PERGAMON

Vision Research 40 (2000) 1839–1855

Vision Research

www.elsevier.com/locate/visres

Comparison of thresholds for high-speed drifting vernier and a matched temporal phase-discrimination task

Ferenc Mechler *, Jonathan D. Victor

Department of Neurology and Neuroscience, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA

Received 14 July 1999; received in revised form 10 December 1999

Abstract

For rapidly translating targets, vernier thresholds correspond to millisecond asynchronies between targets. The 'temporal hypothesis' is that these thresholds reflect the limiting sensitivity of asynchrony detectors. Previous studies showed that temporal thresholds are generally higher than vernier thresholds, but failed to reject the 'temporal hypothesis' because stimuli had differing spatiotemporal characteristics, and temporal thresholds depend strongly on stimulus and task. Here we use matched grating stimuli to test — and reject — the temporal hypothesis. Expressed as asynchrony, temporal phase discrimination was typically 10-fold poorer than vernier thresholds, and differed in dependence on spatial frequency, temporal frequency, contrast, and susceptibility to stroboscopic masks.

Keywords: Spatio-temporal vernier; High-speed asynchrony limit; Temporal phase-discrimination mechanism; Drifting and contrast-reversed sinusoidal grating; Stroboscopic flicker masking

1. Introduction

For high contrast drifting targets, vernier performance falls into two quite distinct regimes, depending on target speed. For low-to-intermediate translation velocities (up to 1–3 deg/s), offset thresholds are similar to thresholds for stationary targets, are within the hyperacuity range, and exhibit immunity to motion blur (Westheimer & McKee, 1975; Morgan, Watt, & McKee, 1983). At velocities above 3 deg/s, vernier acuity degrades in an approximately linear fashion with increasing target velocity, and in excess of the degradation of target visibility (Carney, Silverstein, & Klein, 1995; Chung, Levi, & Bedell, 1996). The fact that vernier thresholds and velocities are approximately proportional is consistent with an apparent temporal limit of asynchrony detectors at threshold performance. This temporal limit on hyperacuity for high-velocity targets was found to be approximately 1 ms in several studies (Burr, 1979; Fahle & Poggio, 1981; Morgan et al., 1983; Carney et al., 1995) in spite of the considerable spatial diversity of the stimuli. While it seems the visual system consistently achieves this high sensitivity for quite diverse stimuli under motion it is not well understood how it does it.

It is generally assumed that the mechanisms that limit vernier performance for moving targets are fundamentally spatial or spatiotemporal. Psychophysical studies such as those of Carney et al. (1995) implicate narrowly tuned oriented filters, similar to those that are thought to underlie stationary vernier acuity at small target separation. In this view, the 1 ms limit is either merely an epiphenomenon, or rather an emerging phenomenon, of threshold performance of the inherently spatiotemporal mechanisms involved in the task.

On the other hand, several lines of evidence indicate that temporal factors contribute to vernier performance for targets with small spatial separations (for a recent review see Victor & Conte, 1999). Moreover, the observation of the '1 ms' limit for high-speed moving vernier and other spatiotemporal hyperacuity tasks gives rise to the temporal hypothesis, according to which 1 ms is the sensitivity limit of temporal (i.e. localized, spatially non-integrating) discrimination mechanisms. This view has been less popular but not fully discredited.

* Corresponding author. Tel.: +1-212-746-6520; fax: +1-212-746-8984.

E-mail address: fmechler@med.cornell.edu (F. Mechler)

0042-6989/00/$ - see front matter
PII: S0042-6989(00)00035-3

An important recent discovery in neurophysiology adds interest to re-examining the temporal hypothesis: central neurons have an intrinsically high — approximately 1 ms — precision both in response to electrical (Mainen & Sejnowski, 1995; Nowak, Sanchez-Vives, & McCormick, 1997) and visual (Bair & Koch, 1996; Berry, Warland, & Meister, 1997; Reich, Victor, Knight, Ozaki, & Kaplan, 1997) stimulation. The correspondence between the psychophysical temporal limit for vernier acuity and the physiological precision limit of spike encoding raises the intriguing possibility that vernier thresholds for high-velocity moving targets are based on the detection of asynchrony in the central spike representation of stimuli, and reflect a fundamental limit on the precision with which visual neurons can encode this temporal information. This correspondence, however, does not imply the existence of asynchrony detectors. Detectors tuned to asynchronies on the order of 1 ms in the stimulus need not exist even if the internal neural representation and processing of visual information has the ability to handle small asynchronies and recognize fine temporal coincidences limited to 1 ms.

Earlier psychophysical studies found that visual temporal asynchrony sensitivity varies over a wide range (3–30 ms) and strongly depends on the task, especially on the spatial configuration of targets (for a brief review see Westheimer & McKee, 1977). The temporal hypothesis for moving vernier, too, was addressed by Westheimer and McKee (1977) when they measured sensitivity to asynchrony between briefly presented stationary lines. The best asynchrony sensitivity they could measure in two observers under optimal conditions (including separations in the range of 2–6 arc min) was no better then 3–5 ms. Based on that and other results it was suggested that the temporal sensitivity of the visual system may be insufficient to account for the millisecond sensitivity observed in certain spatiotemporal hyperacuity tasks. However, a study by Morgan and Watt (1982) of a dichoptic spatiotemporal hyperacuity task revealed a 1 ms asynchrony threshold. Of note, this high temporal acuity required that the stimuli moved across the retina. When movement was blocked (by allowing tracking), these low thresholds could not be obtained. Thus, because the apparent limits to temporal discrimination depend strongly on the task and stimulus characteristics, the argument against the temporal hypothesis, although widely accepted, suffers from the lack of a definitive test.

Since it is hard or impossible to compare sensitivity measures obtained on different subjects using stimuli of different retinal extent, spatial configuration and temporal properties, a definitive test should be based on a controlled comparison of performance of the same observer in a spatiotemporal hyperacuity task and a related temporal task using stimuli of matched spatiotemporal characteristics. The goal of this study was to test the temporal hypothesis via such a controlled matched comparison. If the temporal hypothesis holds for moving vernier, then the 1 ms performance limit should be manifest in modifications of the moving vernier task in which spatial cues have been removed but the temporal cues remain. The latter are expected to activate the critical temporal discrimination mechanisms. With this in mind, we designed a psychophysical task that isolates the temporal component of a specific moving vernier stimulus and eliminates the spatial cues as much as possible. Since detection of misalignment between drifting gratings of the same spatial and temporal frequencies that share a common boundary could be accomplished by a spatial cue (a spatial misalignment or spatial phase offset between gratings at any given time), a temporal cue (asynchrony or a temporal phase offset between the temporal modulation waveforms of the gratings at a given location), or a combination, we will refer to this as the *spatiotemporal task*. To isolate the temporal cue, we introduce a second task, which by design consists of judging only asynchrony: determining temporal phase offset between two spatially aligned gratings that are contrast reversing but slightly out of phase. If the temporal frequencies are identical in the two tasks, the temporal phase offset across the boundary provides a temporal cue equivalent for the two tasks at any location along the boundary between the abutting gratings (or between any two local measuring sites across the boundary). However, in the second task, which grating precedes the other (which has advanced temporal phase) cannot be determined from a snapshot of the stimulus. That is, replacing the drifting gratings of a moving vernier task by contrast-reversed gratings retains the temporal, but not the spatial, cues. We therefore call the second task the *temporal* task. Note that in the temporal task, no single frame suffices to determine which grating leads the other, and the discrimination cannot be made if the individual frames are randomly shuffled in time. Thus, temporal cues are both necessary and sufficient for solving this task.

We compare performance in the spatiotemporal and the temporal tasks. If the limiting mechanisms are similar (i.e. a temporal mechanism limits performance on the spatiotemporal task), then we would expect that performance in the two tasks would be similar. However, we find that offset thresholds measured in the spatiotemporal and temporal tasks, and their dependence on spatial frequency, temporal frequency, and contrast, are quite distinct. We also use a flicker-masking paradigm to provide experimental support that we have indeed isolated a temporal mechanism in the temporal task. These results imply that different mechanisms underlie psychophysical performance in the two tasks and give strong support to the rejection of the

temporal hypothesis. Our approach also allows us to elucidate the possible neuronal mechanisms subserving the tasks.

2. Methods

2.1. Subjects

The first author and three naive observers participated in the experiments. Observers were 25–37 years of age with normal or corrected-to-normal visual acuity.

2.2. Stimuli

Stimuli consisted of horizontally oriented sinusoidal luminance gratings presented in a vertically bisected field on a Tektronix 608 monochrome CRT. At the viewing distance of 200 cm, the 256×256 pixel display subtended $2.7 \times 2.7°$ (1 pixel = 38 s arc), and was viewed through a circular aperture of 2.3° diameter. The mean luminance was 150 cd/m^2; the frame refresh, 270.33 Hz (frame duration 3.7 ms). Across conditions, spatial and temporal frequencies varied in the range of 0.4–11.8 c/deg, and 1–32 Hz, respectively. Except as noted, Michelson contrast was fixed at 0.9 or 1.0. For misalignment and asynchrony threshold measurements, the contrast, and spatial and temporal frequency of the gratings in the two halves of the display were identical. The two gratings differed only in their phase. In experiment 2, threshold contrasts at detection were also measured, but the grating was presented in only one half of the display randomized across trials, with the other half held at mean luminance.

A special purpose stimulus generator (Milkman, Shapley, & Schick, 1978; Milkman, Schick, Rossetto, Ratliff, Shapley, & Victor, 1980) under the control of a PDP-11/93 computer allowed accurate generation of the sinusoidal waveform with arbitrary phases at any frequency by computing the spatial profile of the waveform for each frame. This allowed settings for the spatial or temporal phase offsets between the two gratings to be finer than the pixel size or the frame time of the display.

Dark-adapted subjects were instructed to fixate the center of display. Viewing was binocular. To minimize tracking, a brief presentation time (ranging from 240 to 270 ms) and a fixation aid (a set of four dots arranged in a 0.33° square centered on the field) were used. A cosine envelope was used in a 30 ms on- and off-contrast ramp to reduce transients at the beginning and end of the presentation. The initial absolute phase and the direction of movement were randomized in each trial to rule out the use of these extraneous cues and anticipatory eye movements.

2.3. General procedure

Data collection sessions ($\sim$45 min each) began after 1–2 introductory/practice sessions on a preceding day. Trials were self-paced and unaided by feedback. Thresholds were measured with a two-alternative forced choice (2-AFC) staircase method. The staircase algorithm, using a prefixed scalar factor of 1.3, decreased the step size of the varied parameter (phase or contrast) after two consecutive correct judgments, and increased it after each incorrect judgment. Staircases typically converged and ran through six reversals in 30–60 trials. In each staircase run, six reversals were collected and the geometric average of four reversals (without the two outliers) was used to define a threshold estimate (the 71% correct point). Estimates from five to nine (typically six) staircases were averaged for each data point.

2.4. Spatiotemporal task

In the *spatiotemporal* task (Fig. 1A, left), vertical misalignment thresholds for two abutting drifting gratings (of the same spatial and temporal frequency) were measured. The luminance variation $L(x,t)$ along the vertical positions x at time t in a drifting grating of mean luminance L_0, spatial frequency f_s, temporal frequency f_t, phase ϕ, and contrast C, was $L(x,t) = L_0 (1 + C \sin(2\pi (tf_t - xf_s) - \phi))$. The subject was asked to indicate which half of the display contained the grating that appeared to be above the other. As with all translating vernier stimuli, the task is solvable both spatially (i.e. within one frame, across space) and temporally (i.e. locally, over time), as shown in Fig. 1B, top. This task was readily performed by naïve subjects with typically no practice.

Fig. 1 *(overleaf)*. Illustration of the stimuli. (A) Four frames of a typical stimulus used for the spatiotemporal task (left), and temporal task (right). The frames are separated in time (for illustrative purposes only) by 1/8 of a grating cycle. (B) The spatial cue (left) and temporal cue (right) for the two tasks. A pair of images illustrates the 2-D spatiotemporal light distribution in the stimulus in the left and right half of the display for corresponding stimuli in either the spatiotemporal (top), or temporal task (bottom). Temporal cues are illustrated by a comparison of the time-varying signals along a slice through the spatiotemporal light distribution at an arbitrary vertical position. Similarly, spatial cues are illustrated by a comparison of the spatial variation along a slice through the spatiotemporal light distribution at an arbitrary moment in time. Spatial and temporal cues are meaningful in the spatiotemporal task. The corresponding stimulus in the temporal task preserves only the temporal cue. The spatial cue, however, is ambiguous: the relative contrast and polarities of the gratings oscillate over time.

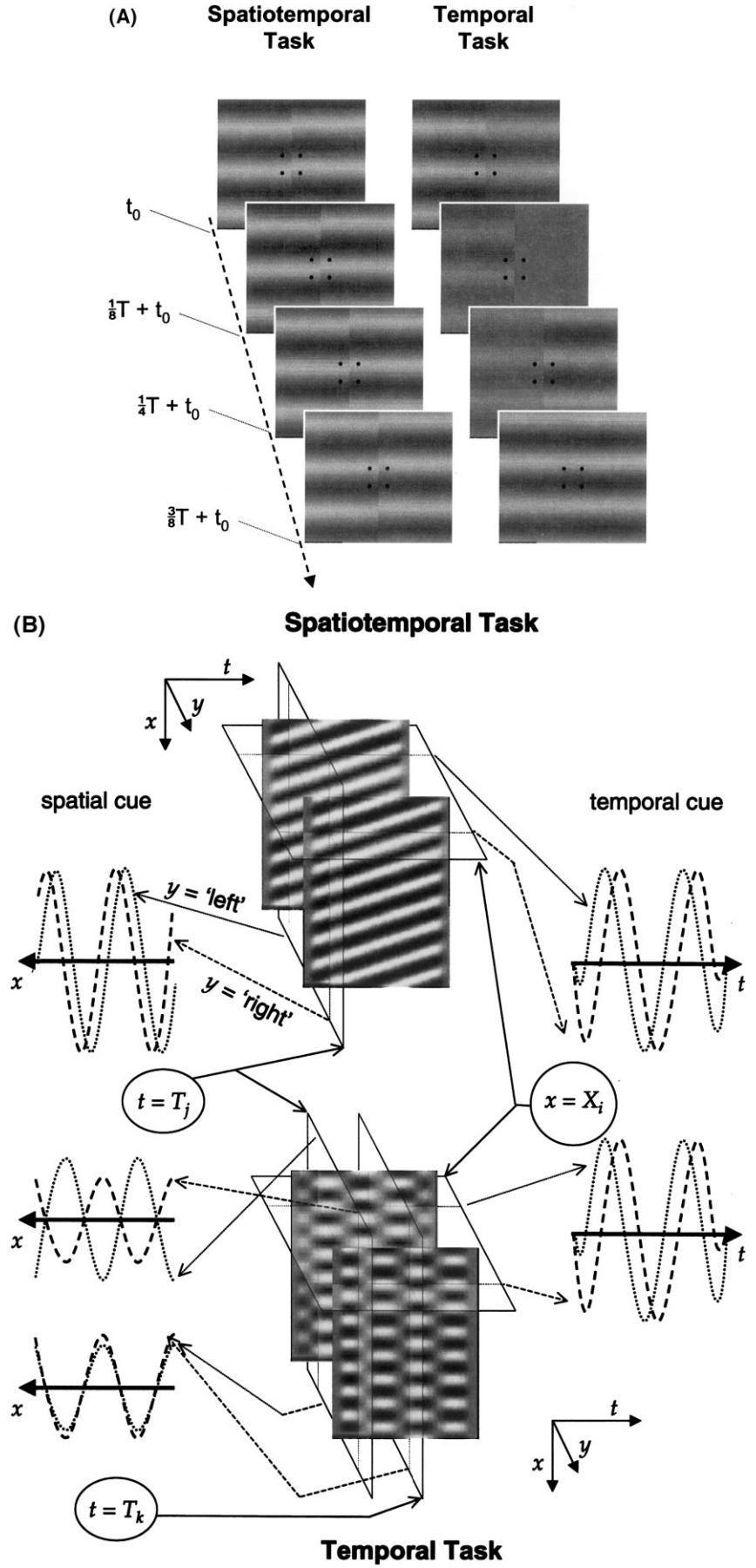
(A)
Spatiotemporal Task
Temporal Task
t_0
$\frac{1}{8}T + t_0$
$\frac{1}{4}T + t_0$
$\frac{3}{8}T + t_0$
(B)
Spatiotemporal Task
t
x
y
spatial cue
temporal cue
y = 'left'
y = 'right'
x
t
$t = T_j$
$x = X_i$
x
t
x
t
y
$t = T_k$
Temporal Task

Fig. 1.

2.5. Temporal task

In the *temporal* task, (Fig. 1A, right), the same (but stationary) gratings were spatially aligned, and sinusoidally modulated in contrast-reversing mode slightly out of phase. The luminance variation $L(x,t)$ along the vertical positions x at time t in a contrast-reversed grating of mean luminance L_0, spatial frequency f_s, temporal frequency f_t, spatial phase ϕ_s, temporal phase ϕ_t, and contrast C, was $L(x,t) = L_0\,(1 + C\,\sin(2\pi\,tf_t - \phi_t)\,\sin(2\pi\,xf_s - \phi_s))$. The subject was asked to indicate which half of display contained the grating with the leading temporal phase. This task is only possible to solve in time, i.e. across two or more consecutive frames. While the temporal cue in this task (Fig. 1B, bottom) is preserved from the corresponding stimulus in the spatiotemporal task, the spatial offset cue has been removed.

For every subject we began by asking 'which grating is leading'. All subjects found it difficult to do the temporal task that way and some soon realized the existence of the apparent motion cue. Therefore, we subsequently suggested to subjects that they make the temporal phase discrimination by judging the direction of apparent motion across the field division (The apparent motion cue is inevitable in a temporal comparison task in which space is used to tag the stimulus. That is, our temporal task can be formally considered to be a different kind of spatiotemporal task. This inevitable intrusion of spatial cues when two stimulus components are presented simultaneously could only result in lower thresholds than an ideal temporal task, and thus, could not lead to an underestimate of asynchrony sensitivity). The apparent motion percept cycles between the two horizontal alternatives (but with a bias towards one of them), and observers usually learned readily which percept corresponded to the leading grating. Subjects required practice sessions for performance to stabilize, and were given a few 'refresher' trials before each data collection session.

Phase offsets are modulo one half of a grating cycle and this is the upper bound to phase offset thresholds (in space or time). At a quarter cycle offset, the phase cue is strongest, and at a half cycle offset, the task becomes ambiguous. For this reason, the staircase method is likely to diverge for thresholds that are near half a cycle, and to be unstable if no phase offset suffices for better-than-chance performance.

3. Results

Performances in the spatiotemporal and temporal tasks were compared in three types of experiments: Experiment 1 examined the dependence of misalignment and asynchrony thresholds on spatial and temporal frequency. Experiment 2 examined the dependence on contrast. Experiment 3 examined the influence of a stroboscopic mask as function of mask frequency.

3.1. Experiment 1: spatial and temporal frequency dependence

We first measured misalignment thresholds in the spatiotemporal task as a function of translation velocity of drifting sine gratings. Velocity for a grating of fixed spatial frequency and high contrast (0.9–1.0) was varied by varying the temporal frequency from 1 to 32 Hz (a 1.5 log unit range). The spatial frequency chosen for each subject was near the subject's optimum for detection ($\sim$3 c/deg, Bradley & Skottun, 1987). For two subjects, threshold measurements were also made at spatial frequencies below or above that optimum in log steps, including one near optimum for static grating vernier ($\sim$12 c/deg, Bradley & Skottun, 1987). For each spatial frequency, offset thresholds were also measured for static targets. Representative data from three subjects are plotted in Fig. 2.

For static targets, offset thresholds were in the typical hyperacuity range of 4–20 s arc (isolated symbols plotted on the ordinate at 0 deg/s grating velocity). For both subjects in whom measurements were made at several spatial frequencies, these thresholds decreased with increasing spatial frequency over the examined range (up to $\sim$12 c/deg), in agreement with the results of Bradley and Skottun (1987). With slow-moving gratings, acuity remains relatively constant up to 1–3 deg/s. At higher velocities, a second regime begins, in which the response functions have a slope of unity, i.e. lie on a locus described by a constant temporal limit. These regimes are consistent with studies (Westheimer & McKee, 1975; Morgan et al., 1983; Carney et al., 1995) that used spatial targets other than gratings (i.e. dots or lines). Within observers, the temporal limit defined by the slope of the curves of Fig. 2 in the high-speed range varied with spatial frequency from 0.5 to 8 ms, even in the restricted range of stimulus parameters we tested. This variation of limits suggests that more than one temporal mechanism is necessary to account for the data of Fig. 2.

Across the explored range of drift velocity, there is a consistent pattern of ordering of the curves for different spatial frequencies. At low velocities (constant threshold regime), the ordering of the curves is inversely related to the spatial frequency, i.e. offset thresholds are lowest for highest spatial frequencies, and highest for the lowest spatial frequency, for the examined range of spatial frequencies. The ordering, and to some extent the spacing, can be explained by the systematic difference of target visibility as determined by the spatial contrast sensitivity to static targets, multiplied by the

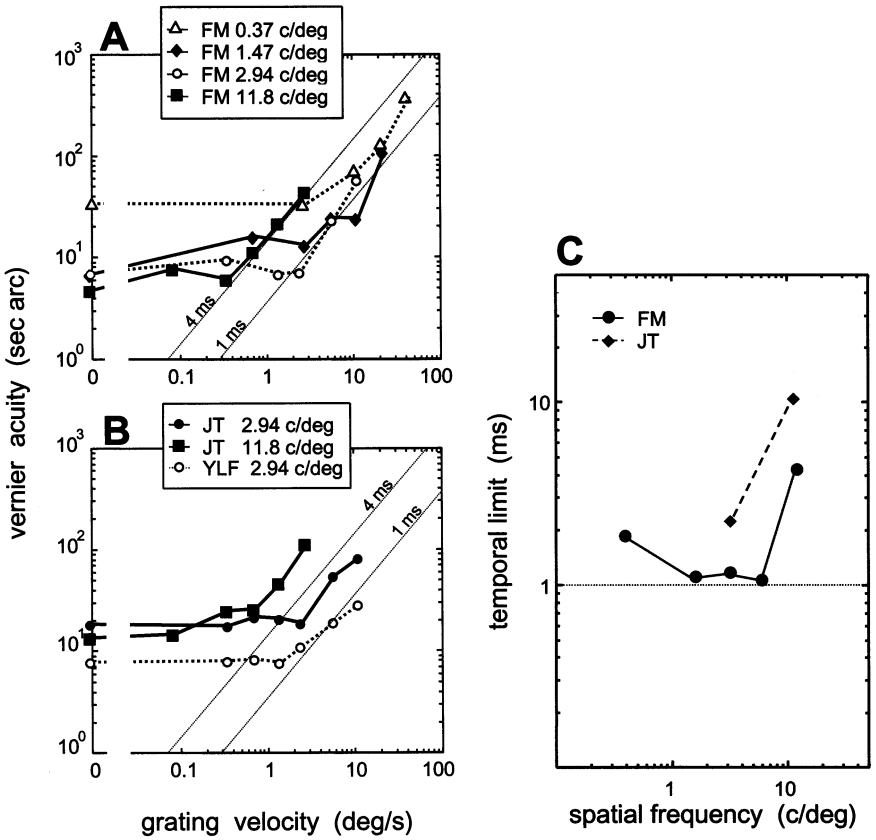

Fig. 2. Displacement thresholds for drifting gratings measured in the spatiotemporal task. For each spatial frequency, there are two distinct regimes. At very high or very low speeds, target visibility could fall below threshold and thus prohibit the task. For a given spatial frequency and observer, a constant asynchrony limits thresholds at velocities above 1–3 deg/s, as shown by the 1 and 4 ms slanting dotted lines. (A) Data from subject FM, four spatial frequencies, contrast 1.0. (B) Data from subjects JT, two spatial frequencies, and YLF, one spatial frequency; contrast 0.90. Error bars were omitted for clarity. For a typical data point in this plot, 1 SEM is comparable to the symbol size. (C) Dependence of the temporal limit on grating spatial frequency. The temporal limit at each spatial frequency was defined as the constant asynchrony best fitting the thresholds measured at the three or four highest speed points (i.e. the high-speed asymptotes on data sets such as shown in A and B). For both subjects, the temporal limit is lower at 3 c/deg than at 12 c/deg grating spatial frequency. The horizontal dotted line indicates 1 ms asynchrony that, for subject FM, forms a lower bound to asynchrony sensitivity. The latter approaches 1 ms only in the intermediate spatial frequencies.

spatial frequency (Bradley & Skottun, 1987), a quantity that is proportional to the contrast of the 'local sign' *across* the axis of misalignment. At high velocities (regime of constant temporal limit), the ordering of the curves is the reverse. Here the determining factor is not target visibility but a shift in the spatial scale of analysis from high to low spatial frequencies (Chung et al., 1996; Levi, 1996). The offset threshold curves thus cross over at intermediate velocities.

A dependence of temporal limit on grating spatial frequency is shown in Fig. 2C for the subjects in whom the temporal limit was assessed at least for two different spatial frequencies (i.e. high-speed asymptote extracted from two or more threshold functions shown in Fig. 2A,B). For both subjects, the temporal limit is lower at 3 c/deg than at 12 c/deg grating spatial frequency. The data from subject FM indicate an intermediate range of spatial frequencies (1–6 c/deg) in which the apparent temporal sensitivity is rather constant near the minimum millisecond.

To compare offset thresholds in the spatiotemporal and temporal tasks, we transformed the independent variable of our analysis from velocity to temporal frequency. This allows us to express thresholds in a comparable fashion in equivalent asynchrony. Each panel in Fig. 3 (closed symbols) shows data from the spatiotemporal task transformed in this way from Fig. 2 (Thresholds for stationary targets, the 0 Hz data points, cannot be transformed in this fashion). Data were obtained at the fixed spatial frequency indicated at the top. At low temporal frequencies, threshold asynchrony varied inversely with temporal frequency, consistent with a constant spatial phase threshold (0.03–0.3 rad, depending on the spatial parameters and the subject's sensitivity). This is indicated in Fig. 3 by the parallel course of the data curves and the dotted lines. At intermediate frequencies (4–8 Hz), each threshold function turns into an approximately constant asynchrony plateau that holds for high temporal frequencies. The level of this plateau is exactly the slope of the corre-

sponding curves seen at high speeds in Fig. 2. In two subjects, under optimal conditions (e.g. FM at 3 c/deg and 8 Hz, Fig. 3A; YLF at 3 c/deg and 32 Hz, Fig. 2B), asynchrony at threshold in the spatiotemporal task fell below 1 ms, indicating even better temporal sensitivity than the '1 ms' limit reported earlier with different stimuli by Carney et al. (1995), but comparable to the equivalent asynchrony limit that can be calculated from other published data (Fahle & Poggio, 1981; Morgan et al., 1983).

Each panel in Fig. 3 also shows data from the temporal task (open symbols), obtained with contrast-reversed gratings of matched spatial and temporal parameters. These functions have similar shapes to the corresponding ones from the spatiotemporal task but lie consistently above them. In the constant temporal phase regime, asynchrony thresholds in the temporal task are two to four times greater than the asynchrony thresholds in the spatiotemporal task. The transition between the two regimes in the temporal tasks often occurs at different temporal frequencies from those in the corresponding spatiotemporal task (e.g. lower in Fig. 3A, C; or higher in Fig. 3B). The asynchrony thresholds in the temporal task never fell below 5 ms, but more typically remained in the vicinity of 10–15 ms. In some cases, at high temporal frequencies, temporal phase offsets increased above the maximum measurable threshold (missing symbols at 32 Hz in Fig. 3B,D above the slanting dotted line marked π) in the temporal task, but not in the spatiotemporal task.

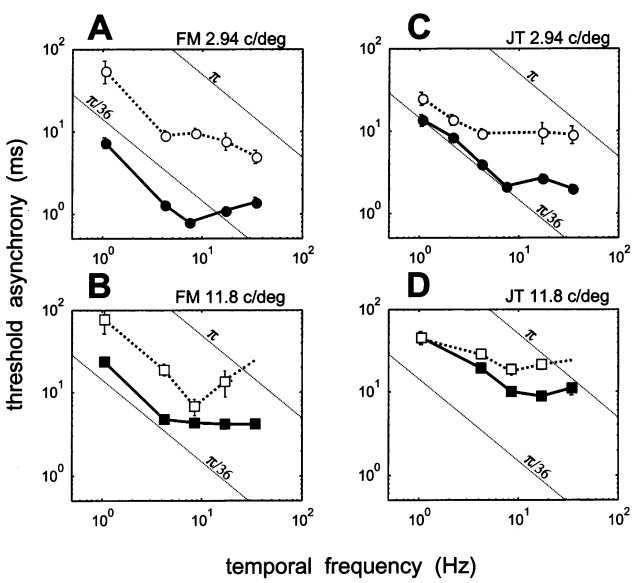

Fig. 3. Comparison of displacement thresholds measured in the spatiotemporal task (closed symbols), and temporal phase offset thresholds measured in the temporal task (open symbols), both plotted as asynchrony at threshold (ms) vs temporal frequency of the grating (Hz). Data for the spatiotemporal task were re-plotted from Fig. 2. Michelson contrast of the gratings was as in Fig. 2. (A, B) Subject FM at 2.94 and 11.8 c/deg, respectively. (C, D) Data from subject JT arranged similarly. The slanting dotted lines indicate the loci of constant temporal phase offsets, π and $\pi/36$ radians, respectively; π is the upper limit of measurable offsets. Error bars here and in subsequent figures are ± 1 SEM.

The consistently greater thresholds measured in each subject for the temporal task than for the spatiotemporal task correspond to the subjective reports of greater difficulty experienced in that task and the larger error bars; both consistent across subjects and stimulus conditions. Subjects who could perform the temporal task required extensive practice for acquiring a reliable cue, but typically did not require any practice for the spatiotemporal task. Moreover, two subjects who both had corrected-to-normal vision and who had no difficulty performing the spatiotemporal task, could not perform the temporal task. One of these subjects, YLF, an experienced psychophysical observer otherwise naïve to the purpose of these experiments, had normal vernier acuity with an apparent temporal acuity limit of 0.5 ms in the spatiotemporal task (Fig. 2B) but we could find no grating parameters at which she could perform the temporal task, despite hours of practice with feedback, prolonged stimulus presentation, and adjusted staircase parameters. Another observer, JT, could perform the temporal task under the conditions of experiment 1 (and his data are plotted in Fig. 3), but he was unable to do the task with a temporal mask (experiment 3 below).

In summary, the observed dependence of alignment sensitivity in the spatiotemporal task on spatial and temporal frequency, as measured here with translating sine gratings, is consistent with the known properties of the detection mechanisms involved in moving vernier. However, we failed to find a single asynchrony limit for thresholds in the high velocity range. The distinctly larger offset thresholds and the greater difficulty of the temporal task, consistently observed across subjects and over a range of grating parameters, together support the notion that this detection mechanism differs from the one employed in moving vernier.

3.2. Experiment 2: contrast dependence

A difference in visibility (i.e. detection threshold contrasts) for drifting and contrast-reversed gratings of otherwise identical spatiotemporal parameters, could account for at least part of the difference between the asynchrony thresholds in the spatiotemporal and temporal tasks shown above, since these experiments were performed with stimuli that were equated for peak (Michelson) contrast, but not for the space-time averaged (RMS) contrast or for visibility. The space-time averaged (RMS) contrast of a drifting grating is $\sqrt{2}$ times that of a contrast-reversed grating when they are of equal Michelson contrast, since a drifting grating is a sum of two quadrature-phase reversing gratings. For the same reason, unequal visibility of gratings of the

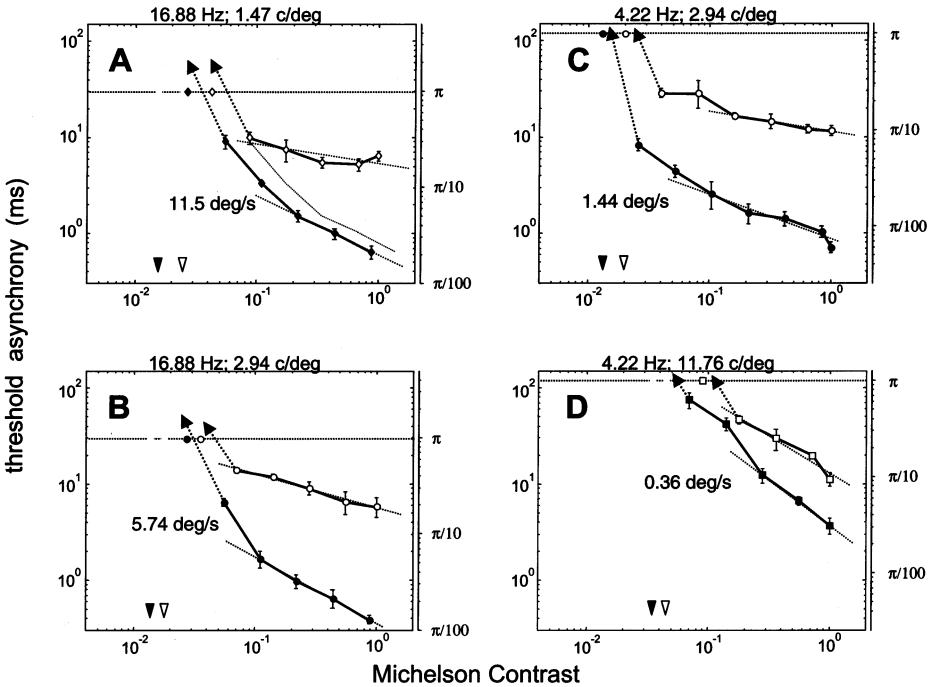

Fig. 4. Dependence of offset thresholds on contrast in subject FM for four different gratings. Thresholds from the spatiotemporal (closed symbols) and temporal tasks (open symbols), expressed in units of asynchrony (ms), are plotted against Michelson contrast. Right vertical axes indicate threshold in units of phase offsets (radians). The horizontal dotted line (at π radians) indicates the half-cycle offset. The broken line continuation of threshold curves with arrows indicate undetermined levels of phase offset thresholds larger than π. Contrast at detection threshold, and exponent of power function fitted to the data, are as follows: (A) detection thresholds: 0.016 ± 0.002 (▼), 0.025 ± 0.005 (▽); fitted exponents: -0.63 (◆), -0.32 (◇). (B) 0.014 ± 0.002 (▼), 0.018 ± 0.009 (▽); -0.69 (●), -0.38 (○); (C) 0.013 ± 0.003 (▼), 0.021 ± 0.004 (▽); -0.41 (●), -0.2 (○); (D) 0.035 ± 0.012 (▼), 0.045 ± 0.014 (▽); -0.95(■), -0.78 (□).

same type but different parameters could account for some of the threshold variation within each task. Therefore, we studied the contrast dependence of offset thresholds in both tasks. One subject (FM) was tested with an extensive set of stimuli in these experiments. Results from a second subject using fewer spatiotemporal parameters for the gratings were in agreement with FM's.

Data from subject FM are shown in Fig. 4. Offset thresholds measured for drifting gratings as a function of the Michelson contrast of the gratings in the spatiotemporal task, and converted as in Fig. 3 in to units of asynchrony (closed symbols), are compared with asynchrony thresholds measured for contrast-reversed gratings in the temporal task (open symbols) over the same range of contrasts. For each of the four sets of comparison (Fig. 4A–D), gratings used in both tasks were of the same fixed spatiotemporal parameters as indicated. The chosen frequencies defined a large range of velocities (0.4–12 deg/s) for the drifting gratings in the spatiotemporal task and ensured that the offset thresholds measured for them at high contrast, as shown in Fig. 2, would fall in or near the region with the '1 ms' limit.

Fig. 4A shows data obtained with gratings of 16.88 Hz and 1.47 c/deg, values for which the drifting grating defined a high velocity (11.5 deg/s). The offset threshold obtained with this grating at high contrast fell significantly below the 1 ms limit (Fig. 2A). Arrowheads indicate on the abscissa the contrast at detection threshold for these gratings (Fig. 4, filled arrowhead, for the drifting grating, open arrowhead for the contrast-reversed grating). Offset thresholds could not be determined for either task at contrasts near detection threshold. At 3.5 times detection threshold, offset thresholds in the two tasks are similar; in terms of asynchrony (approximately 10 ms). As visibility of the gratings improve, offset thresholds in the spatiotemporal task rapidly decrease by an order of magnitude. Above about 0.10–0.20 contrast, thresholds are well approximated by a power function with an exponent of -0.63 (slope of the dotted regression line fitted to the last three data points). In the temporal task, over the same range of increasing visibility, improvement of sensitivity to temporal phase offsets is only half as fast as that in the spatiotemporal task (exponent -0.32 of power function fitted to the last five points). Therefore, the threshold ratio between the two tasks increases 10-fold over the full range of equal visibility. This comparison is made explicit with a copy (Fig. 4A, dotted curve) of the contrast response curve of the spatiotemporal task shifted to match the contrast re-

sponse curve of the temporal task at the first point of measurable threshold.

The same dependence of thresholds on contrast is present in Fig. 4B–D. Contrast ratio at detection threshold was tightly distributed around the average of 1.64 (range 1.2–2, $n=18$) and was not significantly different ($P>0.05$) from $\sqrt{2}$, the ratio of Michelson contrasts (temporal vs spatiotemporal) for which the RMS contrasts are identical. This indicates that the two types of grating have similar thresholds when expressed in terms of space–time average rather than peak contrast, and that the former is probably the critical quantity for detection.

Across conditions within each task, equal visibility produced similar phase offset thresholds (Fig. 4A–D, left ordinate). In all conditions tested, there was a shallower dependence on contrast of thresholds in the temporal task (characteristic exponent -0.2 to -0.3) than the spatiotemporal task (characteristic exponent -0.6). Also, power law behavior set in at lower visibility in the temporal task (at approximately four times detection vs eight to ten times detection for the spatiotemporal task). These differences were especially prominent for spatial and temporal frequencies that correspond to grating velocities above 3 deg/s in the moving vernier stimulus, i.e. the range with an apparent temporal limit (e.g. Fig. 4A–C), and least prominent for highest spatial frequency (and for the spatiotemporal task, lowest velocity) of the gratings (Fig. 4D). These results indicate not only that the difference in visibility is insufficient to account for the difference in asynchrony thresholds observed in the two tasks, but also (especially for high temporal frequencies and low spatial frequencies) that the two tasks must rely on mechanisms with different contrast dependence.

3.3. Experiment 3: influence of multiplicative flicker mask

To ascertain that we isolated a temporal mechanism in the 'temporal' task, and to further distinguish the mechanisms involved in the two tasks, we asked two subjects to perform both tasks in the presence of a multiplicative (stroboscopic) flicker mask. We reasoned that substantial degradation of performance is expected in the presence of a stroboscopic interference signal only if performance relies on a temporal mechanism. One may think of the stroboscopic mask, especially at low temporal frequencies of the mask, as a condition of extreme sampling. Infrequent sampling is expected to introduce more ambiguity in the temporal or motion cue that relies on two or more temporally ordered samples than in the vernier offset cue, for which one snapshot may suffice.

We measured offset threshold in the spatiotemporal task and the temporal task as a function of the temporal frequency of a spatially uniform multiplicative flicker mask. To create this mask, a mask period M equal to a specific number of display frames (3.7 ms each) was chosen. For the first N of these frames, the intensity signal sent to the Tektronix 608 was unchanged (i.e. $\times 1$). For the remaining $M-N$ of the frames, the intensity signal sent to the display monitor was set to zero. This allowed for digital control of the temporal period and duty cycle (N/M) of the mask up to the precision of the frame duration (3.7 ms).

A multiplicative rather than an additive mask was chosen because stimulus contrast is preserved by a multiplicative mask more so than by an additive mask. The latter forces the reduction of the maximum available stimulus contrast in proportion to the mask contrast and at the same time it elevates the luminance in proportion to the mask luminance. With the multiplicative mask we used, contrast was preserved. To make the various masked stimuli as similar as possible in time-averaged luminance, we kept the duty cycle as nearly constant as possible given the constraints of the discrete choices for N and M. Finally, to minimize the slight effect of the mask on temporal phase (because of the end effects at stimulus onset and disappearance), gratings were presented for two full periods (270 ms with a 7.5 Hz grating), with randomization of the initial phase over a the full range across trials.

Fig. 5 shows, for two subjects, that multiplicative flicker mask degrades acuity in the temporal task (open symbols) more than in the spatiotemporal task (filled symbols). Grating frequencies were chosen from the '1 ms' performance range in the spatiotemporal task. Subject BD's performance in the spatiotemporal task (Fig. 5A, filled symbols) indicated only slight interference by the temporal mask above about 20 Hz. Thresholds hovered between 11 and 15 s arc, corresponding to asynchronies between 1.2 and 1.6 ms, only slightly higher than the values obtained in the absence of mask (plotted at 270 Hz, the monitor refresh). The significant threshold elevation near 54 Hz could be accounted for by a sampling artifact (see Appendix). For the temporal task (Fig. 5A, open circles), this subject's asynchrony thresholds were in the neighborhood of 10 ms without masking. Above 40 Hz, there was little or no effect of masking (except for the aforementioned sampling artifact). However, for lower mask frequencies, thresholds in the two tasks had a very different dependence on masking. In the spatiotemporal task, the masking was still moderate below 20 Hz: BD's offset threshold no more than doubled as the mask frequency decreased from 17 to 11 Hz, and did not increase further at lower mask frequencies. In the temporal task, masking became significant below 40 Hz, and asynchrony thresholds increased in an accelerating fashion with decreasing mask frequencies. Below 17 Hz, asynchrony threshold increased to more than 6-fold its unmasked

level, beyond the measurable level of half-period of the grating (~ 68 ms, horizontal dotted line). A vertically shifted copy of the function for the spatiotemporal task (Fig. 5A, dotted curve) is drawn for comparison. Note the very different sensitivities of the thresholds in the two tasks to the temporal mask below 30 Hz.

Subject FM's data showed a similar pattern (Fig. 5B). He exhibited insignificant temporal masking above 40 Hz in both tasks (except for the sampling artifact), but at least a 6-fold and accelerating increase in asynchrony thresholds for the temporal task at low mask frequency. Two additional subjects (YLF and JT, data not shown) were not able to perform the temporal task with this flicker mask. Both, however, could perform the spatiotemporal task under the flicker mask, and their thresholds exhibited a similarly moderate sensitivity to masking in that task to those shown in Fig. 5.

In summary, the stroboscopic mask degrades acuity in the temporal task much more than in the spatiotemporal task, despite its already higher threshold in the absence of the mask. This implies the involvement of a temporal mechanism in the temporal task, but not the spatiotemporal task.

4. Discussion

4.1. Matched comparison of spatiotemporal and temporal tasks

We compared psychophysical performance in detecting offset thresholds in two related tasks. In the 'spatiotemporal' task, vernier thresholds for drifting sine

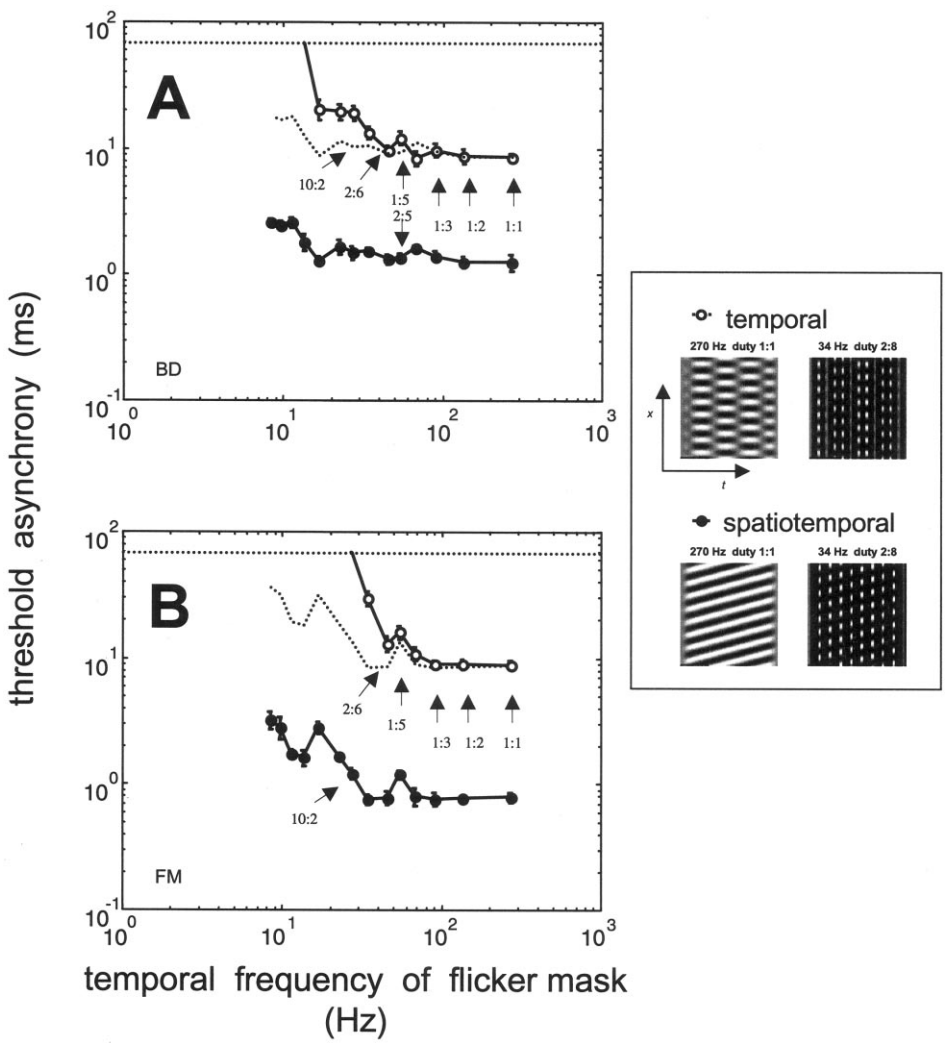

Fig. 5. Asynchrony at thresholds (ms) obtained from subject BD (A) and FM (B) are plotted as a function of the temporal frequency of the mask (Hz). The horizontal dotted line indicates asynchrony at half-cycle offsets. Gratings: 7.4 Hz, 2.94 c/deg, and 0.95 contrast (0.90 for BD in the spatiotemporal task). Insets illustrate the spatiotemporal light distribution in the stimulus, for both tasks, at two different mask frequencies. For most mask frequencies, the duty cycle of the mask was 25%, or n:4n in terms of frame ratios (frames 'on': frames in mask period). Deviations from this at 270 Hz (1:1), 135 Hz (1:2), 90 Hz (1:3), 54 Hz (1:5 or 2:5), 45 Hz (2:6), 27 Hz (2:10 or 3:10) caused a systematic variation in the luminance and contrast that explains the 'bump' in the data near 50 Hz (see text).

gratings were examined. In the 'temporal' task, temporal phase offset thresholds for stationary contrast-reversed sine gratings that were aligned and modulated slightly out-of-phase were examined. This task retained the temporal cue from the corresponding spatiotemporal task but eliminated the spatial offset cue. Our working hypothesis — the temporal hypothesis — was that the apparent temporal limit observed for vernier acuity for translating targets at speeds above 3 deg/s is imposed by one or more asynchrony detection mechanisms which would remain active in the temporal task. Comparison of thresholds in the two tasks, and the dependence of thresholds on stimulus parameters and temporal mask, forces rejection of the temporal hypothesis. Our results are in agreement with earlier studies that argued against the temporal hypothesis (e.g. Westheimer & McKee, 1977), but our experimental design allows a more convincing rejection of this hypothesis, because we match the spatiotemporal properties of stimuli ('front-end'). Moreover, both tasks are two-alternative forced choice tasks and the stimuli have a similar spatial organization (homogeneous bipartite fields). For both stimuli, the cues (spatial phase offset and temporal phase difference and/or apparent motion) are consistent at every location along the field division. Thus, the decision-making paradigm ('back-end') is matched as well.

The temporal hypothesis for moving vernier was addressed earlier by Westheimer and McKee (1977) who measured asynchrony sensitivity by using briefly flashed stationary thin lines and staggering the on- or offset of one of the targets. The asynchrony thresholds they found (3–4 ms under optimal conditions) fell short of the 1 ms limit that the same authors found for moving line vernier stimuli in an earlier study (Westheimer & McKee, 1975). They interpreted the results of this unmatched comparison as evidence against the temporal hypothesis. Our temporal-task stimuli were not intended to be optimal, but rather to be matched to the spatiotemporal task. For these matched stimuli, the apparent temporal thresholds were somewhat worse, thus making the temporal hypothesis even less supportable.

At first glance, the study by Morgan and Watt (1982) in which a temporal threshold of 1 ms (and a spatial hyperacuity) was found, seems to support the temporal hypothesis, in that the temporal threshold was comparable to what is required for drifting vernier. In this study, stimuli were dichoptically presented translating bars whose luminances were sinusoidally modulated out of phase, and which formed the percept of a single bar in depth. Consistent with the findings of Westheimer and McKee (1977) who found that temporal asynchrony sensitivity to dichoptic stimuli is 10 ms at best under near optimal conditions, the authors found that the depth percept, together with high temporal phase acuity, is abolished in the presence of target tracking. This led to the conclusion that 'the cause of the depth percept cannot be temporal differences between the eyes *per se*' (Morgan & Watt, 1982) and thus can also be taken as evidence against the temporal hypothesis. Morgan and Watt's task under tracking can be thought of as a dichoptic temporal phase discrimination task. By eliminating retinal target motion, tracking removes a spatial cue, and is thus analogous to our strategy. However, the two tasks (without and with tracking) are not fully matched because tracking also reduces the retinal extent of the stimulus.

In summary, the visual system achieves high temporal sensitivity not by temporal detectors but by means of highly tuned spatiotemporal integrating mechanisms which are optimally activated by motion signals, i.e. genuine spatiotemporal signals, in which the spatial aspect is essential.

4.2. The motion cue in the temporal tasks

The apparent motion cue is inevitable in a temporal comparison task in which space is used to tag the stimulus. In this way our temporal task can be considered to be a spatiotemporal task, albeit one with the spatial offset cue eliminated. Observers not only perceive apparent motion but need to rely on this spatiotemporal cue: only by using this cue could subjects perform the task with measurable thresholds. Westheimer and McKee (1977) noted in their study of visual asynchrony sensitivity that an apparent motion cue was often present but not essential to achieve low asynchrony thresholds. In other words, multiple cues may give rise to multiple strategies, the most efficient of which — sometimes more than one — will determine psychophysical thresholds. That is, our subjects' ability to make comparisons based only on temporal phase can be no better than their ability to make judgments based on the motion cue — otherwise, they would have achieved lower thresholds. Conversely, the motion cue strategy likely resulted in an underestimate of the temporal thresholds (overestimate of temporal phase sensitivity); were it possible to eliminate this spatiotemporal cue as well, the gap between thresholds on the two tasks could only increase.

The reader might be concerned that the 'temporal task' engages spatiotemporal mechanisms at intermediate stages of processing, particularly because a standing flickering grating can be decomposed into a sum of two gratings drifting in opposite directions. However, the above threshold argument applies here too — engagement of an unintended spatiotemporal cue in the 'temporal task' could only result in a higher threshold than observed, not a lower one.

4.3. Dependence on spatial and temporal frequency of the two tasks

Westheimer and McKee (1975) showed that displacement thresholds for moving Landolt C and for moving line vernier targets remains at the level measured for static targets as horizontal or vertical translation velocity increases up to 2.5 deg/s. Other researchers confirmed this first report (Morgan & Watt, 1983; Morgan & Benton, 1989). At higher target velocities, offset thresholds for vernier line and 3-dot targets become proportional to target velocity (Carney et al., 1995; Chung et al., 1996). With grating targets, experiment 1 reproduced these distinct regimes of behavior.

Carney et al. (1995), using translating 3-dot vernier targets, have found a temporal floor to threshold in the high-velocity range and termed it the '1 ms limit', although their data suggest variation in this limit across a range of dot separations even when target contrast was high. We too find that this temporal limit is dependent on stimulus parameters. Our range (0.4–8 ms) is similar to the range of temporal limits one can cull from the published data by other investigators who used lines or thin bars (0.25–4 ms, Burr, 1979; Fahle & Poggio, 1981; Morgan & Watt, 1983; Carney et al., 1995). All those targets used by the other investigators were spatially limited but spectrally broadband stimuli (containing a mixture of frequencies), while we used gratings (narrowband stimuli). Our results show a dependence of the temporal limit of vernier acuity on the spatiotemporal parameters. Broadband stimuli are 'likely to activate a range of spatial frequency selective mechanisms, and overall performance may be dominated by whichever mechanism is most sensitive under the given experimental conditions' (Whitaker, 1993). This may explain the diversity of results of the other studies.

In terms of asynchrony at threshold, temporal phase offsets in the temporal task are consistently higher (typically by a factor of four to five) than the corresponding vernier thresholds, across a wide range of spatial and temporal frequencies. The dependence of offset thresholds on temporal frequency in the two tasks follows a similar course (constant phase offset regime at low temporal frequencies, followed by a constant asynchrony regime at high temporal frequencies) at each examined spatial frequency. However the transition point separating the frequency ranges corresponding to the constant phase and constant asynchrony regimes may be different in the two tasks.

The absence of a single temporal asymptotic limit to acuity in the spatiotemporal task, and the consistently larger offset thresholds in the temporal task, suggest that detection in the two tasks is limited by different mechanisms. Our results also indicate that more than one mechanism is involved in each task, depending on the stimulus parameters. In regard to the moving vernier task, we note that a recent study (Victor & Conte, 1999) of the effect of the relative temporal phase of sinusoidally flickering bars on offset thresholds in a static vernier task found a cut-off frequency for the effect near 8 Hz, above which vernier acuity had little dependence on the relative temporal phase of the bars. Fendick and Swindale (1994), using counter-phase flickered edges, also found evidence for a switch near 8 Hz of mechanisms in his edge alignment task. In experiment 1, we found a similar cut-off point for phase sensitivity in the spatiotemporal task, where performance shifts from a constant phase offset regime to a constant asynchrony regime.

Consistent with the finding that the limiting mechanism involved in spatiotemporal vernier depends on spatial frequency is the observation that the task is possible to perform even at very high target velocities — up to 1000 deg/s — if the spatial frequency is sufficiently low (Chung et al., 1996; Levi, 1996). Levi and colleagues argue that at very high speeds the spatial analysis switches to low spatial frequencies. A similar observation — spatial scaling of optimal sensitivity in proportion to velocity, essentially maintaining peak sensitivity at a temporal frequency of about 10 Hz — was made earlier by Burr and Ross (1982) using drifting gratings and lines. This seems to be a fundamental property of all spatial tasks, including hyperacuity tasks, that involve moving stimuli. Our results also show that the constant offset regime in the spatiotemporal task extends to progressively higher velocities as the spatial frequency of the sine gratings is lowered (Fig. 2).

4.4. Different dependence of the two tasks on contrast

We found that sensitivity to temporal phase offsets in the temporal task saturated at low contrast (~ 0.05–0.10) while sensitivity to vertical offsets in the spatiotemporal task continued to improve with increasing contrast up to the highest contrasts (0.8–1.0). This reduction in vernier threshold was reasonably well described by a power law function of contrast, with an exponent near -0.6. The dissimilar contrast dependence of thresholds in the two tasks further supports the notion that the detection mechanisms involved in them are different.

For the spatiotemporal task, our results for the dependence of vernier threshold on target contrast are in general agreement with previous studies involving static gratings. Bradley and Skottun (1987) found exponents of -0.77 to -0.84 for stationary sine gratings within the range of 0.25–10.0 c/deg. Whitaker (1993) found a stronger dependence of vernier thresholds on contrast for low than for high spatial frequencies (exponent -1.0 at 1 c/deg; ~ -0.5 above 4 c/deg). This differ-

ence may reflect the vignetted nature of Whittaker's displays.

Studies using spatially localized stationary vernier targets found a similar dependence of vernier acuity on contrast. For Gabor patches, Krauskopf and Farell (1991) found that offset thresholds were inversely proportional to the square root of contrast (i.e. exponent -0.5) for both luminance and chromatic stimuli. For thin lines, Waugh and Levi (1993a) found exponents about -1.0. Fendick and Swindale (1994), using counter-phase edge alignment, found various exponents for the dependence of alignment threshold on contrast: ~ -0.5 below 8 Hz, but ~ -1 above 8 Hz flicker frequency. Our results, obtained with band-limited stimuli, do not confirm this influence of temporal frequency on contrast dependence, but the broadband stimuli they used may account for the difference in results. Wehrhahn and Westheimer (1990) studied offset thresholds for stationary luminance edges. Although the data were described as an exponential improvement of edge vernier acuity with increasing contrast up to 0.22 contrast followed by a plateau, only one point above 0.22 contrast (0.39) was measured. Replotting their data reveals that a power law provides a good fit over the entire range of contrasts, with exponents in the range -0.35 to -0.7 (average -0.53). Monotonic improvement of vernier acuity (without a plateau) with increasing contrast was also found by other investigators: Morgan and Regan (1987), for thin Gaussian bars; and Carney et al. (1995), for translating 3-dot or line targets.

The monotonic contrast dependence of vernier acuity is predicted by short-range models of vernier (Wilson, 1986; Morgan & Regan, 1987; Levi & Waugh, 1996) in which signals are pooled from local spatial or spatiotemporal filters that resemble oriented cortical receptive fields. The similar contrast dependence for static and moving targets suggest that the underlying mechanisms are very similar. These mechanisms were shown (Chung et al., 1996) to be distinct from motion mechanisms that are involved in target detection and that may mediate but not analyze signals for offset detection in the moving vernier task

4.5. Different effects of flicker masking in the two tasks

Further evidence against the hypothetical limiting role of a temporal mechanism in moving vernier acuity comes from the effects of the full-field, stroboscopic (multiplicative flicker) mask. This masking paradigm does not introduce a temporal phase difference between the two targets to be discriminated but it affects temporal integration by introducing a temporal gap between successive samples. This mask profoundly affects performance on judgments of temporal phase offset for contrast-reversed gratings, thus demonstrating that a temporal cue has indeed been isolated. Nevertheless, it has little effect on judgments of misalignment for drifting gratings (the spatiotemporal task). The results imply that performance on the first task, but not the second, depends on a temporal discrimination.

Fahle and Poggio (1981) studied translating line vernier targets in stroboscopic presentation and showed that, with the exception of a narrow range of strobe timing and target velocities, equivalent asynchrony between the target elements did not fully compensate for the percept of spatial offset, indicating that these stimulus dimensions of moving vernier are not equivalent for the mechanism underlying offset judgment. Furthermore, Burr (1979), also studying apparent motion of vernier targets under stroboscopic presentation, showed that the perceptual equivalence between spatial and temporal offsets occurs only if a motion percept is already formed prior to offset judgment. These psychophysical interpolation and extrapolation experiments suggest offset judgment of moving vernier stimuli utilizes spatiotemporal mechanisms that involve the establishment and smooth continuation of a movement trajectory. The neural mechanism underlying filling in (and the anticipation) of the motion trajectory was suggested to be cortical (Fahle & Poggio, 1981), but these could be set up as early as the retina, as demonstrated by a convincing recent study of the population dynamics of retinal ganglion cells in the isolated retina in the presence of a moving visual stimulus (Berry, Brivanlou, Jordan & Meister, 1999).

Previous studies using counter-phase flickering vernier targets found an elevation of alignment thresholds with asynchrony (Westheimer & Hauske, 1975; Wehrhahn & Westheimer, 1993) or temporal phase staggering (Morgan & Watt, 1982; Fendick & Swindale, 1994; Victor & Conte, 1999) between the two halves of the targets. With short-range targets, the thresholds also depend on presentation time, since exposure influences visibility (Morgan et al., 1983; Waugh & Levi, 1993b). However, effects on visibility cannot account for our results, since the mask-induced threshold elevation is disproportionately higher for the temporal task, and thresholds for the latter task have a shallower dependence on visibility than thresholds for the spatiotemporal task.

Although we can only speculate on the nature of the mechanism that solves the temporal task, we do know that a minimum of two snapshots (samples or frames) are necessary, and that information about the temporal separation of snapshots is also necessary. With the flicker mask, snapshots are taken with increasing separation in time as the temporal frequency of the mask is decreased but duty cycle remains fixed. At the adopted 25% duty cycle, the longest possible snapshot remains short (e.g. only 30 ms at 8.5 Hz, the lowest mask

frequency used) and may not be long enough to allow more than a single independent estimate of the stimulus. Even if two estimates were made within 30 ms, those estimates would be highly redundant. More likely, solution of the temporal task requires comparison of two or more of these snapshots across successive flicker periods, and the inefficiency of this comparison is the basis of the degradation of performance by the flicker mask.

4.6. Possible neuronal mechanisms underlying the two tasks

Some retinal ganglion cells possess the two functional properties — high contrast sensitivity and high positional accuracy — that are necessary to meet the extreme demands of spatial vision present in hyperacuity tasks. Cat retinal ganglion X-cells can signal misalignment of gratings better than expected from behavioral hyperacuity (Shapley & Victor, 1986). A majority of neurons in the primate M pathway — M retinal ganglion cells (Kaplan & Shapley, 1986) and their target magnocellular geniculocortical neurons (Shapley, Kaplan & Soodak, 1981; Kaplan & Shapley, 1982) — have similar properties to those of cat X-cells (for a review see Shapley & Perry, 1986; Shapley, 1990), including good spatial resolution, linear spatial summation, high contrast gain, and a tendency to saturation in the contrast response function.

In a sequence of systematic studies (Lee, Wehrhahn, Westheimer & Kremers, 1993, 1995; Ruttiger & Lee, 1998, 1999), Lee and colleagues measured the accuracy and reliability of retinal M and P ganglion cells' in vivo responses to various static and moving stimuli commonly used in psychophysical studies of human hyperacuity. Their results strongly suggests that psychophysical performance is limited mainly by the uncertainty in the responses of retinal M cells, not in neurons at higher levels of the visual pathway. Based on ideal observer analysis of the ganglion cell responses, they argue that pooling signals from only a few M cells (but not P cells) would support human hyperacuity, especially at low contrasts. At higher contrasts, where M cells' responses may saturate (but see Ruttiger & Lee, 1998, 1999), increasing recruitment of the P (parvocellular) neurons could account for the continuing improvement of hyperacuity up to the highest contrasts.

The most sensitive neurons in the visual cortex, in cat area 17 (Swindale & Cynader, 1986) and in monkey V1 (Parker & Hawken, 1985), retain the positional accuracy of the responses of subcortical neurons and support psychophysical hyperacuity. However, the relative M and P pathway contribution to the signal processed in a hyperacuity task by neurons in various layers of V1 and at higher visual cortical areas in the monkey is not known.

The known scaling properties (Mussap & Levi, 1996) and the dependence on orientation cues (Carney et al., 1995; Mussap & Levi, 1997) of psychophysical vernier thresholds under short separations suggest the involvement of spatiotemporal filters resembling receptive fields of V1 neurons that are sharply oriented and tuned to temporal frequency. In the local or short-range models of hyperacuity (Klein & Levi, 1985; Wilson, 1986), signals from local orientation selective filters that scale with size, separation, and spatial frequency, are pooled by a collator mechanism (Mussap & Levi, 1996).

What can we say about the possible cellular mechanisms underlying temporal phase discrimination in the temporal task? The cue in the temporal task is equivalent to perceived direction of apparent motion. Asynchrony thresholds in the temporal task are only weakly dependent on contrast above 0.05–0.10. Psychophysical studies of various direction discrimination tasks (Burr & Ross, 1982; Thompson, 1982; Nakayama & Silverman, 1985; McKee, Silverman & Nakayama, 1986; Derrington & Goddard, 1989; Stone & Thompson, 1992; Hawken, Gegenfurtner & Tang, 1994; Edwards, Badcock & Nishida, 1996) all documented the same distinctive feature we observed in the temporal phase discrimination tasks, namely a remarkable contrast independence of performance above low contrasts (0.05–0.10).

The neuronal substrate for the early vision mechanisms that mediate the motion signal defined by luminance contrast is formed by directionally selective neurons in the primary visual cortex (Hubel & Livingstone, 1987; Reid, Soodak & Shapley, 1987, 1991; Albrecht & Geisler, 1991; Emerson, Bergen & Adelson, 1992) and the extrastriate area MT (Newsome, Britten & Movshon, 1989; Britten, Shadlen, Newsome & Movshon, 1992; Salzman & Newsome, 1994). Converging lines of evidence from neurophysiology (Maunsell, Nealey & DePriest, 1990), neuroimaging (Eden, VanMeter, Rumsey, Maisog, Woods & Zeffiro, 1996), psychophysics (Burr, Morrone & Ross, 1994), as well as modeling (Morgan, 1992) have suggested that this set of directionally selective visual cortical neurons in the primate forms the partially segregated cortical continuation of the M pathway, probably strongly relying on its Y subtype (Kaplan & Shapley, 1982, 1986). The magnocellular afferents endow this low-level motion pathway with a high contrast gain and saturating contrast response function that extends to area MT (Sclar, Maunsell & Lennie, 1990; Gegenfurtner, Kiper, Beusmans, Carandini, Zaidi & Movshon, 1994).

These lines of evidence converge to suggest that the magnocellular pathway is the primary carrier of the signals that support performance in direction discrimination tasks, and in the temporal phase discrimination task. However, the functional segregation of the M and

P pathways breaks down in various parts of the visual cortex (Merigan & Maunsell, 1993). There is ample evidence for some, if not dominant, contribution of the P pathway to motion processing, e.g., for motion and color (Maunsell et al., 1990; Gegenfurtner et al., 1994; Hawken et al., 1994; Gegenfurtner & Hawken, 1996). The anatomical substrate for this exists, perhaps, in the mixed M and P input from neurons in and on the border between layers 4Cα and β, or from neurons of layer 4B, to neurons in layer 2/3 in V1 (Yabuta & Callaway, 1998) that in turn project to higher level motion areas like MT. Since the cortex may pool over signals relayed by both M and P pathways for fine spatial analysis as well as motion processing, especially at high contrasts, it seems that the contrast dependence of offset thresholds in the two tasks can not be explained solely on the basis of neuronal properties in the M pathway. In both tasks, P cell contributions are expected to be most prominent under conditions of low temporal frequency, high spatial frequency, and high contrast (e.g. the relative lack of saturation with increasing contrast in Fig. 4D).

Does the apparent high temporal sensitivity displayed in the moving vernier task reflect the limiting precision with which spiking visual neurons process temporal signal? Here one must distinguish between the ability of the visual system to signal and detect very small asynchronies in the visual stimulus and the ability to process and detect small asynchronies in the internal representation of the stimulus. Our experiments with the temporal task were only designed to answer the first question. Our results indicate that the lower bound to stimulus asynchronies the visual system is able to detect is 5–10 ms at best. Interestingly, this temporal range is at or below the lower end of the 20–60 ms range of temporal coding precision with which visual cortical neurons are known to signal various aspects of the visual stimulus in vivo (Richmond & Optican, 1987; Victor & Purpura, 1996), but the stimuli used in those experiments may not have been optimal for the purpose of temporal coding precision. In principle, the known properties of the neuronal machinery of the visual system allow for reliable detecting and relaying asynchronies in the internal (spike) signal that may be much smaller than those in the external signals. Very high temporal precision — on the order of 1–2 ms or even sub-millisecond — was demonstrated in vitro for retinal ganglion cells (Berry et al., 1997; Berry & Meister, 1998), as well as for neocortical neurons (Mainen & Sejnowski, 1995), and specifically for neurons of the visual cortex (Nowak et al., 1997). The temporal precision demonstrated in vivo for neurons of the primate's subcortical visual pathway (2–5 ms; Reich et al., 1997) was better than the range found for cortical neurons in vivo but somewhat poorer than the in vitro measures.

These studies, together with the aforementioned studies by Lee and his colleagues, suggest that neurons in the early visual pathway can signal fine details in the visual stimulus, both spatial and temporal, with high precision. The psychophysical millisecond sensitivity in high-speed vernier tasks may arise from an early spatiotemporal pooling mechanism provided that a motion trajectory is set up prior to the fine spatial analysis. How this mechanism maintains and detects the high-precision signals it receives from the retina needs further investigation.

Acknowledgements

This work was supported by NIH grants EY7977 and EY9314 to JDV. Part of it was presented at the 1998 Annual Meeting of the Association for Research in Vision and Ophthalmology (Mechler & Victor, 1998). We thank Mary Conte, Daniel Reich, and the reviewers for their critical reading of the manuscript and their many useful suggestions.

Appendix A. Note on the sampling artifact in temporal masking

There is a small but statistically significant and systematic variation of thresholds seen with the variation of the mask frequency above 25 Hz in experiment 3. This is a manifestation of a kind of sampling artifact, due to the use of different duty cycles (needed to accommodate different temporal frequencies), combined with the dependence of RMS contrast and luminance on the duty cycle.

Since the stimulus is digitally sampled in time (in frames), the realizable duty cycles compatible with a mask frequency are integer ratios N/M, where M is the number of frames in the mask period and N is the number of non-masked frames per cycle. The N/M ratio, especially at high frequencies (low M), may be an imprecise approximation to a desired duty cycle (e.g. 25%). For example, to approximate a 25% duty cycle at the mask frequencies 90, 67.5, 54 and 45 Hz, we chose the N/M ratios 1/3, 1/4, 1/5, $2/6 = 1/3$, respectively. RMS contrast necessarily covaries with the duty cycle of the mask. The dependence is similar for both the drifting and contrast-reversed gratings. Additionally, the (spatiotemporal) mean luminance also varies — it is proportional to the duty cycle. Since offset thresholds depend both on contrast (as shown in experiment 2) and on retinal illuminance (Waugh & Levi, 1993a), this discretization necessarily introduces threshold variations across mask frequencies. For these reasons, one expects a threshold elevation — a 'bump' — for the above set of mask frequencies at 54 Hz where the duty

cycle has a local minimum ($N/M = 1/5$), as shown in Fig. 5A, upper curve. But when the N/M ratio at 54 Hz mask frequency was taken to be 2/5 (Fig. 5A, lower curve), the bump was observed at 67.5 Hz because the duty cycle now had its local minimum there with $N/M = 1/4$.

References

Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience*, *7*(6), 531–546.

Bair, W., & Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Computation*, *8*(6), 1185–1202.

Berry, M. J. II, Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature*, *398*(6725), 334–338.

Berry, M. J. II, & Meister, M. (1998). Refractoriness and neural precision. *Journal of Neuroscience*, *18*(6), 2200–2211.

Berry, M. J. II, Warland, D. K., & Meister, M. (1997). The structure and precision of retinal spike trains. *Proceedings of the National Academy of Sciences of the USA*, *94*(10), 5411–5416.

Bradley, A., & Skottun, B. C. (1987). Effects of contrast and spatial frequency on vernier acuity. *Vision Research*, *27*(10), 1817–1824.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*(12), 4745–4765.

Burr, D. C. (1979). Acuity for apparent vernier offset. *Vision Research*, *19*(7), 835–837.

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature*, *371*(6497), 511–513.

Burr, D. C., & Ross, J. (1982). Contrast sensitivity at high velocities. *Vision Research*, *22*(4), 479–484.

Carney, T., Silverstein, D. A., & Klein, S. A. (1995). Vernier acuity during image rotation and translation: visual performance limits. *Vision Research*, *35*(14), 1951–1964.

Chung, S. T., Levi, D. M., & Bedell, H. E. (1996). Vernier in motion: what accounts for the threshold elevation? *Vision Research*, *36*(16), 2395–2410.

Derrington, A. M., & Goddard, P. A. (1989). Failure of motion discrimination at high contrasts: evidence for saturation. *Vision Research*, *29*(12), 1767–1776.

Eden, G. F., VanMeter, J. W., Rumsey, J. M., Maisog, J. M., Woods, R. P., & Zeffiro, T. A. (1996). Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature*, *382*(6586), 66–69.

Edwards, M., Badcock, D. R., & Nishida, S. (1996). Contrast sensitivity of the motion system. *Vision Research*, *36*(16), 2411–2421.

Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, *32*(2), 203–218.

Fahle, M., & Poggio, T. (1981). Visual hyperacuity: spatiotemporal interpolation in human vision. *Proceedings of the Royal Society of London B – Biological Sciences*, *213*(1193), 451–477.

Fendick, M. G., & Swindale, N. V. (1994). Vernier acuity for edges defined by flicker. *Vision Research*, *34*(20), 2717–2726.

Gegenfurtner, K. R., & Hawken, M. J. (1996). Interaction of motion and color in the visual pathways. *Trends in Neurosciences*, *19*(9), 394–401.

Gegenfurtner, K. R., Kiper, D. C., Beusmans, J. M., Carandini, M., Zaidi, Q., & Movshon, J. A. (1994). Chromatic properties of neurons in macaque MT. *Visual Neuroscience*, *11*(3), 455–466.

Hawken, M. J., Gegenfurtner, K. R., & Tang, C. (1994). Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature*, *367*(6460), 268–270.

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience*, *7*(11), 3378–3415.

Kaplan, E., & Shapley, R. M. (1982). X and Y cells in the lateral geniculate nucleus of macaque monkeys. *Journal of Physiology* (*London*), *330*, 125–143.

Kaplan, E., & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences of the USA*, *83*(8), 2755–2757.

Klein, S. A., & Levi, D. M. (1985). Hyperacuity thresholds of 1 s: theoretical predictions and empirical validation. *Journal of the Optical Society of America*, *A*, *2*(7), 1170–1190.

Krauskopf, J., & Farell, B. (1991). Vernier acuity: effects of chromatic content, blur and contrast. *Vision Research*, *31*(4), 735–749.

Lee, B. B., Wehrhahn, C., Westheimer, G., & Kremers, J. (1993). Macaque ganglion cell responses to stimuli that elicit hyperacuity in man: detection of small displacements. *Journal of Neuroscience*, *13*(3), 1001–1009.

Lee, B. B., Wehrhahn, C., Westheimer, G., & Kremers, J. (1995). The spatial precision of macaque ganglion cell responses in relation to vernier acuity of human observers. *Vision Research*, *35*(19), 2743–2758.

Levi, D. M. (1996). Pattern perception at high velocities. *Current Biology*, *6*(8), 1020–1024.

Levi, D. M., & Waugh, S. J. (1996). Position acuity with opposite-contrast polarity features: evidence for a nonlinear collector mechanism for position acuity? *Vision Research*, *36*(4), 573–588.

Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, *268*(5216), 1503–1506.

Maunsell, J. H., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience*, *10*(10), 3323–3334.

McKee, S. P., Silverman, G. H., & Nakayama, K. (1986). Precise velocity discrimination despite random variations in temporal frequency and contrast. *Vision Research*, *26*(4), 609–619.

Mechler, F., & Victor, J. D. (1998). The limits of temporal contribution to spatiotemporal Vernier acuity. *Investigative Ophthalmology and Visual Science* (*Supplement*), *39*(4), 1109.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, *16*.

Milkman, N., Schick, G., Rossetto, M., Ratliff, F., Shapley, R., & Victor, J. D. (1980). A two-dimensional computer-controlled visual stimulator. *Behavioral Research Methods and Instrumentation*, *12*, 283–292.

Milkman, N., Shapley, R., & Schick, G. (1978). Experimental applications. A microcomputer-based visual stimulator. *Behavioral Research Methods and Instrumentation*, *10*, 539–545.

Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature*, *355*(6358), 344–346.

Morgan, M. J., & Benton, S. (1989). Motion-deblurring in human vision. *Nature*, *340*(6232), 385–386.

Morgan, M. J., & Regan, D. (1987). Opponent model for line interval discrimination: interval and vernier performance compared. *Vision Research*, *27*(1), 107–118.

Morgan, M. J., & Watt, R. J. (1982). Hyperacuity for luminance phase angle in the human visual system. *Vision Research*, *22*(7), 863–866.

Morgan, M. J., & Watt, R. J. (1983). On the failure of spatiotemporal interpolation: a filtering model. *Vision Research*, *23*(10), 997–1004.

Morgan, M. J., Watt, R. J., & McKee, S. P. (1983). Exposure duration affects the sensitivity of vernier acuity to target motion. *Vision Research*, *23*(5), 541–546.

Mussap, A. J., & Levi, D. M. (1996). Spatial properties of filters underlying vernier acuity revealed by masking: evidence for collator mechanisms. *Vision Research*, *36*(16), 2459–2473.

Mussap, A. J., & Levi, D. M. (1997). Vernier acuity with plaid masks: the role of oriented filters in vernier acuity. *Vision Research*, *37*(10), 1325–1340.

Nakayama, K., & Silverman, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America*, *A*, *2*(2), 267–274.

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, *341*(6237), 52–54.

Nowak, L. G., Sanchez-Vives, M. V., & McCormick, D. A. (1997). Influence of low and high frequency inputs on spike timing in visual cortical neurons. *Cerebral Cortex*, *7*(6), 487–501.

Parker, A., & Hawken, M. (1985). Capabilities of monkey cortical cells in spatial-resolution tasks. *Journal of the Optical Society of America*, *A*, *2*(7), 1101–1114.

Reich, D. S., Victor, J. D., Knight, B. W., Ozaki, T., & Kaplan, E. (1997). Precise neuronal spike times coexist with large response variability in vivo. *Journal of Neurophysiology*, *77*, 2836–2841.

Reid, R. C., Soodak, R. E., & Shapley, R. M. (1987). Linear mechanisms of directional selectivity in simple cells of cat striate cortex. *Proceedings of the National Academy of Sciences of the USA*, *84*(23), 8740–8744.

Reid, R. C., Soodak, R. E., & Shapley, R. M. (1991). Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. *Journal of Neurophysiology*, *66*(2), 505–529.

Richmond, B. J., & Optican, L. M. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. *Journal of Neurophysiology*, *57*(1), 147–161.

Ruttiger, L., & Lee, B. B. (1998). Vernier signals derived from primate ganglion cells: Effects of motion speed and contrast. *Investigative Ophthalmology and Visual Science* (*Supplement*), *39*(4), S564.

Ruttiger, L., & Lee, B. B. (1999). A comparison of vernier acuity to moving targets with the positional accuracy of ganglion cell signals. *Investigative Ophthalmology and Visual Science* (*Supplement*), *40*(4), S41.

Salzman, C. D., & Newsome, W. T. (1994). Neural mechanisms for forming a perceptual decision. *Science*, *264*(5156), 231–237.

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

Shapley, R. (1990). Visual sensitivity and parallel retinocortical channels. *Annual Review of Psychology*, *41*, 635–658.

Shapley, R., Kaplan, E., & Soodak, R. (1981). Spatial summation and contrast sensitivity of X and Y cells in the lateral geniculate nucleus of the macaque. *Nature*, *292*(5823), 543–545.

Shapley, R., & Perry, V. P. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences*, *9*, 229–235.

Shapley, R., & Victor, J. (1986). Hyperacuity in cat retinal ganglion cells. *Science*, *231*(4741), 999–1002.

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, *32*(8), 1535–1549.

Swindale, N. V., & Cynader, M. S. (1986). Vernier acuity of neurones in cat visual cortex. *Nature*, *319*(6054), 591–593.

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, *22*(3), 377–380.

Victor, J. D., & Conte, M. M. (1999). Short-range vernier acuity: Interactions of temporal frequency, temporal phase, and stimulus polarity. *Vision Research*, *39*, 3351–3371.

Victor, J. D., & Purpura, K. P. (1996). Nature and precision of temporal coding in visual cortex: A metric-space analysis. *Journal of Neurophysiology*, *76*(2), 1310–1326.

Waugh, S. J., & Levi, D. M. (1993a). Visibility, luminance and vernier acuity. *Vision Research*, *33*(4), 527–538.

Waugh, S. J., & Levi, D. M. (1993b). Visibility, timing and vernier acuity. *Vision Research*, *33*(4), 505–526.

Wehrhahn, C., & Westheimer, G. (1990). How vernier acuity depends on contrast. *Experimental Brain Research*, *80*(3), 618–620.

Wehrhahn, C., & Westheimer, G. (1993). Temporal asynchrony interferes with vernier acuity. *Visual Neuroscience*, *10*(1), 13–19.

Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research*, *15*, 1137–1141.

Westheimer, G., & McKee, S. P. (1975). Visual acuity in the presence of retinal-image motion. *Journal of the Optical Society of America*, *65*(7), 847–850.

Westheimer, G., & McKee, S. P. (1977). Perception of temporal order in adjacent visual stimuli. *Vision Research*, *17*(8), 887–892.

Whitaker, D. (1993). What part of a vernier stimulus determines performance? *Vision Research*, *33*(1), 27–32.

Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research*, *26*(3), 453–469.

Yabuta, N. H., & Callaway, E. M. (1998). Functional streams and local connections of layer 4C neurons in primary visual cortex of the macaque monkey. *Journal of Neuroscience*, *18*(22), 9489–9499.