimuli used. Notice that, in every case, the same speeds and disparities were present. The “pure disparity” grating is built from moving dots; there are leftward and rightward dots everywhere in the stimulus, but the depth of the dots alternates as a function of vertical position. Similarly the “pure motion” grating contains two transparent depth planes, but the direction of motion of dots in the two planes alternates. The joint disparity/motion grating (A) cannot be detected from either pure disparity or pure motion information. If viewed with one eye, so removing disparity information, both directions of motion are present everywhere in the stimulus, making it indistinguishable from the noise (D). On any one frame (i.e., removing motion information), dots in both depth planes are present everywhere in the stimulus, again making it indistinguishable from noise. In contrast, the pure disparity grating (B) becomes indistinguishable from noise if disparity information is removed but not if motion information is removed, and the opposite is true for the motion grating (C).

The three different types of grating

Noise

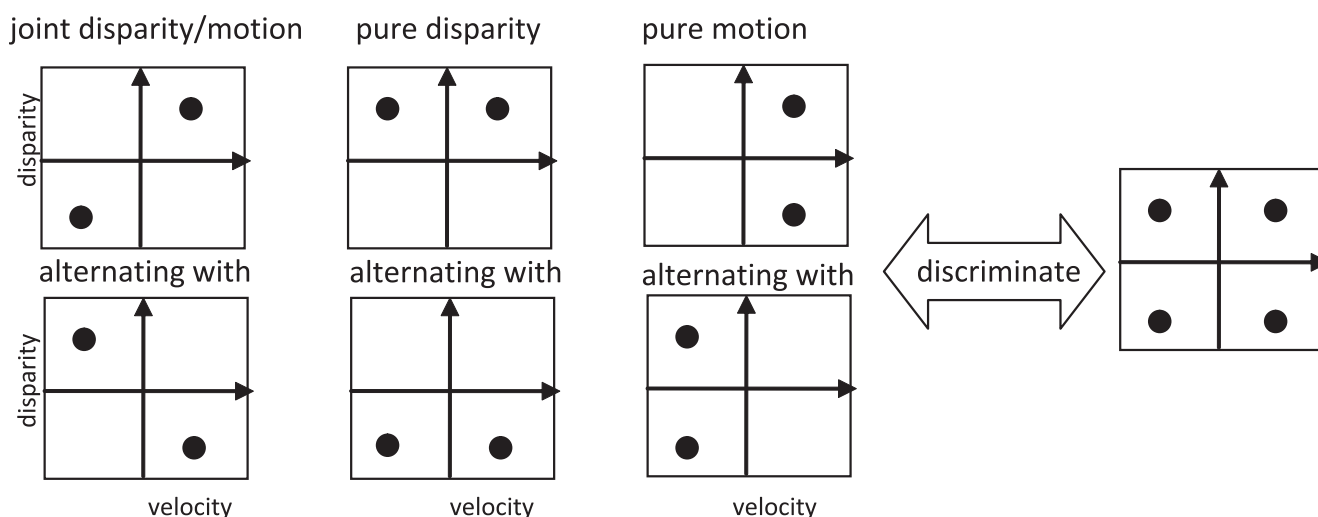


Figure 2. The task in our grating discrimination experiment, sketched in disparity/velocity space. In the gratings, the dot disparities and velocities alternate as a function of vertical position in the image. The noise contains the same velocities and disparities, but without the spatial structure.

jects to discriminate the “signal” grating from “noise,” shown in Figure 1D. Figure 2 represents the stimuli in disparity/velocity space. To a system that detects only disparity, or to one that detects only motion, the joint motion/disparity grating is indistinguishable from noise. Thus, this task requires mechanisms that extract both motion and disparity and the correlations between them (Qian and Andersen, 1997; Anzai et al., 2001; Read and Cumming, 2005c; Qian and Freeman, 2009).

In the same subjects, we probed the spatial resolution for each of these three types of gratings, using correlation thresholds to equalize task difficulty, and obtain an unbiased estimate of spatial resolution. Using a signal detection theory model, we extracted estimates of the receptive field size and internal noise with which the brain detects each type of grating.

Materials and Methods

Equipment. The experiments were performed in a dark room. Stimuli were projected on a projection screen (300 × 200 cm; Stewart Filmscreen 150; www.stewartfilm.com; supplied by Virtualis), which the observers viewed from a distance of 160 cm. The subject’s head was stabilized using a chin rest (UHCOTech HeadSpot). Two projectors, projecting through polarizing filters, were used to separate the two eye’s images. The interocular cross talk was <2%. White had a luminance of 4 cd/m² and black

had a luminance of 0.07 cd/m². The projected image was 71 × 53 cm subtending 25 × 19°. The stimuli were presented in the central region of the image and had a size of 500 × 500 pixels (9 × 9°). The dot size was 2 × 2 pixels (2.1 × 2.1 arcmin).

Stimuli. Stimuli were presented using MATLAB (MathWorks; www.mathworks.com) with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). The stimuli used were random-dot stereograms with equal numbers of dots moving to the left and to the right with equal speed, depicting either a grating or a noise pattern. Three different kinds of gratings were used. The first type of grating had two transparent depth planes and was made up of horizontal strips of equal width, where in each strip all dots moving to the left were in one depth plane and all dots moving to the right were in the other and where the direction of movement in the different depth planes was alternated between adjacent strips (Fig. 1A). The second type of grating was a horizontal square-wave in depth made up of equal numbers of dots moving in both directions (Fig. 1B). The third type of grating consisted of two transparent planes in depth with horizontal strips, in which all dots in a single strip moved in the same direction and the direction of motion alternated between adjacent strips (Fig. 1C). The noise patterns consisted of two transparent depth planes with an equal number of dots moving in both directions in both planes (Fig. 1D). Any individual monocular frame of any stimulus was simply a structureless random-dot pattern with 150 dots per degree². Critically, all the stimuli contained both directions of motion and both

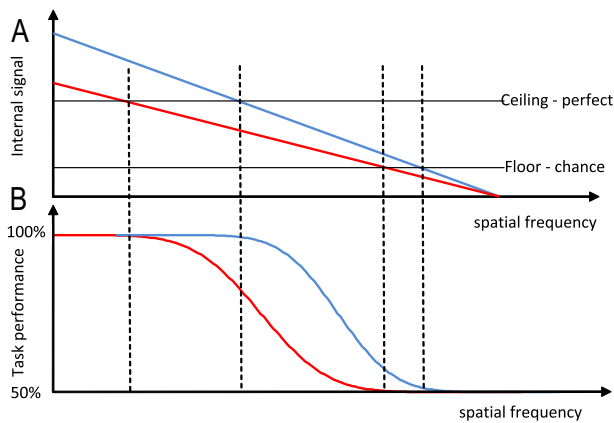


Figure 3. Diagram of a possible relationship between internal signal and performance for two different tasks, represented in red and blue, which could lead to erroneous conclusions about spatial resolution. See text for details.

depth planes. This ensures that any differences in performance are solely due to the task-relevant differences between the stimuli. That said, in pilot experiments we found no difference between performance on static disparity gratings and moving disparity gratings with the speeds used here. We also did not find any difference between performance on motion gratings with one or two depth planes. We therefore expect that our results would not have been significantly different if we had used static disparity gratings and/or motion gratings with a single depth plane.

A problem with comparing resolution for different grating types is that one task may be harder than another. For example, detecting a joint motion/disparity grating requires information from two visual modalities to be combined, and thus arguably requires a more challenging judgment than, say, detecting a motion grating. This could lead to erroneous conclusions regarding resolution. For example, consider the toy example sketched in Figure 3. Figure 3A shows the internal signal for two hypothetical tasks, red and blue. These both have the same resolution, in that the signal is maximal for DC (0) and falls to zero at the same frequency. However, the red task is “harder,” in that, at any frequency, its signal is lower than the blue signal by a constant factor. Now suppose there is some nonlinearity converting this signal into perceptual judgments. In particular, there is a “floor” (when the signal falls below this level, perceptual performance on the relevant task is chance) and a “ceiling” (when the signal falls above this level, performance is perfect). Figure 3B shows the resulting performance. Performance falls at much lower frequencies for the red task, despite the fact that the dependence of the underlying signal on frequency is the same in both cases.

To avoid this problem, we used decorrelation to reduce the strength of the internal signal available for each task. This removed the ceiling effect, at least: if the internal signal was above ceiling, so that performance was perfect, we simply decreased correlation until the performance fell to 82%. In this way, we ensured that the difficulty of each task was equal.

For motion, “decorrelation” means reducing motion coherence; for disparity, it means reducing interocular correlation. Thus, for the pure motion gratings, we measured the motion coherence threshold at each frequency. The motion coherence was varied by, at each frame, giving each dot a probability p of being randomly repositioned rather than displaced in its direction of motion. The coherence level is defined as $1 - p$, such that for example a coherence level of 0.6 means that at any frame each dot had a 40% probability of being randomly repositioned.

For the pure disparity gratings, we measured the interocular correlation threshold at each frequency. The interocular correlation was varied by, in the first frame of the stimulus, giving each dot a probability p of being positioned randomly in both eyes, instead of randomly in one eye and then offset horizontally by the desired disparity in the other eye. In subsequent frames, interocularly uncorrelated dots moved smoothly with the specified motion until they vanished off the edge of the stimulus. For the joint motion/disparity gratings, we measured both correlation and coherence thresholds.

Observers. Ten human observers participated in the experiments: one of the authors (male) and nine inexperienced observers (seven females and two males). Observer CB was unable to perform the interocular correlation threshold parts of Experiment 2. Two of the observers (one male and one female) also participated in the short duration control experiment, although one of the two was only able to perform the interocular correlation threshold part of the experiment.

Tasks. To obtain the speed and disparity amplitude for which the subjects could best detect the joint motion/disparity gratings at high frequencies (Experiment 1), we used a one-interval task as well as a two-interval task. Amplitude is defined as one-half the peak-to-trough range of the waveform, $(\max - \min)/2$. For the one-interval task, in each trial either a grating or a noise pattern was presented and the task was to report, by a button press, whether a grating had been presented or not. The subjects were allowed to view the stimuli for as long as they desired before making a decision. For the two-interval task, one interval contained a grating and the other a noise pattern, and the task was to report, by a button press, which interval contained the grating. The interval length was 750 ms with a 200 ms blank between intervals. Subject PFA was tested with the one interval task and all other subjects with the two-interval task. Once the optimal speed and disparity amplitude had been determined for a subject, that speed and disparity amplitude were used in all further testing of that subject.

To obtain coherence and interocular correlation thresholds once the optimal speed and amplitude had been determined, we used adaptive QUEST staircases (Watson and Pelli, 1983) converging to 82% correct with a two-interval forced-choice task in which one interval contained a grating and the other interval contained a noise pattern and the task was to report, by a button press, which interval contained the grating. The interval length was either 500 or 750 ms with a 200 ms blank between intervals. The 500 ms interval length was used for subject PFA, who is an author and an experienced psychophysical observer, and the 750 ms interval length was used for all other subjects. Each staircase was repeated three times in the same session.

Results

Experiment 1: obtaining optimal stimulus parameters for each subject

In this paper, we wanted to detect the finest resolution with which motion and disparity information is represented. Obtaining four correlation/coherence thresholds at many different spatial frequencies was a long and demanding experiment, and it was not feasible to also examine dependence on speed and disparity amplitude at each frequency. We therefore began by measuring each subject’s performance as a function of speed and disparity only for a single, high frequency. In this way, we aimed to identify a pair of values in which the subject is able to perform well. Because the different subjects stopped being able to do the task at somewhat different frequencies, we had to choose a different “high” frequency for each subject. This frequency was based on initial testing (data not shown), which identified a frequency at which the subject could perform significantly above chance for at least one combination of speed and amplitude without performing close to 100% over too large a region.

Figure 4 shows performance on the joint motion/disparity grating detection task as a function of disparity amplitude and speed for all subjects, for perfectly correlated stimuli. In each case, there is a region of high performance surrounded by a region in which performance was lower. The amplitudes and speeds used in Experiment 2 were chosen for each subject individually to be approximately in the center of the region of high performance for that subject (Fig. 4, white crosses). Table 1 shows the values used for each subject in the subsequent experiments.

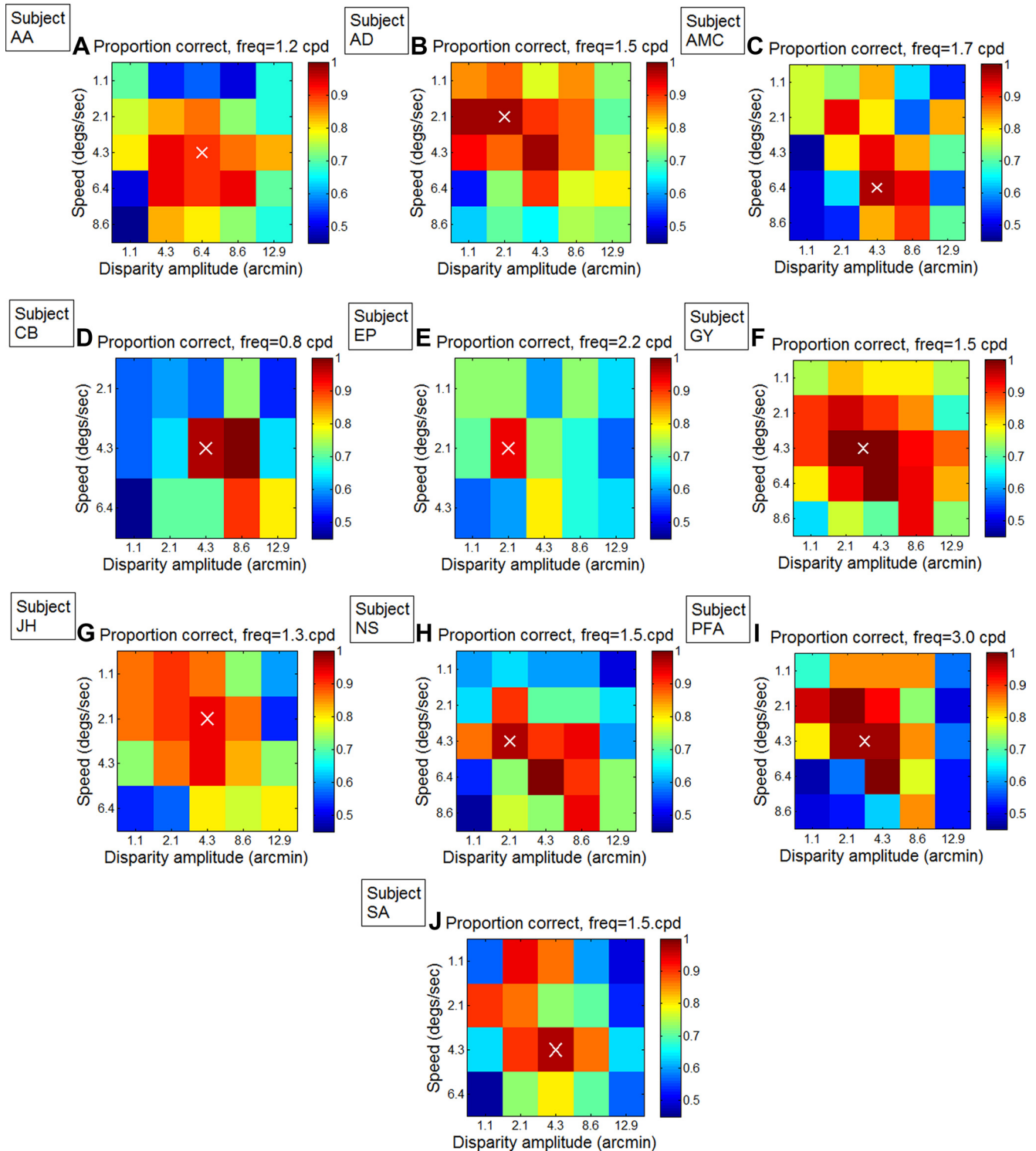


Figure 4. Performance on the 100%-correlated joint motion/disparity grating as a function of speed and disparity amplitude for all subjects. The white crosses show the values used in the subsequent experiments (Table 1).

Experiment 2

We now proceeded to measure coherence and correlation thresholds for the three different types of gratings. Figure 5 shows the motion coherence thresholds measured at different frequencies for both the motion/disparity gratings and the pure motion gratings. The error bars show ± 1 SE based on the three repetitions of each staircase. At low frequencies, subjects are able to perform the tasks at relatively low coherence; as the frequency increases, sub-

jects require progressively more coherence to be able to reach threshold. All subjects can detect motion gratings even at very low coherences, down to 20% at the lowest frequencies. For some subjects, there is little difference between the thresholds for the two types of gratings at low frequencies; PFA, for example, is equally good at detecting both sorts of grating. However, for some subjects, such as AMC in Figure 5, the coherence thresholds are far higher for the joint motion/disparity grating, even at the

Table 1. Speed of dot motion and disparity amplitude used in Experiment 2 chosen based on the results of Experiment 1

| | AA | AD | AMC | CB | EP | GY | JH | NS | PFA | SA |
|------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Speed (degrees/s) | 4.3 | 2.1 | 6.4 | 4.3 | 2.1 | 4.3 | 2.1 | 4.3 | 4.3 | 4.3 |
| Disparity amplitude (arcmin) | 6.4 | 2.1 | 4.3 | 6.4 | 2.1 | 3.3 | 4.3 | 2.1 | 3.3 | 4.3 |

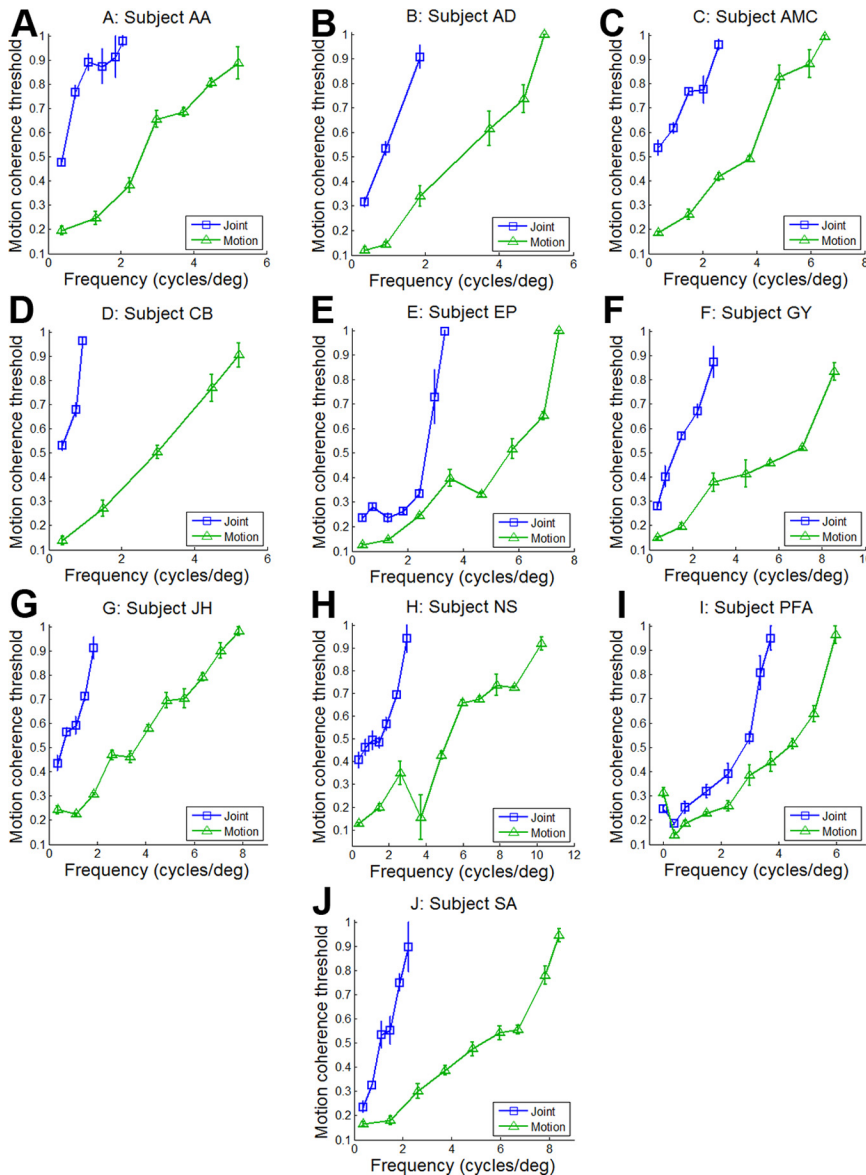


Figure 5. Motion coherence threshold as a function of frequency for the motion/disparity and pure motion gratings for all subjects. The speed and disparity amplitudes used for the gratings were set individually for each subject; values are in Table 1.

very lowest frequencies. This indicates that, for this subject, detecting joint motion/disparity gratings is a genuinely harder task than detecting motion gratings, regardless of their respective spatial resolutions. Thus, without the use of a coherence threshold, one could seriously misestimate the relative resolution in this subject (Fig. 3).

At higher frequencies, the thresholds become increasingly different for all subjects. The pure motion gratings can be detected up to frequencies at which the joint motion/disparity gratings are invisible, even at 100% motion coherence.

Figure 6 shows the interocular correlation thresholds measured at different frequencies for both the motion/disparity grat-

ings and the pure disparity gratings. Here, there is much less difference between the thresholds for the two different types of the gratings at low frequencies. For some subjects, this remains true at high frequencies, while for others, such as JH, there is a large difference at the highest frequencies.

In Figures 5 and 6, we have presented two different types of threshold for the joint motion/disparity gratings: interocular correlation and motion coherence thresholds. Figure 7 compares these two threshold measurements. For some subjects, the thresholds are comparable in the two cases, but where there is a systematic difference such that the thresholds all differ in the same direction at least up to some frequency close to the highest one tested (as for subject PFA in Fig. 7), it is the interocular correlation thresholds that are higher. This suggests that, despite their conceptual similarity, the two manipulations are not equivalent perceptually: reduction in interocular correlation has a more disruptive effect than reduction in motion coherence.

Constructing the model

To turn these measurements of coherence and correlation thresholds into a quantitative estimate of receptive field size, we used a model based on signal detection theory. We assumed that, for 100% correlated stimuli, the internal signal was proportional to the RMS of the unit-amplitude grating waveform after convolution by a Gaussian with SD σ . Recent work has suggested this is a good model for the detection of disparity gratings (Serrano-Pedraza and Read, 2010), in which the disparity signal can be computed by a population of energy model-like disparity-selective cells with Gaussian receptive field envelopes of diameter 2σ (Banks et al., 2004; Filippini and Banks, 2009; Allenmark and Read, 2010, 2011). We shall therefore refer to the parameter σ as “RF size,” taking 2σ as the RF diameter. Note that this model assumes each signal is initially encoded by neurons whose preferred signal value is constant across their receptive fields. For disparity, this is known to be the case (Nienborg et

al., 2004). Our model does not rule out neurons with RF subregions tuned to opposite signal values; indeed, such neurons would be ideal for the perceptual “readout” that detects the grating. However, in our model, these readout neurons would be limited by the resolution with which the signal was initially encoded; their RF subregions could not be smaller than 2σ .

Figure 8 shows how the RMS of the convolution between the Gaussian and the square wave varies as a function of the ratio between the SD of the Gaussian and the wavelength $1/f$ of the square wave. We write this function $RMS(f\sigma)$.

Reducing interocular correlation or motion coherence must reduce this internal signal. As we have seen, a decrease in intero-

cular correlation generally increases task difficulty more than the same decrease in motion coherence. This difference can apply only at intermediate values, since at 100% the interocular correlation and motion coherence versions of the stimulus were exactly the same, while at 0% there is no signal. This can be modeled by assuming that effective signal depends on the correlation/coherence level raised to some power, κ , allowing different values of κ for the correlation and coherence. We refer to κ as the “decorrelation parameter,” since it describes how seriously the available signal is degraded by decoherence/decorrelation. $\kappa = 1$ means that the signal degrades linearly with decoherence/decorrelation; $\kappa > 1$ gives faster degradation. With these assumptions, the internal signal available for performing the task is as follows:

$$\text{sig} = C^\kappa * \text{RMS}(f\sigma).$$

We then used signal detection theory to predict performance on the task. Since a two-interval task was used, the signal detection theory prediction is the following:

$$\text{PC} = 0.5 + 0.5 * \text{erf}\left(\frac{\text{sig}}{\sqrt{2}N}\right), \quad (1)$$

where PC is the proportion of correct answers, erf is the error function, sig is the signal, and N is the internal noise. At the 82% threshold, this yields the following:

$$0.82 = 0.5 + 0.5 * \text{erf}\left(\frac{C_{\text{thresh}}(f)^\kappa * \text{RMS}(f\sigma)}{\sqrt{2}N}\right), \quad (2)$$

from which we obtain the following:

$$C_{\text{thresh}}^{-1}(f) = \left(\frac{\text{RMS}(f\sigma)}{\sqrt{2}N * \text{erf}^{-1}(0.64)}\right)^{\kappa^{-1}}. \quad (3)$$

This means that a scaled version of the RMS curve from Figure 8 can be fitted to the coherence and correlation thresholds from Figures 5 and 6 by finding appropriate values of σ , N , and κ , giving us estimates of the receptive field diameter (2σ) and internal noise levels (N) relevant to each task. The effects on the correlation thresholds of changes in these parameters are illustrated in Figure 9. The RF size σ acts as a horizontal gain, with a doubling in σ halving the cutoff frequency. For $\kappa = 1$, noise N acts as a vertical gain, with a doubling in N doubling the correlation threshold. κ changes the shape of the curve, notably how rapidly sensitivity declines as frequencies increase from zero. Recall that the parameter κ was introduced to account for the difference between $C_{\text{decorr,joint}}$ and $C_{\text{decoh,joint}}$ (Fig. 7), given that the same RF size σ_{joint} and noise N_{joint} apply to both these datasets.

Within a subject, the RF and noise parameters for the two different sets of data on the joint motion/disparity task (i.e., the interocular correlation thresholds and the motion coherence thresholds) were assumed to be the same. Similarly, within each subject, the decorrelation parameter was kept the same for both motion coher-

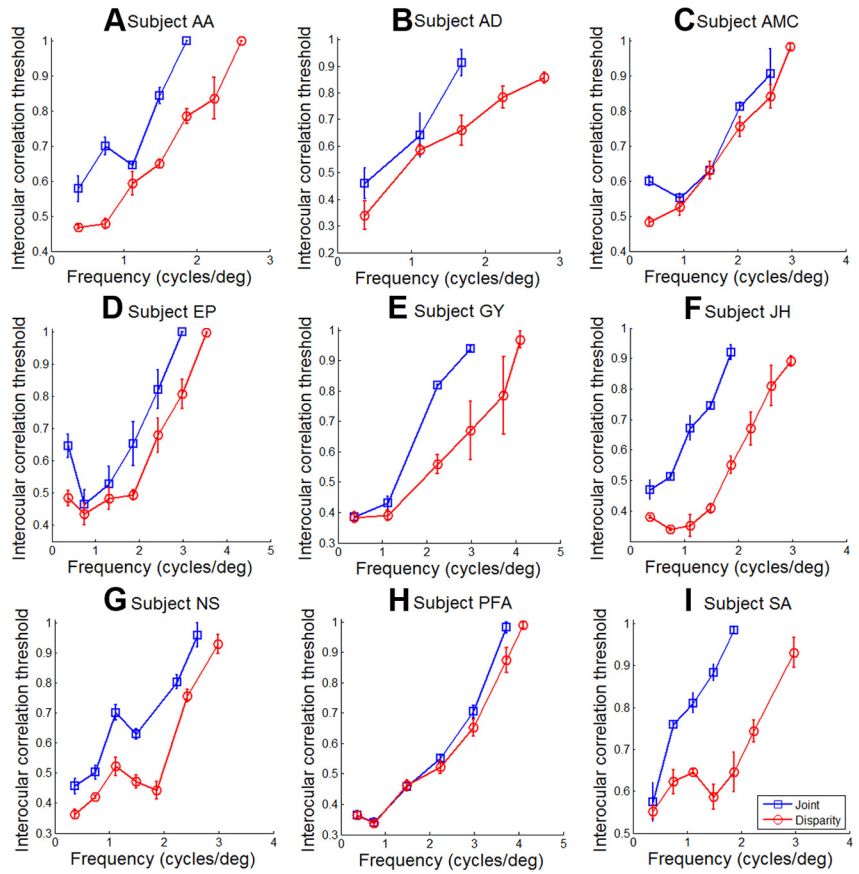


Figure 6. Interocular correlation threshold as a function of frequency for the motion/disparity and pure disparity gratings for all subjects.

ence datasets, and for both interocular correlation datasets. Therefore, there were, for each subject, eight parameters in total: RF diameters and noise parameters for the motion, disparity, and joint motion/disparity data and decorrelation parameters for the interocular correlation and motion coherence thresholds. These eight parameters were fitted to the four experimentally measured datasets $C_{\text{stimulus}}^{-1}(f_{\text{stimulus}})$ according to the following equations:

$$C_{\text{decorr,joint}}^{-1}(f_{\text{decorr,joint}}) = \left(\frac{\text{RMS}(f_{\text{decorr,joint}}\sigma_{\text{joint}})}{\sqrt{2}N_{\text{joint}} * \text{erf}^{-1}(0.64)}\right)^{\kappa_{\text{decorr}}^{-1}} \quad (4)$$

$$C_{\text{decoh,joint}}^{-1}(f_{\text{decoh,joint}}) = \left(\frac{\text{RMS}(f_{\text{decoh,joint}}\sigma_{\text{joint}})}{\sqrt{2}N_{\text{joint}} * \text{erf}^{-1}(0.64)}\right)^{\kappa_{\text{decoh}}^{-1}} \quad (5)$$

$$C_{\text{decorr,disp}}^{-1}(f_{\text{decorr,disp}}) = \left(\frac{\text{RMS}(f_{\text{decorr,disp}}\sigma_{\text{disp}})}{\sqrt{2}N_{\text{disp}} * \text{erf}^{-1}(0.64)}\right)^{\kappa_{\text{decorr}}^{-1}} \quad (6)$$

$$C_{\text{decoh,motion}}^{-1}(f_{\text{decoh,motion}}) = \left(\frac{\text{RMS}(f_{\text{decoh,motion}}\sigma_{\text{motion}})}{\sqrt{2}N_{\text{motion}} * \text{erf}^{-1}(0.64)}\right)^{\kappa_{\text{decoh}}^{-1}}. \quad (7)$$

We optimized the fit by minimizing the sum of squared errors over all four fits plus an additional term $\max(\kappa, \kappa^{-1})$ for each of the two decorrelation parameters. The additional term was included to keep either decorrelation parameter from growing too small/large. We used resampling to obtain error bars on the parameters by repeating the fitting 10,000 times, each time simulating a new repetition of each staircase by running a new staircase

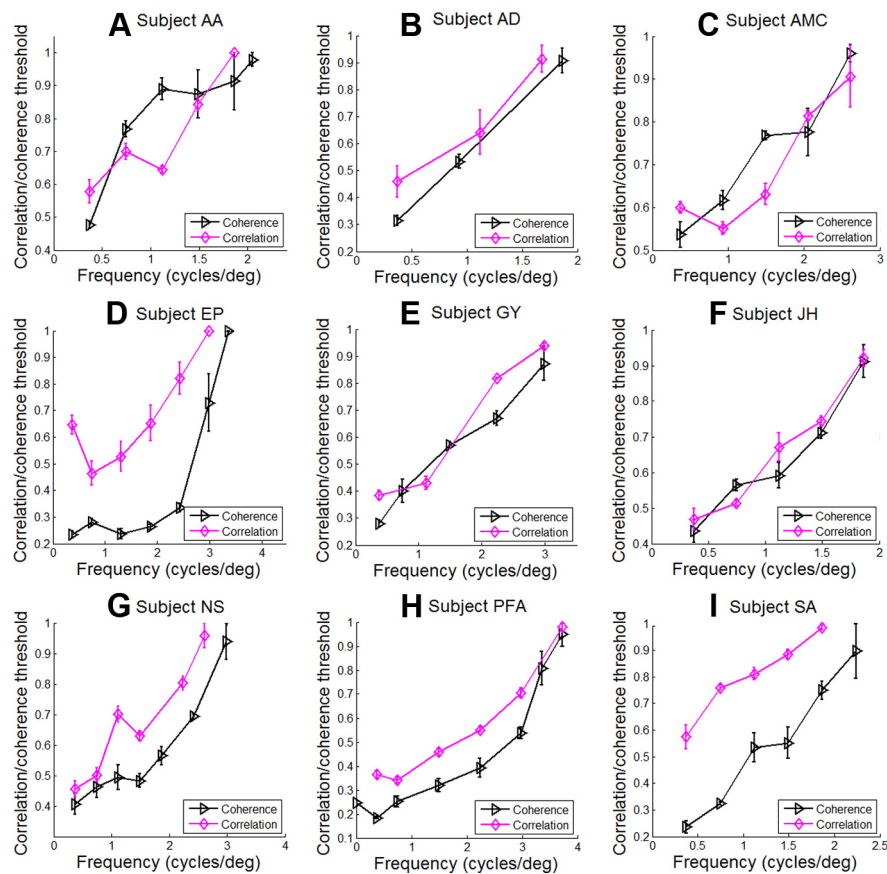


Figure 7. Intercorrelation thresholds and motion coherence thresholds for the motion/disparity gratings.

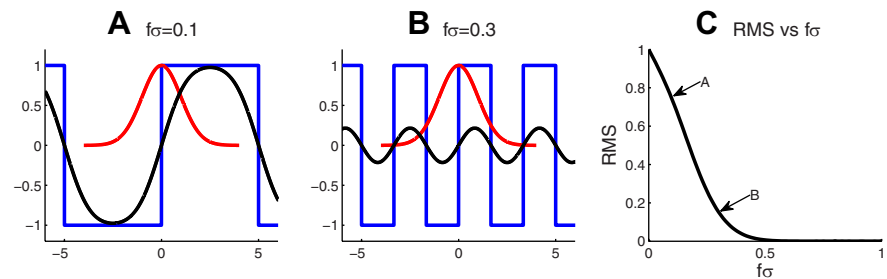


Figure 8. The resulting curves (black) when a square wave (blue) is convolved with a Gaussian (red) for a lower frequency square wave (A) and a higher frequency square wave (B) of the same amplitude, and a plot of the RMS of these resulting curves for a range of ratios between the SD of the Gaussian and the wavelength of the square wave (C).

with a simulated observer with the experimentally measured threshold.

Given the fit parameters, we can then use Equation 3 to estimate f_{max} , the highest grating frequency at which the task could be performed. At this frequency, performance is only at threshold even when the stimulus is perfectly coherent/correlated [i.e., $C_{thresh}(f_{max}) = 1$]. Thus f_{max} is given by the solution of the following:

$$RMS(f_{max}\sigma) = \sqrt{2N} * \text{erf}^{-1}(0.64). \quad (8)$$

Figures 10 and 11 show the inverted coherence and correlation thresholds along with the fits. The fits are generally good, validating the assumptions used in producing our model. The percentage of variance explained was at least 70% and at average 85% for the motion fits, at least 78% and at average 90% for the disparity

fits and at least 40% and at average 84% for the joint fits. Note from Equations 4–7 that each parameter affects more than one curve, so fits are not necessarily optimal for any individual curve.

Motion is encoded with smaller receptive fields and lower noise than disparity

Table 2 and Figure 12 show the parameters that gave the best fits for each subject. The receptive field sizes limiting detection are estimated at around 6 arcmin for the pure motion task and 8 arcmin for the pure disparity, similar to, although slightly larger than, the 6 arcmin previously estimated by Banks and colleagues (Banks et al., 2004; Filippini and Banks, 2009). Figure 13 shows the RF diameters and noise parameters from Figure 12, A and B, normalized by the values for the pure motion data. We see immediately that the RF diameter and neuronal noise estimated for the pure motion task are both smaller than for either the pure disparity or the joint motion/disparity task. This statement holds for all subjects individually, apart from subject AD in which the motion fit is poor (Fig. 10B). At a population level, the RF diameter for pure motion is significantly smaller than for pure disparity ($p < 0.05$, paired t test, $n = 9$, comparing σ_{motion} to $\sigma_{disparity}$; i.e., triangles vs circles in Fig. 12A) and for joint motion/disparity ($p < 0.01$, paired t test, $n = 10$, comparing σ_{motion} to σ_{joint} ; i.e., triangles vs squares in Fig. 12A). Similarly, the noise affecting pure motion judgments is significantly smaller than for pure disparity ($p < 0.01$, paired t test, $n = 9$, comparing N_{motion} to $N_{disparity}$; i.e., triangles vs circles in Fig. 12B) or for joint motion/disparity ($p < 0.01$, paired t test, $n = 10$, comparing N_{motion} to N_{joint} ; i.e., triangles vs squares in Fig. 12B). These two effects, smaller receptive fields and lower noise, combine to make motion gratings detectable up to higher frequencies than gratings defined by disparity. All subjects including AD can detect motion gratings up to higher frequencies than either pure disparity or joint motion/disparity gratings ($(f_{max}^{motion} > f_{max}^{disparity}, f_{max}^{motion} > f_{max}^{joint})$). Thus, our results show clearly that motion is encoded with higher resolution than disparity information, and also that it is affected by less neuronal noise.

Pure disparity is encoded with less noise than joint motion/disparity, but with similar-sized receptive fields

In contrast, there is no such clear difference between “joint” and “disp” (i.e., spatial resolution for pure disparity compared with spatial resolution for conjunctions between motion and disparity). Pure disparity gratings remain detectable up to slightly higher frequencies than joint motion/disparity gratings [3.3 vs 2.5 cycles per degree (cpd)], but this does not seem to reflect a

significant difference in receptive field size. The relative RF diameters estimated for the joint and for the pure disparity gratings show no consistent difference across our population. At the population level, the mean RF diameter is larger for the joint motion/disparity task than for the pure disparity task, but this difference is not significant either for the raw RFs (Fig. 12; $p = 0.07$, paired t test, $n = 9$) or after normalizing by the motion noise parameters (Fig. 13; $p = 0.15$, paired t test, $n = 9$). In contrast, the estimated noise level is larger for the joint than for the pure disparity wherever there is a significant difference (five of nine subjects), and this difference is significant on the population level both for the raw noise parameters (Fig. 12; $p < 0.05$, paired t test, $n = 9$) and after normalizing by the motion noise parameters (Fig. 13; $p < 0.05$, paired t test, $n = 9$). Thus, our analysis suggests that pure disparity and joint motion/disparity gratings are encoded with very similar spatial resolution. The pure disparity encoding is, however, subject to lower effective noise, meaning that pure disparity gratings can be detected up to somewhat higher frequencies than joint motion/disparity gratings, despite the similar RF sizes.

Possible fitting artifacts

These conclusions depend on the ability of our model to extract separate, reliable estimates of RF size and noise. There is some trade-off between RF size and noise: small increases in RF size, which tend to worsen performance, can be offset by small decreases in noise, which improve it. To quantify this trade-off, we performed bootstrap resampling on simulated data generated from the model. Model parameters were chosen to be the averages across subjects (i.e., the values shown in the mean column of Table 2). Simulated average data was generated by the model predictions (i.e., using Eq. 3) at a set of frequencies including all frequencies for which we have human data. Data for 1000 simulated experiments were then generated by, at each frequency, drawing a value from a normal distribution with the average data point as mean and the average SD in the human data, based on the three repetitions of each staircase, as SD. Curves based on the model were then fitted to the data from each simulated experiment, which resulted in 1000 sets of model parameters.

Figure 14 shows scatterplots of the relative noise and RF parameters obtained as described above. There is a highly significant negative correlation, Pearson’s correlation coefficient of -0.70 , between noise parameter and RF size parameter. Since

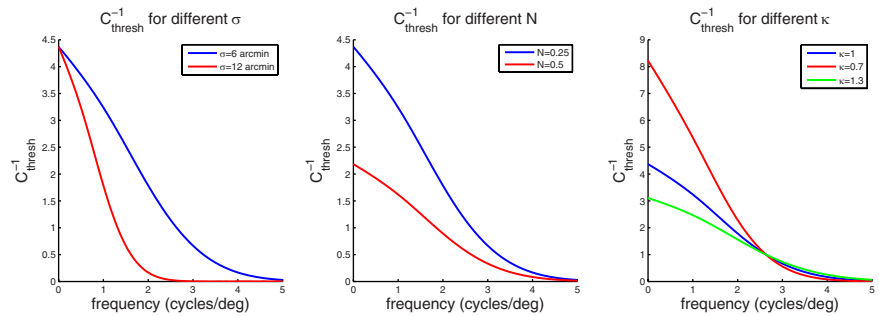


Figure 9. The effects of changing the different parameters. Except where stated otherwise, $\sigma = 6$ arcmin, $N = 0.25$, and $\kappa = 1$.

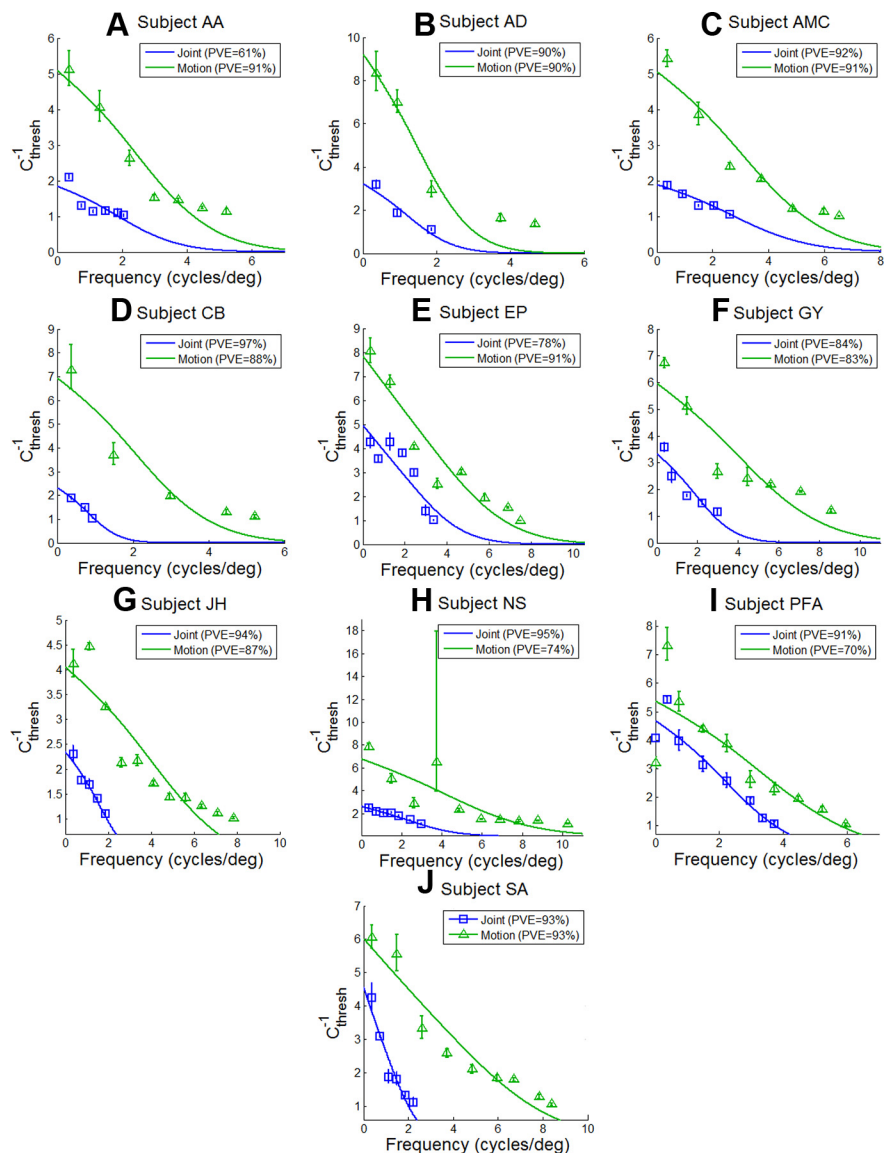


Figure 10. Inverted motion coherence thresholds as a function of frequency for the pure motion gratings (green) and joint motion/disparity gratings (blue) and model fits (see text) for all subjects.

this correlation is across simulated experiments in resampled data based on the data of a single “simulated average subject,” it has to be a consequence of a trade-off between the two parameters in the fitting process. This trade-off is most likely the reason why there is a small negative correlation between relative RF size

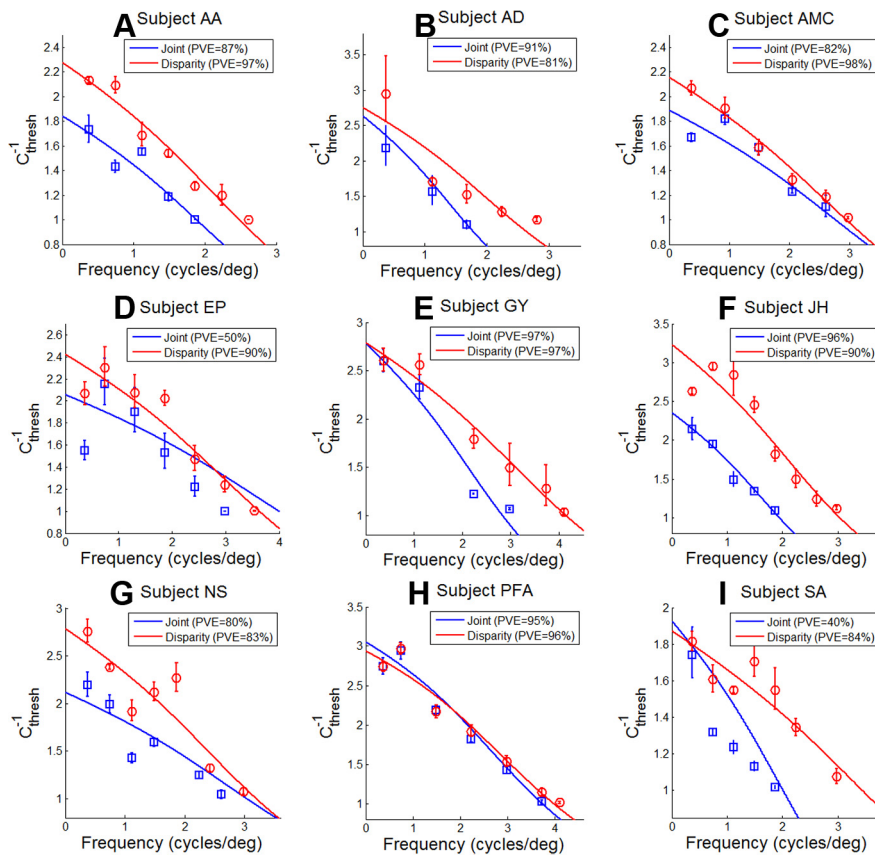


Figure 11. Inverted interocular correlation thresholds as a function of frequency for the pure disparity gratings (red) and joint motion/disparity gratings (blue) and model fits (see text) for all subjects.

and relative noise parameter across subjects, visible in Figure 13. However, this trade-off does not invalidate our conclusions. The effect of the trade-off is captured in the confidence intervals on the estimated model parameters (Figs. 12, 13), which were also generated by bootstrap resampling. Therefore, all conclusions drawn from the model parameters and their confidence intervals remain valid.

The trade-off is limited because, as Figure 9 shows, RF size and noise have qualitatively very different effects on the shape of the fitted curve. Thus, while small changes in RF size and noise can be traded off against one another, large changes cannot, due to their very different predictions across the frequency range. To confirm this, we checked that variations in RF size alone could not account for our data. We repeated the fitting with a single noise parameter for each subject (i.e., using the same noise parameter for disparity, motion, and joint). The decrease in fit quality was not worth the saving of two parameters (likelihood-ratio test, $p < 0.05$ for every subject). This confirms that both noise and RF size are necessary variables to explain our data, while the small confidence intervals obtained from resampling confirm that our estimates are reliable.

Effect of eye tracking

In the experiments described above, the participants were allowed to move their eyes freely during the 750 ms presentation of the stimulus. Because the stimuli consisted of moving dots, it is likely that the participants made tracking eye movements to follow one of the directions of motion present in the stimuli. This also applies to previous studies of motion resolution (Anderson and Burr, 1987; Georgeson and Scott-Samuel, 2000). In this

control experiment, we therefore asked whether similar results are obtained under conditions where no tracking eye movements were made. This was achieved by reducing the presentation time to 117 ms, which is shorter than the time needed to initiate tracking eye movements (Robinson, 1965). It was very difficult to make demanding judgments regarding grating structure in stimuli visible for such a short amount of time, especially for the motion and joint gratings in which task-relevant information was not available in any individual frame but had to be acquired over time. Most subjects who participated in the original experiment were unable to do the task at all at these short durations. Author PFA was the only subject who had had enough practice with these stimuli to perform all four tasks at a stimulus duration of 117 ms; subject JH was able to perform when we varied interocular correlation but not motion coherence, as she required 100% motion coherence to be able to perceive the gratings. Figures 15B and 16B show short-duration data for these two subjects. Each session was started with 10 practice trials during which the duration was gradually decreased from 450 to 117 ms and which were not part of any staircase. Apart from this, the methods were exactly the same as in Experiment 2.

We used our model to extract RF and noise parameter estimates from the control experiment data. We assumed that changes in stimulus duration do not alter RF size but may change the effective noise level. We therefore fitted data collected with short and long durations simultaneously, using a total of 13 fit parameters: 3 noise parameters N and 2 correlation parameters κ for each of the long- and short-duration datasets, and 3 RF size parameters σ shared between both durations. This enabled us to check that data collected without tracking eye movements are consistent with the RF size estimated in the main experiment.

Figure 15 shows the coupled fits to the data from the main experiment and the new short-duration control experiment data on motion coherence thresholds; Figure 16 shows the same for interocular correlation thresholds. Notice that very nearly the exact same fits are obtained for the long-duration data, even though 10 of the 13 fit parameters are now fitted to the short-duration data as well. Table 3 shows the fit parameters for the coupled fits; comparison with Table 2 confirms that the RF size and correlation parameters are virtually identical. Thus, our model can quantitatively account for the poorer performance at short durations by assuming that short duration increases the effective noise level. Constraining κ to be the same for short and long durations produced slightly poorer fits, with slightly lower noise estimates for the short-duration data, but did not alter the estimates of RF size or long-duration noise (maximum change in fitted RF size when both σ and κ are shared between durations vs when only σ is shared: 6%). We also fitted the short-duration data on its own, independent of the long-duration results. For subject PFA, this resulted in a larger estimate of motion RF diameter, $2^* \sigma_{\text{motion}}$ (10 vs 6 arcmin) and a smaller N_{motion} , indicating

Table 2. Fit parameters and derived quantities for all subjects

| Subject | AA | AD | AMC | CB | EP | GY | JH | NS | PFA | SA | Mean | SD |
|---|------|------|------|------|------|------|------|------|------|------|-------|------|
| Fitted parameters | | | | | | | | | | | | |
| RF diameters (arcmin) | | | | | | | | | | | | |
| $2\sigma_{\text{motion}}$ | 7.9 | 11.8 | 6.3 | 9.3 | 3.8 | 4.3 | 4.9 | 3.9 | 5.8 | 3.0 | 6.1 | 2.80 |
| $2\sigma_{\text{disp}}$ | 9.2 | 9.9 | 7.5 | | 6.8 | 6.3 | 9.2 | 8.1 | 7.9 | 5.7 | 7.8 | 1.42 |
| $2\sigma_{\text{joint}}$ | 10.1 | 14.2 | 7.1 | 23.2 | 5.5 | 9.1 | 12.0 | 7.1 | 8.5 | 10.0 | 10.7 | 5.07 |
| Effective noise | | | | | | | | | | | | |
| N_{motion} | 0.22 | 0.17 | 0.22 | 0.16 | 0.42 | 0.24 | 0.27 | 0.25 | 0.21 | 0.51 | 0.26 | 0.11 |
| $N_{\text{disparity}}$ | 0.49 | 0.40 | 0.52 | | 0.43 | 0.40 | 0.34 | 0.40 | 0.26 | 0.60 | 0.42 | 0.10 |
| N_{joint} | 0.60 | 0.42 | 0.59 | 0.48 | 0.52 | 0.40 | 0.47 | 0.53 | 0.24 | 0.58 | 0.47 | 0.11 |
| Decorrelation parameter | | | | | | | | | | | | |
| κ_{decoh} | 1.0 | 0.83 | 1.0 | 1.0 | 0.48 | 0.85 | 1.0 | 0.78 | 0.97 | 0.43 | 0.83 | 0.22 |
| κ_{decorr} | 1.0 | 1.0 | 1.0 | | 1.1 | 1.0 | 1.0 | 1.0 | 1.3 | 1.0 | 1.04 | 0.10 |
| Derived quantities | | | | | | | | | | | | |
| Relative RF size | | | | | | | | | | | | |
| $\sigma_{\text{disparity}}/\sigma_{\text{motion}}$ | 1.16 | 0.83 | 1.19 | | 1.79 | 1.46 | 1.88 | 2.06 | 1.36 | 1.94 | 1.46* | |
| $\sigma_{\text{joint}}/\sigma_{\text{motion}}$ | 1.29 | 1.2 | 1.14 | 2.5 | 1.45 | 2.13 | 2.48 | 1.82 | 1.48 | 3.41 | 1.77* | |
| Relative noise | | | | | | | | | | | | |
| $N_{\text{disparity}}/N_{\text{motion}}$ | 2.23 | 2.28 | 2.34 | | 1.04 | 1.63 | 1.26 | 1.60 | 1.20 | 1.16 | 1.59* | |
| $N_{\text{joint}}/N_{\text{motion}}$ | 2.76 | 2.38 | 2.66 | 2.99 | 1.24 | 1.63 | 1.72 | 2.10 | 1.14 | 1.13 | 1.86* | |
| Maximum detectable frequency (cycles per degree) | | | | | | | | | | | | |
| $f_{\text{max}}^{\text{motion}}$ | 4.22 | 3.03 | 5.28 | 3.94 | 6.68 | 7.47 | 6.32 | 8.10 | 5.78 | 7.55 | 5.84 | 1.70 |
| $f_{\text{max}}^{\text{disparity}}$ | 2.49 | 2.65 | 2.95 | | 3.64 | 4.14 | 3.05 | 3.20 | 3.97 | 3.45 | 3.28 | 0.57 |
| $f_{\text{max}}^{\text{joint}}$ | 1.88 | 1.78 | 2.77 | 1.00 | 4.00 | 2.83 | 1.94 | 3.04 | 3.73 | 2.01 | 2.50 | 0.94 |

The asterisks (*) indicate ratios significantly greater than 1 (t test on the log ratios, $p < 0.01$). "Mean" is the arithmetic mean except for the four rows showing ratios, where it is the geometric mean.

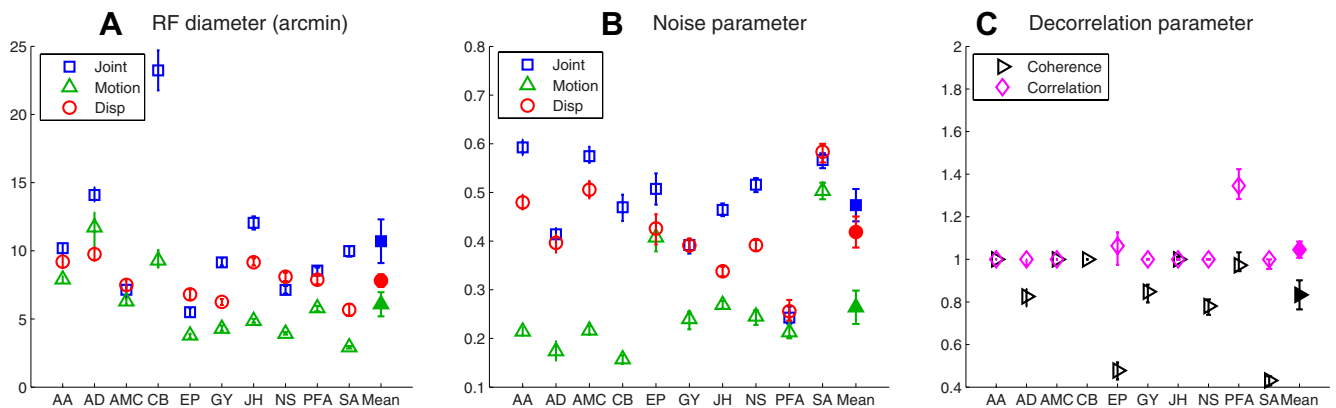


Figure 12. The parameters that gave the best fits to the data. The filled symbols on the right show the averages across subjects. Error bars on individual subjects' results show the 95% confidence intervals obtained by resampling, as described in the text; error bars on the population averages show ± 1 SE of the results from individual subjects.

that the short-duration motion data did not adequately constrain these parameters. However, for both subjects, σ_{disp} and σ_{joint} were similar to the original, long-duration estimates, and once again σ_{joint} was just slightly larger than σ_{disp} [JH, $\sigma_{\text{disp}} = 4.6'$ (long), $5.0'$ (short), $\sigma_{\text{joint}} = 6.0'$ (long), $5.9'$ (short); PFA, $\sigma_{\text{disp}} = 3.9'$ (long), $4.7'$ (short), $\sigma_{\text{joint}} = 4.7'$ (long), $5.6'$ (short)].

Thus, the RF sizes originally obtained from the long-duration dataset in the main experiment can also account for data collected independently at short durations. In addition, essentially the same σ_{disp} and σ_{joint} are obtained from each dataset fitted independently, meaning that our original results are reproduced in a completely independent dataset collected after many months and with a substantial increase in task difficulty. This control experiment rules out the possibility that the small RF estimates for the joint data were an artifact caused by tracking eye movements.

Discussion

Spatial resolution is a key tool for relating visual perception to underlying cortical activity. Our results confirm that resolution

for disparity is limited by the size of V1 receptive fields (Banks et al., 2004; Nienborg et al., 2004; Filippini and Banks, 2009; Allenmark and Read, 2010, 2011), with $2\sigma_{\text{disp}} \sim 8$ arcmin. We also find that resolution for motion is still finer, encoded with an effective RF size of just $2\sigma_{\text{motion}} = 6$ arcmin. Interestingly, our results and those of previous workers (Anderson and Burr, 1987, 1989; Georgeson and Scott-Samuel, 2000) imply that the "motion area" MT is not involved in the detection of motion gratings. MT receptive fields are large, typically around 4° at small eccentricities (Raiguel et al., 1995). Thus, to extract motion gratings with the observed resolution, MT RFs would have to be made up of many small subregions tuned to opposite directions of motion. Motion integration in pattern-selective MT cells may indeed occur at a scale smaller than the entire receptive field (Majaj et al., 2007). However, there is no evidence that MT neurons have subunits tuned to opposite directions and disparities, nor are they selective for motion boundaries (Marcar et al., 1995). Similarly, human brain imaging studies have found no evidence that MT is in-

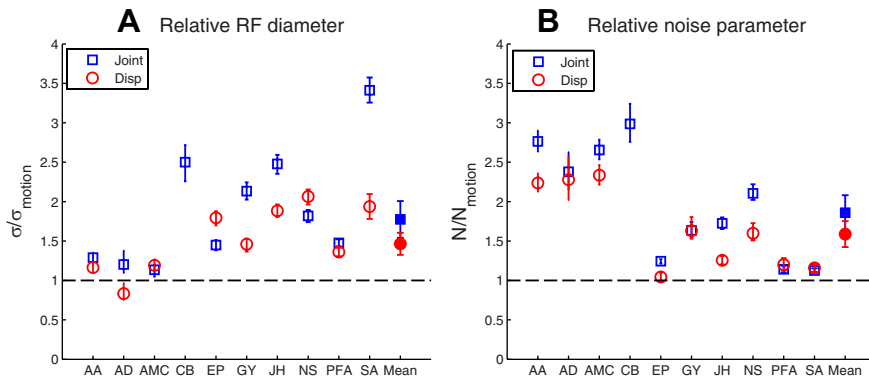


Figure 13. The data from Figure 12, *A* and *B*, normalized to be one for the pure motion data. The filled symbols on the right show the averages across subjects. Error bars on individual subjects' results show the 95% confidence intervals on these ratios, obtained by resampling as described in the text; error bars on the population averages show ± 1 SE of the ratios from individual subjects.

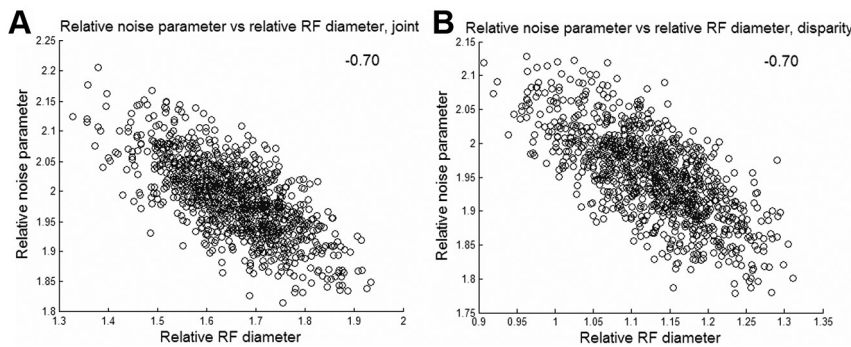


Figure 14. Scatterplot of relative RF size and noise parameters obtained from bootstrap resampling on simulated data based on average model parameters (see text) for joint motion/disparity (*A*) and pure disparity (*B*). Notice that there is a negative correlation between the noise parameter and the RF size parameter. The Pearson correlation coefficients are shown in the top right corners of the scatterplots.

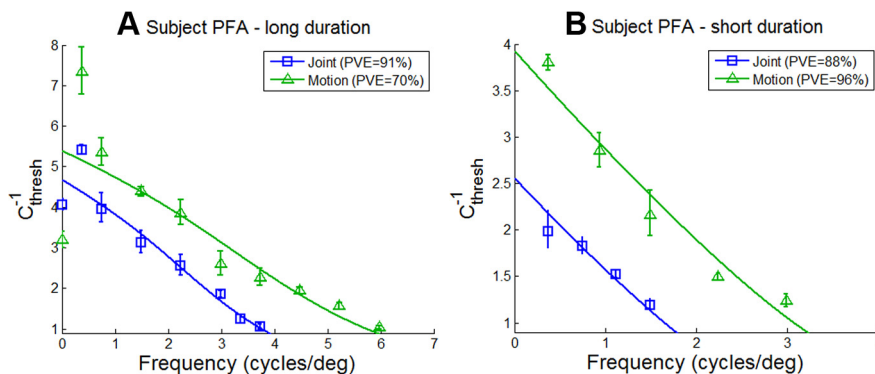


Figure 15. Coupled fits (see text) to long-duration (*A*) and short-duration (*B*) motion coherence data for subject PFA. Notice that, while the actual data in the long-duration plot are exactly the same as in Figure 10, the fits are done slightly differently (see text for explanation).

involved in the perception of motion boundaries (Orban et al., 1995; Reppas et al., 1997), instead identifying a different area (area KO), with no clear counterpart in the monkey visual system (Orban et al., 1995). Thus, we conclude that cortical area MT does not limit the perception of motion gratings.

We therefore introduced a novel disparity/motion conjunction task, designed to require MT. This task requires the observer to extract not only the local motion and disparity in the stimulus but also the conjunctions between them. The literature suggests that MT would be ideally suited for this task. MT contains many

neurons that are sensitive both to motion and disparity. MT neurons are typically suppressed by motion in opposite directions within the same depth plane (Snowden et al., 1991; Qian and Andersen, 1994), as in our noise stimulus. However, they respond well to transparent motion in opposite directions in two different depth planes, as in our grating (Bradley et al., 1995), and some of them are selective for the relative disparity between the two depth planes (Krug and Parker, 2011). Indeed, the transparent motion/disparity random-dot patterns from which we built our joint motion/disparity gratings were originally introduced to study MT neurons (Bradley et al., 1995, 1998; Dodd et al., 2001). Thus, MT neurons should respond more strongly to the signal interval containing the joint motion/disparity than to the noise interval.

If observers use this difference in MT activity to perform the task, we can make a strong prediction about the resulting spatial resolution. The physiological literature suggests that MT neurons respond best when the conjunction between motion and disparity (e.g., left–near/far–right) is the same all over the receptive field (Qian and Andersen, 1994; Bradley et al., 1995). Therefore, if MT limits our joint motion/disparity task, the spatial resolution should be nearly an order of magnitude lower than for pure disparity gratings, in which resolution reflects the much smaller receptive fields found in V1 (Banks et al., 2004; Nienborg et al., 2004; Filippini and Banks, 2009; Allenmark and Read, 2010, 2011). Perhaps the most obvious reason to detect motion/disparity conjunctions is to extract observer self-motion. Here, the sign of the conjunction is the same all over the visual field: objects beyond fixation move in the same direction as the observer, while objects nearer than fixation move in the opposite direction. Thus, very low resolution might well be ecologically sufficient.

Our results comprehensively disprove this prediction. Joint motion/disparity gratings could be detected up to frequencies only slightly lower than pure disparity, at f_{\max} of ~ 2.5 cpd compared with 3.3

cpd. Our analysis suggests that conjunctions between motion and disparity are detected with similar spatial resolution to disparity itself. The effective RF diameter was slightly higher for joint, at 10 arcmin compared with 8 arcmin for pure disparity, but the difference was not significant; the lower frequency limit for joint motion/disparity gratings also reflects a significantly higher noise level. Our results suggest, therefore, that spatial resolution for motion/disparity conjunctions is mainly limited by spatial resolution for each component in isolation. The effective resolution is therefore that of disparity, the lower-resolution component.

Our small estimates of receptive field size suggest that the resolution for motion, disparity, and for conjunctions between the two are all limited by V1 receptive fields. Recently, there has been much debate over whether the ability to detect conjunctions between motion and disparity requires V1 neurons that are specifically tuned to both motion and disparity (Qian and Andersen, 1997; Anzai et al., 2001; Qian and Freeman, 2009), or whether V1 neurons that are tuned solely to motion or solely to disparity can also contribute, if correlations between their activity are read out subsequently (Read and Cumming, 2005a,b; Neri and Levi, 2008). Approximately 15–20% of disparity-selective cells in macaque V1 are also selective for direction of motion (Grunewald and Skoumbourdis, 2004; Read and Cumming, 2005b), and these cells could support performance on the present task. If these cells were solely responsible, it is perhaps slightly surprising that the level of internal noise deduced for the joint task was only 1.12 times higher than for the pure disparity task, given the physiological data implying over four times as many pure disparity cells as jointly tuned cells in early visual cortex. Perhaps performance was supported also by cells selective to motion or disparity alone. Such cells would, individually, be blind to the difference between the joint grating and the noise stimulus, but the presence of the grating could be revealed by correlations in their activity (Read and Cumming, 2005c). The emerging consensus seems to be that both mechanisms contribute (Neri and Levi, 2008), and our results are consistent with that.

V1 is not believed to be a neuronal correlate of perception, implying that higher cortical area(s) read out V1 activity to perceive the joint motion/disparity gratings. Our results show that this readout involves no loss of resolution. The physiological arguments laid out above therefore strongly imply that the readout is not performed in MT. Cortical area MST is also unlikely. MST contains “disparity-dependent direction-selective” neurons that respond to different directions of motion depending on the sign of the disparity of the stimulus (Roy and Wurtz, 1990; Roy et al., 1992). However, MST receptive fields are large, and there is no evidence that they have subregions tuned to motion/disparity conjunctions with opposite signs. The most promising candidate to date may be human cortical area KO, which integrates motion and disparity cues to depth (Ban et al., 2012), as well as detecting motion-defined contours (Orban et al., 1995).

It is perhaps surprising that information in V1 regarding joint motion and disparity can be read out perceptually with no loss of resolution. As noted, if conjunctions between motion and disparity were used primarily to deduce self-motion from motion parallax, a very coarse encoding would suffice. Natural scenes often contain rapid local variations in both motion and depth, but it is not clear why joint motion/disparity encoding would help us perceive such scenes. These scenes could be accurately represented by extracting motion alone and disparity alone, and then overlaying the representations of the two quantities. Our results imply the additional ability to represent different motions and disparities at the same point in space. This more subtle ability

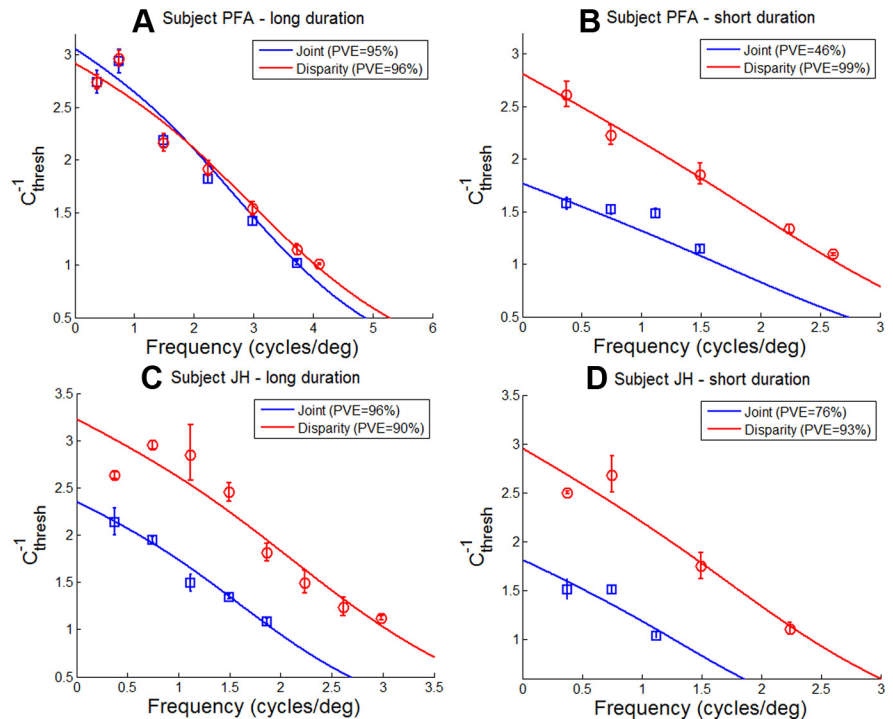


Figure 16. Coupled fits (see text) to long-duration (A, C) and short-duration (B, D) motion coherence data for subject PFA (A, B) and subject JH (C, D). Notice that while the actual data in the long-duration plots are exactly the same as in Figure 11, F and H, the fits are done slightly differently (see text for explanation).

Table 3. Fit parameters from the coupled fits to the data from the main experiment and the short-duration control experiment data for the two subjects who performed the control experiment

| Subject | JH | PFA |
|---|------|------|
| Motion RF diameter (arcmin) | | 5.7 |
| Disparity RF diameter (arcmin) | 9.2 | 7.5 |
| Joint RF diameter (arcmin) | 12.2 | 8.3 |
| N_{motion} (long duration) | | 0.23 |
| $N_{\text{disparity}}$ (long duration) | 0.34 | 0.28 |
| N_{joint} (long duration) | 0.46 | 0.26 |
| κ_{decoh} (long duration) | | 0.93 |
| κ_{decorr} (long duration) | 1 | 1.3 |
| N_{motion} (short duration) | | 0.66 |
| $N_{\text{disparity}}$ (short duration) | 0.51 | 0.58 |
| N_{joint} (short duration) | 0.72 | 0.78 |
| κ_{decoh} (short duration) | | 0.38 |
| κ_{decorr} (short duration) | 0.72 | 0.63 |

benefits scenes with transparency (e.g., a flock of birds in flight), or the branches of a tree moving in the wind, or a shoal of fish under the reflective surface of the water. The remarkable human ability to resolve fine conjunctions between motion and disparity information, revealed in this paper, may reflect the importance of such scenes during our evolution.

References

Allenmark F, Read JC (2010) Detectability of sine- versus square-wave disparity gratings: a challenge for current models of depth perception. *J Vis* 10:17.
 Allenmark F, Read JC (2011) Spatial stereoresolution for depth corrugations may be set in primary visual cortex. *PLoS Comput Biol* 7:e1002142.
 Anderson SJ, Burr DC (1987) Receptive field size of human motion detection units. *Vision Res* 27:621–635.
 Anderson SJ, Burr DC (1989) Receptive field properties of human motion detector units inferred from spatial frequency masking. *Vision Res* 29:1343–1358.

- Anzai A, Ohzawa I, Freeman RD (2001) Joint-encoding of motion and depth by visual cortical neurons: neural basis of the Pulfrich effect. *Nat Neurosci* 4:513–518.
- Ban H, Preston TJ, Meeson A, Welchman AE (2012) The integration of motion and disparity cues to depth in dorsal visual cortex. *Nat Neurosci* 15:636–643.
- Banks MS, Gepshtein S, Landy MS (2004) Why is spatial stereoresolution so low? *J Neurosci* 24:2077–2089.
- Bradley DC, Qian N, Andersen RA (1995) Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* 373:609–611.
- Bradley DC, Chang GC, Andersen RA (1998) Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392:714–717.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- Bredfeldt CE, Cumming BG (2006) A simple account of cyclopean edge responses in macaque V1. *J Neurosci* 26:7581–7596.
- Britten KH, Newsome WT, Shadlen MN, Celebri S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100.
- Cumming BG, Parker AJ (1999) Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *J Neurosci* 19:5602–5618.
- DeAngelis GC, Uka T (2003) Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol* 89:1094–1111.
- Dodd JV, Krug K, Cumming BG, Parker AJ (2001) Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J Neurosci* 21:4809–4821.
- Filippini HR, Banks MS (2009) Limits of stereopsis explained by local cross-correlation. *J Vis* 9:8.1–18.
- Gattass R, Gross CG (1981) Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. *J Neurophysiol* 46:621–638.
- Georgeson MA, Scott-Samuel NE (2000) Spatial resolution and receptive field height of motion sensors in human vision. *Vision Res* 40:745–758.
- Grunewald A, Skoumbourdis EK (2004) The integration of multiple stimulus features by V1 neurons. *J Neurosci* 24:9185–9194.
- Janssen P, Vogels R, Orban GA (1999) Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proc Natl Acad Sci U S A* 96:8217–8222.
- Kleiner M, Brainard DH, Pelli DG (2007) What's new in Psychtoolbox-3? *Perception* 36:ECVP Abstract Supplement.
- Krug K, Parker AJ (2011) Neurons in dorsal visual area V5/MT signal relative disparity. *J Neurosci* 31:17892–17904.
- Majaj NJ, Carandini M, Movshon JA (2007) Motion integration by neurons in macaque MT is local, not global. *J Neurosci* 27:366–370.
- Marcus VL, Xiao DK, Raiguel SE, Maes H, Orban GA (1995) Processing of kinetically defined boundaries in the cortical motion area MT of the macaque monkey. *J Neurophysiol* 74:1258–1270.
- McKee SP, Harrad RA (1993) Fusional suppression in normal and stereo-anomalous observers. *Vision Res* 33:1645–1658.
- McKee SP, Verghese P, Farell B (2004) What is the depth of a sinusoidal grating? *J Vis* 4:524–538.
- Movshon JA, Newsome WT (1996) Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J Neurosci* 16:7733–7741.
- Neri P, Levi DM (2008) Evidence for joint encoding of motion and disparity in human visual perception. *J Neurophysiol* 100:3117–3133.
- Newsome WT, Paré EB (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J Neurosci* 8:2201–2211.
- Nguyenkim JD, DeAngelis GC (2003) Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *J Neurosci* 23:7117–7128.
- Nienborg H, Bridge H, Parker AJ, Cumming BG (2004) Receptive field size in V1 neurons limits acuity for perceiving disparity modulation. *J Neurosci* 24:2065–2076.
- Orban GA, Dupont P, De Bruyn B, Vogels R, Vandenberghe R, Mortelmans L (1995) A motion area in human visual cortex. *Proc Natl Acad Sci U S A* 92:993–997.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Prince SJ, Pointon AD, Cumming BG, Parker AJ (2000) The precision of single neuron responses in cortical area V1 during stereoscopic depth judgments. *J Neurosci* 20:3387–3400.
- Qian N, Andersen RA (1994) Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *J Neurosci* 14:7367–7380.
- Qian N, Andersen RA (1997) A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Res* 37:1683–1698.
- Qian N, Freeman RD (2009) Pulfrich phenomena are coded effectively by a joint motion-disparity process. *J Vis* 9:24.1–16.
- Raiguel S, Van Hulle MM, Xiao DK, Marcar VL, Orban GA (1995) Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. *Eur J Neurosci* 7:2064–2082.
- Read JC, Cumming BG (2005a) The stroboscopic Pulfrich effect is not evidence for the joint encoding of motion and depth. *J Vis* 5:417–434.
- Read JC, Cumming BG (2005b) Effect of interocular delay on disparity-selective V1 neurons: relationship to stereoacuity and the Pulfrich effect. *J Neurophysiol* 94:1541–1553.
- Read JC, Cumming BG (2005c) All Pulfrich-like illusions can be explained without joint encoding of motion and disparity. *J Vis* 5:901–927.
- Reppas JB, Niyogi S, Dale AM, Sereno MI, Tootell RB (1997) Representation of motion boundaries in retinotopic human visual cortical areas. *Nature* 388:175–179.
- Robinson DA (1965) The mechanics of human smooth pursuit eye movement. *J Physiol* 180:569–591.
- Roy JP, Wurtz RH (1990) The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. *Nature* 348:160–162.
- Roy JP, Komatsu H, Wurtz RH (1992) Disparity sensitivity of neurons in monkey extrastriate area MST. *J Neurosci* 12:2478–2492.
- Sakata H, Taira M, Kusunoki M, Murata A, Tsutsui K, Tanaka Y, Shein WN, Miyashita Y (1999) Neural representation of three-dimensional features of manipulation objects with stereopsis. *Exp Brain Res* 128:160–169.
- Salzman CD, Murasugi CM, Britten KH, Newsome WT (1992) Microstimulation in visual area MT: effects on direction discrimination performance. *J Neurosci* 12:2331–2355.
- Serrano-Pedraza I, Read JC (2010) Multiple channels for horizontal, but only one for vertical corrugations? A new look at the stereo anisotropy. *J Vis* 10:10.
- Snowden RJ, Treue S, Erickson RG, Andersen RA (1991) The response of area MT and V1 neurons to transparent motion. *J Neurosci* 11:2768–2785.
- von der Heydt R, Zhou H, Friedman HS (2000) Representation of stereoscopic edges in monkey visual cortex. *Vision Res* 40:1955–1967.
- Watson AB, Pelli DG (1983) QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys* 33:113–120.