

SHORT NOTE

OPTIMAL DISPLACEMENT IN APPARENT MOTION AND QUADRATURE MODELS OF MOTION SENSING

ANDREW B. WATSON

Vision Group, NASA Ames Research Center, Moffett Field, CA 94035, U.S.A.

(Received 20 October 1989; in revised form 14 February 1990)

Abstract—A grating appears to move if it is displaced by some amount between two brief presentations, or between multiple successive presentations. A number of recent experiments have examined the influence of displacement size upon either the sensitivity to motion, or upon the induced motion aftereffect. Several recent motion models are based upon quadrature filters that respond in opposite quadrants in the spatiotemporal frequency plane. Predictions of the quadrature model are derived for both two-frame and multiframe displays. Quadrature models generally predict an optimal displacement of 1/4 cycle for two-frame displays, but in the multiframe case the prediction depends entirely on the frame rate.

Psychophysics Motion perception Apparent motion Motion aftereffect Stroboscopic motion
Temporal frequency Spatial frequency Quadrature Direction selectivity Phase Filters

INTRODUCTION

In a recent paper, Baker, Baydala and Zeitouni (1989; see also Turano & Pantle, 1985) explored the dependence of the motion aftereffect (MAE) upon the spatial displacement between successive presentations of a grating. They expressed their results in terms of a D_{opt} , the displacement producing the largest aftereffect, expressed as a fraction of the spatial period of the grating. They found that, for a range of spatial frequencies between 0.21 and 4.8 c/deg, this optimal displacement was slightly less than one quarter of a spatial period. They interpret this result as supportive of quadrature models of motion sensing, in which direction selectivity is produced through combination of receptive fields in spatial and temporal quadrature phase (Watson & Ahumada, 1983, 1985; Watson, Ahumada & Farrell, 1986; van Santen & Sperling, 1984, 1985; Adelson & Bergen, 1985). The purpose of this note is to analyze predictions of the quadrature model for this multiple-frame display, and also for the two-frame display (Nakayama & Silverman, 1985; Westheimer, 1978). To summarize the results, the quadrature models does not predict a 1/4 period optimal displacement in multiframe displays, but does in the case of a two-frame display.

THEORY

The stimulus of Baker et al. is a grating in "staircase motion": the position of the grating

is a staircase function of time. This is closely related to a stroboscopic presentation, in which the grating is flashed briefly at each position, rather than for the full interval between jumps. The behavior of the quadrature model is most easily understood in the frequency domain (Watson & Ahumada, 1983), so we first consider the spatiotemporal Fourier spectra of stroboscopic and staircase motion. These have been derived previously (Watson et al., 1986). For the present case, where the displacement is a fraction k of the spatial period, and where the strobe rate is ω_s , the spectra are as shown in Fig. 1.

The stroboscopic spectrum is a pair of impulses at $\{u_0, -\omega_0\}$ and $\{-u_0, \omega_0\}$, where

$$\omega_0 = k\omega_s. \quad (1)$$

This impulse pair is replicated at intervals of ω_s . In the main experiment of Baker et al. $\omega_s \approx 15$ Hz.

The staircase stimulus is the stroboscopic stimulus convolved with a time pulse of duration equal to the sampling interval $1/\omega_s$, and thus its spectrum is the stroboscopic spectrum multiplied by $\text{sinc}(\omega/\omega_s)$, as drawn on the left side of Fig. 1. This acts primarily as a lowpass filter, and attenuates the spectral replicas. For the qualitative predictions discussed here, this lowpass filtering is of little consequence.

Superimposed on this spectrum is a schematic indication of the *spectral receptive fields* for quadrature motion sensors for rightward and

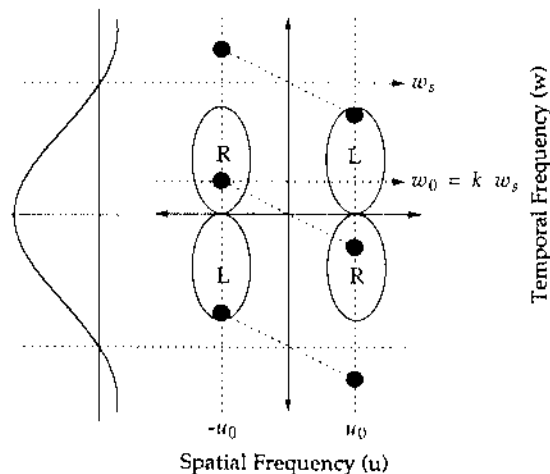


Fig. 1. Fourier spectrum of a stroboscopically presented rightward-moving grating. The spatial frequency is u_0 , and the displacement between frames is a fraction k of the spatial period. The frame rate is w_s . Each large dot represents an impulse. The fundamental spectrum lies at w_0 and $-w_0$. The fundamental spectrum is replicated at intervals of w_s . To obtain the staircase spectrum, the stroboscopic spectrum is multiplied by the sine function drawn on the left. The shaded oval regions indicate the spectral receptive fields of leftward- and rightward-selective quadrature motion sensors.

leftward motion (Watson & Ahumada, 1983). The spectral receptive field is the Fourier transform of the spatiotemporal receptive field, and describes those spectral regions in which a sensor responds. Note that each sensor is selective for two bands of frequency arrayed diametrically about the origin, each band being narrow in spatial frequency and broad in temporal frequency. The details of bandwidth are not important for this discussion.

In the configuration shown, the fundamental spectral components vigorously excite the rightward sensor, while the replicas just tickle the leftward sensor, and consequently a robust rightward apparent motion (and presumably leftward MAE) would result. With this introduction, we can now ask the following questions.

(1) Does the quadrature model predict $k = 1/4$?

Clearly not. The quadrature model predicts that a robust rightward excitation (and presumably leftward MAE) will occur when the fundamental spectrum (at $\pm w_0$) falls within the rightward spectral receptive field, and not too much of the replicas falls within the leftward spectral receptive field (as in Fig. 1). If the peak of the spectral receptive field is at w_m Hz, then the best k will be w_m/w_s . Thus the optimal k can assume just about any value, and equals $1/4$

only for a particular value of the frame rate w_s . As an illustration, consider repeating Baker et al.'s experiment at a frame rate of 1 kHz. If $w_m = 5$ Hz (Baker et al., 1989; Pantle, 1974), then the quadrature model predicts an optimal displacement of $k = 0.005$. On the other hand, if $k = 1/4$, the stimulus would have no energy below 250 Hz! In short, despite the fact that the quadrature model employs submechanisms in quadrature phase, it has no special preference for quadrature phase in successive frames of a strobed or staircased grating.

(2) Why do the data of Baker et al. show an optimal displacement of nearly $1/4$?

The main experimental manipulation of Baker et al. was to vary k , which is equivalent to varying w_0 . In Fig. 1, increasing k amounts to increasing the slope of the dotted line connecting each pair of impulses. A preference for a particular value of k is equivalent to a preference for a particular temporal frequency. In the data of Baker et al., the preference for $k \approx 0.2$ indicates a preference for stimulation at $w_0 \approx 14.7 \times 0.2 \approx 3$ Hz. Thus the most natural explanation for the optimum at $k = 0.2$ is a peak sensitivity of the sensor at 3 Hz.

However, Baker et al. were aware of a possible confounding of their result with a preference for a particular temporal frequency. To control for this, they repeated their experiment with frame rates of 7.5 Hz and 60 Hz. In the latter case, they found *no optimum* in the range of frequencies between 0.44 and 4.8 Hz. They take this as evidence against a strong preference of the motion sensor for a particular temporal frequency, and thus an argument against the theory presented in the previous paragraph for the optimum at around $k = 0.2$.

But this is peculiar. Why should the same experiment produce a clear optimum around 3 Hz when the frame rate is 15 Hz, but not when it is 60 Hz? The only difference is in the location of the spectral replicas, as shown in Fig. 2.

This figure provides an explanation: as k is increased, the replicas invade the spectral receptive field of the opposite direction, the ambiguity of the sensor responses increase, and the MAE declines. On the other hand, when the frame rate is much higher, the replicas are out of harm's way, and so long as the fundamental spectrum falls within the rightward spectral receptive field, an MAE will result (Fig. 2). Thus not only is the optimum at $k = 0.2$ not due to the "quadrature" nature of the sensor, but it is

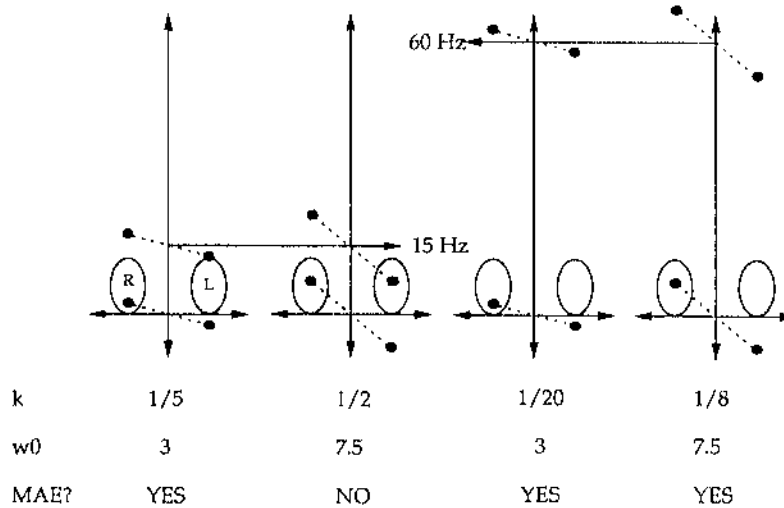


Fig. 2. An explanation of why there is an optimal value of w_0 when the frame rate is 15 Hz, but not when it is 60 Hz. In the second panel, MAE does not occur because the aliasing replica stimulates the leftward sensor. In all other cases, only the rightward sensor is stimulated.

also not due to the temporal tuning of the motion sensor. Instead it is due to the sampling artifacts produced by a low frame rate. A puzzle remains as to why *reducing* k produces a decline in MAE at 15 Hz, but not at 60 Hz frame rate. However, this decline does not occur in the very similar experiments of Turano and Pantle (1985).

Little is reported about the results of the 7.5 Hz frame rate, except that there was "almost no effect" on the optimal displacement. Given that this low frame rate will produce rather serious aliasing (Fig. 1), it is difficult to interpret this result.

(3) *What about optimal displacement in two-frame experiments?*

Nakayama and Silverman (1985) measured contrast sensitivity to a stimulus consisting of two successive pulses of grating, each of

duration T , with some displacement between them. In one experiment, observers were asked to discriminate the direction of the displacement. Sensitivity was found to follow a sinusoidal function of the displacement, expressed as a phase angle, with a peak at 90 deg (1/4 cycle). Does the quadrature model predict $k = 1/4$ in this case of two-frame apparent motion?

We represent the two-frame stimulus as a pair of time pulses at times 0 and T , each of duration T , respectively multiplying gratings of phases 0 and ϕ . For consistency with the earlier discussion, we note that $k = \phi/2\pi$. This has a Fourier transform

$$\frac{1}{2} [\delta(u - u_0) + \delta(u + u_0)] \times [1 + e^{-i2\pi(uT + k \text{sgn}(u))}] \text{sinc}(wT). \quad (2)$$

This consists of two line impulses at u_0 and $-u_0$, modulated along their length (in the temporal frequency dimension) by 1 plus a complex exponential, multiplied by a sinc function. This modulation is shown in Fig. 3.

Because the spectrum falls in both rightward and leftward spectral receptive fields, we must go a little beyond the qualitative predictions of the multiframe case. In essence, we must consider how later stages of the motion sensing apparatus make use of the output of the quadrature filters. We consider two cases. The first assumes that direction is indicated by the difference between "energies" (squared magnitudes) of the signal passed by rightward and leftward filters (Adelson & Bergen, 1985; van

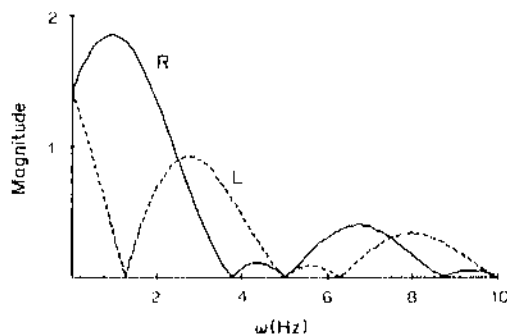


Fig. 3. Examples of magnitude spectra in rightward (R) and leftward (L) quadrants of the Fourier plane, as computed from equation (2). In this example, $T = 0.2$ sec and $k = 1/4$.

Santen & Sperling, 1984, 1985),

$$\frac{1}{4} [\text{sinc}(wT)]^2 \left[|1 + e^{-i2\pi wT - k}|^2 - |1 + e^{-i2\pi wT + k}|^2 \right]. \quad (3)$$

This simplifies to

$$[\text{sinc}(wT)]^2 \sin(2\pi wT) \sin(2\pi k). \quad (4)$$

This clearly indicates that directionality should be maximized at $k = 1/4$, as found by Nakayama and Silverman (1985). These authors also found that the contrast sensitivity for discrimination of direction had a sinusoidal dependence upon k . However, the energy is proportional to contrast squared, so this model predicts a contrast sensitivity that is proportional to the square root of equation (4). This prediction is shown by the outermost curve in Fig. 4.

The second case we consider is what I will call here the *magnitude* model (Watson & Ahumada, 1985). It consists of finding the temporal frequency at which the largest magnitude occurs, and subtracting the magnitudes at that frequency for left and right sensors. Computed values for this algorithm result in the inner solid curve in Fig. 4. As can be seen, they are very close to the nearly sinusoidal data found by Nakayama and Silverman (1985). The slight rightward skew in the predictions of this model agree with a similar skew in the data, though this is of questionable significance.

More generally, it is clear from equation (4) that a model which relies on magnitudes, rather than squared magnitudes ("energies") will better approximate the sinusoidal form of these data. This is evidently an argument against so-called energy models. However, it is clear that both models, which employ quadrature

filters, predict an optimal displacement of about 1/4 period in two-frame motion.

This conclusion is somewhat dependent upon the details of the stimulus. If the two flashes are very brief, the optimum moves towards 0.5. For example, when $T = 0.01$ sec, the optimum is at about 0.4 for the magnitude model, but remains at 1/4 for the energy model. This may provide another method of distinguishing between the models.

SUMMARY

In this note I have discussed predictions of quadrature motion models (Watson & Ahumada, 1983, 1985; Watson et al., 1986; van Santen & Sperling, 1984, 1985; Adelson & Bergen, 1985) for optimal grating displacement. In multi-frame presentations, the optimal displacement depends entirely on the frame rate, and has no particular preference for 1/4 cycle. In two-frame displays, the quadrature model usually does predict a value near to 1/4, but this may depend upon the stimulus or upon details of the model that go beyond the quadrature nature of the motion filter.

Acknowledgements - I thank Lee Stone for useful comments and Jeff Mulligan for suggesting a derivation of equation (4). I also thank the participants in the lunchtime discussion group for allowing me to distract them from their usual subjects.

REFERENCES

- Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2), 284-299.
- Baker, C. L. Jr, Baydala, A. & Zeitouni, N. (1989). Optimal displacement in apparent motion. *Vision Research*, 29(7), 849-859.
- Nakayama, K. & Silverman, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America A*, 2(2), 267-274.
- Pantle, A. (1974). Motion aftereffect magnitude as a measure of the spatiotemporal response properties of direction-sensitive analyzers. *Vision Research*, 14, 1229-1236.
- van Santen, J. P. H. & Sperling, G. (1984). A temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, 1(5), 451-473.
- van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2(2), 300-321.
- Turano, K. & Pantle, A. (1985). Discontinuity limits for the generation of visual motion aftereffects with sine- and square-wave gratings. *Journal of the Optical Society of America A*, 2(2), 260-266.

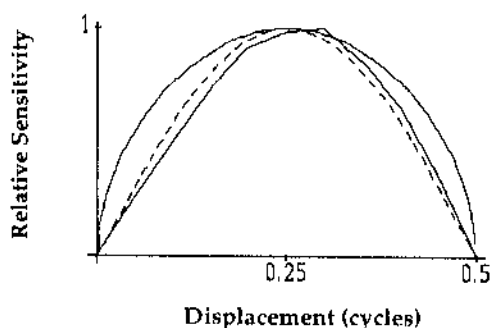


Fig. 4. Contrast sensitivity for direction discrimination as a function of grating displacement in cycles for two models. The outer solid curve is the "energy" model, the inner curve is the magnitude model. The dashed curve is a sinusoid for comparison.

- Watson, A. B. & Ahumada, A. J. Jr (1983). A look at motion in the frequency domain. In Tsotsos, J. K. (Ed.), *Motion: Perception and representation*. New York: Association for Computing, Machinery (also published as NASA Technical Memorandum 84352, 1983).
- Watson, A. B. & Ahumada, A. J. Jr (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2(2), 322-342.
- Watson, A. B., Ahumada, A. J. Jr & Farrell, J. (1986). Window of visibility: Psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A*, 3(3), 300-307 (also published as NASA Technical Paper 2211, 1983).
- Westheimer, G. (1978). Spatial phase sensitivity for sinusoidal grating targets. *Vision Research*, 18, 1073-1074.