NASA Technical Memorandum 84352

A Look at Motion in the Frequency Domain

Andrew B. Watson and Albert J. Ahumada, Jr.

April 1983

Also published as "A look at motion in the frequency domain" by Andrew B. Watson and Albert J. Ahumada, Jr., in *Motion: Representation and Perception*, ACM, Baltimore, 1983 (ISBN 0-89791-100-8).

National Aeronautics and
Space Administration

A LOOK AT MOTION IN THE FREQUENCY DOMAIN

Andrew B. Watson and Albert J. Ahumada, Jr.

NASA Ames Research Center, Moffett Field, California 94035

When a phenomenon seems particularly puzzling, it often helps to view it from a new perspective. The puzzle we will examine is the visual perception of motion, and the new perspective will be the frequency domain. Our precedent here is of course the work of Robson and others [Robson, 1966; Blakemore and Campbell, 1969; Campbell and Robson, 1968] who showed that many aspects of spatial vision could be better understood in the frequency domain.

Our plan is as follows. First we will examine the frequency spectra of moving images and note some of their essential properties. Second we will show an example of how this perspective can provide simple solutions to long-standing problems in motion perception. Finally we will construct a candidate motion detector whose behavior is most easily understood in the frequency domain.

1. Frequency transforms of moving images

Consider some image moving at constant velocity along a straight path. The image contrast distribution can be written

$$c(x,y,t) = c(x,y)\,\delta(x - r_x t)\,\delta(y - r_y t) \tag{1}$$

where x and y are horizontal and vertical image coordinates and t is time. The spatial contrast distribution at time 0 is $c(x,y)$, and r_x and r_y are the velocity components in x and y dimensions. To obtain the spatiotemporal frequency spectrum of this image we take the three-dimensional Fourier transform of Eq. 1,

$$C(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(x,y,t)\,e^{-i2\pi(xu+yv+tw)}\,dx\,dy\,dt \tag{2}$$

where u, v, and w are horizontal spatial, vertical spatial, and temporal frequency respectively. A little effort will show that this spectrum is

$$C(u,v,w) = C(u,v)\,\delta(w + ur_x + vr_y) \tag{3}$$

This function is simple, but somewhat difficult to picture, so we consider for the moment the case in

which motion and spatial variation occur only along the x axis, in which case we can ignore the y dimension. The contrast distribution and frequency spectrum for this case are shown in Fig. 1.

Note the following properties of this spectrum:

(1) it lies along a line of slope $-1/r$ in the u,w space. Thus the higher the spatial frequency component of the image, the higher its temporal frequency. The higher the velocity, the shallower is the slope of the spectrum.

(2) It is symmetric about the origin, and occupies two diagonally opposite quadrants. If motion is to the right, these are the even quadrants, if it is to the left, they are the odd quadrants.

We will now show how this spectrum provides insight into motion phenomena.

2. Stroboscopic apparent motion

If an image is presented briefly and rapidly at a sequence of closely spaced positions, it may appear indistinguishable from a smoothly moving image. This phenomenon of stroboscopic apparent motion underlies movies and television, and is a subject of enduring interest in perceptual psychology. We will show that it is easily explained in the frequency domain.

The strobed image can be viewed as a time-sampled version of a corresponding smooth motion. For example, the contrast distribution of a thin line in smooth horizontal motion at velocity r can be written

$$c(x,t)=\delta(x-rt) \tag{4}$$

and its spectrum,

$$C(u,w)=\delta(w+ru) \tag{5}$$

These two functions are shown in Fig. 2. In keeping with our earlier observations, the spectrum lies along a line with slope $-1/r$ in the u,w space. Since the line is thin, we describe it as an impulse function. This simplifies the spectrum, making it into a line impulse function. A stroboscopic version of this moving line is accomplished by presenting a sample every Δt seconds with contrast Δt,

$$c_s(x,t)=\delta(x-rt)\Delta t \sum_{i=-\infty}^{\infty} \delta(t-i\Delta t) \tag{6}$$

This function is pictured in Fig. 3. Note that the time interval between samples is Δt, and the distance

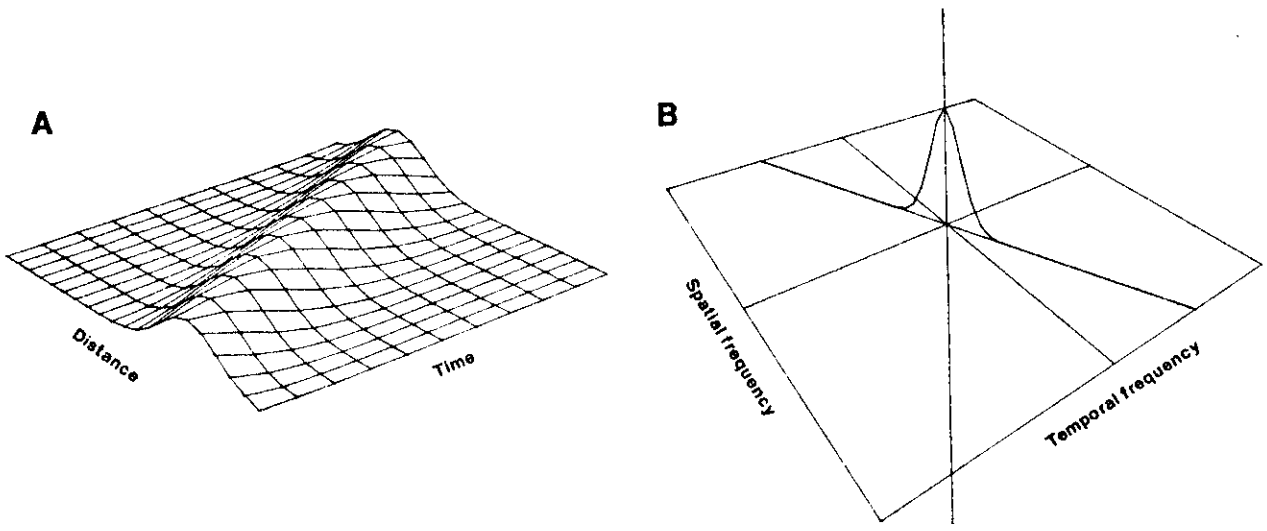

Fig. 1. Contrast distribution (a) and frequency spectrum (b) for a moving image. In this example, a Gaussian bar moves to the right.

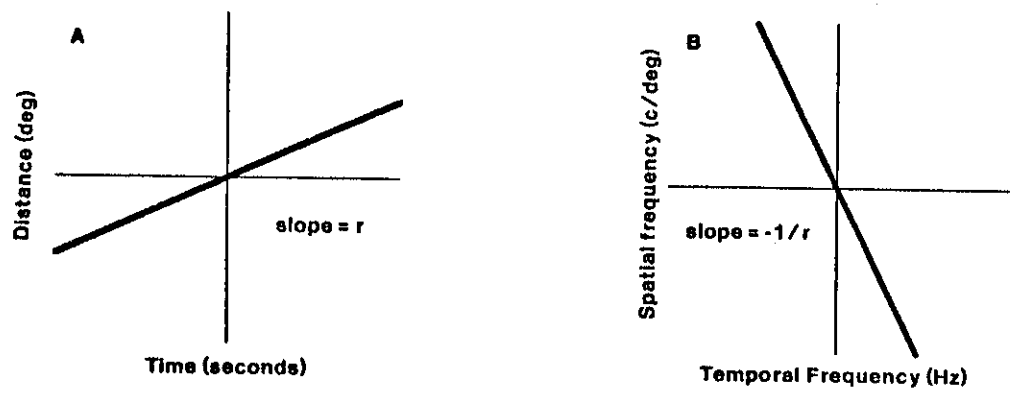

Fig.2. Contrast distribution (a) and frequency spectrum (b) for a thin vertical line moving smoothly to the right at velocity r. Both functions are line impulses, whose amplitude dimension should be imagined to project out of the page.

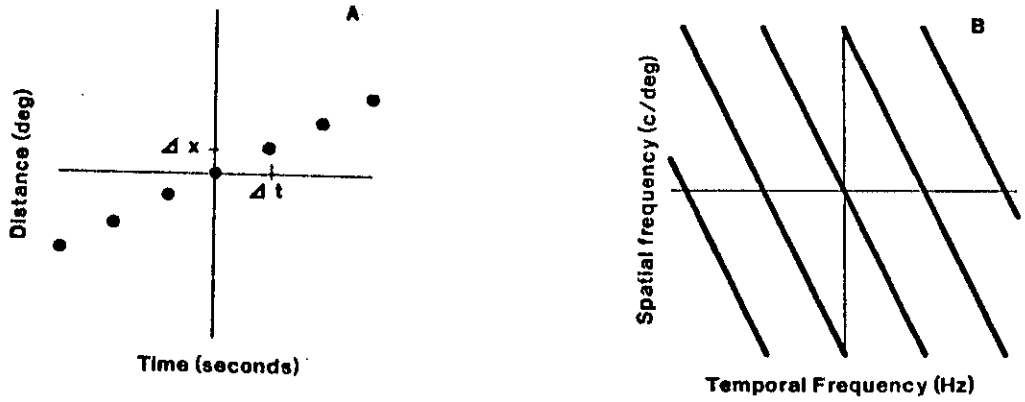

Fig.3. Contrast distribution (a) and frequency spectrum (b) for a thin vertical line in stroboscopic motion to the right at velocity r.

between samples is $r\Delta t$. The *sampling frequency* w_s is the inverse of the time between samples, $w_s = 1/\Delta t$. The spectrum of stroboscopic motion is given by

$$C_s(u,w) = \sum_{i=-\infty}^{\infty} \delta(w + ur - iw_s) \tag{7}$$

and is shown in Fig. 0b. Note that is is identical to the spectrum for smooth motion except for the addition of replicas of the spectrum at intervals of w_s.

At this point we consider the sensitivity of the human eye to stimuli of particular spatial and temporal frequencies, that is to stimuli that lie at particular points in u,w space. It has been known since Shade's work [Shade, 1956] that the human eye is not equally sensitive to all spatial frequencies, and that sinusoidal targets above a critical spatial frequency are invisible. Similarly, de Lange [de Lange, 1954] showed that temporal fluctuations more rapid than a critical temporal frequency are not seen. We will call these limits to spatial and temporal frequency sensitivity u_l and w_l respectively. These two limits have been shown to be relatively independent of each other: the spatial limit does not depend too much upon the temporal frequency of the stimulus, and vice versa [Robson, 1966, Koenderink and van Doorn, 1979]. This permits us to roughly characterize the limits of human visual sensitivity to spatial

and temporal frequencies by a rectangular *window of visibility*, as pictured in Fig. 4. Components that lie within the window may be more or less visible, but those that lie outside the window are invisible. This observation suggests the following hypothesis: *two stimuli will appear identical if their spectra, after passing through the window of visibility, are identical.*

Returning to our stroboscopic stimulus, note that the spectrum of the sampled line differs from that of the smooth line only by the addition of the parallel replicas at intervals of the sampling frequency. Thus the conjecture implies that if these replicas lie outside the window of visibility, then the smoothly moving line and the sampled line will be indistinguishable. The replicas may be moved outside the window of visibility by either increasing the sampling frequency (which moves the replicas farther from the origin), or reducing the velocity (which makes the replicas more nearly vertical). More precisely, we note that for any velocity, the critical sampling frequency will be achieved when the first spectral replica is just touching the corner of the window of visibility, as shown in Fig.4. This will occur at a *critical sampling frequency* w_s, given by

$$w_c = w_l + ru_l \tag{8}$$

Thus the predicted critical sampling frequency is a linear function of velocity, with intercept given by the temporal frequency limit and slope given by the spatial frequency limit.

2.1. An experiment

We tested this prediction by means of a two-interval forced-choice experiment. One interval contained a vertical line which moved smoothly to the right or left, the other interval contained a line moving at the same velocity but sampled at a rate w_s. The observer attempted to pick the interval containing the sampled version.

Fig. 5 shows the results. In each case, the critical sampling frequency increases approximately linearly with velocity, as predicted by Eq. 8. For both observers, the intercept is at about 30 Hz which is a good estimate for the temporal frequency limit (w_l) under these conditions. The slope of the curve, which according to theory is an estimate of the spatial frequency limit (u_l), is 6 c/deg for one observer and 13 c/deg for the other. These are somewhat low for estimates of the spatial frequency limit, but are not unreasonable given the low contrast and brief duration of the frequency component presumably serving to distinguish between smooth and sampled versions. Thus the data in Fig.5 support our hypothesis that smooth and sampled motion are visually indistinguishable when the spectral components that differ between them lie outside the window of visibility.

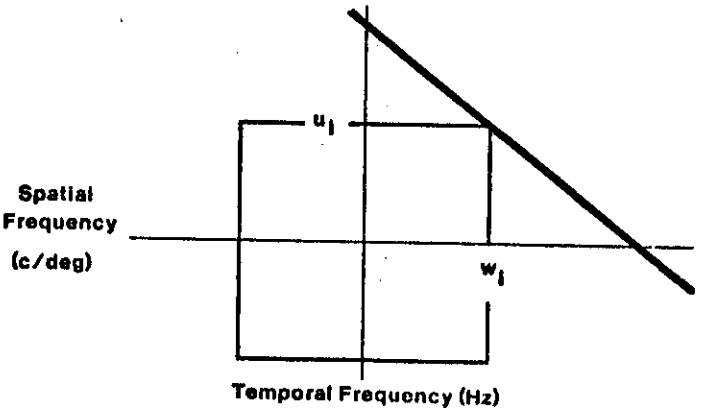

Fig.4. The window of visibility. Stimulus components lying outside the window are invisible. One spectral replica is shown lying just outside the window, to illustrate a condition in which smooth and sampled images would be just indistinguishable.

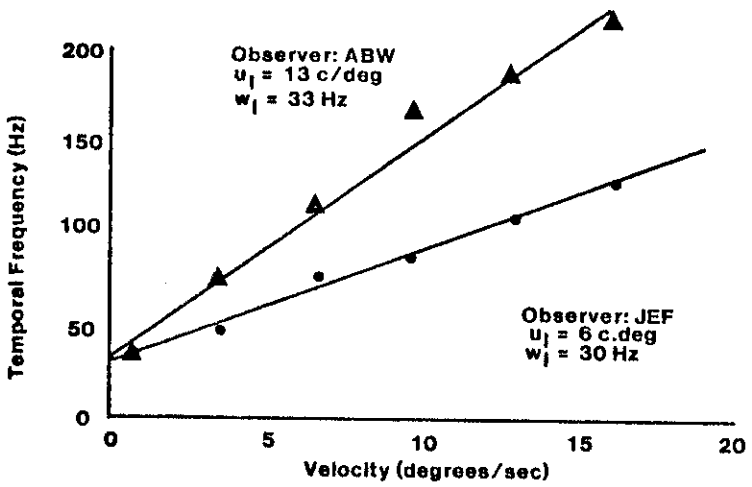

Fig.5. Critical sampling frequency as a function of velocity for two observers. The stimulus was a vertical line 50 minutes long by 0.65 minutes wide which moved horizontally. Observers fixated a point at the center of the path of travel. Line contrast was 200%, and background luminance was 50 cd m^{-2}. A session consisted of 25 trials at each of five sampling frequencies, all at a single velocity. Critical sampling frequency was defined as the frequency at which the observer was correct 75% of the time. The straight lines are fitted by eye.

This theory is not confined to the case of a vertical line moving horizontally with fixed velocity. We have considered this case only because it gives rise to particularly simple predictions and because it is a case much considered in the literature of apparent motion [Kolers, 1972; Morgan, 1979]. This general notion can be extended to an arbitrary spatial image undergoing an arbitrary transformation over time, and the sampling process can be extended to the two spatial dimensions, as well as time. We believe that consideration of the visual filtering of sampled displays provides answers to some long-standing puzzles in perceptual psychology, and to some modern problems in advanced visual displays [Watson, Ahumada, & Farrell, 1983].

3. A linear motion sensor

As a second illustration of the utility of frequency descriptions of motion phenomena we will construct a mechanism capable of sensing motion in a particular direction. Our argument in the preceding section was that what cannot get through the window of visibility cannot influence perception. Now we turn to the issue of how information within the window might be used to sense motion.

The psychophysical literature on motion perception provides good evidence for the existence of mechanisms that are selective for direction of motion. For example a stimulus consisting of the sum of two oppositely moving images is little more visible than either image alone, even though the peak contrast of the sum is twice as great [Levinson & Sekuler, 1975; Watson, Thompson, Murphy & Nachmias, 1980]. Similarly, prolonged viewing of motion in one direction raises thresholds much more for targets moving in that direction than in the opposite direction [Pantle & Sekuler, 1969], and patterns moving in opposite directions can be discriminated at detection threshold [Watson, Thompson, Murphy & Nachmias, 1980]. But there have been few efforts to construct plausible models of motion sensors capable of accounting for human performance. What models there are generally rely on autocorrelation, requiring a nonlinear multiplicative process. These models have also largely neglected the spatial dimension. In contrast, the model we propose is linear, and is explicitly described in both spatial and temporal dimensions. We offer it as a candidate model for human motion sensors.

3.1. The hyperbolic filter

Before constructing our sensor we digress in order to introduce the notion of a *hyperbolic filter,* with impulse response

$$h(x)=\frac{1}{\pi x} \tag{9}$$

The transfer function of this filter is

$$H(u)=-i\,\mathrm{sgn}(u) \tag{10}$$

where sgn is the sign function. This filter has the interesting property of having constant unit gain, and a constant phase lag of $\pi/2$. We will use this filter twice in the succeeding development, once in the space domain, and once in the temporal domain. It is important to realize that although this function is neither causal nor physically realizable, an approximation can be constructed which has the appropriate properties, yet is causal and realizable. This is illustrated by the approximate hyperbolic filter shown in Fig.6.

3.2. Constructing the sensor

We have seen that the spectra of moving images lie along lines through the origin in u,w space. To sense this motion one might construct a matched filter passing only energy lying along that line. We take a different approach. We base our model largely upon the properties of certain well-studied visual cells in the cortex of the cat and monkey. These *simple cells* have a spatial weighting function that is reasonably well described by the product of a 2D sinusoid and a radially symmetric Gaussian function. We assume the diameter of the Gaussian at half height is 1.324 times the the period of the sinusoid [Watson, 1983]. This diameter determines the selectivity of the cell for spatial frequency, and the number we have chosen gives the cell a spatial frequency bandwidth (at half height) of one octave. It has also been observed that simple cells occur in pairs [Pollen & Ronner, 1981] with spatial phases $\pi/2$ apart. If the spatial impulse response of one cell is $s(x)$, we can construct the paired cell by application of the hyperbolic filter, yielding $h(x)*s(x)$. (For simplicity we consider only the horizontal dimension of a vertically oriented weighting function.) If we arbitrarily assume $s(x)$ is an even function

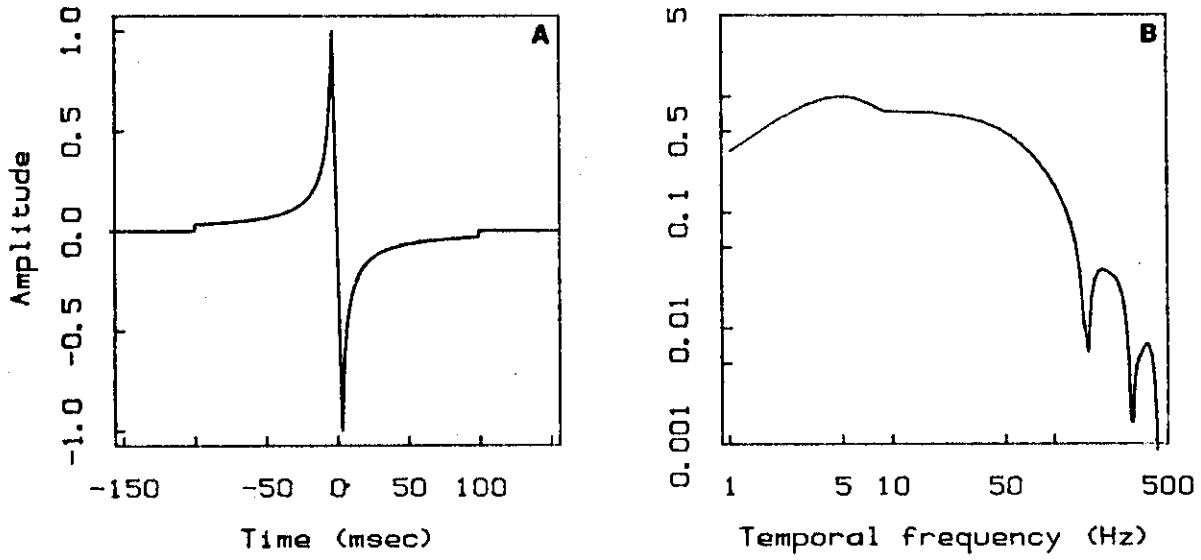

Fig.6. Impulse response (a) and amplitude response (b) of the hyperbolic temporal filter. The example shown is an approximation to a hyperbolic filter that is truncated at its onset and near to the origin. Its frequency behaviour is not significantly different from the true hyperbolic filter over the range of frequencies of interest (1-60 Hz).

(this is not critical), then $h(x)*s(x)$ is an odd function. These even and odd impulse responses are shown in Fig.7.

We assume that the two cells share a common separable temporal impulse response $f(t)$. In the absence of better information, we model this by a function fit to human temporal sensitivity, as shown in Fig. 8. We now apply a hyperbolic temporal filter to the odd pathway. To insure that the filter is causal, we must delay its impulse response by an appropriate amount τ. The delay of τ in the odd pathway introduces an additional phase lag which must be matched by the even path, so we put an equivalent delay in there. All of the preceding steps are diagrammed in Fig. 9.

What have we accomplished by these manipulations? The impulse response of our sensor is obtained by convolving the impulse responses of all the cascaded elements, and adding those in parallel,

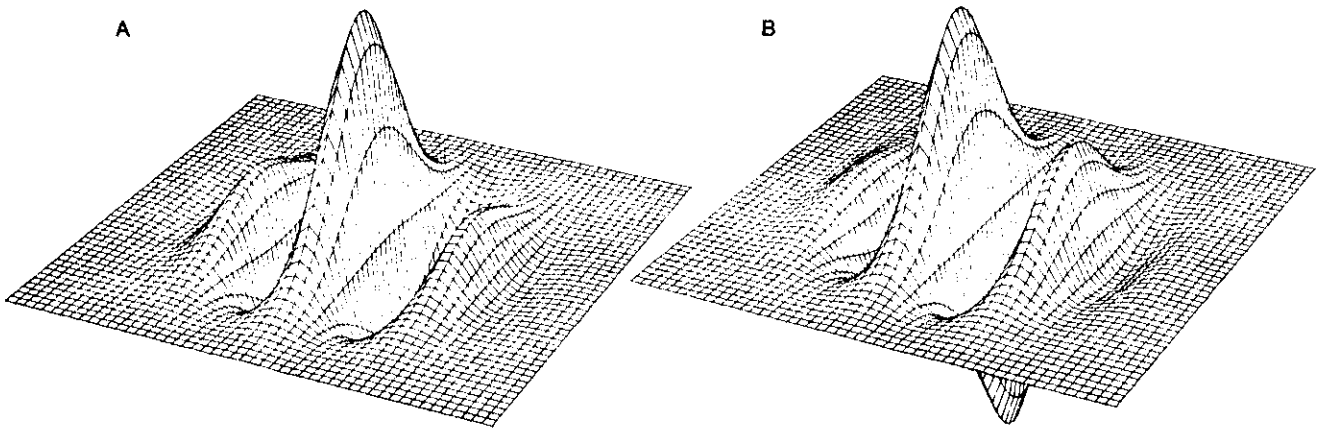

Fig.7. Spatial impulse responses of even (a) and odd (b) pathways. Each is the product of a sinusoid and a Gaussian with diameter at half height of 1.324 times the period. The phase in (a) is 0, in (b) it is $\pi/2$.

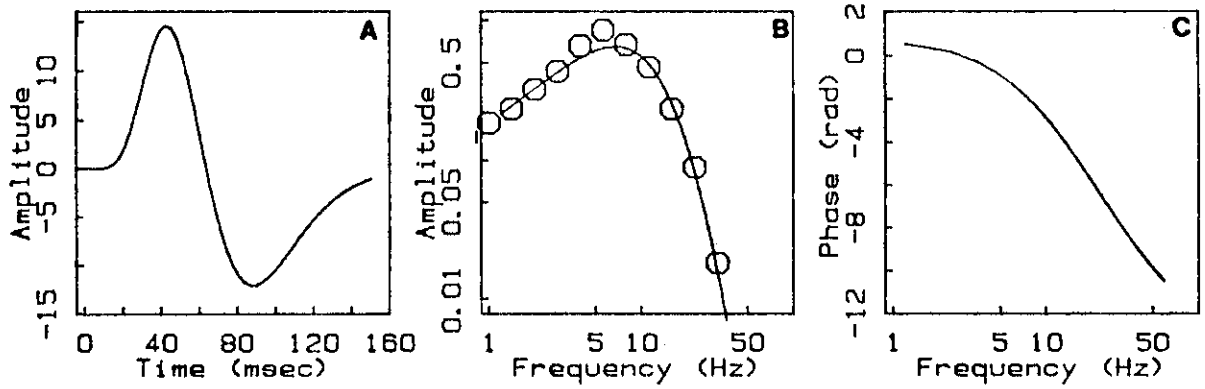

Fig.8. The impulse response (a), amplitude response (b), and phase response (c) of the initial temporal filter. The amplitude response in (b) has been fit to sensitivities to temporal modulation of a sinusoidal 0.5 c/deg grating as measured by Robson [1966].

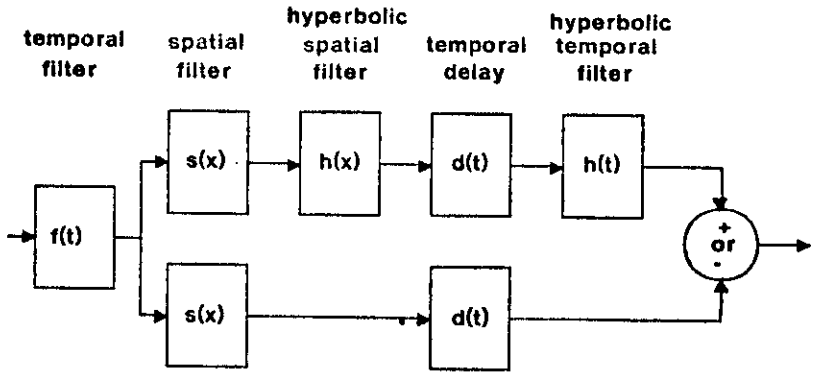

Fig.9. Block diagram of the linear motion sensor.

$$g(x,t) = f(t) * [s(x) * \delta(t-\tau) + s(x) * h(x) * h(t) * \delta(t-\tau)] \tag{11}$$

$$= f(t) * s(x) * \delta(t-\tau) * [1 + h(x) * h(t)]$$

The frequency spectrum of our sensor is obtained by Fourier transforming Eq. 11,

$$G(u,w) = F(w)S(u)e^{-i2\pi w\tau}[1 - \mathrm{sgn}(u)\mathrm{sgn}(w)] \tag{12}$$

The modulus of this spectrum is shown in Fig.10. Note that it is non-zero only in the odd quadrants of w,u space. Recall that these are the quadrants occupied by the spectrum of a leftward moving image, so our sensor will respond to motion exclusively in a leftward direction.

An intuitive explanation for this behavior is as follows. As a sinusoidal grating moves over the sensor, the response of the even and odd spatial filters will be temporal sinusoids differing in phase by plus or minus $\pi/2$, for right and left motion respectively. The hyperbolic temporal filter introduces a further shift in the odd pathway of $\pi/2$, for a resulting phase difference of π or 0. When even and odd pathways are summed, the result is twice the amplitude of either path for left motion, and 0 for right motion.

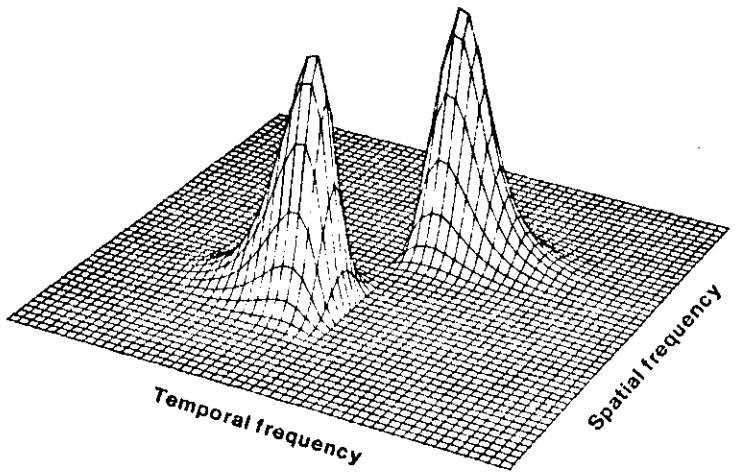

Fig.10. Amplitude response of the motion sensor.

We draw attention to the following features of our sensor.

(1) The sensor is selective for direction of motion in 2D space, but is not selective for speed. This agrees with psychophysical data that indicate that humans are poor at speed discrimination at threshold [Watson & Robson, 1981; Thompson, 1982]. Above threshold, speed can be judged by noting the temporal frequency of the sensor response, or by examining the pattern of activity across many sensors. This is in contrast to other models of speed discrimination which assume mechanisms tuned for speed.

(2) The sensor is selective for spatial frequency and orientation. This is consistent with a wide range of psychophysical and physiological results [Movshon, Thompson & Tolhurst, 1978a,b; Schiller, Finlay & Volman, 1976; De Valois, Albrecht & Thorell, 1982; De Valois, Yund & Helper, 1982; Campbell & Robson, 1968; Blakemore & Campbell, 1969; Watson, 1983]. In common with recent models [Watson, 1983; Sakitt & Barlow, 1982], we assume the visual system to be populated with a many sensors selective for different frequencies, positions, and orientations. To include our sensor in such a model is only a minor elaboration, rather than the addition of a whole new parallel pathway, since the even and odd spatial filters are already present.

(3) The preferred direction of motion of the sensor is orthogonal to the preferred orientation of the sensor. This is because the spatial phase shift between even and odd weighting functions is along the axis of sinusoidal modulation. For a given sensor orientation, sensors for the two opposite directions are created by either adding or subtracting the odd and even pathways.

(4) The direction selectivity of the sensor is established by the spatial impulse response $s(x)$. The spatial frequency bandwidth, the orientation bandwidth, and the direction bandwidth of the sensor are all inversely proportional to the diameter of the spatial Gaussian.

We are not aware of any physiological evidence for or against our hypothetical hyperbolic temporal filter. It should be noted that it can be approximated by more commonplace mechanisms such as a delay or a differentiator. We hope in the future to examine in further detail the extent to which our sensor can account for the behaviour of direction-selective visual neurons and for the sensory performance of the human observer.

4. Summary

We have attempted to show how motion phenomena can be usefully examined in the frequency domain. Our first example was a demonstration of how the known spatiotemporal frequency limits of vision provide a simple explanation of stroboscopic apparent motion. In our second example, we showed how the constraints on the spectra of moving images lead to a sensor capable of sensing motion in a particular direction. We recommend this perspective to other investigators of visual motion perception, not as a theoretical panacea, but as another useful analytic tool.

References

C. Blakemore and F. W. Campbell On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol., Lond.*, 1969 , *203* , 237-260 .

F. W. Campbell and J. G. Robson Application of fourier analysis to the visibility of gratings. *J. Physiol., Lond.*, 1968 , *197* , 551-566 .

J. J. Koenderink and A. J. van Doorn Spatiotemporal contrast detection threshold surface is bimodal. *Optics Letters*, 1979 , *4* , 32-34 .

P. A. Kolers *Aspects of motion perception.* New York: Pergammon, 1972.

H. De Lange Relationship between critical flicker frequency and a set of low frequency characteristics of the eye. *J. opt. Soc. Am.*, 1954 , *44* , 380-389 .

E. Levinson and R. Sekuler The independence of channels in human vision selective for direction of movement. *J. Physiol., Lond.*, 1975 , *250*, 347-366 .

M. J. Morgan Perception of continuity in stroboscopic motion: A temporal frequency analysis. *Vision Res.*, 1979 , *19* , 491-500 .

J. A. Movshon, I. D. Thompson, and D. J. Tolhurst Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol.*, 1978, *283* , 53-77 .

J. A. Movshon, I. D. Thompson, and D. J. Tolhurst Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *J. Physiol.*, 1978, *283* , 101-120 .

A. Pantle and R. W. Sekuler Contrast response of human visual mechanisms sensitive to orientation and direction of motion. *Vision Res.*, 1969, *9*, 397-406.

D. A. Pollen and S. F. Ronner Phase relationship between adjacent simple cells in the visual cortex. *Science*, 1981, *212*, 1409-1411.

J. G. Robson Spatial and temporal contrast sensitivity functions of the visual system. *J. opt. Soc. Am.*, 1966 , *56* , 1141-1142 .

B. Sakitt and H. B. Barlow A model for the economical encoding of the visual image in cerebral cortex. *Biological Cybernetics*, 1982, *43*, 97-108.

P. H. Schiller, B. L. Finlay, and S. F. Volman Quantitative studies of single cell properties in monkey striate cortex. III. Spatial frequency. *J. Neurophysiol.*, 1976, *39*, 1334-1351.

O. H. Shade Optical and photoelectric analog of the eye. *J. opt. Soc. Am.*, 1956 , *46* , 721-739 .

P. Thompson Detection and discrimination of moving gratings. *Acta Psychologica*, 1982, *In press.*.

R. L. De Valois, D. G. Albrecht, and L. G. Thorell Spatial frequency selectivity of cells in the macaque visual cortex. *Vision Res.*, 1982, *22*, 545-559..

R. L. De Valois, E. W. Yund, and H. Helper The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res.*, 1982, *22*, 531-544.

A. B. Watson Detection and recognition of simple spatial forms. In O. J. Braddick and A. C. Slade (Eds.), *Physical and biological processing of images* Berlin: Springer-Verlag, 1983.

A. B. Watson, A. J. Ahumada Jr., and J. Farrell Sampling considerations in visual motion displays. *NASA Technical Paper*, 1983, *In preparation.*.

A. B. Watson and J. G. Robson Discrimination at threshold: labelled detectors in human vision. *Vision Res.*, 1981 , *21*, 1115-1122.

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

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-84355		

4. Title and Subtitle	5. Report Date
A LOOK AT MOTION IN THE FREQUENCY DOMAIN	April 1983
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Andrew B. Watson and Albert J. Ahumada, Jr.	A-9304
	10. Work Unit No. T-4330

9. Performing Organization Name and Address	11. Contract or Grant No.
Ames Research Center Moffett Field, California 94035	

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
National Aeronautics and Space Administration Washington, D.C. 20546	Technical Memorandum
	14. Sponsoring Agency Code 505-35-31

15. Supplementary Notes

Point of Contact: Andrew B. Watson, Ames Research Center, M/S 239-2, Moffett Field, Calif. 94035, (415) 965-6584, FTS 448-6584

16. Abstract

A moving image can be specified by a contrast distribution, $c(x,y,t)$, over the dimensions of space x,y, and time t. Alternatively, it can be specified by the distribution $C(u,v,w)$ over spatial frequency u,v and temporal frequency w. The frequency representation of a moving image is shown to have a characteristic form. This permits two useful observations. The first is that the apparent smoothness of time-sampled moving images (apparent motion) can be explained by the filtering action of the human visual system. This leads to the following formula for the required update rate for time-sampled displays.

$$w_c = w_l + ru_l$$

where w_c is the required update rate in Hz, w_l is the limit of human temporal resolution in Hz, r is the velocity of the moving image in degrees/sec, and u_l is the limit of human spatial resolution in cycles/deg.

The second observation is that it is possible to construct a linear sensor that responds to images moving in a particular direction. The sensor is derived and its properties are discussed.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Vision Motion perception Human factors	Unlimited Subject Category - 53

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	12	A02

*For sale by the National Technical Information Service, Springfield, Virginia 22161