

Convergent evidence for the visual analysis of optic flow through anisotropic attenuation of high spatial frequencies

Horace B. Barlow

Physiological Laboratory, University of Cambridge
Cambridge, UK



Bruno A. Olshausen

Center for Neuroscience & Department of Neurobiology,
Physiology, and Behavior, UC Davis, Davis, CA, &
Redwood Neuroscience Institute, Menlo Park, CA



Photoreceptors strongly attenuate high *temporal* frequencies. Hence when an image moves, high *spatial* frequency components are lost if their direction of modulation coincides with the direction of movement, but not if it is orthogonal. The power spectra of natural images are remarkably consistent in having a $1/f^2$ falloff in power in all directions. For moving images, the spatial power spectra will be distorted by becoming steeper in the direction corresponding to modulation in the direction of motion, and the contours of equal power will tend to become elliptical. This study demonstrates that the mammalian visual system is specifically sensitive to such anisotropic changes of the local power spectrum, and it is suggested that these distortions are used to determine patterns of optic flow. Convergent evidence from work on Glass figures, motion streaks, and sensitivity to non-Cartesian gratings is called on in support of this interpretation, which has been foreshadowed in several recent publications.

Keywords: glass patterns, motion, motion blur, optic flow, power spectrum, phase spectrum

Introduction

It has been suggested by Geisler (1999) and Burr (2000) that the “motion streaks” that result from movement and the “speed-lines” that artists use to depict motion are related to the curious streakiness or “flow” seen in the moiré patterns of overlaid random dots described by Glass (1969). This work follows up these suggestions by analyzing the distortions of the power spectrum that occur when the image of a natural scene is moved. Because the visual system responds poorly to high temporal frequencies, movement introduces characteristic anisotropies; therefore, it is proposed that the streaky flow seen in Glass patterns results from activation of mechanisms that normally detect the distortions of the local power spectrum caused by motion. It is shown that the streaks can be abolished by adding noise that evens out the power spectrum of a Glass pattern, or opposite-polarity dot pairs that cancel out the autocorrelation of same-polarity pairs. It is also shown that the orientation or direction of flow of the streaks is determined by the local power spectrum, rather than the phase component of their local Fourier transform. Finally, it is pointed out that the known properties of neurons in V1 would produce approximate representations of the local power spectrum (Hubel & Wiesel, 1962; Campbell & Robson, 1968; Maffei & Fiorentini, 1973; Movshon, Thompson, & Tolhurst, 1978), and that neurons in V4 have been described that would detect the types of pattern that result from optic flow (Gallant, Braun, & Van Essen, 1993; Gal-

lant, Connor, Rackshit, Lewis, & Van Essen, 1996). Wilson and Wilkinson (1998) proposed models of the mechanism underlying the detection of Glass patterns that would also respond to the patterns resulting from motion blur, and according to our hypothesis, this is their main functional role.

To examine this argument, we must first consider how the effective power spectrum of a natural image changes when it moves over the retina. We then discuss the autocorrelation functions and power spectra of Glass patterns, showing how these share characteristics with the power spectra of natural images blurred by movement. We also show that modifying these characteristics modifies the visibility of the patterns. Finally, we discuss the neural mechanisms that may bring about this analysis.

The Fourier transform and power spectrum of natural images

The discrete Fourier transform of a picture contains all the information of the original, but instead of being specified by the intensities of, say, 512 x 512 pixels in an image, it is specified by an equal number of sine and cosine coefficients in a plane defined by spatial frequencies periodic in the *x* and *y* directions. The full transform has two parts, the power spectrum formed by squaring and summing the sine and cosine coefficients at each locus in the *X,Y* frequency plane, and the phase spectrum, which is \tan^{-1} of the

ratio of each pair of coefficients. Neither of these alone contains all the information in the Fourier transform and both are required to reconstruct the original picture accurately, but the phase spectrum is more critical for the overall visual appearance of a reconstituted image (Oppenheim & Lim, 1981; Piotrowski & Campbell, 1982), though there are exceptions (Tadmor & Tolhurst, 1993). By contrast, we shall show that the local power spectrum is important, and the phase spectrum unimportant, for the image characteristics we are concerned with here.

The effect of motion on the power spectrum

The top row of Figure 1 shows an image of natural scenery and its power spectrum. Two rings have been superimposed on the power spectrum; if the original picture subtends an angle of 10^0 at the observer's eye, the inner ring lies at a spatial frequency of 2 cycles/deg at all orientations, whereas the outer ring is at 10 cycles/deg; these two frequencies straddle the peak of the spatial contrast sensitivity curve of humans, and it is the region between these two rings that is most important for the present problem.

The lower row of Figure 1 is derived from the top row and shows (center and right) the power spectrum of the image that would, when stationary, produce the same excitatory effects on the visual system at all spatial frequencies

and orientations as the original image would when moving diagonally at 2 deg/s. The lower left figure shows the image reconstituted by computing the inverse Fourier transform of the amplitude-modified spectrum (phase spectrum was left unaltered).

For several reasons, it is unexpectedly difficult to derive this. First, the effects of spatial and temporal modulation of contrast on the visibility of sinusoidal gratings have been shown to be inseparable (Robson, 1966; Koenderink & van Doorn, 1979; Watson & Ahumada, 1985; Kelley, 1985), as are the influences of area and duration on spatial summation (Barlow, 1958). This means that one needs information about human contrast sensitivity over the whole spatio-temporal plane, rather than being able to treat space and time separately. Koenderink and van Doorn have made such measurements, although only for a single luminance level. A moving grating will induce tracking eye-movements if it is above threshold, so to minimize this problem, they obtained these results using counterphase flickering gratings. To predict the visual sensitivity to a single moving grating, one must assume that the excitatory effects of two gratings moving in opposite directions are additive.

To show the effect of motion, it is convenient to define a filter that gives the additional attenuation of all components of the 2D-power spectrum when an image is moved at 2 deg/s. First, the spatial contrast sensitivity function at

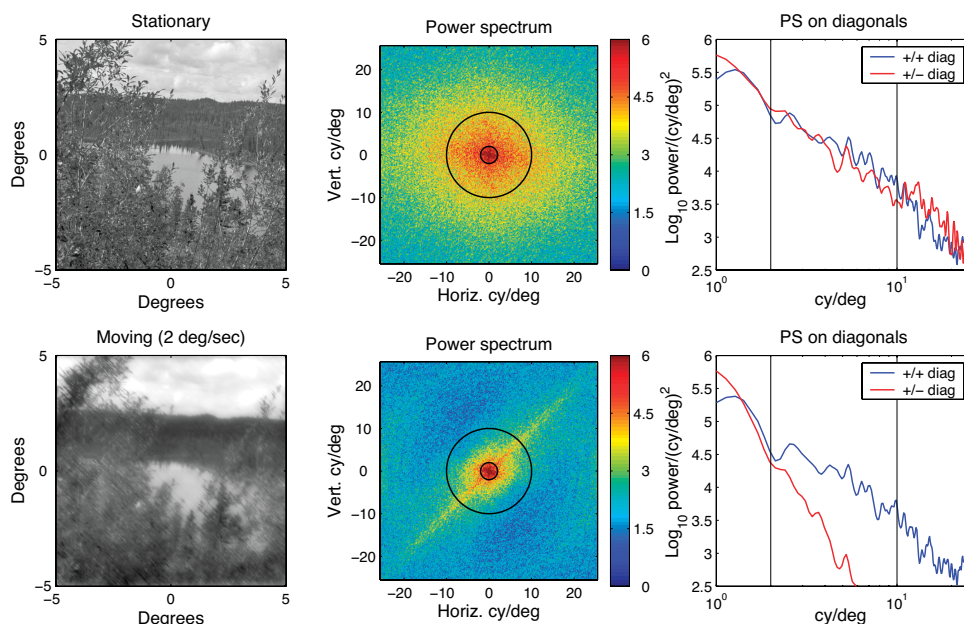


Figure 1. The influence of motion on the effective power spectrum and appearance of an image, assuming the whole image subtends 10×10 deg and moves diagonally down and to the left at 2 deg/sec. Top row shows the original static image, its 2-D spatial power spectrum, and sections through this power spectrum along the two diagonals. The lower row shows the corresponding three figures for the moving image, obtained by attenuating the 2-D spatial Fourier Transform of the static image by factors derived from the contrast sensitivity measurements of Koenderink and van Doorn. These calculations are described in Figure 2 and in the text. The effect of motion on the power spectrum is brought out by the cross sections (right pair of figures) along the diagonals in the direction of motion (red) and orthogonal to it (blue). Circles in the center pair and lines in the right-hand pair are drawn at 2 and 10 cy/deg to indicate the range where contrast sensitivity is high in the human visual system, and it will be seen that motion at 2 deg/sec severely attenuates spatial frequencies in the upper part of this range.

0.1 Hz is obtained from Koenderink and van Doorn's measurements; this is the lowest temporal frequency they used, and here it is taken as an approximation to the values expected at zero velocity of movement. Next, a corresponding contrast sensitivity function for motion at 2 deg/s is obtained by reading off the appropriate attenuations corresponding to each spatial-frequency in the 2D-Fourier plane. For spatial-frequencies in the direction of motion, the attenuations are given by interpolating between the contours along the solid line in Figure 2 (left). Spatial-frequencies at oblique orientations with respect to the direction of motion are effectively moving at a speed reduced by $\cos\theta$, and so the attenuations are obtained along lines corresponding to slower speeds (e.g., dashed line in Figure 2, left). The attenuations thus obtained for each spatial-frequency in the 2D-Fourier plane are then divided by the attenuations due to the spatial CSF at 0.1 Hz. This approximates the additional attenuation just due to motion (without the static portion of the CSF) and is shown in Figure 2 (right).

Going back to Figure 1, the middle figure in the lower row shows the power spectrum obtained by multiplying the middle top power spectrum by the 2 deg/s motion filter, calculated as described above and shown in Figure 2 (right). The left-hand figure shows the corresponding motion-blurred image, obtained as the inverse Fourier transform (including phase) of the filtered power spectrum. In the lower center power spectrum, it will be seen that motion steepens the flanks of the unfiltered power spectrum in the direction of motion, and the contours of equal power become strongly elliptical rather than nearly circular. The right-hand figure shows two sections through it, one at right angles to the direction of motion, which is like the one above for the unmoved image, and the other along the direction of motion, which now declines much more

steeply. The top image of the pair shows the same sections through the static image. It is clear that motion blur differentially attenuates spatial components modulated along the direction of motion compared with orthogonal to it. This causes the power spectrum to become strongly anisotropic, especially in the crucial range between 2 and 10 cycles/deg.

We think this differential attenuation is used by the visual system as a hallmark of motion and an aid in its analysis. To understand better how the effect of motion on the power spectrum might be detected, one needs to look at the shape of the filter (Figure 2, right) that produces this anisotropy. This will be considered briefly in the "Discussion."

The power spectrum of local patches

The image in Figure 1 was blurred by moving all parts of it at the same velocity down and to the right. This could have resulted from rotating a camera, or the eye, about the appropriate axis, but the patterns of image movement (optic flow) that occur naturally are usually more complicated. For example, if the eye fixates on a distant part of the scene while the observer moves at right angles to this line of sight, the images of near objects move fast while those of distant objects move slowly or not at all. If an observer moves toward a point while fixating it, its image remains stationary while those at other positions move along lines radiating from this focus of expansion, their velocities depending both on their angular distance from the focus, and their true distance from the observer. It is clear that if power spectra are to be used to recover motion, they must be computed separately for different local patches of the image, not globally for the whole image.

The size of such patches can only be approximately specified at the moment. Assuming as above that the over-

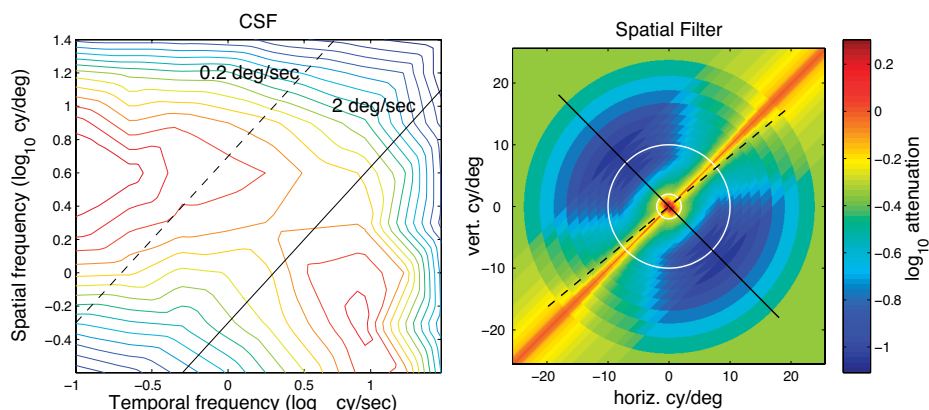


Figure 2. The construction of the filter in the spatial frequency domain that would attenuate high frequencies to the same extent as motion at 2 deg/sec down and to the right. At left is shown a replot of the contrast sensitivity measurements of Koenderink and van Doorn (1979) using counterphase flickering gratings. Contour spacing is 1 dB. At right are shown the extra attenuations at all orientations and spatial-frequencies caused by moving an image at 2 deg/sec in the direction of the negative (+/-) diagonal. The solid line shows the attenuations for spatial-frequency components oriented in the direction of motion; the dashed line shows the attenuations for spatial-frequency components oriented at 84 deg anticlockwise to the direction of motion, which have an effective speed of 0.2 deg/sec. Attenuations at all spatial-frequencies are relative to the value at 0.1 Hz (the lowest frequency available from Koenderink and van Doorn) rather than the value for a truly stationary image.

all image size is $10^0 \times 10^0$, it would be reasonable to take about 40 patches each about $1.5^0 \times 1.5^0$, and this would be consistent with the results of Morrone, Burr, and Vaina (1995), with the model for the detection of Glass patterns of Wilson and Wilkinson (1998), and with the neurophysiological results on macaques of Smith, Bair, and Movshon (2002). However, for two reasons, we think it is premature to suppose that there is a single fixed size for the patches over which the computations are performed. First, visual resolution declines rapidly with eccentricity from the fovea, and it is very probable that patch size also changes. Second, there are advantages in analyzing images simultaneously at different scales (Marr, 1982), and this may be true for the types of analysis envisioned here. Notice that using small patches would improve the spatial detail in an analysis of optic flow, but the power spectrum of a small patch would be noisy and anisotropies from motion blur might become undetectable.

Glass patterns

Glass patterns (Glass, 1969) are the streaky patterns that are seen when marks (usually randomly positioned dots, as in Figure 3) on a sheet of paper are copied, shifted by translation, rotation, or expansion, and then superimposed on the original set of marks. On close inspection, one sees that each of the original marks now has a companion close to it displaced in a direction that depends on the type of shift used, but even without close examination, one is aware of a streakiness or flow that lies along the directions of these displacements. It is not easy to convince oneself that this appearance results solely from the pairing of each dot; one has the feeling that something extra, like harmony in music, is giving order to the experience. We now show that their power spectra have properties shared with those of moving natural images and might, therefore, be expected to excite the visual system similarly.

The auto-correlation of Glass patterns

While it is often suggested that the appearance results from the pairs of dots preferentially stimulating the elongated, orientationally selective, receptive fields of neurons in the visual cortex (Hubel & Wiesel, 1962), this explanation is not as straightforward as it seems. The average number of dots falling in an elongated patch depends solely on the mean density of dots in the pattern and the area of the patch, and is completely unaffected by the orientation of the pairs relative to the elongation of the patch. What does vary with relative orientation is the distribution of the numbers per elongated patch around the mean number, for it is more probable that both members of a pair will fall in the same patch when the orientation of patch and pair are the same. We thus get our first insight into the origin of the streaks: It is unlikely to result from the greater *mean* excitation of those orientation selective neurons that are aligned with the orientation of the Glass pairs, but it may result from increased *variation* in the excitation of these neurons. To develop this insight further, consider the auto-correlation function of a Glass pattern, which can be deduced rather directly from the way they are made. Power spectra contain exactly the same information as auto-correlation functions and can be derived from them.

The simplest of the Glass patterns is shown in Figure 4 (top left). This was formed by linearly displacing a "seed" set of N -random dots downward and rightward to form a "daughter" set, and then superimposing them (here $N=400$). A representation of its auto-correlation function is also shown (Figure 4, top center). To see how this function is derived, visualize the whole Glass figure (the seed set with its daughter set) being shifted over itself with all possible combinations of vertical and horizontal displacement. For each size and direction of shift, the value of the auto-correlation function is formed by multiplying the contrasts (defined as $(L - L_{mean})/L_{mean}$) of corresponding pixels in shifted and unshifted images, and summing the products over the whole image. This gives a figure proportional to the total number of coincidences between dot positions in

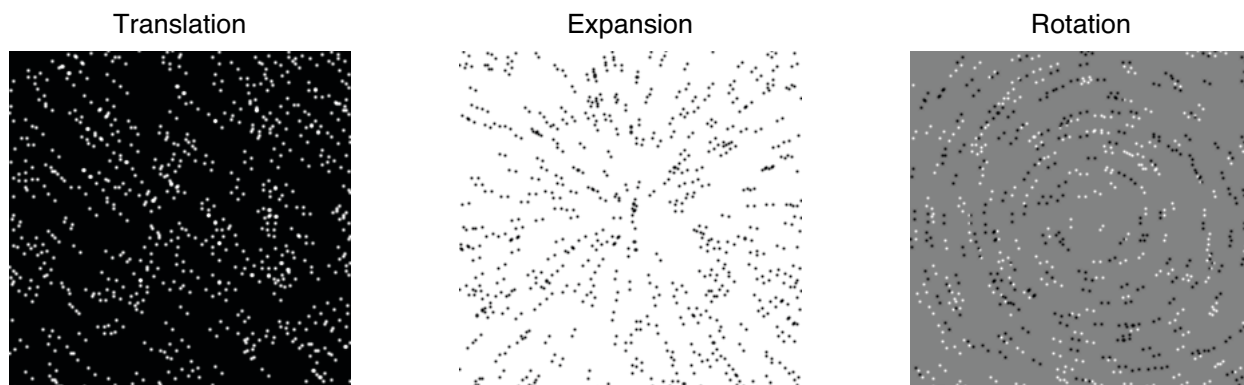


Figure 3. Glass figures for linear translation, expansion and rotation. Note that if each dot of a pair has the same contrast, both positive and negative contrasts contribute to the effect. When the dots of a pair are of opposite contrast the effects are very different (see Figures 4 and 8).

shifted and unshifted figures. The center corresponds to zero shift, and the peak here is caused by every dot coinciding with itself, seed with seed and daughter with daughter, so each of the $2N$ dots coincides with itself and there are $2N$ coincidences. The peak down and to the right is caused by shifted seed coinciding with unshifted daughter, and the peak above and to the left by coincidence of unshifted seed with shifted daughter; each of these contributes N coincidences. Everywhere else there is a low and variable number of coincidences resulting from the shifts happening to correspond to the separations of dot pairs in the combined pattern. The number of these coincidences depends on the fraction of the available positions that is occupied by a dot, and half the expected number of accidental coincidences per position will be added to the N coincidences at each of the lesser peaks.

Note that for this figure, the dots were formed as small Gaussian blobs and as a result the peaks of the autocorrelation function have some spread.

The power spectrum of Glass patterns

From a well-known mathematical result, the power spectrum of the Glass pattern is obtained by Fourier transforming its auto-correlation function, and the result is shown in Figure 4 (upper right). The coordinates of the plot are spatial frequencies in the X and Y directions, with the DC component (frequency = 0) at the center. Thus the

ridge running through the center from top right to bottom left represents high power in all spatial frequencies periodic at right angles to the direction of dot displacement. Each dot by itself produces uniformly distributed power at all spatial frequencies and orientations; for those spatial frequencies in the same direction as the dot displacement, the waves generated by the dot and its pair will exactly coincide for wavelengths that are an integer fraction of the dot separation, thus producing the banding pattern. The attenuation at high spatial-frequencies is due to the fact that the dots themselves are represented by Gaussian blobs, and this causes a Gaussian taper in the spatial-frequency domain. (Note that the banding pattern is not unique to Glass patterns, but would be produced by any image that is translated and added to itself.)

The circles at 2 and 10 cycles/deg demarcate the region of visual importance in the power spectrum, and within this region, the troughs between the central ridge and its two neighboring ridges are the prominent features. These indicate lack of power in the corresponding spatial frequencies and orientations, and the minima occur where the components generated by seed dot and its daughter dot are in antiphase, so their amplitudes cancel each other.

The power spectrum derived from the stationary Glass pattern resembles the power spectrum of the moving natural scene in being anisotropic in the region of visual importance: They both fall off more rapidly along the top-

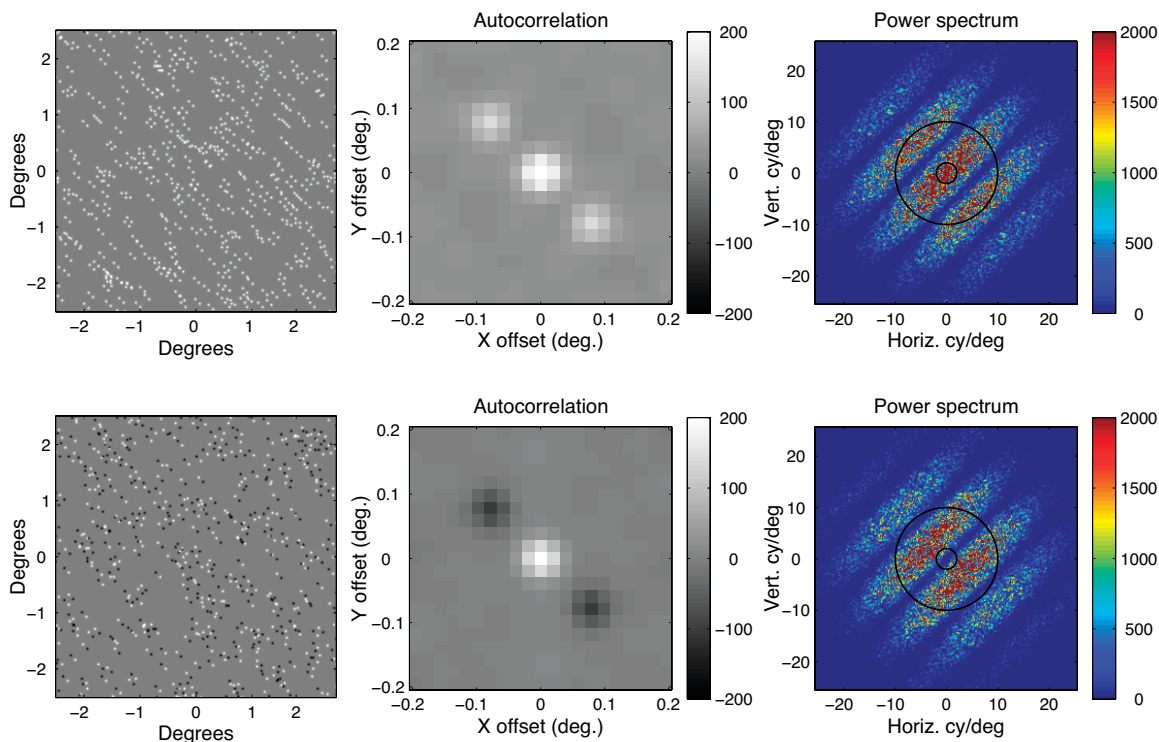


Figure 4. A Glass pattern (top left) is formed from a randomly positioned array of seed dots each of which is shifted down and to the right to form a daughter dot that is added to the original array. Upper middle shows its autocorrelation function and upper right its power spectrum. The lower row shows corresponding figures for an array of “anti-pairs” in which a black dot is consistently paired with a white dot and vice versa; in this case the polarity of the pairs was randomized. Such anti-pairs are used in the experiment of Figure 8.

left/bottom-right axis than along the top-right/bottom-left axis. The main difference between the two power spectra is that the top-right/bottom-left axis is nearly flat for the Glass figure, but falls off for the moving natural scene, though less steeply than along the line of motion.

Because of this similarity in their power spectra, we think Glass figures stimulate a system normally used in the analysis of optic flow through motion blur, and hence they give rise to an appearance of streakiness or flow. This hypothesis can be tested as follows.

Three tests of the anisotropy hypothesis

1. Power spectrum not phase spectrum controls flow

A very direct test of the assertion that the streakiness is determined by the power spectrum, not the phase spectrum, is shown in Figure 5. We computed the full Fourier transforms of a pair of Glass patterns, one like the top row of Figure 4, the other having a different random set of parent dots and the orthogonal direction of displacement, and hence the orthogonal orientation of its apparent flow. The transforms were divided into their phase spectrum and amplitude spectrum (square root of the power spectrum), and then the phase spectrum of one was combined with the

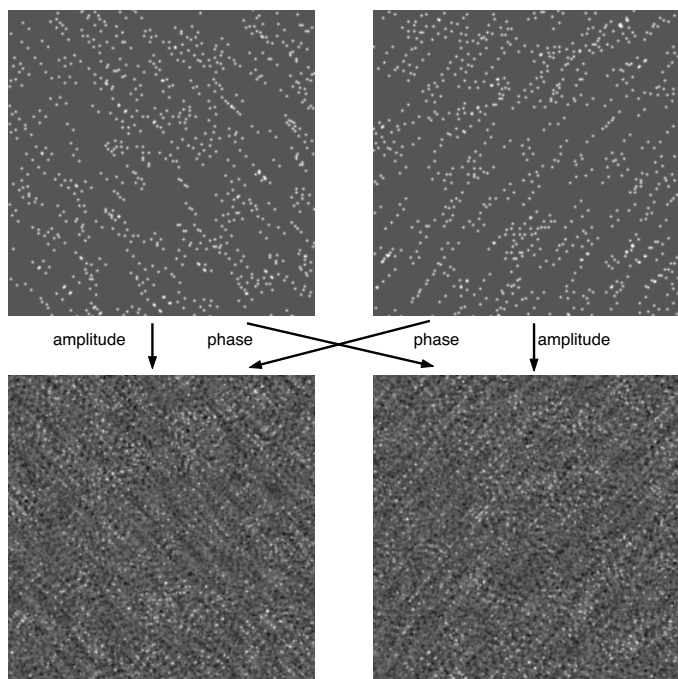


Figure 5. Power spectrum, not phase spectrum, determines flow in Glass patterns. Fourier transforms were made of the two translational Glass patterns in the top row, then recombined in the bottom row with their phase and amplitude components switched.

amplitude spectrum of the other, and vice versa. Although this procedure does not reconstitute the original appearance, streaks are clearly visible, and their orientations follow the amplitude, not the phase, component. This is in contrast with the usual state of affairs when recognizing the subject matter of an image (Oppenheim, & Lim, 1981; Piotrowski & Campbell, 1982).

2. Adding complementary noise reduces streakiness of Glass patterns

If anisotropy in the power spectrum is the cause of streakiness or flow in Glass patterns, then filtered noise that has a complementary power spectrum to that of Figure 4 (top row) should reduce or eliminate the streakiness when added to a Glass pattern. The top row of Figure 6 shows a Glass pattern, the complementary noise described above, and their combination; the bottom row shows their power spectra. The power spectrum of the combination is uniform, and as predicted by the hypothesis, little streakiness is detectable in the figure shown here. We have found, however, that under other conditions, some streakiness remains, and we consider this after describing some additional observations.

To control for the possibility that the reduced streakiness is simply due to the addition of noise to the image, we made the additional experiments shown in Figure 7. The top row shows three types of noise: left, with power spectrum complementary to that of the Glass figure (same as in Figure 6 center), center, with a flat power spectrum, and right, with noise having the same spectrum as the Glass figure. All these noise figures had the same total power as the original Glass pattern. In the lower row, these have each been added to the original Glass image, and it will be seen that the complementary noise is more effective than flat-spectrum noise in reducing the streakiness, whereas the similar spectrum noise, not unexpectedly, increases the streakiness. In the top row, it is worth noting that the complementary spectrum noise has a granular appearance unlike that of white noise, though it lacks any obvious oriented component. The right-hand figure shows that noise with the same power spectrum as the Glass figure has the characteristic streaky appearance of a translational Glass figure.

These observations are all in broad agreement with the hypothesis, though the appearance of the complementary noise was not predicted by it, and this suggests an important modification: We should not expect all anisotropies of power spectra to produce streakiness or flow, but only those patterns of anisotropy that mimic the ones occurring in moving images.

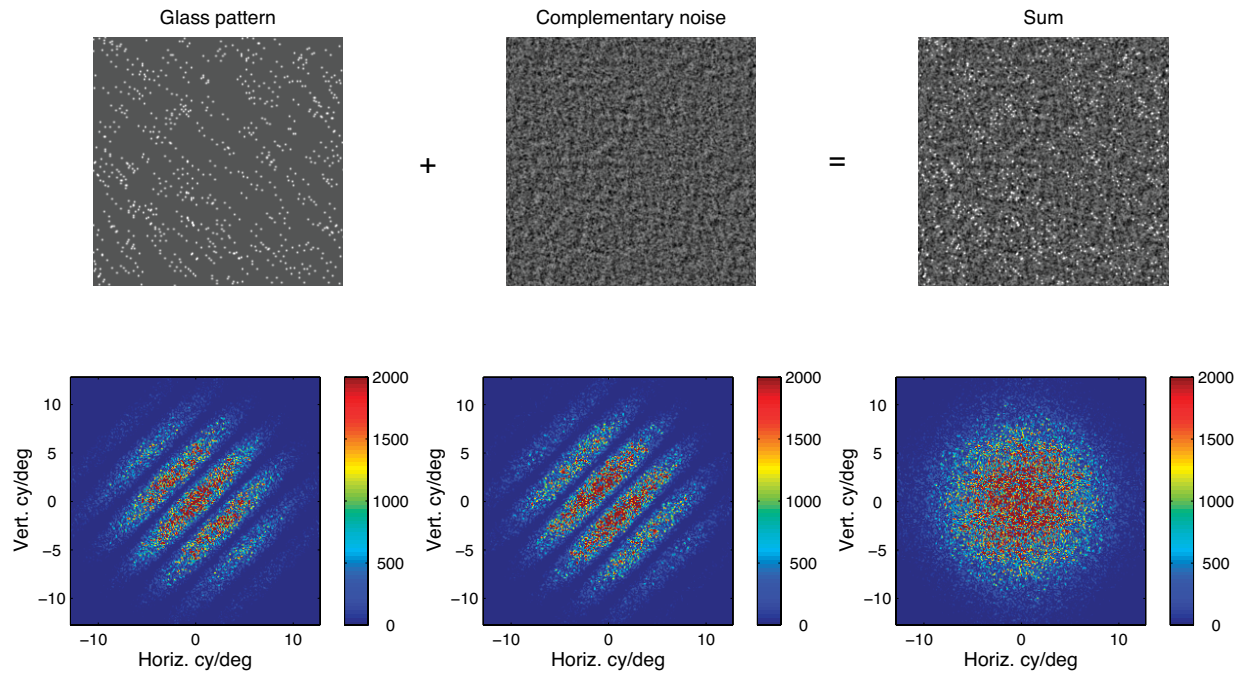


Figure 6. Glass pattern, complementary noise generated by filtering 2D white noise, and sum of the two, with their respective power spectra. Note that the Glass pattern has a maximum at (0,0), whereas the complementary noise has a trough here. The complementary noise eliminates much of the streakiness visible in the Glass pattern.

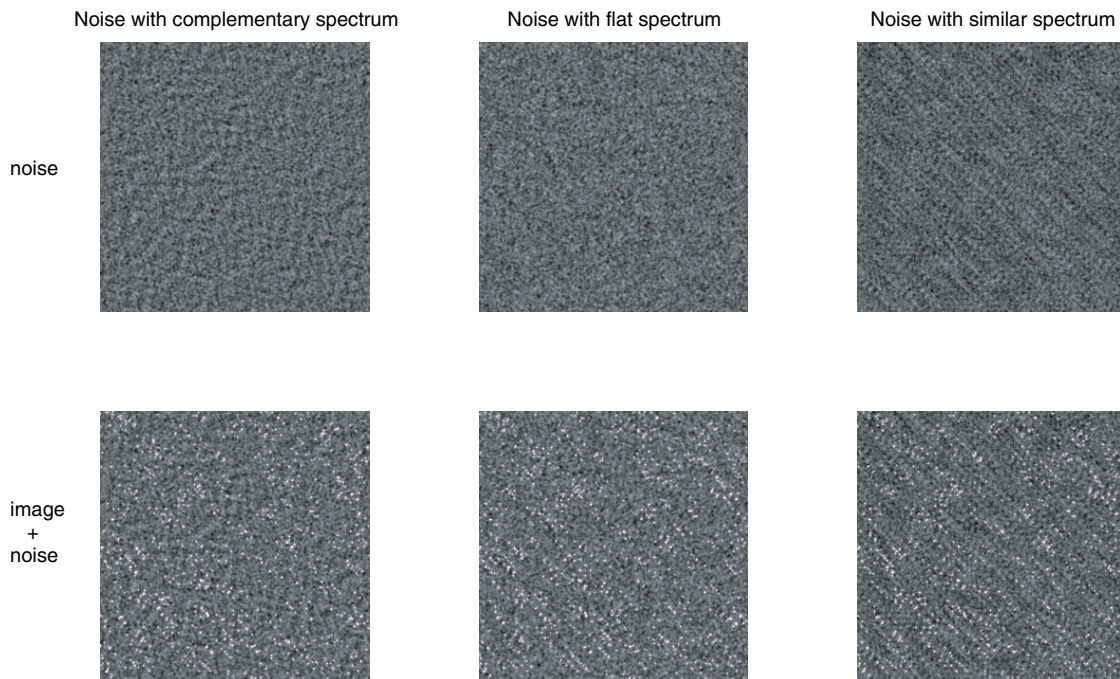


Figure 7. Three types of noise (top row) added to a translational Glass pattern (bottom row). The three noise-alone patterns have the same power as the Glass figure and they had power spectra that were complementary to it (left), flat (centre), and similar to the Glass figure (right). The complementary noise is clearly most effective in reducing the streakiness, whereas similar spectrum noise increases it. Note that the complementary noise pattern, viewed alone, has a granular texture but no salient oriented component, and that noise with the same power spectrum as the Glass figure has streaks similar to those of the Glass figure, all as predicted by the hypothesis.

We now return to the point mentioned above, that the streakiness of the original Glass figure was sometimes incompletely suppressed by complementary noise and seemed to “shine through.” We noted that suppression was more effective when dot density was high, whereas for low dot densities, the added noise, even though it evens out the power spectrum, fails to eliminate streakiness completely. We think the reason is that, as pointed out above, the visual system computes power spectra over local patches of the image, whereas in the above experiment, the power spectrum was made isotropic when averaged over the whole image. This leaves room for random variations where the anisotropy caused by the original Glass pattern is not exactly matched by the additional anisotropy that was supposed to annul it. Thus the locally computed power spectra will not be uniform in all parts of the combined image, allowing local regions to show flow or streakiness.

3. Canceling streakiness with negative auto-correlations

Operations on the power spectrum of the whole image work with translational Glass patterns because the power spectrum is similar all over them, but they will not work with rotational or expansion Glass figures because the local power spectra vary over the image, and according to the hypothesis, this is responsible for the varying directions of the streakiness in different parts. There is, however, a neat manipulation (suggested to us by members and visitors to

the Olshausen lab) that flattens the power spectrum appropriately in each local part.

The left panels of [Figure 8](#) are rotational Glass figures that contain pairs of black dots and pairs of white dots on a grey background, as can be seen by detailed inspection. When the local auto-correlation function of this is formed, coincidences of seed with daughter will form at the same shifts for both black and white dots, and the three-peaked auto-correlation function is very similar to one having only white dots on a dark ground, or only black on a white ground. But in the mixed case, it is possible to arrange for a black dot in the seed pattern to be paired with a white dot in the daughter pattern, and vice versa, forming opposite-signed pairs we call “antipairs.” The contrast of a dot is positive if it is above the mean luminance, and negative if it is below, so in these cases, the signs of the coinciding dots are different, + with - or - with +, and because the contributions to the auto-correlation function are formed by multiplication, these reverse coincidences contribute negatively to the auto-correlation function. The autocorrelation function and power spectrum of a translational, antipair Glass pattern is shown in the bottom row of [Figure 4](#). It will be seen that the power spectrum is complementary to that of the normal paired Glass pattern, shown in the row above in [Figure 4](#). This means that if an antipair Glass pattern is added to a normal Glass pattern with the same number of dots, the streakiness should be annulled. The nice feature of using antipairs is that it should work with forms of Glass pattern other than the translational one, for if the antipairs

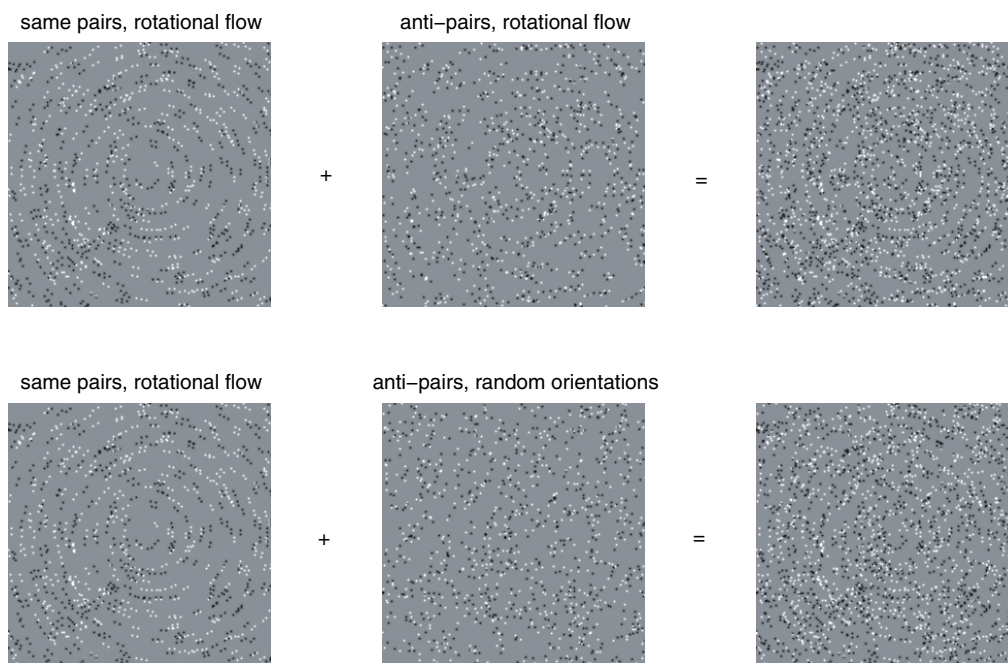


Figure 8. Canceling Glass patterns with anti-pairs. Top left is a rotational Glass pattern formed with paired white dots of positive contrast together with paired black dots of negative contrast. Top center is the same except that one dot of each pair was reversed in polarity to convert each pair into an anti-pair (see Fig. 4), and it was formed from a different set of random seed dots. When this is combined with the normal Glass figure it greatly reduces the rotational “flow” (top right). In the bottom row the anti-pairs are randomly oriented and are much less effective in reducing the rotational flow, so we do not think the effect is simply due to the addition of noise.

have the same rotational or expansion pattern as the pairs, they should have the correct orientation in each part of the image to flatten the local power spectrum all over the image.

Figure 8 tests whether this happens. The left figure in the top row is a rotational Glass pattern with pairs of white and pairs of black dots on a gray background. The center figure is composed of antipairs and is made to the same rotational plan, but using different sets of random positions for the black and white parent dots. Notice that it lacks the strong appearance of concentric rings seen in the left image, though it is not completely structureless. The right image shows the sum of the two, and it will be seen that, as predicted, the concentric rings have been much reduced.

Again it might be suggested that this is just a nonspecific effect of adding noise, so in the bottom row, the left image is another example of the same type of Glass figure as that shown in the top row, whereas the center image is formed of antipairs as in the top row, but in this case each pair is set at a random orientation, not according to the rotational pattern that was used in the top row. The lower right-hand figure shows that these randomly oriented antipairs are much less effective in reducing the concentric appearance than the concentrically organized antipairs of the top center pattern.

Discussion

Our hypothesis says that the visual system analyses optic flow, not only by detecting the spatio-temporal correlations that result from movement, but also by detecting the anisotropies of local spatial power spectra that result from motion blur. We think this hypothesis fits much of the work that preceded it in this area by Geisler and his colleagues (Geisler, 1999; Geisler, Albrecht, Crane, & Stern 2001), Burr (2000) and Ross (Burr & Ross, 2002; Ross, Badcock, & Hayes, 2000), Wilson and Wilkinson (1998), and others, and it has survived the three simple tests we have applied. Obviously, further tests are desirable, some of which will become evident in what follows, but first we must point out a gap in the evidence and an objection to the hypothesis.

A gap in the evidence

Our calculations of the extent and severity of motion blur depended on measurements of human spatio-temporal contrast sensitivity made by Koenderink and van Doorn (1979), but they used counterphase flickering gratings in order to minimize the problem caused by the tendency of the eye to track moving gratings. Although, mathematically, counter-phase flicker is exactly the same as two gratings of the same frequency moving in opposite directions at the same velocity, it is not certain that the threshold modulation for the sum of these two gratings when moving in opposite directions is the same as the threshold modulation

of their sum when they are moving, in phase, in the same direction. If there are differences, this will require recalculation of the effects of motion blur, but we doubt if our main argument will be much weakened.

An objection to the hypothesis

If anisotropies of local power spectra resulting from motion blur provide us with information about movement, why do we not actually see movement when we are presented patterns containing such anisotropies? First notice that the streakiness seen in Glass patterns is sometimes described as "flow," which might be taken to imply movement. On the other hand, if one pressed a person who described the effect as flow by asking, "Does it actually appear to move?" few if any would answer "Yes," so the objection still retains force.

An important point to make is that motion blur could give only the axis of the motion that caused the blur - it could not give the direction of motion along that axis. In the absence of that information, would one really expect to experience true motion? Perhaps the evidence from motion blur, and the motion streaks (Geisler, 1999) that movement should produce, is only used when true spatio-temporal correlations are also present, and it adds sensitivity and orientational precision to impressions gained from true motion detectors, as the results of Burr and Ross (2002) suggest.

An interesting illusion

Ross, Badcock, and Hayes (2000) have shown that viewing a succession of independently created Glass patterns gives the vivid impression of fast motion in one of the directions of the flow described above, but the direction of this motion is unstable, tending to reverse suddenly as a whole, or occasionally to fragment into parts moving in opposite directions. Now, although it is true (as Ross et al., 2000, point out) that these sequences do not contain a consistent excess of spatio-temporal correlations indicating motion in one direction rather than another, such sequences of Glass patterns do nevertheless contain plentiful spatio-temporal correlations in all directions and velocities resulting from "correspondence noise" (Barlow & Tripathy, 1997). These are the spurious motion signals that result from ignorance about which dots in successive frames of a random dot kinematogram are to be considered pairs. Correspondence noise must activate many elements of true motion detecting systems, and it is a safe bet that presentation of such a stimulus would cause a massive increase in the firing rates of almost all units of MT in rhesus monkey cortex. This is a prediction that ought to be tested: Is it possible that such tests would reveal transient, unstable, increases of activity of units tuned to the directions of motion in which humans experience transient apparent motion? Krekelberg, Dannenberg, Hoffmann, Bremmer, and Ross (2003) have made such observations in MT and MST, and we think our hypothesis fits well with most of their

ideas and their results. Possibly evidence about the axis of motion derived from motion blur reaches MT, and it might be responsible for the sharpening of directional selectivity that occurs in the first few hundred milliseconds of the response (Pack & Born, 2001).

It is interesting to note that detectors of spatio-temporal correlations use high temporal frequency information, whereas the motion blur mechanism would use low temporal frequency information, so to this extent, the two mechanisms provide complementary, independent, information about motion.

Physiological mechanisms

The hypothesis that motion blur is used to detect optic flow requires (1) a system for forming a representation of the local power spectra of different parts of the image, (2) a system for detecting anisotropies in these local spectra, and (3) a system for detecting the patterns formed by the different directions of anisotropy in the different parts. Before the properties of single units in visual cortex had been explored neurophysiologically, the whole suggestion would, justifiably, have seemed very implausible, but this is no longer the case because the results of these explorations seem, at least at first glance, to fit the requirements eerily well. This is not the place for detailed discussion, but we shall point to evidence that each of the above requirements is met.

The first requirement is a representation that resembles a local Fourier transform, that is, one whose elements are activated by sinusoidally modulated patterns of light at many orientations and spatial frequencies, covering small patches of the visual field. Arguing from psychophysical evidence, Campbell and Robson (1968) suggested that Fourier analysis had a role in understanding visual processing, and for the next 20 years this idea was pursued by many. It was realized early that global Fourier analysis makes little sense (see Westheimer, 2001), but Maffei and Fiorentini (1973) and Marcelja (1980) pointed out that the description that Hubel and Wiesel gave of the structure and organization of the primary visual cortex in cat and monkey (Hubel & Wiesel, 1962, 1968) fitted the idea of local Fourier analysis or wavelet decomposition, and Andrews and Pollen (1979) showed that the profiles of simple cell receptive fields are compatible with this idea. On this view, the projection of a hypercolumn in the visual field determines the size of the local patch, the orientation of a simple cell's elongated receptive field determines the orientation of the component detected, and the separation of its excitatory and inhibitory parts determines its spatial frequency. To form the power spectrum from this representation, the outputs of simple cells of similar orientation and spatial frequency would need to be squared and summed, ignoring their phase and exact position, and as shown by Movshon, Thompson, and Tolhurst (1978), complex cells have many of the required properties.

We think the first two requirements of our hypothesis are well met. Next a mechanism is needed for detecting the patterns of anisotropy that are characteristic of motion blur. These patterns typically cover large parts of the visual field and are unlikely to be mediated by connections confined to V1. In macaque, there is evidence that information from the appropriate subsets of neurons in V1 project on to single neurons in V4, where Gallant et al. (1993, 1996) described units that respond selectively to non-Cartesian gratings, that is circular, star-shaped, and hyperbolic patterns that are similar to the patterns of deformation that result from optic flow (Koenderink & van Doorn, 1975). In addition to the minimum set of rotation, expansion, and deformation, they also found units responding to spiral patterns that would result from combinations of the first two. Notice that these units must result from a non-topographic mode of projection (Barlow, 1981) in which differently oriented elements of the pattern occurring at different positions in the visual field converge on to single V4 neurons.

Saito et al. (1986) and Orban et al. (1992) found units in MST responsive to moving stimuli that were selective for the same patterns of motion. It is likely, then, that there are two mechanisms for analyzing motion, a slow mechanism based on detecting the local anisotropies of power spectra resulting from motion blur, as well as a fast mechanism based on the detection of true directional motion through spatio-temporal correlations. Both are specialized to detect the characteristic patterns of these anisotropies that result from optic flow, and they are complementary because they use different temporal frequencies.

Using prior probabilities

It has been suggested by the authors, among others (Olshausen & Field, 1996; Barlow, 2001), that the visual system exploits the redundancy of natural images when interpreting sensory information, and we think the proposed use of anisotropies of local power spectra to detect motion blur is a nice example of this: Without the expectation that local power spectra are isotropic, it would not be possible to make the Bayesian inference that anisotropy is due to motion. It should be objected here that local power spectra are not in fact strictly isotropic, even in images of the natural environment, and signs of this are visible in the power spectrum shown in Figure 1, where the region of high power tends to be diamond-shaped rather than truly circular. The model for detecting Glass figures that Wilson and Wilkinson (1998) proposed had a curious feature required to make the model fit the psychophysical evidence. For detecting radial and circular Glass patterns, they found it necessary to attach greater weight to the orientations of pairs lying along or close to the diagonals rather than the horizontal or vertical axes of the patterns. It looks then as if the mechanism discounts evidence from horizontal and vertical orientations in a way that is appropriate because anisotropies along these axes occur more frequently in

natural images. When details like this fit a hypothesis, one begins to have more confidence in it.

The present study is incomplete and demands further investigation at many points. Nevertheless, we think the idea that the visual system uses motion blur in a constructive way brings together many previously disconnected topics of recent research. Furthermore, it may make a contribution toward understanding how we see so well in spite of the constant motions of the visual image.

Acknowledgments

The rapid development of the ideas and observations in this work stemmed largely from lively discussions among members of and visitors to the Olshausen laboratory at UC Davis; in particular, we would like to thank Phil Sallee, Dan Ruderman, Jeff Johnson, and Konrad Kording. We would also like to thank David Burr and John Ross for showing us unpublished manuscripts on the interactions of Glass figures and motion. Horace Barlow visited the Olshausen lab while he was Regent's Lecturer at UC Davis, and Bruno Olshausen is supported by NIMH grant MH57921.

Commercial Relationships: None.

Corresponding author: Bruno Olshausen.

Address: Center for Neuroscience & Department of Neurobiology, Physiology, and Behavior, UC Davis, Davis, CA.

Email: baolshausen@ucdavis.edu

References

- Andrews, B. W., & Pollen, D. A. (1979). Relationship between spatial frequency selectivity and receptive field profile of simple cells. *Journal of Physiology*, *287*, 163-176. [PubMed]
- Barlow, H. B. (1958). Temporal and spatial summation in human vision at different background intensities. *Journal of Physiology, London*, *141*, 337-350.
- Barlow, H. B. (1981). Critical limiting factors in the design of the eye and visual cortex: The Ferrier lecture, 1980. *Proceedings of the Royal Society of London B*, *212*, 1-34. [PubMed]
- Barlow, H. B. (2001). Redundancy reduction revisited. *Network: Computation in Neural Systems*, *12*, 241-253.
- Barlow, H. B., & Tripathy, S. P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. *Journal of Neuroscience*, *17*, 7954-7966. [Article]
- Burr, D. (2000). Motion vision: Are 'speed lines' used in human visual motion? *Current Biology*, *10*, R440-443. [PubMed]
- Burr, D. C., & Ross, J. (2002). Direct evidence that 'speedlines' influence motion perception. *Journal of Neuroscience*, *22*, 8661-8664. [PubMed]
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, *197*, 551-566. [PubMed]
- Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for polar, hyperbolic and cartesian gratings in macaque visual cortex. *Science*, *259*, 100-103. [PubMed]
- Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. L., & Van Essen, D. C. (1996). Neural responses for polar, hyperbolic and cartesian gratings in area V4 of macaque monkey. *Journal of Neurophysiology*, *76*, 2718-2739. [PubMed]
- Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature*, *300*, 323-325. [PubMed]
- Geisler, W. S., Albrecht, D. A., Crane, A. M., & Stern, L. (2001). Motion direction signals in the primary visual cortex of cat and monkey. *Visual Neuroscience*, *18*, 501-516. [PubMed]
- Glass, L. (1969). Moiré effect from random dots. *Nature*, *223*, 578-580. [PubMed]
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106-154.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195*, 215-243. [PubMed]
- Kelley, D. H. (1985). Visual processing of moving stimuli. *Journal of the Optical Society of America A*, *2*, 216-225. [PubMed]
- Koenderink, J. J., & van Doorn, A. J. (1975). About decomposing optic flow. *Optica Acta*, *22*, 773-791.
- Koenderink, J. J., & van Doorn, A. J. (1979). Spatiotemporal contrast detection threshold is bimodal. *Optics Letters*, *4*, 32-34.
- Krekelberg, B., Dannenberg, S., Hoffmann, K.-P., Bremner, F., & Ross, J. (2003). Neural correlates of implied motion. *Nature*, *424*, 674-677. [PubMed]
- Maffei, L., & Fiorentini, A. (1973). The visual cortex as a spatial frequency analyser. *Vision Research*, *13*, 1255-1267. [PubMed]
- Marcelja, S. (1980). Mathematical description of the response of simple cortical cells. *Journal of the Optical Society of America*, *70*, 1297-1300. [PubMed]
- Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

- Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, 376, 507-509. [[PubMed](#)]
- Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology*, 283, 79-99. [[PubMed](#)]
- Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381, 607-609. [[PubMed](#)]
- Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, 69, 529-541.
- Orban, G. A., Lagae, L., Verri, A., Raiguel, S., Xiao, D., Maes, H., & Torre, V. (1992). First order analysis of optical flow in the monkey brain. *Proceedings of the National Academy of Sciences U.S.A.*, 89, 2595-2599. [[PubMed](#)] [[Article](#)]
- Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, 409, 1040-1042. [[PubMed](#)]
- Piotrowski, L. N., & Campbell, F. W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception*, 11, 337-346. [[PubMed](#)]
- Robson J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America*, 56, 1141-1142.
- Ross, J., Badcock, D. R., & Hayes, A. (2000) Coherent global motion in the absence of coherent velocity signals. *Current Biology*, 10, 679-682. [[PubMed](#)]
- Smith, M. A., Bair, W., & Movshon, J. A. (2002). Signals in Macaque striate cortical neurons that support the perception of Glass patterns. *Journal of Neuroscience*, 22, 8344-8345. [[PubMed](#)]
- Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of macaque monkey. *Journal of Neuroscience*, 6, 145-157. [[PubMed](#)]
- Tadmor, Y., & Tolhurst, D. J. (1993). Both the phase and the amplitude spectrum may determine the appearance of natural images. *Vision Research*, 33, 141-145. [[PubMed](#)]
- Watson, A. B., & Ahumada, A. J. (1985). Model of human visual motion sensing. *Journal of the Optical Society of America A*, 2, 322-341. [[PubMed](#)]
- Westheimer, G. (2001). The Fourier theory of vision. *Perception*, 30, 531-541. [[PubMed](#)]
- Wilson, H. R., & Wilkinson, F. (1998). Detection of global structure in Glass patterns: Implications for form vision. *Vision Research*, 38, 2933-2947. [[PubMed](#)]