Observations on spatial-frequency doubling

C W Tyler¶
Department of Psychology, University of Bristol, Bristol 8, England
Received 7 March 1974

Abstract. The involvement of disparity detectors in the perception of spatial-frequency doubling in a flickering grating was tested by viewing the grating horizontally. Frequency doubling was unimpaired, and must therefore occur independently of disparity detection. A distortion model of frequency doubling was compared with a flicker-detection model, in predicting the appearance of nonsinusoidal gratings. The results support the flicker-detection model but not the distortion model. Physiological considerations suggest that spatial-frequency doubling may be mediated by the transient response of the amacrine cells in the retina.

1 Effects of disparity

1.1 Introduction

If a sinusoidal grating is alternated with an antiphase grating of the same frequency, spatiotemporal flicker is produced. At low spatial frequencies and high flicker rates an impression of a grating of double the original spatial frequency is perceived (Kelly, 1966).

Richards and Felton (1973) have used disparity-specific adaption to spatiotemporal gratings to investigate the physiological locus of spatial-frequency doubling. They report that disparity-specific adaptation occurs to the fundamental but not the doubled frequency of a spatiotemporal grating, even though the doubled frequency is predominantly perceived. They interpret this result to imply either that frequency doubling occurs more centrally than disparity processing, or that the two processes are independent.

I suggest a simple test to distinguish between these two hypotheses. If the spatiotemporal grating is viewed with the bars horizontal rather than vertical, it contains no elements that can stimulate the disparity system. If spatial-frequency doubling is linked with the disparity system, it should disappear or be less marked for a horizontal than a vertical spatiotemporal grating. This test was set up on an oscilloscope, using sinusoidal luminance gratings with a contrast of 0·6, at a spatial frequency of 1 cycle deg^{-1} multiplied electronically by a temporal sinusoidal modulation at a rate of 20 Hz. The possibility of distortion in the equipment was eliminated by ensuring that above flicker fusion the stimulus appeared uniform.

1.2 Results

Four observers reported that horizontal and vertical gratings gave an equal degree of spatial-frequency doubling. This result is also implicit in the original report of Kelly since the patterns he used were circularly symmetric Bessel functions. Furthermore, doubling was subjectively unaffected by monocular viewing of either horizontal or vertical gratings, which is unlikely if doubling requires disparity processing.

The observation of Richards and Felton that the doubled grating appeared at the oscilloscope face rather than at a disparate fixation point is not hard to explain. A repetitive pattern can appear at any depth corresponding to displacement by an

integral number of the repetition periods. If one of these depths coincides with other cues such as the edge of the oscilloscope screen, it is likely to be preferred over other depths. Thus the frequency-doubled grating probably appeared on the screen rather than at the depth of the fixation point because the edges of the screen made the best resolution of its otherwise ambiguous depth cues.

2 Models of spatial-frequency doubling

2.1 *The distortion model*
Richards and Felton (1973) accept the model of spatial-frequency doubling presented by Kelly (1966) in detailed mathematical form. This model essentially proposes that the doubled frequency is the residual from incomplete cancellation between the two opposite-phase waveforms alternated in time. Cancellation is incomplete because the spatial waveforms have been distorted from the pure sinusoidal input by the logarithmic transformation proposed by de Lange (1957). The form of a sinusoidal grating stimulus $s(x)$ is given by

$$s(x) = 1 + m \cos \alpha x , \tag{1}$$

where m is the modulation depth and α is a constant determining spatial frequency.

The integrated visual response $r(x, t)$ to spatiotemporal alternation in the distortion model is given by Kelly [1966, equation (6c)] as

$$r(x, t_0) + r(x, t_\pi) = -m^2 \cos^2 \alpha x + e \approx -\tfrac{1}{2}m^2(1 + \cos 2\alpha x) = -\tfrac{1}{2}m^2 - \tfrac{1}{2}m^2 \cos 2\alpha x , \tag{2}$$

e being a negligible factor resulting from higher harmonics.

Four predictions may be derived from this model, each of which can be tested experimentally:
(i) The first term in equation (2) determines the average brightness level of the spatiotemporal grating. If the fact that this is negative is ignored for a moment, because the sign of a neural response is difficult to interpret, there is nevertheless a remarkable prediction for the absolute magnitude of the brightness. As the modulation is reduced the average brightness of the stimulus is also reduced, and the entire stimulus would become invisible. When $m = 0$, $r(x, t) = 0$. This is a strong prediction which can be tested by direct observation. In fact my observers reported little perceptible change in the average brightness level as modulation depth was varied. At high modulation depths the judgement was difficult, for the flickering area seemed paradoxically both light and dark simultaneously, but the average brightness did not appear reduced at low modulation depths.
(ii) Modulation amplitude is given by the multiplier of the cosine term, $\tfrac{1}{2}m^2$. Sensitivity for spatial-frequency doubling may thus be derived from that for static gratings. If we assume that modulation threshold is unaffected by the perceived flicker and occurs at a contrast of $0 \cdot 01$, doubling should become invisible when the modulation depth m reaches $0 \cdot 14$ [equation (2)], which is more than a log unit above threshold. But both Kelly and Richards and Felton report that doubling persists right down to threshold.
(iii) The waveform of doubling is predicted by equation (2) and is always sinusoidal. But my observations concur with those of Richards and Felton that the waveform is usually a distorted sinusoid, with narrow nodes and wide antinodes.
(iv) By symmetry the model predicts the appearance of temporal frequency doubling, but neither Kelly nor my observers were able to observe such a phenomenon.

2.2 *A flicker-detection model*
The failure of these predictions led me to consider an alternative model of spatial-frequency doubling based on the detection of flicker. Kelly describes the doubling both in terms of flicker nodes and antinodes, and in terms of contrast variations.

A model such as the one he proposes can account for contrast variations at large modulation depths, which do seem to have the characteristics required by predictions (ii) and (iii).

However, the spatial distribution of flicker in spatiotemporal gratings has frequency-doubling characteristics which are more simply explained by a simple flicker detector. Assume that by small-signal linearisation the spatial waveform of the response $r(x)$ approximates the input modulation $[m \cos(\alpha x)]$ and that in the steady state the response to the mean level of illumination is adapted out to zero. When the waveform is alternated in phase form $r_1(x)$ to $r_2(x)$ (figure 1a) the amplitude of flicker $f(x)$ in retinal space x is given by

$$f(x) = |r_1(x) - r_2(x)| = |m \cos(\alpha x) - m \cos(\alpha x + \pi)| = |2m \cos(\alpha x)| . \tag{3}$$

The modulus of a sine wave (figure 1b) has peaks at twice the frequency of the original wave.

Human observers are able to estimate the modulation depth m of flicker with good accuracy, conforming to an approximately linear function (or a power law with an exponent close to $1 \cdot 0$; Regan and Beverley, 1973). This suggests the existence of a flicker detection mechanism in the brain with an output $f(x)$ which is linear to a first approximation. The output of such a mechanism is shown in figure 2 for

(a)   (b)
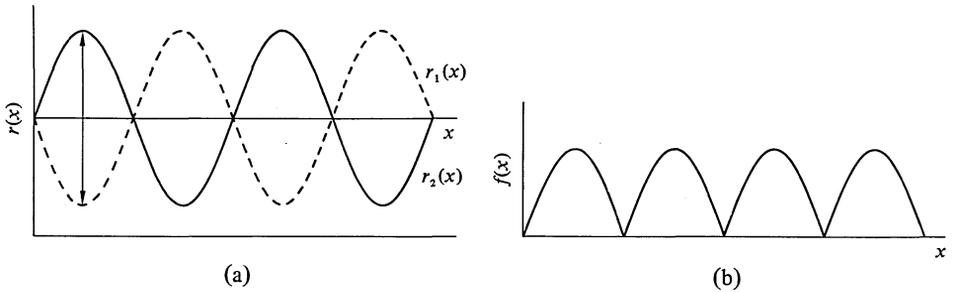
Figure 1. (a) Waveform in retinal space x of visual response $r(x)$ at the site of temporal combination. Waveform is plotted at time t_1 (dashed line) and half a period later at time t_2 (full line). Note that literal 'frequency doubling' occurs in the combined envelope of the waveform. (b) Waveform in retinal space of the output $f(x)$ of a hypothetical flicker detection mechanism whose input is $r(x)$. The output $f(x)$ is a full-wave rectification of the instantaneous input $r(x)$, whose fundamental frequency is double the input frequency. $f(x)$ also has a characteristic waveform which corresponds well with the perceived waveform.

(a)   (c)

(b)
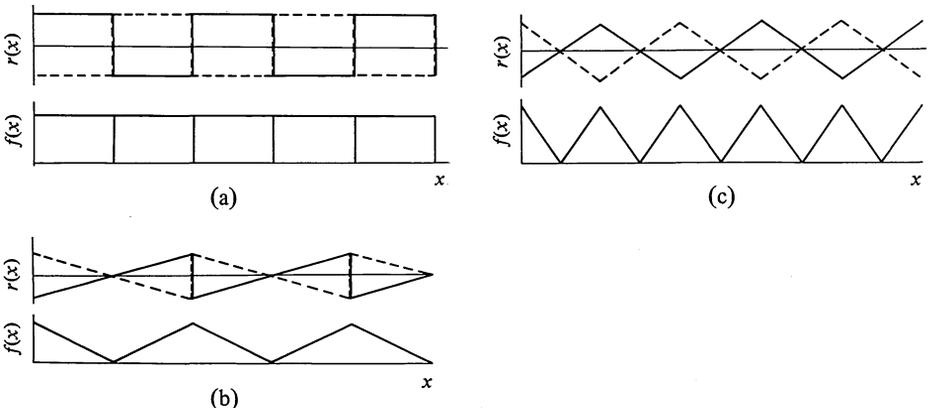
Figure 2. Waveforms of the input $r(x)$ and output $f(x)$ of a hypothetical flicker detector, as in figure 1: (a) rectangular wave stimulus; (b) triangular wave stimulus; (c) sawtooth stimulus. Note that the flicker detector output differs markedly from the instantaneous stimulus input.

temporally-alternated rectangular, triangular, and sawtooth wave inputs. Note that this output corresponds to the difference between the antiphase inputs, rather than their sum as in a distortion model.

3 Experimental observations

The same procedure and apparatus as described above was used to test five predictions from the flicker-detection model. Observers were asked for spontaneous descriptions of relevant aspects of each stimulus alternated at 20 Hz in successive comparison with the static waveform. An adhesive marker was used to determine the spatial phase of phenomena in the alternated stimulus relative to the static stimulus.
(i) The rectangular wave input gives predictions which are similar to those of the distortion model, namely a uniform flicker, with slight perturbations at the boundaries of the changeover if the waveform is not perfectly square. Kelly does not seem to have observed this stimulus, but my observers confirm these predictions.
(ii) A sinusoidal input to the flicker detector should produce a 'wicker-fence' output of predominantly double the input frequency as is characteristic of full-wave rectification (figure 1b). This output should be relatively unaffected by logarithmic distortion of the input waveform, because in the differencing operation of flicker detection the even harmonics resulting from logarithmic distortion cancel to zero. Observation of sinusoidal flickering gratings suggested that the wicker-fence waveform with narrow bands of steady field and wide bands of flicker is a good description of the doubled waveform. It was noted that the flicker was somewhat more uniform across the top of the waveform than predicted by the model, but such an effect might well occur if flicker detection had a saturation limit.
(iii) For a triangular input a contrasting prediction is obtained from the two models. The flicker-detection model predicts a difference waveform that is also triangular, although it has twice the spatial frequency of the stimulus. This means that the spatial rate of increase in flicker in any cycle is more gradual than for a sinusoidal waveform. All subjects reported that the bands of invariant brightness appeared much wider and the edges of the flickering bands looked much smoother. The distortion model does not predict a difference in appearance between sine and triangular waveform.
(iv) For the sawtooth waveform the two models make opposite predictions, although both specify that frequency doubling should not occur. In the flicker-detection model the maxima occur around the crossover of the fast phase, whereas in the distortion model the maxima occur at the crossover of the slow phases, since the logarithmic distortion bulges both curves upward maximally at this point. The observers all reported flicker maxima at the crossover of the fast phases, and little variation in contrast.
(v) Finally, the flicker detection model does not postulate the existence of a spatial-grating detector as well, and therefore does not imply that temporal-frequency doubling should occur. Since temporal-frequency doubling has not been observed in alternating gratings despite effort to observe it, the theory is in accord with the observations.

4 Discussion

4.1 *Independence of fundamental- and doubled-frequency channels*

The experimental observations conform well to the flicker-detection model of frequency doubling, which in turn suggests an explanation for Richards and Felton's results. The doubled frequency failed to adapt disparity detectors because flicker detection occurs in a separate channel from static contrast detection, and *ex hypothesi* only static contrast can act as an input for disparity detectors. No physiological locus for flicker detection, and hence spatial-frequency doubling, is

specified by either set of experiments. One classic test of physiological locus is to measure binocular summation for the phenomenon of interest. Sherrington (1906) established that the fusion frequency for binocular flicker was less than 10% greater when the flicker was in phase than when it was in antiphase in the two eyes. Similarly, Tyler and Torres (unpublished observations) have found that the de Lange curve of flicker sensitivity is largely unaffected by the interocular phase relationship. Clearly if brightness summation occurred after binocular combination, antiphase flicker should be essentially invisible. On the other hand, if flicker detection occurred prior to binocular combination no effect of phase would be predicted. Binocular flicker observations therefore support the hypothesis that flicker detection and therefore spatial-frequency doubling occurs in monocular channels.

4.2 Physiological correlates
The proposed flicker detector has two salient properties: (a) its response to the mean level of illumination should be zero in the steady state, and (b) its response to changing illumination should rectify the input modulation [equation (4)]. In considering physiological correlates it is simpler to translate these properties of responses to step changes in illumination, for which condition (a) is met if the response is transient rather than sustained, and hence adapts out quickly to steady illumination, and condition (b) is met if the response is similar for both increases and decreases with illumination.

One need look no further than the ganglion cells of the retina to find responses which meet conditions (a) and (b). Hartline (1938) found that of the three categories of frog optic nerve fibres (which are ganglion cell axons) one showed similar ('on-off') responses to light onset and offset, and the responses were transient. Werblin and Dowling (1969) showed by means of intracellular recording in the mudpuppy retina that this type of rectified, transient response originates from the amacrine cells in the inner nuclear layer of the retina. Since the structure of the retina differs little from mudpuppy to monkey, and since 'on-off' responses have been recorded in higher species such as cat (Kuffler, 1953) and spider monkey (Hubel and Wiesel, 1960), it is not unreasonable to suppose that similar retinal organisation for flicker detection exists in man (although other responses may be elicited by a flickering stimulus under some conditions).

4.3 Psychophysical corollaries
The identification of flicker detection with the transient response of the amacrine cells gives rise to some interesting predictions with regard to the spatiotemporal limits of the doubling. The response should be attenuated at low temporal frequencies, as these approximate a steady illumination to which adaptation will occur. The spatial extent of the amacrine cells across the retina is large, so that the response should have poor spatial resolution. The predicted range of spatial-frequency doubling is therefore limited to low spatial frequencies but high temporal frequencies, which is exactly the range in which Kelly (1966, figure 5) reports its occurrence.

Of course a flickering stimulus will also tend to stimulate the sustained response system to an extent depending on spatiotemporal frequency (Kulikowski and Tolhurst, 1973; Tyler, 1974). But even if the sustained response can still follow the stimulus at the high rates of flicker involved, this will merely add in a non-doubled component to the response, which may well supply an explanation for the somewhat paradoxical quality of the doubled-frequency percept. It may also underlie the gradual change in perceived spatial frequency with increased temporal frequency observed by Richards and Felton (1973).

4.4 Temporal-frequency doubling?
An apparent problem arises with the fact that the rectified amacrine cell potential gives a response at twice the frequency of the input variation. This immediately

seems to imply the presence of temporal-frequency doubling, which was specifically excluded in the model. However, it is clear that the observer would only perceive this as temporal-frequency doubling if as a comparison there were conditions under which such doubling did not occur. The natural comparison to make is between the spatiotemporal grating and a uniform flickering field. But these are in general both appropriate to stimulate the amacrine cells, and indeed the flicker looks similar in both conditions.

One can go further and turn the question of temporal-frequency doubling upside down. If the flicker detector involves a rectifying response, uniform field flicker should always be perceived at twice the frequency of the input. But by designing conditions to eliminate the amacrine cell response it should be possible to obtain temporal-frequency *halving* relative to uniform field flicker. Temporal-frequency halving should then be predicted in two regions, either when the spatial frequency is so high as to be beyond the resolution of the amacrine field, or when the temporal frequency is low enough to allow transient response to adapt out.

Referring to Kelly's (1966, figure 5) diagram of percepts in spatiotemporal gratings, apparent motion is actually perceived in the regions in which temporal-frequency halving is predicted. The test is therefore not a good one, but observers report that even when fixating a stationary spot the perceived movement gives much smoother and slower impression than the flicker of spatial-frequency doubling at the same rate of temporal alternation. The observations therefore weakly support the prediction of temporal-frequency halving at high spatial or low temporal frequencies.

References
de Lange, H., 1957, *Attenuation Characteristics and Phase Shift Characteristics of the Human Fovea-cortex Systems in Relation to Flicker-fusion Phenomena,* Thesis, Technical University of Delft.
Hartline, H. K., 1938, "The response of single optic nerve fibres of the vertebrate eye to illumination of the retina", *American Journal of Physiology,* **121**, 400–415.
Hubel, D. H., Wiesel, T. N., 1960, "Receptive fields of optic nerve fibres in the spider monkey", *Journal of Physiology,* **154**, 572–580.
Kelly, D. H., 1966, "Frequency doubling in visual response", *Journal of the Optical Society of America,* **56**, 1628–1633.
Kuffler, S. W., 1953, "Discharge patterns and functional organisation of mammalian retina", *Journal of Neurophysiology,* **16**, 37–68.
Kulikowski, J. J., Tolhurst, D. J., 1973, "Psychophysical evidence for sustained and transient detectors in human vision", *Journal of Physiology,* **232**, 149–162.
Regan, D., Beverley, K. N., 1973, "The relation between the magnitude of flicker perception and evoked potential amplitude in man", *Perception,* **2**, 61–65.
Richards, W., Felton, T. N., 1973, "Spatial frequency doubling: retinal or central?", *Vision Research,* **13**, 2129–2138.
Sherrington, C. S., 1906, *The Integrative Action of the Nervous System* (Yale University Press, New Haven).
Tyler, C. W., 1974, "Analysis of visual modulation sensitivity: two components in flicker perception", *Vision, Research,* in press.
Werblin, F. S., Dowling, J. E., 1969, "Organisation of the retina of the mudpuppy, *Necturus maculosis*: II. Intracellular recording", *Journal of Neurophysiology,* **32**, 339–355.

p