Perception, 1986, volume 15, pages 59-66

A spatial illusion from motion rivalry

David C Burr, John Ross, M Concetta Morrone
Department of Psychology, University of Western Australia, Nedlands, WA 6009, Australia
Received 24 January 1985, in revised form 30 September 1985

Abstract. A new dynamic visual illusion is reported: contrast reversal of a horizontal and vertical plaid pattern (produced by adding two orthogonal sinusoidal gratings) causes the pattern to appear as an array of lustrous diamonds, cut by sharp lines into a diagonal lattice structure. On the basis of computer simulations it is suggested that the illusion results from rivalrous interaction of motion detectors tuned to opposing directions of motion.

1 Introduction

Contour at a given orientation is usually associated with energy at the orthogonal orientation in the frequency domain; and hard-edged contour is usually associated with high spatial frequencies. A new two-dimensional illusion is reported where dynamic modulation of a soft-edged pattern made up from two crossed sinusoids, its power only at the horizontal and vertical orientations, creates a stable hard-edged diagonal lattice with lustrous diamonds as its elements. The illusion was noticed and remarked upon spontaneously by observers (some of them naive) in an experiment on visual evoked potentials where the stimulus producing it was being employed. The purpose of the research reported here was to define more clearly the conditions under which the illusion is seen so as to explain the appearance of sharp diagonal lines and of lustre.

The illusion is seen only at low and medium spatial frequencies and within a middle range of temporal frequencies. Outside that range there is nothing unexpected about the appearance of the pattern. We suggest that both aspects of the illusory structure, sharp diagonal lines and lustre, are created by motion-sensitive mechanisms in the visual system.

2 Methods

2.1 *Stimuli*

The stimulus for this study was produced by adding two sinusoidal gratings, one horizontal and one vertical. Figure 1a is a photograph of the stimulus, illustrating its static appearance. To modulate it the whole pattern was multiplied by a third waveform which varied sinusoidally over time, so that the contrast periodically reversed. The instantaneous luminance profile $L(x, y, t)$ of the waveform is given by

$$L(x,y,t) = L_m + a\sin(2\pi f_t t)[\sin(2\pi f_s x) + \sin(2\pi f_s y)], \qquad (1)$$

where x and y are spatial coordinates, t is time, L_m the mean luminance, a the amplitude, f_t the temporal frequency of contrast reversal, and f_s the spatial frequency of both of the crossed sinusoids. For some of the initial observations the two spatial sinusoids were made to counterphase in variable temporal phase relationships. The more general equation of luminance profile which incorporates a variable phase offset is:

$$L(x,y,t) = L_m + a[\sin(2\pi f_t t + \theta)\sin(2\pi f_s x) + \sin(2\pi f_t t)\sin(2\pi f_s y)], \qquad (2)$$

where θ is the relative temporal phase relationship between the modulation of the two sinusoids.

The waveforms were generated by a computer (Cromemco Z-2D) and displayed on the face of an oscilloscope (Hewlett Packard 1317A) by means of a modified raster technique, at 300 frames s^{-1} and 256 lines $frame^{-1}$. The slow raster, a 300 Hz sawtooth wave, was supplied by the computer, and the fast raster, a 1 MHz triangle wave, by a function generator. On alternate frames, the x and y axes of the oscilloscope were exchanged electronically, giving the appearance of simultaneous orthogonal sinusoidal gratings. The P31 phosphor of the oscilloscope decays to 1% brightness in 0.4 ms, substantially less than the interframe interval of 3.3 ms, thus eliminating the possibility of frame smear. Phase reversal was achieved by multiplying the waveforms by a third sinusoid (also generated by the computer).

Eight cycles of each grating were displayed on the 20 cm × 20 cm screen, giving a spatial frequency of 0.4 cycle cm^{-1}. Spatial frequency at the eye was varied between 0.25 and 10 cycles deg^{-1} by changing viewing distance from 35 cms to 14 m. The mean luminance of the screen was 20 cd m^{-2}.

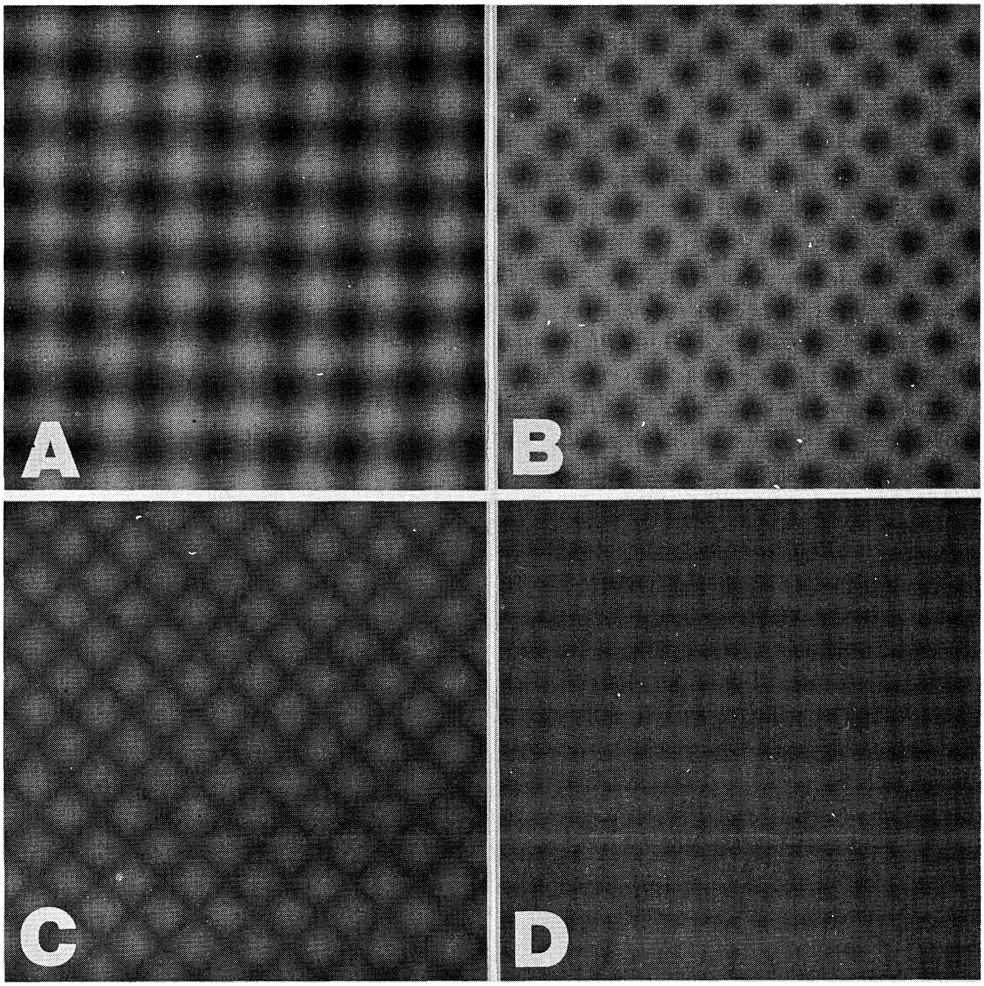

Figure 1. (a) Appearance of the stimulus pattern [of equation (1)] when stationary ($t = \frac{1}{4}f_t$). (b) Luminance profile of the stimulus pattern after performing the logarithmic conversion and temporal integration suggested by Kelly (1966) [equation (3)]. (c) Amplitude of local flicker in the stimulus, expressed as the root-mean-square amplitude [equation (4)]. (d) The same as (c) with θ of equation (2) set to $\frac{1}{2}\pi$.

2.2 *Procedure*

Measurements were made for the minimum contrast necessary to perceive the display as having a diagonal lattice structure (see section 3.1 for a description of the appearance). The minimum contrast required for detection under similar conditions was also measured. Measurements for both conditions were made for seven spatial frequencies, varying from 0.25 to 8 cycles deg^{-1} in octave steps (plus 10 cycles deg^{-1}), and seven temporal frequencies (4, 8, 12, 16, 20, 24, and 30 Hz), by the method of adjustment. Observers turned a handheld potentiometer connected to the computer, which regulated grating contrast by a digital attenuator. If an observer could not see lattice structure at any contrast for a particular condition he pressed an appropriate button. For a given spatial frequency (and therefore viewing distance), the observer was required to make three threshold judgements for each of the seven temporal frequencies, all randomly intermingled. The procedure was repeated five times for each spatial frequency, so that the final judgements represent the mean of fifteen separate observations for each spatial and temporal frequency.

Many observers (some deliberately selected as naive) viewed the display and reported their impressions at various conditions. All agreed on the mean features of the illusion. Extensive threshold measurements were made for two observers, authors DB and JR of this paper. Both observers have corrected 6/6 vision. Additional positive lenses were employed at the close distances to aid accommodation.

2.3 *Simulations*

To illustrate the nature of the illusion, as well as to explore possible explanations, computer simulations were made. These were generated by a Digital computer (PDP-11/60) and displayed on a standard video monitor by means of a frame store (Arluna, TF 4000).

3 Results

3.1 *General observations*

When the pattern defined by equation (1) was caused to counterphase at an appropriate temporal frequency (about 10 Hz), the vertical and horizontal plaid gave way to a diagonal lattice forming diamond-shaped elements. The lattice lines remained continuously in view, and did not jitter or fluctuate in brightness. The diamonds shimmered, could also jitter on one or other diagonal, and were divided from each other by strong stable diagonal lines. It is this dramatic change from a soft-edged vertical–horizontal pattern to a hard-edged diagonal one with lustrous shimmering elements which constitutes the illusion.

What the observers saw is illustrated by the photograph in figure 1c. In the dynamic form of the illusion, however, the diagonal lines dividing the diamonds (depicted as black in figure 1c) were a stable grey colour and sharp, while the diamonds themselves exhibited a shimmering metallic lustre, somewhat akin to binocular rivalry (Helmholtz 1909/1962). It was difficult to judge whether the diamonds or the dividing lines were brighter. Observers' judgements varied considerably, and all complained that the judgement required the comparison of incomparables. Lustre and brightness became confused.

At low temporal frequencies (<4 Hz) the impression of diagonality began to weaken. The pattern was seen instead as a vertical and horizontal plaid, in random apparent motion. The motion was usually oscillatory, varying between vertical, horizontal, and one or other of the diagonals. Similarly, at high spatial frequencies ($\gtrsim 8$ cycles deg^{-1}), whatever the temporal frequency, the diagonality gave way to the plaid pattern. Under some conditions, near the boundary of these two effects, the pattern would alternate between lattice and plaid structure.

When the relative phase of contrast reversal of the two component sinusoids varied, so did the appearance of the display. At a relative phase of $\frac{1}{2}\pi$ [θ in equation (2)] the pattern lost its lattice structure and took on a vertical and horizontal structure, illustrated in figure 1d. The apparent spatial frequency of the pattern, both horizontally and vertically, was double the veridical frequency of the stimulus. An interesting pattern could be generated by counterphasing the two sine waves at slightly different temporal frequencies, say 8 and 8.5 Hz. This created a beat frequency of 0.5 Hz, when the phases of the two temporal waveforms coincided. This pattern alternated from lattice to plaid, at a rate of 1 Hz. Whenever the phase relationship was 0 or θ, it seemed lattice-like, and at $\frac{1}{2}\theta$ and $\frac{3}{2}\theta$ it seemed plaid-like.

3.2 *Thresholds*

To obtain a quantitative measure of the illusion the minimum contrast necessary to see the pattern as lattice-like was measured (for phase θ equal to 0). After some practice observers established a criterion for the contrast necessary for the pattern to have a clear diagonal structure. Measurements were made over a range of spatial and temporal frequencies. When the structure never appeared lattice-like observers pressed a 'no-go' button. For every spatial and temporal frequency, measurements were also made of the minimum contrast for detection of the stimulus.

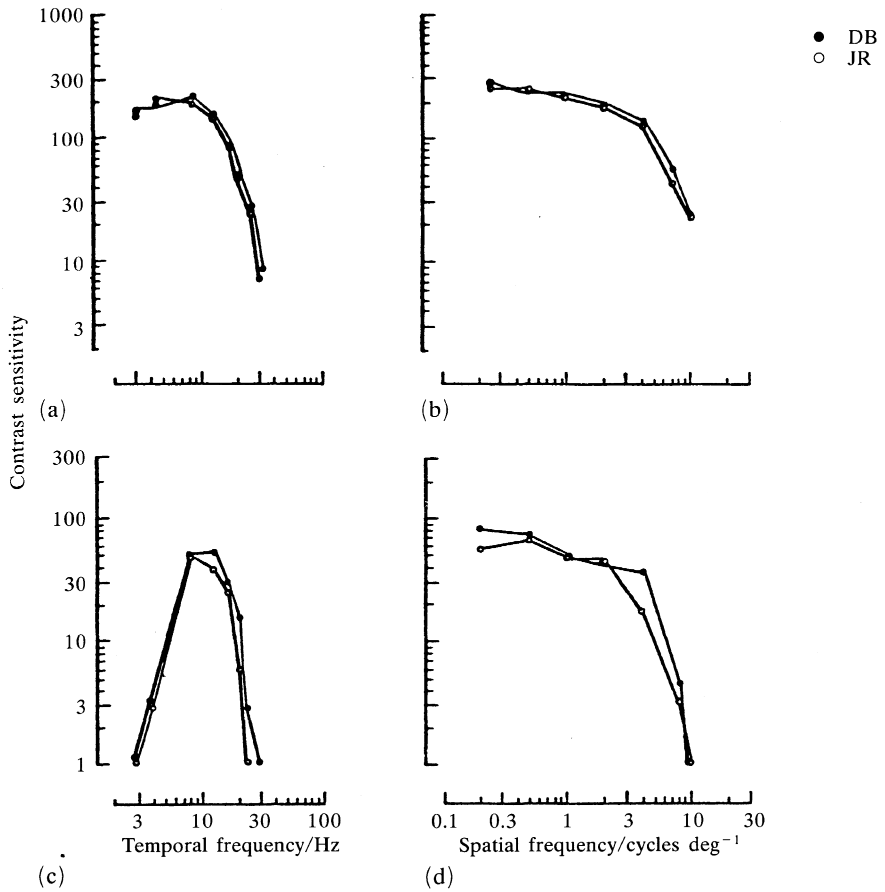

Figure 2. Contrast sensitivity for (a) stimulus detection of two crossed sinusoidal gratings of 1 cycle deg^{-1} as a function of temporal frequency; (b) stimulus detection at 8 Hz as a function of spatial frequency; (c) seeing the lattice illusion at 1 cycle deg^{-1} as a function of temporal frequency; (d) seeing the illusion at 8 Hz as a function of spatial frequency.

Figure 2 shows a sample of the results for the two observers. Figures 2a and 2b show contrast sensitivity for stimulus detection, as a function of, respectively, temporal and spatial frequency (measured at 1 cycle deg^{-1} and 8 Hz respectively). As may be expected, these two curves are similar to the standard temporal and spatial contrast sensitivity curves for single sine wave gratings in counterphase (eg Robson 1966).

Figures 2c and 2d show the contrast sensitivity functions for seeing the lattice illusion. If the illusion was not seen at a particular frequency, that frequency was assigned an arbitrary sensitivity of 1. The temporal curve (figure 2c) is quite sharp, peaking at about 10 Hz and falling sharply for higher and lower temporal frequencies. It differs from the detection sensitivity curve in that it has a pronounced attenuation at low temporal frequencies. The spatial curve (figure 2d) is fairly flat up 4 cycles deg^{-1}, and then falls sharply, far more rapidly than detection sensitivity.

Figure 3 shows the results for all the spatial and temporal frequencies measured. It is given as a contour plot of the suprathreshold contrast (in dB above detection threshold) necessary to see the illusion over a range of spatial and temporal frequencies. The contour lines follow the third-order polynomial interpolation of the experimental data. The lattice structure is best seen in the temporal frequency range 7–20 Hz, and at spatial frequencies below 8 cycles deg^{-1}. Within this range diamonds were visible at a contrast of 15 dB above (six times) detection threshold. Outside this range more contrast was needed to see the effect, which finally vanished above 12 cycles deg^{-1}, above 24 Hz, and below 4.4 Hz.

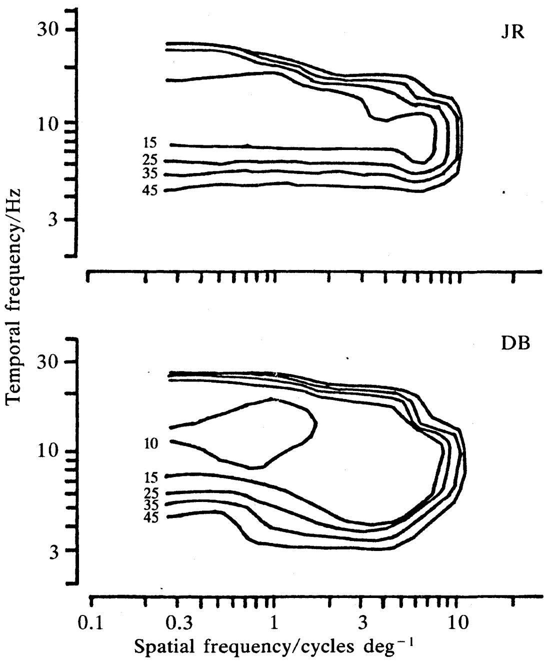

Figure 3. A contour plot of the suprathreshold contrast (expressed in dB) necessary to see the lattice illusion over a range of spatial and temporal frequencies. The averaged raw data were expressed as ratios of threshold for seeing the illusion to detection threshold, converted to dB, smoothed with a third-order polynomial, and plotted as a contour map.

4 Discussion

The illusion reported here is two-dimensional and has two conspicuous elements: the appearance of diagonal lines forming diagonal elements and the lustrous appearance of the elements themselves. To comprehend the illusion we have identified where on the spatiotemporal Fourier plane it occurs. The results make it likely that the illusion shares an explanation with the one-dimensional illusion of frequency doubling described by Kelly (1966), even though the spatiotemporal ranges of the two do not correspond exactly, and one involves a shift of orientation which the other does not. One indication that this is so is that when crossed sinusoidal gratings are independently modulated with a phase difference of $\frac{1}{2}\pi$, they maintain their orientation but appear to double their veridical frequency.

4.1 *Logarithmic compression?*

One possible explanation for both illusions is that the visual system compresses logarithmically then integrates in time, as proposed by Kelly (1966). Figure 1b is a simulation of this model. The luminance profile $L_{\log}(x, y)$ of figure 2b is given by:

$$L_{\log}(x, y) \propto \int_0^{f_t^{-1}} \log[L(x, y, t)]\,\mathrm{d}t\,, \tag{3}$$

where $L(x, y, t)$ is given by equation (1) (with $a/L_m = 0.02$) and f_t is the temporal frequency.

This model is successful qualitatively, in that it introduces the type of spatial distortions we observed. However, there are problems.

First, there is little empirical evidence that the visual system performs an instantaneous logarithmic conversion. Rather, it seems to adapt to varying light levels by changing its gain, but behaves almost linearly having done so [see Shapley and Enroth-Cugell (1984) for an excellent discussion of this point].

Second, although a logarithmic transform will introduce distortion products, they will be quite small for low-contrast stimuli. Figure 2 shows that at the optimal spatial and temporal frequencies the lattice structure was seen with a sensitivity of 50: ie at a contrast of 0.02 (the contrast at which the simulation was performed). The distortion products at this contrast have only one hundredth of the amplitude of the output fundamental, yet sensitivity for the illusion was only about a quarter of that for detection. Thus although a logarithmic transform would introduce distortions of appropriate spatial organization, they would be far too small to be detected at the contrasts where the illusion was apparent (unless one suggests that these are selectively amplified). The same argument applies for frequency doubling: at a contrast of 0.02 the second harmonic distortion products of a single sinusoid also equal one hundredth of the logarithm of the fundamental.

Finally, the observed brightnesses are not consistent with a simple integral of the logarithmic transform. This model suggests that the diagonal lines should appear brighter than the rest of the screen, literally like figure 1b. However, as previously mentioned, the structure was created not by regions of varying brightness, but rather by regions of 'lustre' or 'shimmer' divided by stable lines.

4.2 *Motion rivalry?*

Tyler (1974) has suggested that frequency doubling is due to flicker sensitivity, a suggestion adopted in part by Kelly (1981) as one component of a model containing a linear and a nonlinear element. This provides a second possible explanation for the illusion and one that implicates the motion mechanisms of vision. The temporal sensitivity curve of figure 2c closely resembles temporal tuning functions of motion detectors [as revealed by masking studies (Anderson and Burr 1985)], and the spatial curve of figure 2d indicates that the illusion is strongest at low spatial frequencies, where motion mechanisms prevail (eg Kulikowski and Tolhurst 1973; Burr 1981; Campbell and Maffei 1981).

Sinusoidal phase reversal is mathematically equivalent to the sum of simultaneous motion in opposing directions. Therefore phase reversal of high modulation amplitude should be a strong stimulus for motion detectors. Figure 1c is a description of the amplitude of local modulation. Areas of white indicate high modulation amplitude, and black no modulation. The amount of modulation is expressed as the root-mean-square (RMS) modulation:

$$L_{\mathrm{rms}}(x, y) = f_{\mathrm{t}} \int_0^{f_{\mathrm{t}}^{-1}} \log[L(x, y, t)]^2 \, \mathrm{d}t\,. \tag{4}$$

There is modulation (or flicker in Tyler's terms) everywhere except at the zero-crossings of the crossed sinusoidal pattern. They follow diagonal, not horizontal and vertical lines. The description of local modulation depicts well the spatial structure of the illusion: a lattice of clear diamond shapes divided by strong diagonal lines. However, in this simulation, white signifies maximal modulation amplitude, not maximal luminance. This corresponds to observers' impressions: the diamond-shaped elements do not look brighter than the lines that divide them but more lustrous. The simulation also accurately matches the magnitude of the illusion. The amplitude of the RMS distribution is half the amplitude of the stationary sinusoidal pattern, corresponding well to the amount of suprathreshold contrast necessary to see the illusion.

Figure 1d is a description of the predicted spatial appearance when the two component sinusoids are counterphased with a fixed phase difference of $\frac{1}{2}\pi$ [θ of equation (2) equal to $\frac{1}{2}\pi$]. Again this gives a good match to the appearance of the stimulus, showing the vertical and horizontal structure, and the spatial frequency doubling. Of course, were there only one component sinusoid in the display, this too would appear at twice the veridical spatial frequency.

The simulations show that the magnitude of local modulation predicts the structure of the illusions. The results of the threshold measurements show that the temporal sensitivity functions for the illusion closely resemble the temporal tuning of motion detectors (Anderson and Burr 1985). They also show that the illusion fails at high spatial frequencies, where temporally tuned motion detectors are either absent or have special characteristics (eg Kulikowski and Tolhurst 1973; Watson et al 1980; Burr 1981; Campbell and Maffei 1981).

Flicker stimulates motion detectors of opposite directional preference (eg Levinson and Sekuler 1975). There is now strong physiological and psychophysical evidence that these detectors inhibit each other (eg Barlow and Levick 1965; Goodwin and Henry 1975; Goodwin et al 1975; Dean et al 1980; Stromeyer et al 1984). We suggest that the mutal inhibition generated when detectors of opposing directions are simultaneously stimulated generates a form of 'rivalry', which gives rise to the impression of lustre. Why this should occur is not yet clear, but we note that binocular rivalry produces a similar impression of lustre (Helmholtz 1909/1962). Perhaps it is true more generally that lustre is vision's response to two conflicting signals from one region of the visual field.

References

Anderson S J, Burr D C, 1985 "Spatial and temporal sensitivity of the motion detection system" *Vision Research* **25** 1147 - 1154

Barlow H B, Levick W R, 1965 "The mechanism of directionally selective units in rabbit's retina. II: Directionally selective units" *Journal of Physiology (London)* **178** 447 - 504

Burr D C, 1981 "Temporal summation of moving images by the human visual system" *Proceedings of the Royal Society of London, Series B* **211** 321 - 339

Campbell F W, Maffei L, 1981 "The influence of spatial frequency and contrast on the perception of moving patterns" *Vision Research* **21** 713 - 721

Dean A F, Hess R F, Tolhurst D J, 1980 "Divisive inhibition involved in direction selectivity" *Journal of Physiology (London)* **308** 84

Goodwin A W, Henry G H, 1975 "Direction selectivity of complex cells in a comparison with simple cells" *Journal of Neurophysiology* **23** 1524 - 1540

Goodwin A W, Henry G H, Bishop P O, 1975 "Direction selectivity of simple striate cells: properties and mechanisms" *Journal of Neurophysiology* **38** 1500 - 1523

Helmholtz H von, 1909/1962 *Physiological Optics* volume 3 (New York: Dover, 1962); English translation by J P C Southall for the Optical Society of America (1924) from the 3rd German edition of *Handbuch der physiologischen Optik* (Hamburg: Voss, 1909)

Kelly D H, 1966 "Frequency doubling in visual responses" *Journal of the Optical Society of America* **56** 1628 - 1633

Kelly D H, 1981 "Nonlinear visual responses to flickering sinusoidal gratings" *Journal of the Optical Society of America* **71** 1051 - 1055

Kulikowski J J, Tolhurst D J, 1973 "Psychophysical evidence for sustained and transient neurones in the human visual system" *Journal of Physiology (London)* **232** 149 - 162

Levinson E, Sekuler R, 1975 "The independence of channels in human vision selective for direction of movement" *Journal of Physiology (London)* **250** 347 - 366

Robson J G, 1966 "Spatial and temporal contrast sensitivity function of the visual system" *Journal of the Optical Society of America* **56** 1141 - 1142

Shapley R, Enroth-Cugell C, 1984 "Visual adaptation and retinal gain controls" in *Progess in Retinal Research* eds N N Osborne, G J Chader (Oxford: Pergamon Press) pp 263 - 346

Tyler C W, 1974 "Observations on spatial-frequency doubling" *Perception* **3** 81 - 86

Watson A B, Thompson P G, Murphy B J, Nachmias J, 1980 "Summation and discrimination of gratings moving in opposite directions" *Vision Research* **20** 341 - 347