Contrast thresholds for component motion with full and poor attention

Naotsugu Tsuchiya  
Humanities and Social Sciences, California Institute of Technology, Pasadena, CA, USA

Jochen Braun  
Cognitive Biology, University of Magdeburg, Magdeburg, Germany

We compare luminance-contrast-masking thresholds for fully and poorly attended stimuli, controlling attention with a demanding concurrent task. We use dynamic displays composed of discrete spatiotemporal wavelets, comparing three conditions ("single," "parallel," and "random"). In contrast to static displays, we do not find that attention modulates the "dipper" regime for masks of low luminance contrast. Nor does attention alter direction-selective masking by multiple wavelets moving in random directions, a condition designed to isolate effects on component motion. However, direction-selective masking by multiple wavelets moving in parallel is significantly reduced by attention. As the latter condition is expected to excite both component and pattern motion mechanisms, this implies that attention may alter the visual representation of pattern motion. In addition, attention exhibits its well-known effect of reducing lateral masking between nearby spatiotemporal wavelets.

Keywords: component motion, pattern motion, luminance contrast masking, detection, discrimination, attention, dual-tasks


Introduction

The psychophysical paradigm of “luminance contrast masking” (LCM) was developed to probe the visual representation of luminance contrast, orientation, and spatial frequency of static visual patterns (Foley, 1994; Itti, Koch, & Braun, 2000; Lee, Itti, Koch, & Braun, 1999; Legge & Foley, 1980; Wilson, 1980). Results from this approach agree quantitatively with the dependence of responses of cortical neurons on luminance contrast, orientation, and spatial frequency (Geisler & Albrecht, 1997; Itti et al., 2000). In addition, LCM can uncover how visual representations are altered by attention (Carrasco, Penpeci-Talgar, & Eckstein, 2000; Freeman, Sagì, & Driver, 2001; Lee, Itti, et al., 1999; Morrone, Denti, & Spinelli, 2002; Zenger, Braun, & Koch, 2000). For example, LCM reveals that attention intensifies competitive interactions among visual filters, resulting in a higher effective gain and a sharper effective tuning for static visual patterns (Braun, Koch, Lee, & Itti, 2001; Lee, Itti, et al., 1999).

Here, we ask whether LCM manifests comparable attention effects for dynamic visual patterns. Traditionally, attention is thought to interact, though little, with the perception of visual motion. Manipulations of attention with cueing and visual search paradigms typically produce little or no effect on the perception of visual motion (Raymond, 2000). However, more recent psychophysical work (Chaudhuri, 1990; Raymond, O’Donnell, & Tipper, 1998) as well as neuroimaging (Gandhi, Heeger, & Boynton, 1999; Huk & Heeger, 2000; Saenz, Buracas, & Boynton, 2002; Watanabe et al., 1998) and neurophysiological studies (Martinez-Trujillo & Treue, 2002; Seidemann & Newsome, 1999; Treue & Maunsell, 1996) have established robust attention effects on neural responses to visual motion. We attempt to quantify attention effects on the perception of visual motion with the help of LCM.

An obstacle to achieving this goal is that visual motion is represented at multiple levels in the visual system. Particularly relevant here are representations of “component” and “pattern” motion (Adelson & Movshon, 1982; Simoncelli & Heeger, 1998; Welch, 1989; Wilson & Kim, 1994). Visual filters tuned to a particular spatiotemporal frequency are inherently ambiguous about the true direction and speed of motion (component motion; Adelson & Bergen, 1985). A wide range of spatiotemporal frequencies must be compared to identify the veridical motion vector (pattern motion; Adelson & Movshon, 1982; Welch, 1989). Whereas several areas of visual cortex, including area V1, are tuned to component motion, selectivity for pattern motion appears concentrated in the middle temporal cortex (area MT or V5; Huk & Heeger, 2002; Movshon, Adelson, Gizzi, & Newsome, 1985). The neural circuits underlying this transformation are under active study (Heuer & Britten, 2002; Movshon & Newsome, 1996). The distinct representations of component and pattern motion were first studied with displays (“moving plaids”).
that superimpose two moving gratings (Adelson & Movshon, 1982). However, intersections between gratings are perceptually conspicuous and complicate the interpretation of results (Stoner & Albright, 1992; Stoner, Albright, & Ramachandran, 1990; Wilson & Kim, 1994). Schrater, Knill, and Simoncelli (2000) filtered dynamic noise to distribute motion energy in a manner comparable to moving plaids but without introducing conspicuous features. Adopting a similar approach, we combined discrete “wavelets” of spatiotemporal luminance variation to create dynamic textures of spatially uniform appearance.

To distinguish the respective contributions of component and pattern mechanisms to the perception of visual motion, we took advantage of known properties of pattern-selective neurons in middle temporal area MT. The response of such neurons to a preferred motion is reduced and, in some cases, even suppressed by the simultaneous presence of motion in the opposite direction. This nonlinear interaction between different motion components is known as “motion opponency” (Heeger, Boynton, Demb, Seidemann, & Newsome, 1999; Qian & Andersen, 1994; Snowden, Treue, Erickson, & Andersen, 1991). In fact, area MT as a whole responds only minimally to multiple motion components in random directions (Britten, Shadlen, Newsome, & Movshon, 1993; Rees, Friston, & Koch, 2000). Presumably, the response to any one component is inhibited by the simultaneous presence of the other components (Simoncelli & Heeger, 1998). Accordingly, dynamic patterns containing all directions of motion should drive component mechanisms far better than pattern mechanisms.

We conducted LCM experiments with just such a stimulus to probe the visual representation of component motion. Our results confirmed and extended several earlier studies on motion masking (Anderson & Burr, 1985, 1989; Anderson, Burr, & Morrone, 1991; Ferrera & Wilson, 1987; Lu & Sperling, 1995, 1996). To measure the effect of attention, we used an established dual-task paradigm (Braun, 1994, 1998; Braun & Julesz, 1998; Lee, Itti, et al., 1999; Lee, Koch, & Braun, 1997, 1999; Li, VanRullen, Koch, & Perona, 2002; Zenger et al., 2000) and compared LCM thresholds for moving patterns that are either fully or poorly attended.

**Methods**

**Subjects and apparatus**

Six naïve subjects participated in the study. Stimuli were displayed on a 19-in. raster monitor (1,280 × 1,024 pixels RGB) with a 74-Hz refresh rate (13 ms/frame). Average display luminance was 40 cd/m², and gamma correction combined with color bit stealing (Tyler, 1997) provided linear luminance steps of 0.07 cd/m². Viewing was binocular (80 pixels per 1° visual angle).

**Peripheral task (LCM)**

The peripheral stimulus was composed of 1 or 23 pairs of “moving wavelets” (see below), each pair comprising a target wavelet and a superimposed masker wavelet. A standard adaptive staircase method was used to establish contrast thresholds for detecting target wavelets (Levitt, 1971). Subjects pressed separate keys to report “presence” or “absence” of target wavelets. Masking wavelets were always present, and their contrast remained fixed during each block of trials. The peripheral stimulus filled a circular region of up to 1.5°, centered at 3.7° eccentricity. In a control experiment, the peripheral stimulus was preceded by a briefly flashed (26 ms) cue, which preceded the stimulus by 104 ms.

**Central task (letter discrimination)**

To draw attention away from the visual periphery, subjects discriminated an array of seven letters (Figure 1a), which appeared at varying positions and orientations within 1.4° eccentricity. Subjects pressed separate keys to report “all the same” or “one letter different.” Letters were masked, and stimulus onset asynchrony (SOA) was adjusted (SOA = 164–250 ms) to obtain a performance of approximately 80%.

When performing both tasks, subjects gave priority to the central task. Dual-task blocks with poor central performance (<75%) were discarded. Both central and peripheral stimuli were always present, and subjects always fixated at display center, ensuring identical visual stimulation under single- and dual-task conditions.

**Log Gabor wavelets**

We used a set of self-similar wavelets of spatiotemporal luminance variation to generate visual motion. To facilitate comparison with neuronal tuning, we chose to use log Gabor wavelets (Field, 1987) instead of the more conventional linear Gabors (Figure 2a; Movie 1). Log and linear Gabors are similar in that both present a drifting contrast phase within a stationary spatiotemporal envelope and in that Fourier energy is concentrated around one particular spatiotemporal frequency (Ωx, Ωy, Ωτ). However, the Fourier energy of log Gabors conforms to Gaussian distributions with respect to the logarithm of spatial frequency and the logarithm of temporal frequency (as well as with respect to linear spatial direction), similar to the spatiotemporal tuning of cortical neurons (Geisler &
Albrecht, 1997). The Fourier amplitude of a log Gabor wavelet is

\[ E(\omega_x, \omega_y, \omega_t) = E(\omega_r, \theta, \omega_t) \]

\[ = \frac{\sqrt{\pi}}{\Omega^2 \Omega_r \Lambda^2 \Lambda_t \Lambda_t \ln^2 2} \exp\left( -\frac{\ln^2 \omega_r}{2\Lambda^2 \ln^2 2} \right) \]

\[ \cdot \left( A^+_{(\theta)} \cdot B^+_{(\omega)} \cdot e^{i\Phi} + A^-_{(\theta)} \cdot B^-_{(\omega)} \cdot e^{-i\Phi} \right) \]

\[ \omega_r = \sqrt{\omega_x^2 + \omega_y^2} \]

\[ \theta = \arctan(\omega_x/\omega_y) \]

\[ \Omega^+ = \Omega_\theta + \pi \]

\[ \Omega^- = \Omega_\theta \]

where the Cartesian coordinates of Fourier space \((\omega_x, \omega_y, \omega_t)\) are replaced by polar coordinates \((\omega_r, \theta, \omega_t)\), \(\Omega_r\) (cpd), \(\Omega_t\) (Hz), and \(\Omega_\theta\) (°) are the peak spatial and temporal frequencies and directions, respectively; \(\Lambda_t\) (octaves), \(\Lambda_r\) (octaves), and \(\Lambda_\theta\) (°) are the standard deviations or bandwidth, and \(\Phi\) is the phase of the wavelet.

A spatiotemporal wavelet \(W(x, y, t)\) was obtained as the inverse Fourier transform of \(E(\omega_x, \omega_y, \omega_t)\). The normalization of \(E(\omega_x, \omega_y, \omega_t)\) was chosen such that \(|W(x, y, t)|\) takes maximal values on the order of unity. The same normalization factor was used for all the 144 wavelets.

The functions \(A^+, A^-\), \(B^+,\) and \(B^-\) denote positive and negative lobes of the Fourier amplitude, which jointly determine the wavelet motion in space–time. \(A^+\) gives the direction dependency \(\theta\),

\[ A^\pm_{(\theta)} = \exp\left( -\frac{(\theta - \Omega^\pm_\theta)^2}{2\Lambda^2_\theta} \right) \]

For example, a horizontally oriented and vertically upward-moving wavelet (going in the 90° direction) has \(A^+\) with \(\Omega_\theta = -90°\) and \(A^-\) with \(\Omega_\theta = +90°\), whereas a downward moving wavelet has \(A^+\) with \(\Omega_\theta = +90°\) and \(A^-\) with \(\Omega_\theta = -90°\).

\(B^\pm\) gives the Gaussian dependency on the logarithm of temporal frequency \(\omega\),

\[ B^\pm_{(\omega)} = \exp\left( -\ln^2 |\pm \omega_{(\omega)}| \right), \text{ where } |x| = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \]

The dimensions of each log Gabor wavelet were 128 pixels × 128 pixels × 16 video frames. The peaks and standard deviations of Fourier amplitude were \(\Omega_r = 2.5\) cpd, \(\Lambda_r = 0.6\) octaves, \(\Lambda_\theta = 13°\), \(\Omega_t = 6.0\) Hz, and \(\Lambda_t = 0.6\) octaves. For comparison, the median values for area V1 neurons of macaque are \(\Omega_r = 4.2\) cpd, \(\Lambda_r = 0.72\) octaves, \(\Lambda_\theta = 15°\), \(\Omega_t = 7.2\) Hz, and \(\Lambda_t = 1.2\) octaves (Geisler & Albrecht, 1997).

Using 3D inverse Fourier transform, we computed 144 wavelets covering 36 directions (0°, 10°, 20°…) and 4 phases (0°, 90°, 180°, and 270°), each in the form of a
A 3D real-valued matrix (128 × 128 pixels × 16 frames). The contrast of individual wavelets was defined as

\[
C_{\text{set}} = \frac{\max_{(x,y,t)} W_{(x,y,t)} - \min_{(x,y,t)} W_{(x,y,t)}}{\max_{(x,y,t)} W_{(x,y,t)} + \min_{(x,y,t)} W_{(x,y,t)}},
\]

where maxima and minima are taken over the entire set of 144 wavelets. Additional contrast values were obtained by linear scaling.

**Pairs of wavelets**

Coextensive target and masker wavelets were superimposed to form wavelet pairs (Figure 3; Movie 2). The phase difference (relative phase) between target and masker was fixed at 0° to maximize interactions. The direction difference (relative direction 0°, 30°, 90°, 150°, or 180°) and the masker contrast (0% to 64%) remained constant during each block of trials.

For a relative direction of 0°, the pair effectively forms a single wavelet of higher contrast. For other relative directions, the pair forms various “interference patterns.” For example, an intersection forms at 90° and a contrast flash at 180° (“counterphase”). Above threshold contrast, these interference patterns provide a cue to the presence of target wavelets. Near threshold contrast, however, interference patterns fade and subjects judge based on contrast information.

**Wavelet composites**

To create wavelet composites, we randomly placed 23 wavelet pairs (but with a minimal center-to-center spacing...
of 0.25°) in a circular area with a diameter of 1.5°, centered at an eccentricity $\varepsilon = 3.7°$. For comparison, the average diameter of central receptive fields at this eccentricity has been estimated as 0.22° in area V1 (Dow, Snyder, Vautin, & Bauer, 1981) and as 3.3° in area MT (Albright & Desimone, 1987).

Three conditions were compared (Figure 4): (i) a single wavelet pair of random direction placed randomly in the 1.5° area (cf. Figure 2b); (ii) 23 parallel wavelet pairs of identical orientation and direction of motion, chosen randomly for each trial (cf. Figure 2c and Movie 3); and (iii) 23 random wavelets pairs with different orientations and directions of motion, randomly assigned to each pair (cf. Figure 2d and Movie 4).

### Results

#### Measurements with full attention

**Contrast-increment thresholds (Figures 5a–5c):**

When target and masker wavelets move in the same direction (relative direction, 0°), the peripheral task involves discriminating two patterns of different contrast (“contrast-increment thresholds”).

Absolute detection thresholds (masker contrast, 0%; relative direction, 0°) were $6.7 \pm 0.2%$ for single, $1.9 \pm 0.2%$ for parallel, and $3.7 \pm 0.1%$ for random wavelets (mean and standard error from four, two, and six observers).

![Figure 3](image)

**Figure 3.** Superposition of target and mask wavelets. Wavelets occupy a volume in space–time ($X = 1, ..., 128, Y = 1, ..., 128, T = 1, ..., 16$) and are represented by slices through this volume. (Top) Mask wavelet of a given contrast (exaggerated for clarity). Instantaneous appearance is illustrated by three $X$–$Y$ slices (at times $T = 4, 8$, and 12) and temporal evolution by one $X$–$T$ slice (at position $Y = 64$). (Middle) Target wavelet of half the masking contrast, here differing by 30° in its direction of motion from mask wavelet. Other directional differences used were 0°, 90°, 150°, and 180° (not shown). (Bottom) Target and mask wavelet superimposed. Both the instantaneous appearance and temporal evolution are affected by the superposition.
Movie 2. (A) 60% mask + 10% target. A target wavelet goes in the 60° direction. The mask goes in the −30° direction (90° mask). (B) 60% mask + 30% target. The target contrast increased to 30%. It is easy to notice the difference from the left, mask wavelet only condition. Note that the presence of the target creates a noticeable “node” at the intersection of two wavelets, which is a problem when one superimposes two suprathreshold gratings. Also, one feels strong percept of “pattern motion” going to the +15° direction. We knocked out this percept using multiple random wavelets. Click on each image to view the corresponding movie.

Figure 4. Instantaneous appearance of wavelet arrays at the time of maximal contrast. (Top) Single wavelets. Observers discriminated between a mask wavelet (leftmost frame) and the superposition of mask and target wavelet (other frames). The relative directions of motion are 0°, 30°, or 90° (150° and 180° not shown). (Middle) Twenty-three parallel wavelet pairs. Relative directions of target and masker are identical for all pairs, in some cases creating the appearance of a “plaid” (rightmost frame). (Bottom) Twenty-three wavelets of random direction. The relative direction of target and masker wavelets was identical for all pairs.
respectively). This rank order of thresholds (parallel < random < single) held for all observers. The difference between configurations was significant (one-tailed $t$ test on single vs. random: $t$ score = 4.73, $df = 8$; random vs. parallel: $t$ score = 2.04, $df = 6$).

Contrast-increment thresholds for a series of masker contrasts (0%, 1%, 2%, 4%, 16%, and 32%) are shown in Figures 5a–5c. As is typical, thresholds improve as masker contrast increases from zero to approximately the level of detection threshold, before rising as masker contrast increases.

Movie 3. (A) Twenty-three parallel wavelets (mask only). The contrast of each wavelet is 60%. The direction of all the wavelets is the same. This configuration resembles the conventional sinusoidal gratings, faced with the problem of "nodes" at the intersection. The movie is 8.5 times slower than the actual stimuli. (B) 60% mask + 30% target. Each target wavelet goes in the opposite direction (180° different) from the paired local mask wavelet. Click on each image to view the corresponding movie.

Movie 4. (A) Twenty-three random wavelets (mask only). The contrast of each wavelet is 60%. The direction of each wavelet is set randomly. This low coherency, yet, allows us to perform the LCM experiments. The movie is 8.5 times slower than the actual stimuli. (B) 60% mask + 30% target. Each target wavelet goes 150° different from the paired local mask wavelet. Click on each image to view the corresponding movie.
contrast increases beyond the detection level. This dependence defines a facilitatory regime (or “dipper”) and an inhibitory regime of mask contrast.

The dipper was pronounced for single wavelets (64% reduction) and for random wavelets (73% reduction) but less so for parallel wavelets (35% reduction). A two-way ANOVA (Subject × Contrast) revealed a main effect of contrast ($P < 1 \times 10^{-6}$, $F = 66.76, 42.07,$ and $179.55$ for the single, parallel, and random configurations, respectively). An interaction (Configuration × Contrast) was significant at low (0–4%) but not at high (＞4%) mask contrast ($F = 5.24$ and $F = 1.03$, respectively).

To assess the extent to which our results were contaminated by “positional uncertainty” (Foley & Schwarz, 1998; Solomon, Lavie, & Morgan, 1997) at low mask contrast, we repeated some measurements with positional cueing (see the Methods section). Although absolute detection thresholds were reduced, they remained significantly above the minimal thresholds in the dipper regime (not shown).

**Contrast-masking thresholds (Figures 5d–5i)**

At relative directions other than 0°, the nature of the peripheral task changes. It now involves detecting one moving pattern (target wavelets) in the presence of another (masker wavelets). We measured how “contrast-masking thresholds” depend on relative direction of target and masker wavelets for facilitatory masker contrasts (1% or 4%) and for inhibitory masker contrasts (16% or 32%).

Threshold facilitation by low-contrast maskers is depicted in Figures 5d–5f (filled symbols). In general, thresholds increased for relative directions 0° to 30°, reaching a plateau for relative directions of 90° and above. The particular results for each configuration mirror contrast-increment thresholds: The lowest point and the

---

![Figure 5](https://example.com/figure5.png)

Figure 5. Contrast-increment and contrast-masking thresholds measured with full attention (filled symbols) and poor attention (open symbols). (a–c) Thresholds as a function of mask contrast (log–log plot). (d–f) Thresholds as a function of relative direction of target and mask wavelets; low masking contrast (1% or 4%). (g–i) Same for high masking contrast (16% or 32%). Results for single, parallel, and random wavelets are shown in the left, middle, and right columns, respectively.
plateau level in Figures 5d–5f correspond to, respectively, the lowest point of the dipper and the absolute detection threshold in Figures 5a–5c.

Interestingly, no significant facilitation occurred at 180° (opponent or counterphase motion), although target and masker wavelets shared the same spatial orientation, demonstrating that facilitation is mediated by motion-specific mechanisms.

Threshold elevation by high-contrast maskers is shown in Figures 5g–5i (filled symbols). For all configurations, the lowest thresholds were observed for a relative direction of 90°, rising to higher levels for relative directions that are less than or greater than 90°. The details of this rise suggest qualitative differences between wavelet configurations (see the Discussion section).

For single wavelets, there was a pronounced asymmetry between relative directions of 0° and 30° on the one hand and at 150° and 180°, on the other, with parallel motion masking more effectively than opponent motion. In contrast, the random wavelet configuration produced a more symmetric pattern of thresholds, with comparable values at 0° and 30° and at 150° and 180°, suggesting an inhibition specific for spatial orientation rather than for direction of motion. This inhibition appears weaker for maskers of identical (0°) and exactly opponent (180°) direction.

In the case of parallel wavelets, no clear pattern emerged with full attention, and the results at 0° and 30° and at 150° and 180° may well reflect a combination of orientation-selective (symmetric) and direction-selective (asymmetric) local inhibition.

With three subjects, we conducted additional experiments with single and random configurations and high-contrast maskers. The resulting threshold patterns were consistently symmetric for random wavelets and consistently asymmetric for single wavelets (results not shown).

**Measurements with poor attention**

**Attentional strategy**

Performance of the central task was nearly constant across configurations (81.1%, 83.4%, and 81.1% correct for single, parallel, and random wavelets, respectively). To rule out the possibility that attention might have swerved to the peripheral array in a subset of trials, we analyzed correlations between central and peripheral performance (correct or incorrect responses) for trials near threshold contrast. If attention favors one task on some trials and the other task on others, one can expect a negative correlation between successes (failures) for trials near threshold contrast. If attention favors one task on some trials and the other task on others, one can expect a negative correlation between successes (failures) for trials near threshold contrast. If attention favors one task on some trials and the other task on others, one can expect a negative correlation between successes (failures) for trials near threshold contrast. If attention favors one task on some trials and the other task on others, one can expect a negative correlation between successes (failures) for trials near threshold contrast. However, for random wavelets, the attention effect was uniform across all relative directions (F = 1.22).

**Contrast-increment thresholds (Figures 5a–5c)**

Under conditions of poor attention, absolute detection thresholds increased by 12%, 46%, and 13% for single, random, and parallel wavelets, respectively. The difference was significant for single and parallel wavelets (F = 11.03 and 10.44) and almost significant (P < .07, F = 3.73) for random wavelets.

For higher masker contrasts, the effect of attention depended strongly on array configuration: Thresholds for single wavelets were elevated by 22% on average, whereas for parallel and random wavelets, thresholds were elevated by an average of 109% and 94%, respectively.

For all wavelet configurations, threshold elevation was larger for high mask contrasts. A three-way ANOVA (Subject × Contrast × Attention) revealed significant main effects of contrast (F = 127.59, 82.37, and 195.91) and attention (F = 15.09, 82.56, and 111.82) and a significant interaction between contrast and attention (F = 4.48, 4.28, and 11.24).

**Contrast-masking thresholds (Figures 5d–5i)**

With low-contrast maskers, poor attention elevated thresholds slightly, but the difference reached significance only in 3 of 15 conditions (Figures 5d–5f, open symbols). Apparently, attention is of little consequence as long as the interaction between target and masker remains facilitatory.

Poor attention had a rather more dramatic effect with high-contrast maskers (Figures 5g–5i, open symbols). Thresholds were 33% higher on average for single wavelets, 216% higher for parallel wavelets, and 82% higher for random wavelets. A three-way ANOVA (Subject × Mask Direction × Attention) revealed significant effects of relative direction (F = 39.28, 6.97, and 5.73) and attention (F = 69.28, 123.52, and 80.56). For single and parallel wavelets, attention and relative direction interacted significantly (F = 10.54 and 4.31). However, for random wavelets, the attention effect was uniform across all relative directions (F = 1.22).

**Discussion**

Our aim was to compare the visual representation of component motion (Adelson & Movshon, 1982; Simoncelli & Heeger, 1998) under conditions of full and poor attention. One of our moving arrays—random wavelets—sought to minimize pattern motion by stimulating all directions of motion equally (cf. Figure 2d). Extensive neurophysiological evidence shows that multidirectional motion is a comparably poor stimulus for pattern-sensitive mechanisms in visual area MT/V5 (Britten et al., 1993; Heeger et al., 1999;
Absolute thresholds

Absolute detection thresholds were substantially lower for multiple wavelets than for single wavelet, as predicted by signal summation. The observed degree of summation was quantitatively comparable to other studies (Bonnehe & Sagi, 1998; Meese & Williams, 2000; Quick, 1974; Tyler & Chen, 2000). However, random wavelets summed somewhat less than expected and parallel wavelets summed somewhat more than expected (Table 1). The comparatively small effects of attention may well be due to decisional factors (i.e., reduced positional uncertainty). Their magnitude confirms the traditional view that motion processing depends only marginally on attention (Raymond, 2000).

Lateral inhibition

Lateral interactions should be evident in the comparison of single and multiple wavelets. For high-contrast maskers, such interactions are expected to be inhibitory (Polat & Sagi, 1993; Zenger et al., 2000). The effect of lateral interactions should be most evident when local interactions are minimal. This is the case for orthogonal target and masker wavelets (relative direction, 90°), where local interactions proved minimal (see below). Table 2 lists threshold elevation by high-contrast maskers for different wavelet configurations and states of attention. Threshold elevation is significantly higher for multiple wavelets, especially in the case of poor attention, presumably reflecting stronger lateral inhibition.

The large attention effect is observed for both configurations of multiple wavelets and for all relative directions, not only for 90° (Figures 5h and 5i). The implication is that attention reduces lateral inhibition by nonoverlapping wavelets of high contrast. The strength of this inhibition and the degree of reduction depend on wavelet configuration (random or parallel). This interpretation—attention reduces lateral inhibition—is consistent with previous findings concerning static stimuli. It is well known that attention modulates lateral interactions between high-contrast stimuli in a configuration-dependent manner (e.g., Freeman et al., 2001; Freeman, Sagi, & Driver, 2004; Zenger et al., 2000). Inhibitory interactions are particularly affected and attention may decrease their effectiveness by a factor of 4 or more (Zenger et al., 2000).

Local facilitation

We observed a threshold reduction by low-contrast maskers (dipper) for all configurations, including random wavelets, the configuration designed to isolate local mechanisms representing component motion (Figures 5a–5c). The reduction is largest when target and masker wavelets move in identical directions (Figures 5d–5f), and therefore, it reflects subthreshold summation by visual filters representing component motion (Levinson & Sekuler, 1975; Stromeyer, Kronauer, Madsen, & Klein, 1984; Wilson, 1985). The comparatively shallow dipper obtained for parallel wavelets echoes previous results with sinusoidal gratings (Bowne, 1990; Lu & Sperling, 1995, 1996).

We found no evidence that attention alters this local facilitation. Neither the depth nor the direction dependence of the dipper function (Figures 5a–5c) was affected by attention. Apparently, the initial contrast response of
component motion mechanisms is not modulated by attention.

This finding differs markedly from our earlier results for static visual patterns. In that case, facilitation by low-contrast maskers was significantly enhanced by attention (Lee, Itti, et al., 1999). Thus, there appears to be a clear difference between dynamic and static visual filters: Attentional feedback seems to reach only the latter class of mechanisms.

Orientation-selective and local inhibition

For high-contrast maskers, we observed elevated thresholds for all wavelet configurations (Figures 5g–5i). We first consider random wavelets, the configuration that probes local interactions between mechanisms representing component motion, and interpret elevated thresholds as evidence for local inhibition among such mechanisms. For random wavelets, this local inhibition is specific to the relative orientation, not to the relative direction, of target and masker wavelets (symmetric “M” shape in Figure 5i).

This may seem surprising, but this corresponds to the organization of orientation and direction columns in area V1, where adjacent columns represent identical orientations but opposite directions (Ohki, Chung, Ch’ng, Kara, & Reid, 2005; Shmuel & Grinvald, 1996; Weliky, Bosking, & Fitzpatrick, 1996). Accordingly, a local but otherwise indiscriminate connectivity would entail interactions that are orientation selective without being direction selective.

We find no evidence that attention modulates this orientation-selective and local inhibition (no significant interaction between attention and relative direction), confirming the lack of attention effects on the representation of component motion.

Direction-selective inhibition

The other two wavelet configurations—single and parallel wavelets—do not distinguish between local and lateral interactions: The relative direction of target and masker wavelets controls how target wavelets are situated relative to both overlapping and nonoverlapping maskers. The observation of elevated thresholds for relative directions of 0° and 30°, but not 150° and 180°, implies an inhibitory interaction that is direction selective (Figures 5g and 5h). This inhibition may be local or lateral or it may be both. Anderson and Burr (1985) reported a direction-selective inhibition of similar magnitude using low-pass-filtered, one-dimensional noise.

The direction-selective inhibition is of particular interest, as it is significantly reduced by attention (Figures 5g and 5h). This inhibition seems to originate from mechanisms selective for pattern motion (significant interaction between attention and relative direction), as it is observed only when these mechanisms are driven (i.e., by single and parallel, but not by random, wavelets). The evidence suggests, therefore, that attention modulates the representation of pattern motion but not that of component motion. This throws a new light on neurophysiological attention effects in visual area MT/V5 (Martinez-Trujillo & Treue, 2002; Seidemann & Newsome, 1999; Treue & Maunsell, 1996).

Presumed neural correlates

According to a model by Simoncelli and Heeger (1998), visual motion is represented at two stages. The first stage (component motion) consists of visual filters selective for particular spatiotemporal frequencies. Our wavelets match the average filter bandwidth at this first stage. The second stage (pattern motion) comprises filters selective for visual motion of a particular direction and velocity. Our
random wavelet array minimizes stimulation of this second stage.

It is thought that the output of first-stage filters is normalized (inhibited) by the summed activity of a subpopulation of filters and half-rectified (Figure 6). Our observation of orientation-selective (rather than direction-selective) inhibition speaks to the composition of this subpopulation. Specifically, it implies that a given filter is normalized (inhibited) by other filters preferring both similar and opposite directions of motion. Our failure to observe orientation-specific attention effects shows that the divisive normalization of dynamic visual filters is not modulated by attention. This stands in sharp contrast to static visual filters, which depend dramatically on attention (Lee, Itti, et al., 1999).

Acknowledgments

We thank Geraint Rees and Christof Koch for helpful comments. N.T. was funded by grants from NIMH, NSF, the Keck Foundation, and the Moore Foundation.

Commercial relationships: none.
Corresponding author: Jochen Braun.
Email: jochen.braun@nat.uni-magdeburg.de.
Address: Cognitive Biology, University of Magdeburg, Leipziger Strasse 44, 39120 Magdeburg, Germany.

References


