Sperling, G., Chubb, C., Solomon, J. A., and Lu, Z-L. (1994). Fullwave and halfwave processes in second order motion and texture. In *Higher-order processing in the visual system.* Chichester, UK: Wiley (Ciba Foundation Symposium, 184). Pp. 287-303.

Full-wave and half-wave processes in second-order motion and texture

George Sperling, Charles Chubb†, Joshua A. Solomon* and Zhong-Lin Lu

Department of Cognitive Science and Institute of Mathematical Behavioral Sciences, University of California, Irvine, CA 92717 and †Department of Psychology, Rutgers University, New Brunswick, NJ 08903, USA

Abstract. A theory of human second-order motion perception is proposed and further applied to the discrimination of texture slant. The computational algorithms for deriving the direction of left–right motion from a sequence of images are equivalent to the algorithms for deriving the direction of slant (e.g. from top left to bottom right or top right to bottom left) in a single 2D image. There is a broad range of phenomena for which Fourier analysis of the image plus a few simple rules gives a good account of human perception. The problem with this first-order analysis is that there exists a broad class of 'microbalanced' stimuli in which the motion or slant is completely obvious to human subjects but is invisible to first-order analysis. Microbalanced stimuli require second-order analysis which consists of non-linear preprocessing (spatiotemporal filtering followed by rectification of the input signal) before standard motion or slant analysis. To determine whether the second-order rectification is half-wave or full-wave, we construct two special microbalanced stimulus types: 'half-wave stimuli' whose motion (or texture slant) is interpretable by a half-wave rectifying system but not by full-wave or a first-order (Fourier) analysis and 'full-wave stimuli' which are interpretable only after full-wave rectification. Such experiments show that second-order texture-slant perception utilizes both half-wave and full-wave processes, second-order motion-direction discrimination depends predominantly on full-wave rectification and second-order spatial interactions such as lateral contrast–contrast inhibition and second-order Mach bands are exclusively full-wave.

1994 Higher-order processing in the visual system. Wiley, Chichester (Ciba Foundation Symposium 184) p 287–308

First-order motion and texture

The stimulus domain

A basic motion stimulus, a vertical grey line that moves from left to right in successive frames, is shown in Fig. 1a. The concern here is only with

*Present address: NASA-Ames Research Center, Moffett Field, CA 94035-100, USA.

monochromatic stimuli, so the full description of such a 2D stimulus is the specification of luminance $l(x,y,t)$ at every point in space x,y and time t. A dynamic 2D stimulus is defined within a cube in x,y,t. Under ordinary illumination, vision is largely independent of the absolute level of luminance; it is more useful to describe stimuli in terms of contrast, the local luminance divided by the average luminance l_0 of the display: $c(x,y,t) = [l(x,y,t) - l_0]/l_0$.

Coordinates: motion (x,t) versus texture (x,y)

Because the moving bar of Fig. 1a does not vary in the vertical direction, the vertical coordinate can be omitted. An x,t cross-section of the 2D space–time stimulus, as in Fig. 1b, fully specifies the stimulus. Motion to the right is represented along a diagonal from upper left to lower right. Motion to the left would be represented by the opposite slant (from right to lower left). We can see at once that the motion-direction problem of determining whether the bar moves from left to right or from right to left in x,t is formally equivalent to the texture-slant problem of determining whether a series of squares forms a pattern that slants from upper left to lower right or vice versa.

Receptive fields, Gabor filters

It is easy to solve the texture-slant problem in terms of receptive fields of neurons in the visual cortex—the 'simple cells' of Hubel & Wiesel (1968). The receptive fields of four such cells are shown in Fig. 1c: line detectors (labelled cos) and edge detectors (labelled sine). When a slanted pattern (as in Fig. 1b) is superimposed on such a receptive field, the pattern produces a large response. The response may be either positive or negative, depending on whether the squares are dark or light and on whether they align with + or − regions. Neurons tend not to communicate negative signals well, if at all, so the normal configuration in the nervous system is a neural pair. For example, for line detectors, a plus-centre neuron would signal a positive response for a light line in its centre or a dark line in the adjacent side lobes. A minus-centre line-detector (−cos) has the plus and minus signs reversed and responds to dark lines in its centre (Fig. 1c, bottom). A push–pull pair (Sperling 1970, p. 529*ff.*) of neurons with oppositely signed receptive fields can convey a full range of both signs of contrast and has the advantage of being silent when there is zero input.

A push–pull pair of oppositely signed neurons can be regarded, to a first approximation, as being equivalent to a linear filter that conveys positive and negative values equally well in its output. Such filters for representing visual sensory neurons are often modelled as Gabor functions (Gaussian windowed sine waves, e.g. Olzak & Thomas 1986).

Either sine or cosine filters could be used to distinguish between +45° and −45° slants. To determine whether the line is slanted +45° or −45°, we compared the outputs of filters slanted +45° and −45°. Because the filter output might be either positive or negative, the comparison is between the *magnitude* or square of the outputs, not between the signed outputs.

First-order perception: quadrature pairs, directional energy, standard motion analysis

When the exact location of the line relative to the receptive field is unknown, the problem is a little more complicated. One scheme would be to consider only the +45° and −45° filters with the largest magnitude outputs and compare their outputs. Another scheme might be to consider the total stimulus power at the +45° and −45° slants. A similar problem occurs in the detection of radar signals (Weinstein & Zubakov 1962); the optimum solution is to sum the squared outputs of a sine and a cosine filter. Such a pair of sine and cosine filters is a *quadrature pair* and it computes the total 'energy' of the input in its direction of slant. Looking at all slant directions and finding the quadrature pair with the greatest output is a solution to the problem of finding the dominant slant of a texture (Knutsson & Granlund 1983).

In motion, directional energy is computed in the x,t domain instead of in the x,y domain, as in texture. In the realm of motion, van Santen & Sperling (1985) proved that the three then-current computational theories of motion-direction perception all, in effect, made a comparison between quadrature pairs: (1) the widely accepted Reichardt model (Reichardt 1957, van Santen & Sperling 1984), (2) the motion-energy model (Adelson & Bergen 1985) and (3) a reasonable elaboration of the Watson–Ahumada (1985) motion filter.

There is a broad range of phenomena for which Fourier analysis of the image plus a few simple rules gives a good account of human motion and texture perception (e.g. Watson et al 1986, Caelli 1982). van Santen & Sperling (1984) tested and verified three incisive, counter-intuitive predictions made by the Reichardt model about complex motion stimuli. Thus, a directional energy theory is particularly well established in motion perception. We shall refer to this entire class of theories in various contexts as *standard motion analysis* or as *first-order motion analysis* or as a *directional energy* computation.

Second-order motion and texture

Second-order stimuli

There are two obvious problems with standard motion analysis. The first is exemplified by the stimulus shown in Fig. 1d. This is a contrast-modulated random noise pattern. The random noise is a stationary 'carrier'; only the modulation moves. The x,t motion of this type of stimulus is invisible to standard motion analysis, but we have yet to find an observer who does not perceive its obvious apparent motion. Similarly, the x,t version of the stimulus (Fig. 1d),

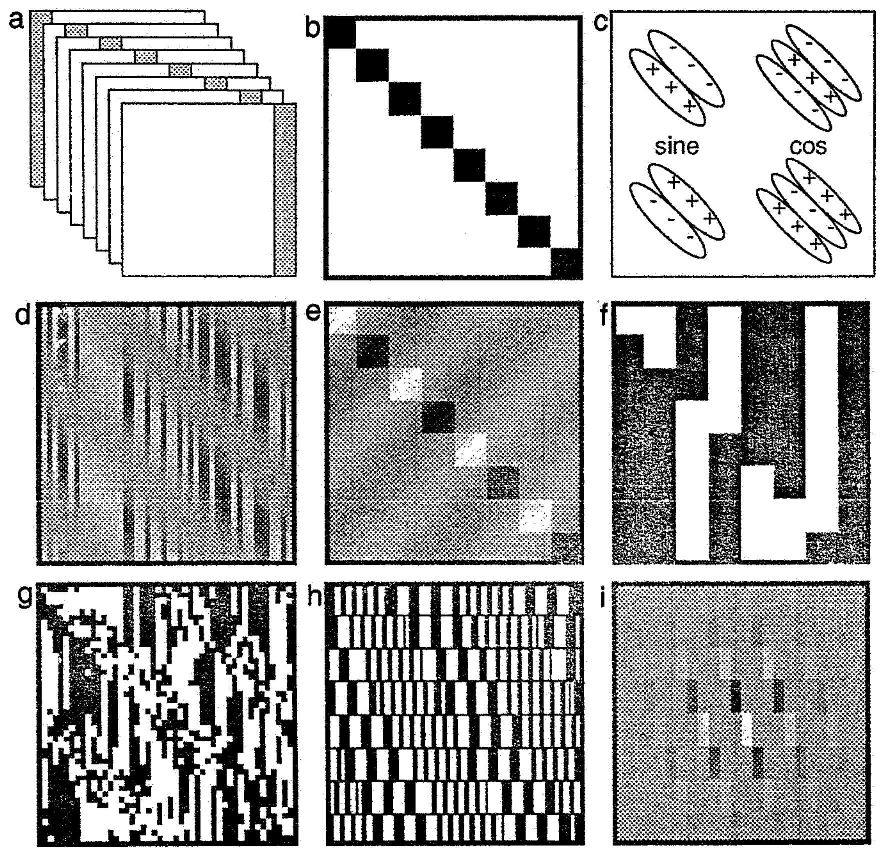


FIG. 1. Motion/texture stimuli. (a) Eight frames of a dark vertical bar moving from left to right. (b) Cross section of (a). The abscissa is horizontal space, x; the ordinate is time, t. Panel b is an x,t representation of a moving bar; equivalently, it is an x,y representation of a slanted line. (c) Four receptive fields of Hubel–Weisel simple cells or, equivalently, linear filters in x,y (or x,t). Areas of positive response are indicated by +, areas of negative response by −. (d–i) x,t representations of moving stimuli or, equivalently, x,y textures. The abscissa is x, the ordinate is t (motion) or y (texture); d,f,g,h are microbalanced. (d) Left–right movement of contrast modulation of static noise. (e) Contrast-alternating moving bar. The primary Fourier motion component (upper right to lower left) is lightly superimposed on the actual pattern. (f) Left–right-moving pixel flip. (g) Left–right-moving rate-modulation of pixel flips. (h) A texture quilt. The eight rows represent eight consecutive frames. (i) Left–right-moving contrast-alternating grating within a Gaussian x,t window. Fourier components slant from upper right to lower left; full-wave components slant from upper left to lower right; half-wave components are ambiguous.

the upper-left to lower-right slant is completely obvious, but it cannot be detected by the Gabor filters or Hubel–Wiesel simple cells shown in Fig. 1c.

The second problem with standard motion analysis is exposed by a stimulus (Fig. 1e) in which the bars in successive frames alternate between black and white (cf. Anstis 1970). Here the slant is obviously from upper left to lower right, but the dominant Fourier component is from upper right to lower left (indicated by slight shading in Fig. 1e). Normally, observers see the x,t (moving) version of the stimulus as moving to the right. However, in peripheral vision or when viewed from far away, the reversed motion is easily seen (Chubb & Sperling 1989).

The above examples indicate that there must be another motion computation in addition to standard motion analysis and another spatial slant computation in addition to a quadrature-based slant analysis. Chubb & Sperling (1988) pointed out that simple rectification (e.g. absolute value or squaring) of local stimulus *contrast* prior to standard motion analysis is a logical candidate for the second computation; they called perception based on such a computation *second-order* perception. Simple rectification would expose the latent x,t motion or x,y slant in the stimuli of Fig. 1d and 1e to standard motion analysis (Fig. 2b versus Fig. 2a). But rectification fails with binary intensity stimuli, such as Fig. 1f and 1g, because it converts them into a completely uniform field.

Drift-balanced and microbalanced stimuli

The stimuli shown in Fig. 1d,f,g,h are all random stimuli that have the property of being drift-balanced and microbalanced (Chubb & Sperling 1988, 1991). *Drift-balanced* means that the expected power in mirror-opposite directions ($+\phi$, $-\phi$) is exactly equal. Summed over the area of a drift-balanced stimulus, a field of Reichardt motion detectors will have zero expected output. Stimuli in Fig. 1d,f,g,h have the further property of being *microbalanced*: that is, they remain driftbalanced when viewed through any space–time separable window. Microbalanced stimuli are particularly useful for neurophysiological investigations where the receptive field (window) of a neuron is not known. The expected output of every individual Reichardt detector is zero for microbalanced stimuli. Nevertheless, no sighted observer has yet failed to see the motion in the stimuli of Fig. 1d,f,g,h.

Texture quilts and texture grabbers

The motion of stimuli of Figs. 1f and 1g might be exposed to standard motion analysis by embellishing simple rectification with temporal differentiation or temporal bandpass filtering (Fig. 2c). The stimulus of Fig. 1h (Chubb & Sperling 1991) provides a counter example to this too-simplistic line of improvement. It is called a *texture quilt* and it produces easily visble apparent motion. In Frame 1,

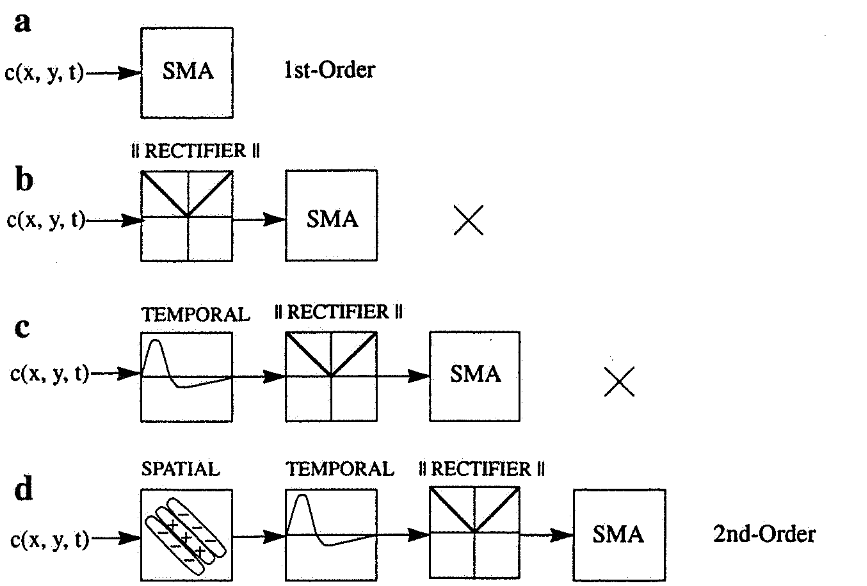


FIG. 2. Successive approximations to a model of motion analysis. SMA, standard motion analysis. (a) The input is contrast as a function of space and time $c(x,y,t)$; the output indicates the direction of movement. (b) A rectifier converts first-order to second-order analysis and extracts latent second-order motion of the stimuli shown in Fig. 1d and 1e, but this scheme fails for binary stimuli (Fig. 1f,g,h). (c) Adding a temporal bandpass filter to (b) enables second-order analysis of the stimuli shown in Fig. 1f and 1g but fails for Fig. 1h. (d) Adding a spatial filter creates a 'texture grabber' which can extract the latent second-order motion in all the stimuli of Fig. 1.

areas of high spatial frequency (A_1) alternate with areas of low spatial frequency (B_1). In each successive frame, the entire *ABABAB* framework shifts rightward 45°, and all the areas A_i and B_i are replaced with new random samples A_{i+1} and B_{i+1} of the same textures. Because they are independent samples, A_i and A_{i+1} are uncorrelated. For the correlation that is needed to extract motion, there must first be recognition that there are two different textures. This is accomplished by a *texture grabber*—a spatial filter with a different response to the *A* and *B* textures plus a rectifier operating on the filter's output (Fig. 2d). The texture grabber parses the stimulus into areas that have more and less of the grabber's preferred texture and therefore its output is amenable to standard motion analysis. Chubb & Sperling (1991) proved the necessity of a texture grabber: no pointwise (intensity or purely temporal) transformation is sufficient to expose the second-order motion of texture quilts.

Sufficiency of full-wave rectification

A basic question is whether rectification in second-order processing is full-wave or half-wave. A full-wave rectifier $F(c)$, is a non-negative, monotonic function of the absolute value of contrast, c. A positive half-wave rectifier, $F^{+}(c)$, equals $F(c)$ for $c \geqslant \epsilon$ and otherwise equals 0 (e.g. Fig. 3b). A negative half-wave rectifier, $F^{+}(c)$, equals $F(c)$ for $c \leqslant \epsilon$ and otherwise equals 0 (e.g. Fig. 3d). Stimulus Fig. 1i (Chubb & Sperling 1989) is a contrast-reversing grating, which is a variant of the contrast-reversing bar of Fig. 1e. The left–right second-order motion of the stimulus of Fig. 1i is completely obvious when seen in central vision, whereas the reverse (first-order) motion can be seen in peripheral vision. The stimulus of Fig. 1i has the property that the black squares by themselves, or the white squares by themselves, are completely uninformative about the direction of motion or of slant. Thus, simple half-wave rectification would be useless in revealing motion or slant to directional energy analysis. Nor do reasonable temporal transformations followed by half-wave rectification succeed. But almost every variant of full-wave rectification exposes the motion of Fig. 1i to directional energy analysis. These observations demonstrate that half-wave rectification is not necessary for second-order motion perception. We now consider the question: is there half-wave motion perception?

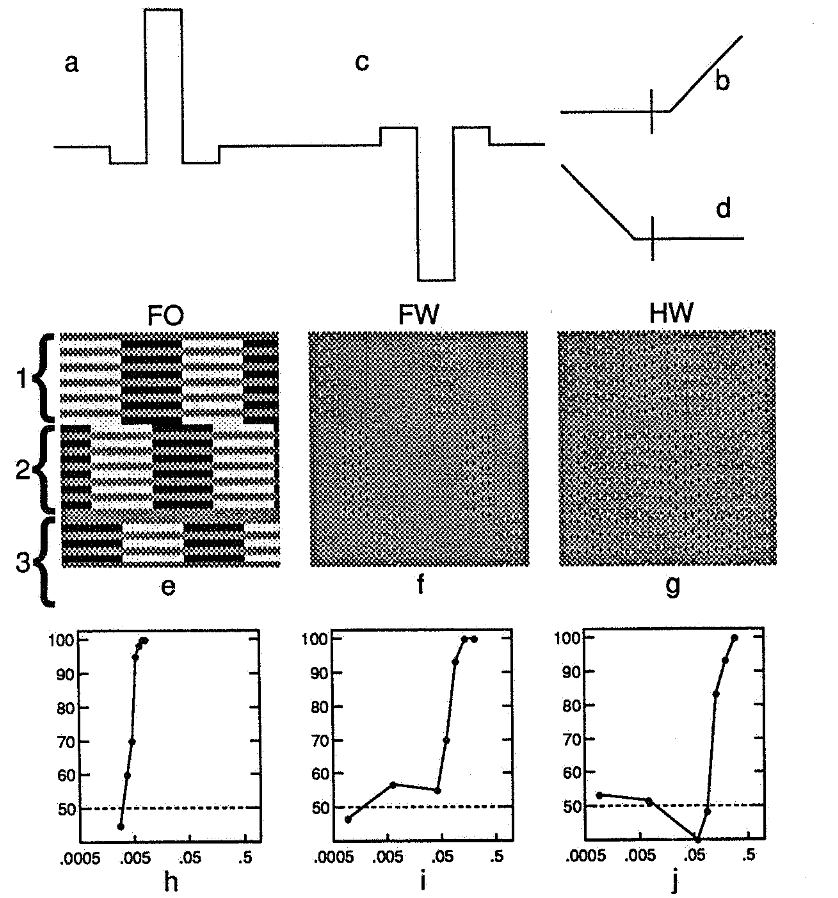
Full-wave and half-wave stimuli

Half-wave hats

Solomon et al's (1993) basic tools for the study of visual half-wave systems are two kinds of half-wave hats, plus hats (Fig. 3a) and minus hats (Fig. 3b). These 2D stimulus elements are called hats because the 3D graph of intensity as a function of space looks like a hat. They are based on the original hat elements of Carlson et al (1980) that were designed to eliminate first-order processes. Plus hats have the property that when passed through a positive half-wave rectifier that has a small threshold (Fig. 3b), they produce a large output; when passed through the symmetric negative half-wave rectifier, their output is zero. It is not necessary for the rectifier to have a threshold to produce this selectivity: a soft threshold non-linearity (such as the square law that is usually assumed for near-threshold processing) works almost as well.

Half-wave stimuli

A half-wave stimulus consists of alternating columns of plus and minus hats (Fig. 3g). Whereas plus hats selectively stimulate positive half-wave detectors and negative hats selectively stimulate negative half-wave detectors, this stimulus would appear as a uniform field to a full-wave detector. In fact, one can go one step further: by varying the amplitude ratio of plus : minus hats, it is possible



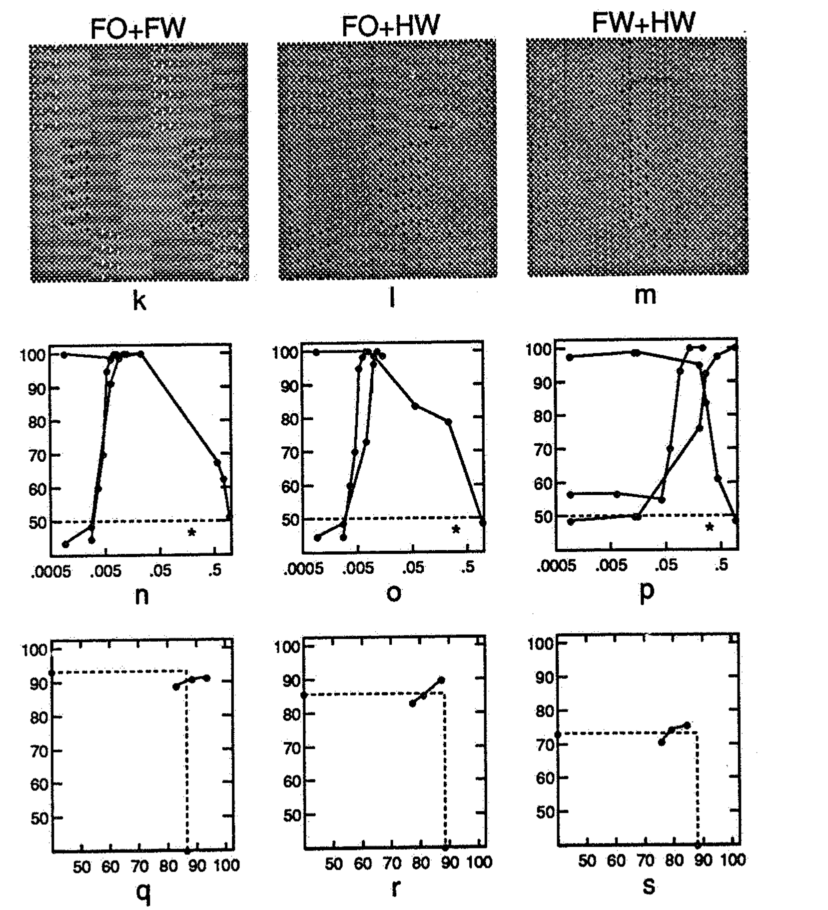


FIG. 3. Fourier, full-wave and half-wave stimuli and data. (a) Graph of contrast versus space $c(x)$ for a *plus hat*. (b) Output versus input for a positive half-wave rectifier with a non-zero threshold (zero is indicated by the vertical bar). The output is zero for minus hats but positive for plus hats. (c) A *minus hat*. (d) A negative half-wave rectifier (with a threshold). (e) The Fourier stimulus (FO). Three frames of a square-wave grating (even rows) on a uniform background (odd rows). Each frame (1,2,3) shows a portion of the image displayed to the subject. Odd- and even-row stimuli alternate rows in successive frames (f) Full-wave stimulus (FW). Contrast-reversing hats move left to right. (g) Half-wave stimulus (HW). A row of alternating-contrast hats moves left to right. (h) Psychometric function (per cent correct motion-direction judgements versus contrast) for the Fourier stimulus, FO. (i) Psychometric function for FW. (j) Psychometric function for HW. (k,l,m) Compound stimuli: even and odd rows contain oppositely directed moving component stimuli as indicated above each panel. (n) Data for one subject viewing FO + FW. The two almost superimposed rising curves (ending in top centre) represent psychometric functions for detecting motion-direction of Fourier stimulus as function of Fourier contrast (abscissa) with full-wave masking stimulus set either to zero (left curve) or to the contrast indicated by the asterisk (right curve). The falling curve indicates the contrast threshold for the full-wave grating (with contrast indicated by asterisk) as a function of Fourier masking contrast. (o) Similar to (n) with Fourier component (contrast = abscissa) plus half-wave component (contrast = asterisk) stimuli. (p) Similar to (n) with full-wave component (contrast = abscissa) plus half-wave component (contrast = asterisk) stimuli. (q) Attention operating characteristic. Ordinate is per cent correct Fourier motion judgements; abscissa is per cent correct full-wave motion judgements in a particular compound stimulus. Performances (averages of three subjects) in three conditions of selective attention (see text) are indicated by points along line. Performance in control conditions (attend to and report only one component) is indicated by points on axes. (r) Attention operating characteristic: Fourier (ordinate) plus half-wave (abscissa). (s) Attention operating characteristic: full-wave (ordinate) plus half-wave (abscissa).

to find a point of minimum visibility of the half-wave stimulus either for motion or for texture, very much the way one might vary the luminance of green and red stripes to find a point of isoluminance. For motion, there is a sharp minimum, different for every observer. The average ratio is 1.17 : 1 indicating that, on average, minus hats are somewhat more effective at stimulating the full-wave system than plus hats, but that they can be perfectly balanced. Once the hats are full-wave balanced, the half-wave stimulus should appear completely uniform to the full-wave system. Of course, the luminances of the hats are physically balanced for invisibility to the first-order (Fourier) analysis at the frequencies of the to-be-detected stripes, so the half-wave stimulus also is neutral to the first-order system.

Every observer easily sees the stripes in a static half-wave stimulus, indicating that there is a half-wave texture system. Indeed, half-wave theories of spatial vision have been proposed by several of the participants in this symposium (e.g. Watt & Morgan 1985, Graham 1994, this volume, Malik & Rosenholtz 1994, this volume). But only ⅓ of Solomon & Sperling's (1994) observers was able to detect the direction of motion of half-wave stimuli. In a few instances, deficient subjects were given extensive practice; it did not seem to improve their half-wave performance. This suggests that for those who do not already possess it, the ability to detect half-wave motion is not easily acquired. The psychometric function (percent correct as function of half-wave contrast) for a subject who has excellent ability to discriminate half-wave motion is shown in Fig. 3j.

Full-wave stimulus

The full-wave analogue to the stimulus of Fig. 1i is shown in Fig. 3f; it becomes ambiguous after half-wave rectification. Because it is composed of hats, it is also invisible to first-order directional energy analysis at the frequencies of the stripes. Contrast threshold for direction discrimination of the full-wave stimulus is about 0.05 (Fig. 3i), almost an order of magnitude less than the contrast threshold for the half-wave stimulus.

Fourier stimulus

The Fourier stimulus is a square-wave grating of the same spatial frequency as the others. Square-wave contrast threshold for direction discrimination is about 0.005, an order of magnitude less than for the full-wave stimulus.

Relative efficiency computation

Most of the sensitivity difference between the Fourier (first-order) stimulus and the second-order stimuli rests in the stimuli themselves. If we assume a square-law contrast rectifier (i.e. a power computation), we find that there is much more stimulus energy in the Fourier stimulus than the others. Relative to the

Fourier stimulus, at threshold, the full-wave and half-wave stimuli, respectively, require 1.9 and 33.2 more contrast power, yielding a full-wave relative power efficiency of 0.52% and a half-wave efficiency of 0.03%. However, if an absolute value rectifier is assumed, then, for all subjects, the efficiency of discriminating full-wave stimuli is actually greater than that of Fourier stimuli (Solomon & Sperling 1994).

Once the stimulus contrast power is taken into account, second-order full-wave motion discrimination is almost as or perhaps more efficient than first-order. But, half-wave motion discrimination lags behind by more than a factor of 10, even for the subset of the population especially chosen for their ability to make this discrimination. The conclusions are that second-order full-wave computation ranks with first-order Fourier motion as a major computation that has been sufficiently important in evolution to have evolved high efficiency. Half-wave motion perception, even for the few who possess it, is a relatively weak process.

In the kinetic depth effect, Dosher et al (1989) showed it was the 2D *motion flowfield* that gave rise to a 3D perceptual structure. However, only first-order motion flowfields transmit the level of resolution required to make 3D shape discriminations (Sperling et al 1989). Although direction discrimination of full-wave motion is an efficient process, the full-wave motion computation is too coarse to be useful in the kinetic depth effect.

Concurrent motion-discrimination tasks

Between-systems transparency

Superimposing two sine wave motion signals of the same spatial and temporal frequency travelling in opposite directions produces a 'counterphase flickering' grating in which no motion is perceived. When the spatial frequencies differ sufficiently (e.g. by more than an octave), superposition fails to produce cancellation of apparent motion—instead, motion transparency is perceived. When first- and second-order systems are involved, the situation is similar to the case of different spatial frequencies: oppositely directed equal-strength stimuli produce transparency, not cancellation.

Solomon & Sperling (1994) systematically studied interactions between superimposed, independently moving, first- and second-order stimuli. A full-wave motion stimulus of contrast 0.16 was presented together with a Fourier square wave that varied in contrast from 0.001 to 0.93 on successive trials. The subject was asked the direction of either the first-order stimulus or the second-order stimulus. The two rising curves in Fig. 3n represent the accuracy of first-order motion-direction judgements as a function of Fourier contrast, with and without the full-wave masking stimulus. The falling curve represents the accuracy of motion-direction judgements of the second-order stimulus.

Two important facts are apparent from the data of Fig. 3n: (1) there is a significant range of contrasts where the subject can correctly report the direction of either stimulus (i.e. there is motion transparency); (2) there is relatively little cross-masking (the two rising curves nearly overlap). The low level of cross-masking is important because it rules out the possibility of detecting second-order motion by spillover into the first-order domain or vice versa—any spillover components would be rendered useless by the large, already present 'other' stimulus. Similarly, the motion transparency also indicates that there must be separate computational systems.

Figure 3o shows results when a fixed-contrast half-wave stimulus is masked by Fourier square-waves of various contrasts. The data are similar to those of Fig. 3n and, again, the conclusion is that half-wave and Fourier stimuli are analysed in largely independent channels.

Figure 3p shows that when a constant half-wave stimulus is masked by various full-wave stimuli, there is no significant region in which both stimuli can be judged independently. The subjects' problem is evident in the subset of stimuli in which full-wave and half-wave stimuli travel in opposite directions. The appearance is one of motion transparency, but subjects are unable to link the directions of motion with the appropriate full-wave and half-wave components.

Attention operating characteristics

Solomon & Sperling (1994) selected a pair of approximately equal-strength, maximally visible components from each of the three pairs of superimposed stimuli (Fourier/full-wave, Fourier/half-wave, full-wave/half-wave). The choices were appropriate for each subject. As before, the subjects were presented stimuli composed of two components moving in randomly chosen directions. Now, subjects were required to report both directions—in effect, to perform two concurrent tasks. In various conditions, they were instructed to give primary attention either to one component, to the other, or to divide attention equally between both. In control conditions, subjects were shown only one component.

Task independence and resources. In general, performance on concurrent tasks reflects the performance on the joint tasks (in this case, accuracy of motion-direction judgements) as a function of the attentional condition. These dual-response data form an attention operating characteristic (Figs 3q,r,s). There is a small improvement when attending to the more difficult stimulus, but there is no selective improvement for the attended stimulus. If the data lay on a diagonal connecting the two control points, detection of each type of stimulus would depend on the same processing resource. However, the data for the three stimulus pairs all lie near the independence point—the intersection of perpendiculars to the axes drawn through the control data points (dotted lines). This indicates that totally separate resources are involved (Sperling & Dosher 1986), in this case, independent motion computations. The conclusion is that Solomon & Sperling's (1994) subjects could perform two independent motion computations. The limiting resource could not be consciously controlled; that is, at least one of the motion computations in each pair of stimuli was automatic and did not require attentional resources.

The full-wave and half-wave stimuli obviously did involve a capacity limit because subjects were unable to judge both motions in equal-strength stimuli. However, giving full attention to one or the other motion component, or attending both equally, all produced equal levels of performance. This means that allocation of the limiting resource in this process was not under conscious control.

Spatial interactions

Contrast–contrast

In a well known brightness illusion, a patch of grey looks brighter when it is surrounded by black (Fig. 4a) than when surrounded by white (Fig. 4b). This is a first-order illusion. The second-order analogue (Chubb et al 1989) is a contrast–contrast illusion in which a patch of random texture surrounded by a high contrast texture (Fig. 4d) looks less contrasty than a similar patch surrounded by 'texture' of zero contrast (Fig. 4c). Although this second-order illusion works in a static illustration, it is even stronger in dynamic displays, e.g. when new random samples of the textures appear every 1/8th second. The contrast–contrast illusion is specific to spatial frequency. When the spatial frequency of the centre and the surround differ by an octave (Figs 4e and 4g) contrast–contrast inhibition is much less than when the spatial frequencies are the same (Fig. 4f). Orientation also matters at frequencies greater than about 3 cycles/degree so that parallel gratings in the centre and surround produce greater contrast–contrast then do perpendicular gratings (Figs 4h and 4i) (Solomon et al 1993, Cannon & Fullencamp 1991).

When the contrast textures are produced by hats rather than by gratings or visual noise, the magnitude of the illusion is completely indifferent to the signs of the hats in the centre and the surrounds (Solomon et al 1993). Plus hat surrounds produce just as much contrast inhibition on minus hat centres as do minus hat surrounds, and vice versa (Figs 4j,k,l,m). That the signs of the inhibiting and inhibited hats are irrelevant means that lateral contrast inhibition is a full-wave process.

Second-order Mach bands and Craik–O'Brien–Cornsweet illusion

Mach bands are illusory brightness enhancements and darkness enhancements (Fig. 5b) at the points where an intensity ramp joins a uniform brightness

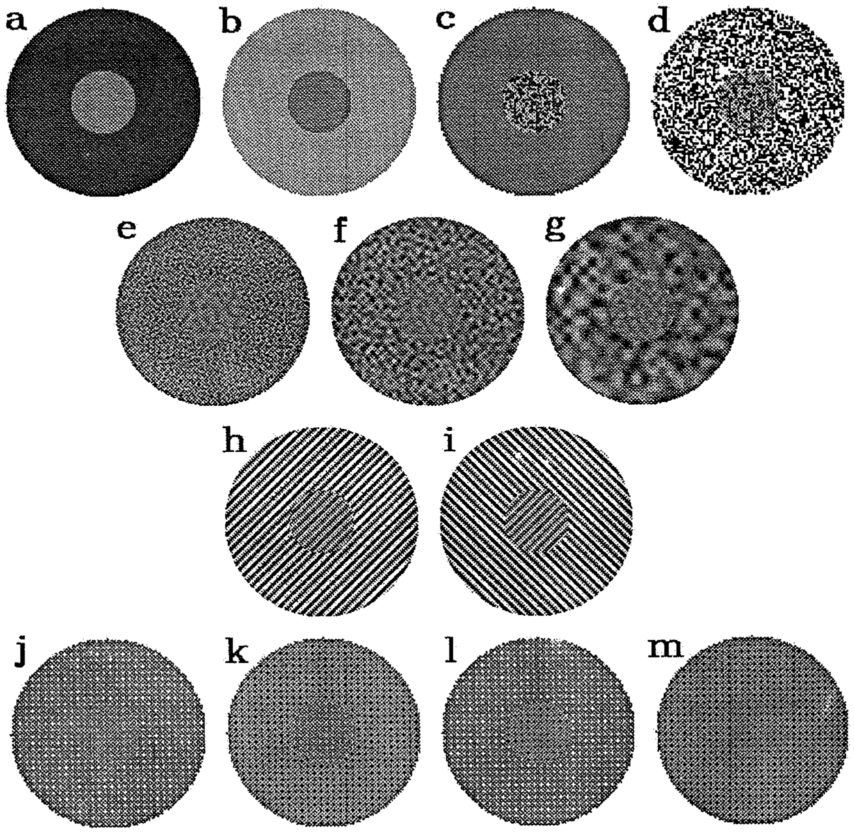


FIG. 4. Spatial interactions. (a,b) Luminance contrast. The central disk has the same contrast in (a) and (b). (c,d) Contrast–contrast. The texture in the central disk has the same contrast in (c) and (d). (e,f,g) The texture in the central disk has the same contrast in each of (e,f,g) and so does the surrounding annulus but, relative to (f), the annular spatial frequency is an octave higher in (e) and an octave lower in (g). (h,i) The effect of the orientation of the grating in the annular surround on the perceived contrast of the grating in the central disk. (j–m) The effect of plus versus minus hats in producing spatial interactions. Centre/surround: (j) +/+; (k) +/−; (l) −/+; (m) −/−.

plateau (Fig. 5a). Mach bands usually are taken to be indicative of the same lateral inhibitory interactions that produced the lightness (Fig. 4a,b) and the contrast–contrast illusions (Fig. 4c,d). Lu & Sperling (1993) replaced the first-order luminance stimulus (Fig. 5a) with second-order contrast modulation to produce a second-order ramp (contrast modulation of a random noise carrier). Psychometric measurements showed that the induced second-order (contrast) Mach bands were equal in magnitude to the first-order bands induced by a luminance ramp.

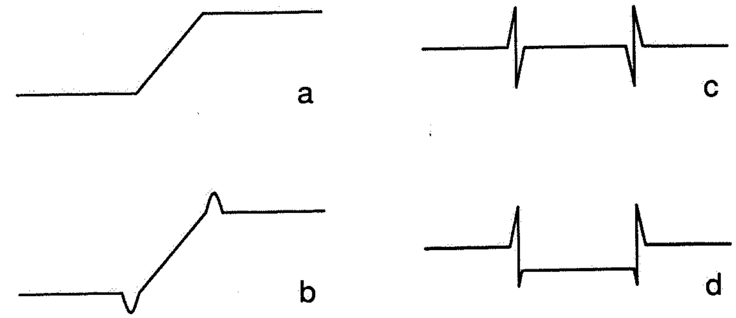


FIG. 5. (a) Spatial luminance ramp for first-order Mach band or spatial contrast ramp for second-order Mach band. (b) Perceived brightness (or contrast) as a function of space for luminance (first-order) or texture-contrast (second-order) stimuli (a). (c) Spatial pattern for the Craik–O'Brien–Cornsweet illusion. (d) The illusion: the perceived brightness (or contrast) of pattern (c).

In the Craik–O'Brien–Cornsweet illusion, the mean brightness of a disk surrounded by two thin rings having higher and lower brightness, respectively, (Fig. 5c) causes the interior patch to appear to have overall lower brightness (Fig. 5d). Again, Lu & Sperling (1993) found the corresponding second-order Craik–O'Brien–Cornsweet illusion to be approximately equal in magnitude to the first-order illusion.

Both the second-order Mach bands and the second-order Craik–O'Brien–Cornsweet illusions can be produced with textures composed of random textures composed of hats as well as of random noise. For hat stimuli, the illusions are indifferent to the signs of the hats. Half-wave versions of the stimuli in which the illusion depended on the sign (instead of the amplitude) of the hats caused the illusions to fail completely. This shows, not surprisingly, like the lateral contrast–contrast inhibition (Fig. 3), that the lateral interactions responsible for second-order Mach bands and the Craik–O'Brien–Cornsweet illusions are entirely full-wave interactions.

Conclusion

Second-order processes are ubiquitous in perception and, after a strong initial non-linearity (full-wave rectification), seem to follow the same principles as first-order interactions.

Acknowledgment

This work was supported by the US Air Force Office of Scientific Research, Visual Information Processing Program, Grant 91-0178.

References

Adelson EH, Bergen JR 1985 Spatio-temporal energy models for the perception of apparent motion. Opt Soc Am A Opt Image Sci 2:284–299
Anstis SM 1970 Phi movement as a subtraction process. Vision Res 10:1411–1430
Caelli T 1982 On discriminating visual textures and images. Percept Psychophys 31:149–159
Cannon MW, Fullencamp SC 1991 Spatial interactions in apparent contrast: inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. Vision Res 31:1985–1998
Carlson CR, Anderson CH, Moeller JR 1980 Visual illusions without low spatial frequencies Invest Ophthalmol & Visual Sci 19:165–166
Chubb C, Sperling G 1988 Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. J Opt Soc Am A Opt Image Sci 5:1986–2006
Chubb C, Sperling G 1989 Two motion perception mechanisms revealed by distance driven reversal of apparent motion. Proc Natl Acad Sci USA 86:2985–2989
Chubb C, Sperling G 1991 Texture quilts: basis tools for studying motion-from-texture. J Math Psychol 35:411–442
Chubb C, Sperling G, Solomon JA 1989 Texture interactions determine perceived contrast. Proc Natl Acad Sci USA 86:9631–9635
Dosher BA, Landy MS, Sperling G 1989 Kinetic depth effect and optic flow. 1. 3D shape from Fourier motion. Vision Res 29:1789–1813
Graham N 1994 Non-linearities in texture segregation. In: Higher-order processing in the visual system. Wiley, Chichester (Ciba Found Symp 184) p 309–329
Hubel DH, Weisel TN 1968 Receptive fields and functional architecture of monkey striate cortex. J Physiol 195:215–243
Knutsson H, Granlund GH 1983 Texture analysis using two-dimensional quadrature filters. In: IEEE Computer Society workshop on computer architecture for pattern analysis and image database management—CAPAIDM. IEEE (Inst Electr Electron Eng) Computer Society, Silver Spring, MD, p 206–213
Lu Z-L, Sperling G 1993 2nd-order illusions: Mach bands, Craik–O'Brien–Cornsweet. Invest Ophthalmol & Visual Sci (ARVO suppl) 34:1289(abstr)
Malik J, Rosenholtz R 1994 A computational model for shape from texture. In: Higher-order processing in the visual system. Wiley, Chichester (Ciba Found Symp 184) p 272–286
Olzak LA, Thomas J 1986 Seeing spatial patterns. In Boff K, Kaufman L, Thomas L (eds) Handbook of perception and performance. Wiley, New York, vol 1:7-1–7-56
Reichardt W 1957 Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems. Zeitschrift Naturforsch 12b:447–457
Solomon JA, Sperling G 1994 Full-wave and half-wave rectification in 2nd-order motion perception. Vision Res 34, in press
Solomon JA, Sperling G, Chubb C 1993 The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation but specific to orientation. Vision Res 33:2671–2683
Sperling G 1970 Binocular vision: a physical and a neural theory. Am J Psychol 83:461–534
Sperling G, Dosher BA 1986 Strategy and optimization in human information processing. In: Boff K, Kaufman L, Thomas J (eds) Handbook of perception and performance. Wiley, New York, vol 1:2-1–2-65
Sperling G, Landy MS, Dosher BA, Perkins ME 1989 The kinetic depth effect and the identification of shape. J Exp Psychol Hum Percept Perform 15:426–440

van Santen JPH, Sperling G 1984 Temporal covariance model of human motion perception. J Opt Soc Am A Opt Image Sci 1:451–473
van Santen JPH, Sperling G 1985 Elaborated Reichardt detectors. J Opt Soc Am A Opt Image Sci 2:300–321
Watson AB, Ahumada AJ 1985 Model of human visual-motion sensing. J Opt Soc Am A Opt Image Sci 1:322–342
Watson AB, Ahumada AJ, Farrell JE 1986 Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. J Opt Soc Am A Opt Image Sci 3:300–307
Watt RJ, Morgan MJ 1985 A theory of the primitive spatial code in human vision. Vision Res 25:1661–1674
Weinstein LA, Zubakov VD 1962 Extraction of signals from noise (translation from Russian by R. A. Silverman). Prentice Hall, London (reprinted by Dover Publications, NY, 1970)

DISCUSSION

Morgan: It's critical for your argument against the half-wave mechanisms to suppose that one sign of 'hat' stimulates uniquely one of the two mechanisms (ON or OFF). Why can't black hats stimulate both ON and OFF channels?

Sperling: Indeed, they might stimulate both.

Morgan: That would give an appearance of full-wave rectification.

Sperling: That would lead to the half-wave stimuli appearing to be homogeneous. But that is not what we observe. In all half-wave stimuli, for all observers, the alternating black-hat and white-hat stripes are completely obvious. For those observers who do see half-wave motion, the motion system as well as the texture system is sensitive to the difference between the white and black hats.

Wilson: Could you argue that the small effects in motion that you get with half-wave rectified stimuli are due to a slight imperfection or asymmetry in the visual apparatus for full-wave rectification?

Sperling: You could argue that, but not successfully. For each observer, we titrated the half-wave black and white hats so that they were perfectly balanced with respect to the full-wave system, and we still got good apparent motion. Furthermore, suppose the visibility of half-wave motion was a result of bleeding over into the full-wave system. Then we would expect that, in the full-wave masking condition, you would be unable to see half-wave motion, because the full-wave mask and the half-wave stimulus would interact to produce only one direction. But our subjects can discriminate the direction of half-wave motion even with masking levels that are orders of magnitude greater than any distortion product would be.

Heeger: Why are these our only options?

Sperling: These are data—you can have any options you like!

Heeger: It seems like you picked two sorts of convenient non-linearities to consider (half-wave and full-wave rectification), but there are many other possibilities.

Graham: There's squaring and half-squaring. But the experiments reported here don't distinguish between linear rectification and squaring.

Sperling: If you are dealing with linear theories, you can resolve the whole issue with a finite number of measurements and you know exactly what they should be. When you have to deal with non-linear theory, you can never resolve such issues with a finite number of measurements unless you state the bounds of the theory. All I can do here is to begin to impose some constraints on the possible theories.

Heeger: Your conclusions are based on particular assumptions. If we made different assumptions (different non-linearities), might we not reach different conclusions?

Sperling: I agree with you, in the sense that one might think of other mechanisms. What I'm working out are the consequences of a particular set of assumptions. Suppose we had a half-wave system and suppose it responded to these stimuli in a certain way. What would we expect?

It turns out the half-wave stimuli are particularly interesting because, in texture, they yield perfectly clear stripes to everyone. In our motion task, these same stripes were useless for more than half our observers. These are facts that any theory will have to explain.

Morgan: It is very striking that some stimuli which are completely visible to static vision through texture are quite invisible to the motion system. What could be the neural basis for this? Is this a fundamental difference between M and P pathways? Is it just the M pathway which doesn't have the half-wave rectified signal available to work upon?

Movshon: From the physiology, bits that plausibly might be in both pathways seem to have all the signals you would want, so the question concerns why they can't be used. If any of these constructions are even vaguely right you can equate half-rectified things with simple cells and full-rectified things with complex or complex-like cells.

Morgan: Have we any animal data? Is there any evidence that cats can see non-Fourier motion?

Movshon: Monkey MT cells can see it (O'Keefe et al 1993).

Bergen: In order to perform a motion computation (in particular, one that involves discriminating the direction of motion), more is required than that whatever component of the stimulus is moving simply be detected. In order to compute direction rather than just change, the dynamic characteristics of the response must be appropriate. I know of no reason to assume that the dynamic characteristics of whatever mechanism is seeing some of the stimuli in question are *not* sufficient to support analysis of direction of motion but, logically, there is a gap between the observation that psychophysically we can see the stimuli as static textures and the expectation that we will be able to discriminate their direction of motion. In other words, there is no *logical* inconsistency in the fact that we can see one and not the other.

Graham: And to be able to do the directional computation, the dynamics have to be right.

Bergen: I don't have any good reason to know why they shouldn't be right, that's the problem.

Spekreijse: I am puzzled by the difference in efficiency of the non-Fourier versus the Fourier pathway, because the only thing you add to the Fourier pathway is rectification. If that slope is 1, I would have expected the same efficiency for both pathways. Why is there a factor of two difference and why is the threshold 0.5% for the Fourier and 5% for the full-wave rectifier?

Sperling: If we assume a square-law rectifier, we find that the full-wave system works with about 50% of the efficiency of the Fourier system; assuming a linear rectifier yields full-wave efficiency that is actually greater than Fourier efficiency. Obviously, there is no problem with the full-wave motion perception. But for every subject, the efficiency of the half-wave motion discrimination is reduced by a factor of at least 10 from full-wave. That's a guaranteed efficiency loss of an order of magnitude even for the best half-wave perceiving subjects. It's the half-wave motion system that's inefficient.

Derrington: Is the large difference in threshold contrast just because the full-wave stimulus occupies fewer pixels?

Chubb: Yes. In fact, only one quarter of the cycle of the square wave that makes up the full-wave stimulus contains texture elements whose contrast differs from zero. On the other hand, the entire cycle of the square wave composing the Fourier stimulus contains texture elements of non-zero contrast. Thus, the Fourier stimulus contains four times as many informative texture elements as the full-wave stimulus.

Spekreijse: You can't compare those numbers directly.

Movshon: The half-wave stimulus should be visible to either a mechanism consisting of only the right half-sign rectifiers or to a full-wave rectifier.

Chubb: You get the same output image if you apply a full-wave rectifier to the contrast modulation function of a negative hat as when you apply it to a positive hat. Thus, for instance, taking the absolute value of stimulus contrast yields the same output image when applied to a field of positive hats as it does when applied to a field of negative hats. However, the half-wave stimulus is a square wave that alternates between fields of positive and negative hats. Therefore, when a full-wave rectifier is applied to the half-wave stimulus, the result is a stimulus that comprises a field of uniform texture. This output stimulus carries no motion information.

Shapley: This box that's called 'standard motion analysis' is assumed to be the same for all of these pathways. Is there any evidence for this? Are there experiments that could test whether one box could do standard motion analysis for both pathways? Some of the results that Andrew Derrington reported (Derrington & Henning 1994, this volume) indicated that there might be dynamic differences between non-Fourier and Fourier motion—maybe there are two different standard motion analyses. Maybe that's not the case and there's one set of motion analysis and there's some filter in front of the non-Fourier pathway.

Sperling: The parameters of motion analysis are different; for some of these stimuli the limiting temporal frequencies are much lower. Although it has been claimed, I don't think the parameters of full-wave motion are necessarily different from those of Fourier motion. But certainly half-wave motion perception is a much slower system. Subjects can't do anything with half-wave stimuli comparable to what they can do with full-wave stimuli.

Georgeson: If you slow the stimulus down, might the performance come back?

Sperling: At some point you are no longer dealing with the motion system. We wanted to be sure that we had a stimulus that was a motion signal and subjects weren't tracking. So these displays were brief displays that went by relatively quickly—they were over in a third of a second.

Chubb: For the short-range motion system, there are a few signature stimuli there that indicate very clearly what is in the standard motion analysis box. A particularly dramatic stimulus (Chubb & Sperling 1989) comprises a small (e.g. 1°) square that appears in an otherwise grey visual field and proceeds to step its own width around five times, rapidly to the right (say, at the rate of 60 steps/second), reversing its contrast with each step. For instance, the square may be white when it first appears, but then changes to black when it next appears (shifted directly to the right of its first location) and continues to alternate between black and white with each successive step. What is interesting about this stimulus is that the bulk of its directional energy is in Fourier components (i.e. drifting sinusoidal gratings) that move in the direction opposite to that of the steps taken by the square itself. That is, the rightward-stepping, contrast-reversing square has almost all of its directional energy in leftward-moving sinusoidal gratings. When you 'foveate' this stimulus, you see decisive motion to the right (in the direction of the steps taken by the square). However, when you view this stimulus either peripherally or from very far away, you see instead an unambiguous jolt of motion to the left. The behaviour of the foveally viewed stimulus can be explained in terms of a second-order motion mechanism that applies standard motion analysis to some rectified transformation of stimulus contrast. For instance, when you take the absolute value of stimulus contrast, the result is a square of uniformly high intensity rapidly stepping across a background field of lower value. The rightward motion of this rectified stimulus is clearly signalled by its directional Fourier energy.

What is more interesting in this current context, however, is the behaviour of the square when it is viewed peripherally. When you look off to the side, you isolate the first-order motion mechanism on the stimulus; under these

conditions you really do see motion in the opposite direction, in spite of the fact that there are literally no positive correlations of non-zero contrasts in that direction. It's a counter-intuitive stimulus—you don't want your motion mechanism to give you this response; you want it to detect the vertical steps of the square. This leads one to think that if the high-order box that's doing the standard motion analysis for non-Fourier sorts of mechanisms is really using the same kind of computation, maybe we can produce some stimuli that will betray this fact by yielding similarly counter-intuitive results.

One thing you can try is to generate an analogue to the contrast-reversing square using, let's say, the amplitude of noise instead of stimulus luminance. Thus, rather than having a square that alternates between white and black on a grey background, you use a square that alternates between maximum contrast noise and uniform grey as it steps across a field of medium contrast noise. If we simply full-wave rectify the contrast of this stimulus, we obtain a spatiotemporal output function akin to the original contrast-reversing square. If the standard motion analysis box into which this output function is fed really uses the same computation as the first-order motion system, then we should expect the noise amplitude analogue to the contrast-reversing square to elicit leftward motion (contrary to the steps it takes). I've fiddled with this and other stimuli and haven't been able to obtain the counter-intuitive motion that I was looking for. It's very hard to get the 'standard motion analysis' box at the far end of the non-Fourier motion mechanism to reveal itself as a Fourier energy analyser.

Heeger: In Hugh Wilson's model (1994, this volume), the Fourier and the non-Fourier signals are recombined. According to Albright (1992) and recent results from Tony Movshon's lab (unpublished), at least some MT cells respond to both Fourier and non-Fourier signals. But George Sperling has been telling us that, in fact, these signals don't recombine, and that you can make independent judgements about Fourier and non-Fourier mechanisms superimposed. Where are we here? Does that mean that the non-Fourier mechanisms that you're looking at, George, are distinct from MT? Does that mean that for some stimuli the signals recombine but for other stimuli they don't?

Wilson: One possible resolution is that if you consider the spatial extent of these stimuli, it's conceivable that Fourier and non-Fourier signals are normally combined. However, some receptive fields are larger than others, and with these 'unusual' stimuli, the bandpass filtering of the black and white hats might sometimes produce motion signals on very different spatial scales. Then one might see two distinct motions at the same point. This feature hasn't been incorporated into any of the models; issues such as receptive field size for pooling of various signals.

Sperling: An answer to Dave Heeger's question is that our procedure concentrates on those signals that are maintained as separate. Neurons that combine Fourier and non-Fourier motion signals, whether they be in MT or elsewhere, cannot serve the simultaneous judgement of two independent motions. Two relatively independent sets of neurons are required. In response to Hugh Wilson, our stimuli (Fourier, full-wave and half-wave) are designed to be of the same spatial frequency at the level of motion analysis. They would produce counterphase flicker to any system that combined them. What distinguishes our Fourier and non-Fourier motions is the carrier (luminance versus beats), not the modulation frequency, which is the same.

References

Albright TD 1992 Form-cue invariant motion processing in primate visual cortex. Science 255:1141–1143

Chubb C, Sperling G 1989 Two motion perception mechanisms revealed by distance driven reversal of apparent motion. Proc Natl Acad Sci USA 86:2985–2989

Derrington AM, Henning GB 1994 Implications of motion detection for early non-linearities. In: Higher-order processing in the visual system. Wiley, Chichester (Ciba Found Symp 184) p 211–226

O'Keefe LP, Carandini M, Beusmans JMH, Movshon JA 1993 MT neuronal responses to 1st- and 2nd-order motion. Soc Neurosci Abstr 19:1283

Wilson HR 1994 Role of second-order motion signals in coherence and transparency. In: Higher-order processing in the visual system. Wiley, Chichester (Ciba Found Symp 184) p 227–244