Vision Res. Vol. 27, No. 6, pp. 1029–1040, 1987
Printed in Great Britain.

0042-6989/87 \$3.00 + 0.00

STATISTICAL EFFICIENCY FOR THE DETECTION OF VISUAL NOISE

DANIEL KERSTEN

Walter S. Hunter Laboratory of Psychology, Brown University, Providence, RI 02912, U.S.A.

(*Received* 11 *June* 1985; *in revised form* 24 *July* 1986)

Abstract—Visual detection thresholds were measured for one-dimensional static noise patterns centered at 2 c/deg on a logarithmic scale, for several widths and spatial frequency bandwidths in the presence and absence of a one-dimensional dynamic visual noise mask. Human performance was compared with the ideal observer's performance for the same task. The ideal's contrast thresholds increase with the fourth root of bandwidth. For thresholds measured in the presence and absence of noise, the fourth-root-law for bandwidth fitted the human observers' data from 1 to 6 octaves for a space constant of 2° and from 2 to 6 octaves for a space constant of 0.25°. Detection efficiencies were quite high (30–60%) for patterns of 1- and 6-octave bandwidths. These results can be interpreted either in terms of a channel with an adjustable spatial frequency bandwidth or as very efficient combination of information between spatial frequency channels.

Detection Efficiency Visual noise Spatial-frequency

INTRODUCTION

This research was motivated by two somewhat independent research traditions. One has been the study of absolute limits to human visual detection imposed by the photon nature of light (Hecht *et al.*, 1942; Rose, 1942; de Vries, 1943; Barlow, 1962) and subsequent extensions to detection limited by luminance noise and to random dot patterns (Sturm and Morgan, 1949; Green *et al.*, 1959; Chambers and Courtney-Pratt, 1969; Barlow, 1978; Burgess and Barlow, 1983). The central idea is to measure human performance at a detection or discrimination task with reference to the ideal for the same task. The ideal way to detect a spot of light is to count the number of photons in the space-time region in which the signal is expected. Because of the inherent stochastic nature of photon absorption (and emission), this strategy can lead to errors. Thus it is possible to compare human performance with the ideal when each have the same error rate. Barlow (1962) defined efficiency as the ratio of the number of photons which the ideal requires to the number that the observer under study requires when they are both at the same performance level. Barlow (1978) modified this definition by replacing "photons" with "dots" in order to compare human performance with that of the ideal observer which is required to detect a change in dot density in a display of random dots. Chambers and Courtney-Pratt (1969) had observers find a patch of higher dot density in a background of dots. The location of the patch was randomized. Relative to the ideal for the same task, the human observers required only about twice the number of dots required by the ideal (efficiency = 44%). Barlow (1978) found similar results for the detection of changes of dot density. For static dot displays, efficiencies were highest (50–60%) for target sizes of 1–4°. For both of these studies, the signals were defined statistically.

A second tradition has been the spatial-frequency analysis of human pattern perception (Schade, 1956; Campbell and Robson, 1968). Detection of sinusoidal gratings in visual noise was first investigated by Greis and Rohler (1970), Pollehn and Roehrig (1970) and then by Stromeyer and Julesz (1972). Stromeyer and Julesz measured contrast thresholds for gratings in one-dimensional dynamic noise with low-pass, high-pass and band-pass spatial spectra. Their conclusion was that only spatial frequencies within a 1 octave range of the signal frequency were effective in masking the signal. In addition to using noise as a mask, noise signal patterns have also been used in the spatial-frequency analysis of vision (Koenderink and van Doorn, 1974; Richards and Polit, 1974; Mitchell, 1976; Mostafavi and Sakrison, 1976; Quick *et al.*, 1976; Mayhew and Frisby, 1978; Quick *et al.*, 1978; Jamar and Koenderink, 1985). None of these studies, however, measured how well noise patterns are detected relative to the ideal observer. To do this requires specification of the non-signal as well as the signal noise properties. This has not been done previously for the detection of visual noise.

Studies of efficiency and spatial-frequency analysis have merged only recently. Van Meeteren and Barlow (1981) measured the efficiency for the detection of sinusoidally modulated stationary dot patterns and found efficiencies as high as 50% for gratings with 5 or less cycles. Burgess *et al.* (1981) found that discrimination of grating contrasts in static noise can be as high as 70%. Kersten (1984) found efficiencies for the detection of gratings in dynamic noise between 10 and 30% for widths of about a cycle. Quantum efficiencies for gratings at photopic levels tend to be much smaller, principally because of intrinsic noise (Pelli, 1981; Kersten and Barlow, 1985). Watson *et al.* (1983) found 0.05% for their best grating stimulus. However, all of these studies compared human performance with the ideal for visual tasks in which the signal was exactly "known". That is, the signal was deterministic rather than stochastic.

To do a well at detecting a change in dot density requires a fundamentally different strategy to that required to do well at detecting a sinusoidal grating in noise. In addition to the fact that the signal in the former is stochastic, the random dot signal is a broad-band signal, whereas a grating is typically narrow-band. For example, if the diameter of dot in a random dot display is 0.5′, the average spatial frequency spectrum extends to about 60 c/deg. In order to attain a high efficiency in the detection of a change of dot density, the observer should monitor a wide range of spatial frequencies. In order to do well at the detection of a grating in white noise, the observer should "look" only at the range of spatial frequencies containing the signal. The spatial-frequency studies argue for relatively narrow bandwidth channels (1 to 2 octaves). On the other hand, the high efficiencies found in the dot density studies argue for much broader bandwidths. Is this a paradox, or do observers have channels of adjustable bandwidth?

Of fundamental importance is the level of uncertainty at which a detection strategy operates. In accordance with an earlier suggestion of Tanner (1961), Nachmias and Kocher (1970) suggested that for increment detection, observers are intrinsically uncertain. This means that even when given prior knowledge about what to look for, observers "look" with many filters or channels, some insensitive to the signal, and say "yes" when any one of them exceeds a threshold criterion. This type of detection strategy is inefficient when signals are specified exactly, in which case the best strategy is to "look" only with that filter which matches the expected signal. Under certain conditions, humans may be uncertain as to signal parameters such as spatial location and time of arrival (Lasley and Cohn, 1981; Pelli, 1981, 1985; Cohn and Wardlaw, 1985). Davis and Graham (1981) have shown that observers can make use of prior knowledge of spatial frequency to improve performance in detecting a grating against a uniform field. However, the improvement can be small when compared to what is possible ideally. A noise signal consists of many spatial-frequency components, with random amplitudes and phases. Because the exact parameters of the signal are unknown, to do well on average, the ideal detector should monitor the entire spectrum. Thus, when signals are specified statistically, we might perform with high efficiency relative to the ideal observer which has to cope with the same degree of uncertainty.

The purpose of this study was to begin to characterize the role of spatial frequency and spatial extent in the detection of stochastic visual signals. The signals to be detected were drawn from space-limited and spatial frequency band-limited noise processes. For comparison with previous studies using deterministic signals (Kersten, 1984) the noise signals were static (random in one spatial dimension), and windowed with Gaussian envelopes, and the noise mask was dynamic (random in one spatial dimension and time). In a given trial, there was no prior information about the relative phases or the spatial frequency amplitudes in the pass-band. However, prior information was assumed for the ideal, (and provided for the observer) regarding the bandwidth, the spatial extent and temporal duration. Noise signals are in some sense at the opposite end of the uncertainty continuum from exactly specified signals.

METHOD

Apparatus

The visual stimuli were produced on the face of a Joyce Electronics CRT display by Z-axis modulation (Campbell and Green, 1965). The display was of the electromagnetic deflection type, with a raster frequency of 100 kHz, and a non-interlaced frame rate of 100 Hz. The display had a P31 phosphor, an unmodulated luminance of 340 cd/m^2, and a dark surround. The viewing distance was 228 cm. The display was 30 cm wide and 16 cm high. Thus, at the viewing distance of 228 cm, the screen subtended 7.5° horizontally by 4° vertically.

The signal luminance waveforms were synthesized digitally with an LSI-11/23 computer. The noise signal was static and randomly varied only in the horizontal direction. It was generated by a software implementation of a parallel set of 12 31-bit shift-registers. Each shift-register used exclusive-or feedback which provided maximal length sequences (described below). Each noise sample was the sum of four 12-bit words generated from the outputs of the shift-registers. The resulting amplitude distribution of this noise is approximately Gaussian ($P > 0.05$ for χ^2 fit). The power spectrum of sample sequences, measured by taking the Fourier transform of the autocorrelation function, was approximately flat up to the Nyquist rate. The ratio of peak contrast to standard deviation (crest-factor) was 3.5. For a true normal distribution, the probability of a sample exceeding 3.5 times the standard deviation is 4.7×10^{-4}. For each 10 msec frame, the computer generated 300 pairs of voltage samples. The voltage samples were routed through two 12-bit digital-to-analog converters (DAC). One DAC generated the spatial envelope, and the other the noise signal. The signal was band-pass filtered with a Krohn-Hite 4-pole Butterworth filter (model 3500R), multiplied by the envelope and routed through a programmable attenuator which temporally modulated the signal voltage. The signal was then added to the noise mask before being applied to the Z-axis input of the CRT.

The luminance noise mask varied in time and one spatial dimension. Binary pseudorandom noise was digitally synthesized by a 31-bit shift register with exclusive-or feedback (Roberts, 1963; Davies, 1970; Horowitz and Hill, 1980). At the maximum rate used (60 kHz), the noise generator would cycle once ($2^{31} - 1$ shifts) in 20 hr. Because of the high frame rate, the binary noise appears Guassian with 20 frames occurring within an assumed integration time of 200 msec of the eye.

Definitions

The following definitions are used in describing the stimulus conditions. The *contrast function* of visual noise is

$$c(x, t) = [L(x, t) - L_0]/L_0$$

where $L(x, t)$ is the luminance at point x and time t sec, and L_0 is the mean luminance. (The signals and noise used in this experiment varied in one spatial dimension and time.) The *contrast power* is equal to the variance of the contrast function

$$c^2 = \text{var}[c(x, t)] = \text{var}[L(x, t)/L_0].$$

For signals varying in one spatial and one time dimension, the *noise spectral density*, N_{dynamic}, is the contrast power in a 1 c/deg by 1 Hz band. The *static* noise spectral density, N_{static}, is the contrast power in a 1 c/deg band, ignoring time (e.g. in one frame of a display). For constant noise spectral density within the passband, the contrast power is equal to the two-sided spatial-frequency bandwidth times the static noise spectral density. To compute the static noise spectral density from the dynamic noise spectral density, compute the contrast power

$$c^2 = B_x B_t N_{\text{dynamic}} = B_x N_{\text{static}}.$$

Thus, the static noise spectral density is equal to the dynamic noise spectral density times the qtwo-sided temporal bandwidth. In order to derive the performance of the ideal (see below), the static noise spectral density of the dynamic noise mask is needed. With a 100 Hz frame rate, the static spectral density is $100\times$ dynamic noise spectral density. For most calculations, it is convenient to use the two-sided bandwidth (denoted B_x and B_t for spatial and temporal frequencies) which is measured over positive and negative frequencies and is thus twice the one-sided bandwidth (denoted W_x for spatial frequency). The number of degrees of freedom of the static signal is $B_x X$, where X is the spatial width. This is also called the logon-content of the signal (Gabor, 1946; MacKay, 1950).

The *noise-equivalent bandwidth* of a filter is the area under its squared transfer function.* The noise-equivalent bandwidth is useful for calculating the power coming out of a filter whose input is noise with a flat power density spectrum with larger bandwidth. If c_i^2 is the power coming into the filter, and the incoming noise has bandwidth B_x, then the output contrast power of the filter is

$$c_0^2 = B_e c_i^2 / B_x$$

where B_e is the two-sided noise equivalent bandwidth of the filter. For constant noise spectral density within the passband, the contrast power is equal to the two-sided spatial-frequency bandwidth times the static noise spectral density.

Contrast energy is used in the Discussion. It is defined as

$$E = \iint c^2(x, t)\, dx\, dt.$$

For a noise signal of spatial width X, with uniform variance over space and time the average contrast energy is given by

$$E = c^2 XT = XTB_x B_t N_{\text{dynamic}}.$$

Efficiency (η) is defined as

$$\eta = S_i / S_0$$

where S_i and S_0 are the threshold signal noise spectral densities of the ideal and the human observer at which they get the same proportion correct in the same task. For dot densities sufficiently high, this definition is also consistent with Barlow's (Barlow, 1978). This is because the reciprocal of dot flux is the noise spectral density (Pelli, 1981). Because the spectral density is equal to the contrast power divided by the bandwidth, the efficiency can also be expressed by

$$\eta = (c_i / c_0)^2$$

where c_i and c_0 are the ideal's and humans' r.m.s. contrast thresholds respectively. This definition is consistent with that of Tanner and Birdsall (1958).†

Stimuli

The luminance of the signal at a point x, y (in degrees) at time t (sec) is given by

$$L(x, y, t) = L_0\, [1 + m(x, t) s(x)]$$

where L_0 is the mean luminance. The term $s(x)$ is a noise signal sample taken from a stationary Gaussian process with a flat noise spectral density. The modulating function $m(x, t)$ is given by the product of Gaussian functions horizontally and temporally

$$m(x, t) = \exp[-(x/s_x)^2 - (t/s_t)^2]$$

where s_t and s_x are the time and horizontal space constants respectively. The noise signal was centered at 2 c/deg on a logarithmic scale. The two-sided noise-equivalent bandwidths of the signals prior to windowing were 1.53, 2.92, 6.16, 15.4 and 32.32 c/deg for 0.25 and 2° width patterns. These bandwidths correspond to one-sided bandwidths of approximately 0.5, 1, 2, 4, and 6 octaves. The time constant (s_t) was 80 msec.

The spectral density of the noise mask was flat to 33 c/deg and 50 Hz. The dynamic noise spectral density was either 0 or 10^{-5} sec deg. Thus the static spectral density was either 0 or 10^{-3} deg at the 100 Hz frame rate. The noise was turned on abruptly 80 msec before the first temporal 1/e point of the signal was reached, and turned off abruptly 80 msec after the second 1/e point. The noise mask extended uniformly across the screen and had rectangular spatial and temporal envelopes.

Procedures

I will describe the procedure for finding the contrast thresholds of the human observers and then the ideal observer. For the human observers, r.m.s. contrast thresholds (c_0) were measured by a temporal two-alternative forced-choice procedure as a function of bandwidth. A maximum-likelihood procedure (QUEST) was

---

*The one-sided noise-equivalent bandwidth of a 4-pole Butterworth filter can be calculated by contour integration and is given by

$$W_x = \frac{\pi f_2^8}{4(f_2^8 - f_1^8)} \times [\sin(\pi/8) + \sin(3\pi/8)] \times (f_2 - f_1)$$

where f_2 and f_1 are the upper and lower corner frequencies. For the bandwidths used in this study, W_x differs from the difference between the corner frequencies by less than 3%.

†Tanner and Birdsall (1958) defined efficiency as the ratio of the ideal observer's threshold energy to that of the observer under study, each at the same performance level. Watson *et al.* (1983) define "contrast energy" which corresponds to energy in Tanner and Birdsall's definition. Because, the signals used in this study are stochastic, this definition is generalized by replacing energy with average (contrast) energy. Further, since average contrast energy is proportional to the r.m.s. contrast squared, we can define efficiency as the ratio of contrasts squared.

used to present the signal at the current "best" estimate of threshold contrast (Watson and Pelli, 1983). The contrast yielding 75% correct was estimated. The step size was 1 dB (0.05 log units of contrast). Typically, a minimum of 4 blocks of 50 trials were used to determine threshold. The intervals between trials were separated by 600 msec. Each interval was marked by an auditory tone. A feedback tone indicated to the observer whether he was right. A sample of the noise signal was shown prior to each trial to familiarize the observer with the bandwidth and width to be expected. The bandwidth and width were fixed for a given block. Although this procedure lacks the advantages of a randomized procedure, it was felt that it resulted in better familiarity with the signal and noise statistics. The observers were given a minimum of 2 hr of practice observations, prior to the experiment. This was thought sufficient to insure asymptotic levels of performance, however a possible exception is noted below.

The ideal's contrast thresholds were computed by three methods. For most conditions in the current study, the low-energy-coherence (LEC) approximation to the ideal is adequate (Appendix 1). Let d' be $\sqrt{2}$ times the normal deviate of the proportion correct in a two-alternative forced-choice task. An expression for d' is given by

$$d'^2 \simeq (\pi/8)(S_0/N_0)^2 B_x B_t s_x s_t$$

where S_0 and N_0 are the static noise spectral densities of the signal and noise respectively. B_x and B_t are the spatial and temporal frequency two-sided bandwidths of the signal. This approximation is good for $S_0/N_0 \ll 1$ and for $B_x s_x \gg 1$. Because the static signal noise spectral density is equal to the r.m.s. contrast squared divided by the two-sided spatial frequency bandwidth, this equation can be re-written as

$$c^2 \simeq d' N_0 \sqrt{[8B_x/(\pi B_t s_x s_t)]}$$

or

$$c \simeq \sqrt{(d' N_0)} \left(\frac{8B_x}{\pi B_t s_x s_t}\right)^{1/4}.$$

Thus for a constant proportion correct, (75% correct corresponds to $d' = 0.953$ in these experiments), contrast threshold is proportional to the fourth root of signal bandwidth for the ideal and inversely proportional to the fourth root of width.

Because of the conditions on the validity of the LEC approximation, the performance of the ideal was examined using two other methods. In the first or closed-form method, d' was computed based on an expression for the likelihood ratio which was valid for all values of S_0/N_0. These values were used in the calculation of efficiency. The worst-case disagreement in contrast threshold between the closed-form solution and the LEC approximation was 8% for the 0.25° 2-octave signal ($S_0/N_0 = 0.45$). This would correspond to a 16% underestimate of the efficiency when the LEC approximation is used. As $B_x s_x$ increases, the error in the LEC approximation decreases. The second method simulated the psychophysical procedure using the expression for the likelihood ratio from the first method. The results of the simulations were used to spot check performance calculations for the LEC and "closed-form" solutions for the larger values of S_0/N_0. For the values checked, the simulated contrast thresholds agreed to within 1 standard error with the closed-form predictions. These methods are described in detail in Appendix 2.

Observers

There were two observers. D.K. is the author and is an emmetrope. P.K. is a myope with 20/20 corrected vision (−0.75 D for both eyes). P.K. was naive to the details of the experiment. Viewing was binocular with natural pupils, and with a fixation point at the center of the screen.

RESULTS

Figure 1(A) and (B) show r.m.s. contrast thresholds for D.K. and P.K. as a function of one-sided noise-equivalent nominal signal bandwidth (i.e. prior to windowing) in c/deg for narrow ($s_x = 0.25°$, represented by triangles) and broad ($s_x = 2°$, represented by circles) patterns. The octave ranges are shown on the upper abscissa. Due to spectral spreading induced by the spatial window which followed the filter, the actual bandwidths are somewhat larger than the nominal bandwidths.* The upper set of points represent data collected in high noise (dynamic spectral density of noise mask $= 10^{-5}$ sec-deg) and the lower points data collected in the absence of a luminance noise mask. The standard

*A Gaussian enveloped sinusoidal has a spectral width between 1/e points of $2/\pi s_x$. It can be argued that the "effective" spread is about $W_x + \sqrt{(\pi/2)/(\pi s_x)}$ (Gabor, 1946). So for a 0.25 degree window, the bandwidths are about $W_x + 1.6$ c/deg and for a 2 deg window, about $W_x + 0.2$. This is discussed further in Appendix 2.

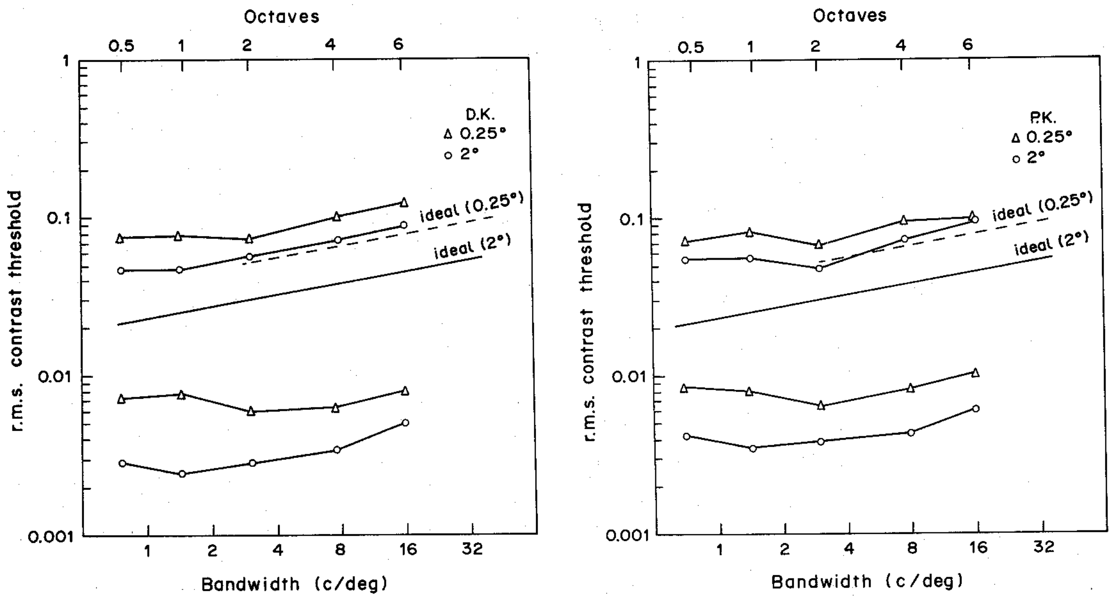

Fig. 1. (A) r.m.s. contrast thresholds of visual noise signals as a function of one-sided nominal bandwidth in c/deg (lower abscissa) and in octaves (upper abscissa). The upper and lower sets of points show thresholds collected in the presence and absence of a noise mask. The open triangles and circles represent data for space constants 0.25° and 2°. The dashed and solid straight lines represent the r.m.s. contrast thresholds as a function of bandwidth for the ideal observer for the noise mask case. Observer D.K. (B) Observer P.K. Other details as for (A).

errors were 10% or less, corresponding to about plus or minus one symbol's height.

Contrast thresholds are higher for the narrower width. For bandwidths greater than 1 octave with space constant 2° and greater than 2 octaves with space constant 0.25°, contrast thresholds tend to increase with increasing bandwidth. The upper dashed and solid lines are the ideal's thresholds for the high noise condition for the 0.25° and 2° windows respectively. Because of the limitation of the LEC approximation, the lines are drawn only where the ideal's signal-to-noise ratio (S_0/N_0) is less than 0.45.

Linear regression lines were fitted over bandwidths from 3 to 16 c/deg for the triangles, and from 1.4 to 16 c/deg for the circles. The slopes and 95% confidence intervals are given in Table 1. The LEC approximation to the ideal has a slope of 0.25. None of the slopes in Table 1 differ significantly from this. The mean slopes are 0.25 ± 0.03 and 0.24 ± 0.02 for D.K. and P.K. respectively. Except for one slope (D.K., $s_x = 0.25°$, no noise), the slopes are all significantly greater than zero ($P < 0.05$).

Figure 2(A) and (B) show efficiency as a function of bandwidth for the patterns detected in noise. First consider Fig. 2(A). The symbols are as for Fig. 1. The efficiencies are on the ordinate and the one-sided bandwidth on the abscissa. Again, the upper abscissa shows the one-sided bandwidth in octaves. For the narrow patterns, efficiencies are high (54–45%) for bandwidths from 2 to 6 octaves for D.K. Efficiencies are higher for the narrow than for the broad patterns, even though the contrast thresholds are higher for the narrow patterns. For the broad patterns, efficiencies remain high (30%) from 1 to 6 octaves with a slight drop to 23% at 0.5 octaves.

Observer P.K. [Fig. 2(B)] has efficiencies of 65, 50 and 69% for octave ranges 2, 4 and 6 for the narrow pattern. Again, efficiencies are lower for the spatially broad patterns. The efficiencies

Table 1. Log contrast threshold vs log bandwidth (c/deg)

Subject	Width (deg)	Noise spectral density (sec deg)	Slope[a] (m)
D.K.	2	0	0.28 ± 0.10
D.K.	0.25	0	0.17 ± 0.20
D.K.	2	10^{-5}	0.27 ± 0.11
D.K.	0.25	10^{-5}	0.28 ± 0.15
P.K.	2	0	0.22 ± 0.09
P.K.	0.25	0	0.27 ± 0.11
P.K.	2	10^{-5}	0.27 ± 0.10
P.K.	0.25	10^{-5}	0.21 ± 0.17

[a]The ± signs indicate 95% confidence intervals.

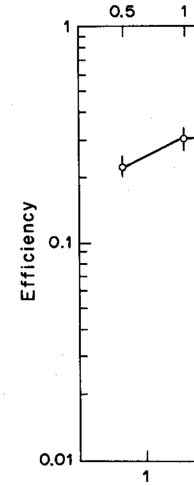

Fig. 2. (A) Tl
and symbols
through a 2-

peak at about 2 o
broad pattern. It
effect of learnin
($s_x = 2°$). Data co
experiment dupl
(within one stan
octave condition
When queried, t
paying more atte
the perceived co

If the observe
wide channels c
expect a drop in
inversely propor
bandwidth in c/d
The threshold sig
ideal is inversely
of the bandwidth

If the observer h
c/deg, then the
density is

The ratio of the
sity (S_i) to the
proportional to
width of the sign
width (B_c) rema

$S_i/$

But this ratio is

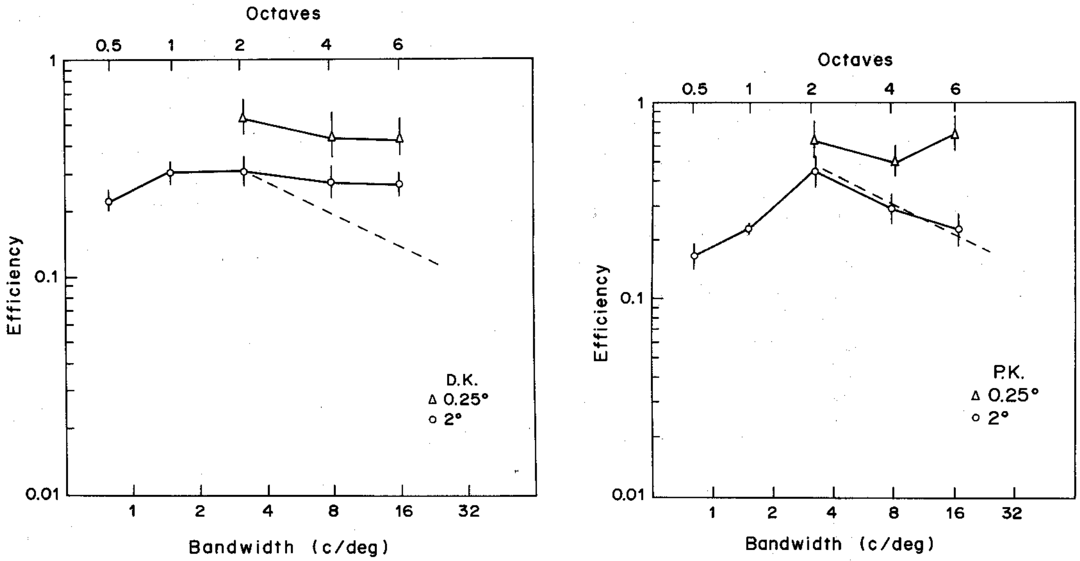

Fig. 2. (A) Thresholds measured in noise from Fig. 1(A) replotted as efficiency vs bandwidth. The abscissas and symbols are as for Fig. 1. The dashed line shows the efficiency expected, if the observer "looked" through a 2-octave wide filter. Observer D.K. (B) Thresholds measured in noise from Fig. 1(B) replotted as efficiency vs. bandwidth. Observer P.K. Other details as for Fig. 2(A).

peak at about 2 octaves for this observer for the broad pattern. It is possible that there was an effect of learning for the 6 octave condition ($s_x = 2°$). Data collected towards the end of the experiment duplicated all of the efficiencies (within one standard error) except for the 6 octave condition which rose from 15 to 35%. When queried, the observer said he was now paying more attention to the texture as well as the perceived contrast.

If the observers "looked" through 2-octave wide channels centered at 2 c/deg, we would expect a drop in efficiency beyond 2 octaves inversely proportional to the square root of bandwidth in c/deg. The reasoning is as follows. The threshold signal spectral density (S_i) of the ideal is inversely proportional to the square root of the bandwidth of the signal (B_i):

$$S_i \propto 1/\sqrt{B_i}.$$

If the observer has a "critical bandwidth" of B_c c/deg, then the observer's threshold spectral density is

$$S_0 \propto 1/\sqrt{B_c}.$$

The ratio of the ideal's threshold spectral density (S_i) to the observer's is thus inversely proportional to the square root of the bandwidth of the signal, the observer's critical bandwidth (B_c) remaining fixed in the numerator

$$S_i/S_0 \propto \sqrt{B_c}/\sqrt{B_i}.$$

But this ratio is equal to the efficiency.

The dashed line shows the drop in efficiency one would expect once the signal bandwidth is outside a critical band of 2 octaves. Except for the 2° data for P.K., both observers behave as if they have a band that increases with signal bandwidth, rather than as if they are looking through only one channel. Another way of interpreting this is that spatial frequency information is being combined with rather high and uniform efficiency from 2 to 6 octaves for detection in noise. This bandwidth is three to six times the channel bandwidth estimated from studies with sinusoidal gratings.

It is interesting to note that although the 0.25° patch is detected with higher efficiency, it is encodable by fewer logons (a maximum of about $32 \times 0.5 = 16$). The maximum logon-content of the 2° patch is about $32 \times 4 = 128$. Although the logon-content at which efficiency begins to drop has not been measured, we would expect it to be limited by the bandwidth of the eye. Thus a fall-off in efficiency should occur at least by 60 logons (0.5×120) and 480 logons (4×120), assuming a one-sided bandwidth of 60 c/deg.

DISCUSSION

The results show that it is possible to integrate spatial-frequency information in noise with high efficiency out to 6 octaves for a center frequency of 2 c/deg. Contrast thresholds grow with the fourth root of bandwidth in the presence and absence of a noise mask for bandwidths greater than 1 or 2 octaves. Efficiencies for wide-band signals are higher for space constants of 0.25° than for 2°.

The finding that contrast thresholds increase with increasing bandwidth is not new. Mitchell (1976) and Mostafavi and Sakrison (1976) report similar findings for two-dimensional noise patterns. However, it has not previously been pointed out that human observers can show the same dependence of contrast threshold on bandwidth as the ideal observer for large bandwidths. Below, we explore some of the implications of this finding. First the energy detector is explained and the notion of adjustable spatial-frequency bandwidths is discussed. Then the energy detector is contrasted with the cross-correlator. It is pointed out how a bank of cross-correlators can be modified to do energy detection, but that efficient combination of the outputs over several spatial scales may be needed to account for these data.

The energy detector. Because the performance of the ideal noise detector for arbitrary spatial and frequency windows is often difficult to compute, it is instructive to look at a simple implementation of the ideal when the windows are rectangular. The energy detector is a realization of the ideal detector for noise signals described by Green and Swets (1974). The energy detector is comprised of a linear filter followed by a square-law device. The output of the square-law device is integrated over space (or time) to give a quantity monotonic with likelihood. The performance d' is given by

$$d' \simeq \frac{\sqrt{2}\,(E/N)}{\sqrt{\{BX + E/N\}}} \quad BX > 10$$

where B is the two-sided bandwidth, X the width, E the contrast energy and N the noise spectral density (over the two-sided bandwidth).* The above equation can be written in terms of contrast as

$$d' \simeq \frac{\sqrt{2c^2X/N}}{\sqrt{\{BX + c^2X/N\}}} \quad BX > 10.$$

When

$$BX \gg c^2X/N.$$

then

$$d' \simeq c^2\sqrt{(2X/B)}.$$

With some rearrangement, this equation produces the fourth root laws for bandwidth and width

$$c \propto \left(\frac{B}{X}\right)^{1/4}.$$

The idea of a square-law operation in vision receives some support from contrast matching data. Perceived contrast of noise patterns seems to be proportional to r.m.s. contrast (Quick *et al.*, 1976; Mayhew and Frisby, 1978). Legge (1984a,b) has proposed a model of binocular interaction, involving summing the squared contrasts to each eye. d' has been known for some time to be an accelerating function of contrast (Nachmias and Kocher, 1970).

Adjustable bandwidths. Quick *et al.* (1976) using a contrast matching paradigm, found no measurable critical band at suprathreshold levels for noise patterns. This conclusion was based on the observation that (after allowing for visual sensitivity based on grating contrast matching results) apparent contrast seemed to be determined by the r.m.s. contrast. Mayhew and Frisby (1978) also report that perceived contrast of noise patterns is proportional to the square root of summed contrast power. More recently, Jamar and Koenderink (1985) measured contrast thresholds for dynamic one-dimensional 4° × 4° noise patterns (filtered to allow for the contrast sensitivity of the eye). They found that threshold r.m.s. contrast was constant as a function of spatial frequency bandwidth up to 4 octaves for various central frequencies. They point out that their results can be interpreted in one of two ways. One can adopt a single channel model in which contrast is detected when the spatial power after filtering exceeds a threshold value. Alternatively, threshold could be determined by probability summation among the outputs of the multiple spatial-frequency channels, assuming a power exponent of 2 (see Quick *et al.*, 1976). Neither of these models explicitly allows for non-signal noise which is essential to interpreting the results presented here. An implicit assumption of the above two models is that any non-signal noise is added just prior to the decision stage. In contrast, the energy detector explicitly assumes additive noise prior to the non-linear squaring operation. If the noise was added after squaring, the contrast threshold would be constant as a

*This derivation is adapted from Green and Swets (1974). The generalization (Appendices 1 and 2) of this equation to include temporal dependence and an arbitrary spatial window was used to compute the efficiencies reported in Results.

function of bandwidth. The finding of fourth root dependence of contrast threshold on bandwidth could be made consistent with the interpretation of Jamar and Koenderink by assuming a particular form for the contrast sensitivity function. However, under the conditions of the experiment reported here, human detection is ideal-like with respect to the stimulus at the screen, without making allowance for the contrast sensitivity of the eye at various spatial frequencies. The similarity of the dependence of contrast threshold in the presence and absence of noise suggests the placement of all noise (internal and external) prior to any hypothesized non-linearity.

There are several ways of interpreting the rise in contrast threshold with increased bandwidth. The observer may adjust his spatial-frequency bandwidth in much the same way as an implementation of the ideal observer might. This is one interpretation which Green (1960) gave for analogous results found in audition. Green (1960) measured the detectability of an auditory noise signal in noise. At a performance level of 75% correct (2AFC), he found that spectral density thresholds decreased with bandwidth (from about 600 to 5000 Hz) as predicted by the ideal observer. The human observers were also "ideal-like" with respect to duration (from 3 to 300 msec). One discrepancy between the ideal for noise signals and human performance was that the human observers had steeper psychometric functions. Efficiencies at the 75% correct point were 25–33%. In neither Green's study, nor in the present one, were the bandwidths randomized from trial to trial. It is possible that "adjustable bandwidths" would be found even if the observer did not know what bandwidth to expect.

Before discussing alternative interpretations to the energy detector, it is useful to contrast the energy detector with the cross-correlator.

The cross-correlator and multiple spatial frequency channels. In contrast to the energy detector, the cross-correlator is an implementation of the ideal detector for a known deterministic signal (Kersten, 1983; Kersten, 1984; Burgess, 1985). The cross-correlator does rather poorly for the detection of noise signals without the introduction of an appropriate non-linearity at some stage. The average response of a cross-correlator to a stimulus with zero mean is zero. Because the noise signal and the noise mask functions have zero means, d' is zero. However, squaring the output as with the energy detector, or taking the absolute value solves this problem by making use of knowledge that the variances differ. One can imagine the latter scheme being accomplished by taking the maximum response of a pair of "on-type" and "off-type" single cell cross-correlators.

An improvement on the single cross-correlator, is to assume a bank of similar cross-correlators distributed across space. This aggregate would comprise a channel of a specific bandwidth. If the outputs are squared and then summed, we are back to an energy detector of a particular bandwidth and spatial extent. However, other non-linearities and combination schemes may work as well. For example, one could take the maximum responses over all the on-center and off-center simple cell cross-correlators. This is an example of a channel uncertainty model (Pelli, 1985). As the dashed line in Fig. 2 shows, for most of the conditions, the observer could not be using a single moderate bandwidth channel. The results could then be interpreted as monitoring the channel appropriate for the bandwidths of the stimulus. One problem with this interpretation is that estimates of spatial-frequency bandwidth in vision are usually at most two octaves.

An alternative explanation which is in line with narrower band channels was also suggested by Green for his auditory results. Detection can still be mediated by relatively narrow-band channels. As bandwidth increases, additional channels are brought into play. However, the high efficiencies and form of the contrast threshold dependence on bandwidth argue that the combination of channel outputs should be efficient.

CONCLUSIONS

Human detection efficiencies for certain deterministic signals in visual noise can be high enough to rule out most models except for a cross-correlator matched to the signal (Burgess *et al.*, 1981; Kersten, 1984). Broad-band noise detection such as found here and in the dot density experiments yield efficiencies too high to be accounted for by a cross-correlator. This indicates that human observers are capable of adjusting detection strategy to the task. This flexibility may be less for contrast detection against a uniform field where high intrinsic uncertainty may be unavoidable, at least for large field viewing (Pelli, 1985; Cohn and Wardlaw, 1985). This suggests that higher quantum efficiencies may be found for noise signal detection than for deterministic pattern detection.

Acknowledgements—I would like to thank Dr G. E. Legge, Dr D. Pelli, Dr G. Rubin and G. Wakefield for many hours of useful discussion. I also wish to thank Drs J. G. Robson, N. Viemeister, D. Burkhardt, R. Purple and H. Gershenson for criticisms and suggestions. M. Bergman, B. Witkofsky, M. Schleske and D. Koenen provided technical support. Joanne Kersten typed and Philip Kersten observed. The experiments reported in this paper formed part of a doctoral dissertation submitted to the University of Minnesota (Kersten, 1983). The research was supported by PHS grant EY02857 to Dr Legge and a University of Minnesota Doctoral Dissertation Grant. This paper was completed at the Physiological Laboratory, University of Cambridge, with the support of an NIH postdoctoral fellowship EY05660.

REFERENCES

Barlow H. B. (1962) Measurements of the quantum efficiency of discrimination in human scotopic vision. *J. Physiol., Lond.* **160,** 169–188.

Barlow H. B. (1962) A method of determining the overall quantum efficiency of visual discrimination. *J. Physiol., Lond.* **160,** 155–168.

Barlow H. B. (1978) The efficiency of detecting changes of density in random dot patterns. *Vision Res.* **18,** 637–650.

Burgess A. (1985) Visual signal detection III. On Bayesian use of prior knowledge and cross correlation. *J. opt. Soc. Am.* **2,** 1498–1507.

Burgess A. E. and Barlow H. B. (1983) The efficiency of numerosity discrimination in random dot images. *Vision Res.* **23,** 811–829.

Burgess A. E., Wagner R. F., Jennings R. J. and Barlow H. B. (1981) Efficiency of human visual signal discrimination. *Science, N.Y.* **214,** 93–94.

Campbell R. W. and Green D. G. (1965) Optical and retinal factors affecting visual resolution. *J. Physiol., Lond.* **181,** 576–593.

Campbell R. W. and Robson J. G. (1968) Application of Fourier analysis to the visibility of gratings. *J. Physiol., Lond.* **197,** 551–566.

Chambers R. R. and Courtney-Pratt J. S. (1969) Experiments on the detection of visual signals in noise using computer-generated signals. *Photographic Sci. Engng* **13,** 286–298.

Cohn T. E. and Wardlaw J.C. (1985) Effect of large spatial uncertainty on foveal luminance increment detectability. *J. opt. Soc. Am.* **A2,** 820–825.

Cohn T. E. (1976) Quantum fluctuation limit in foveal vision. *Vision Res.* **16,** 573–579.

Davies W. D. T. (1970) *System Identification for Self-Adaptive Control.* Wiley-Interscience, New York.

Davis E. T. and Graham N. (1981) Spatial frequency uncertainty effects in the detection of sinusoidal gratings. *Vision Res.* **21,** 705–712.

De Vries H. (1943) The quantum character of light and its bearing upon the threshold of vision, the differential sensitivity and acuity of the eye. *Physica* **10,** 553–564.

Gabor D. (1946) Theory of communication. *J. Inst. Electr. Engrs, Lond.* **93,** 429–457.

Green D. M. (1960) Auditory detection of a noise signal. *J. acoust. Soc. Am.* **32,** 121–131.

Green B. F., Wolf A. K. and White B. W. (1959) The detection of statistically defined patterns in a matrix of dots. *Am. J. Psychol.* **72,** 503–520.

Green D. M. and Swets J. (1966) *Signal Detection Theory and Psychophysics.* Wiley, New York.

Greis U. and Rohler R. (1970) A study of the subjective detectability of patterns by means of spatial frequency filtering. *Optical Acta* **17,** 515–526.

Hecht S., Shlaer S. and Pirenne M. H. (1942) Energy, quanta, and vision. *J. gen. Physiol.* **25,** 819–840.

Horowitz P. and Hill W. (1980) *The Art of Electronics.* Cambridge Univ. Press.

Jamar J. H. T. and Koenderik J. J. (1985) Contrast detection and detection of contrast modulation for noise gratings. *Vision Res.* **15,** 511–521.

Kersten D. (1983) A comparison of human and ideal performance for the detection of visual pattern. Ph.D. thesis, Univ. of Minnesota, Minneapolis, Minnesota.

Kersten D. (1984) Spatial summation in visual noise. *Vision Res.* **24,** 1977–1990.

Kersten D. and Barlow H. B. (1985) Why are contrast thresholds so high? *Invest. Ophthal. visual Sci. Suppl.* **26,** 140.

Koenderink J.J. and van Doorn A. J. (1974) Spatial noise for visual research. *Vision Res.* **14,** 721–723.

Lasley D. J. and Cohn T. E. (1981) Why luminance discrimination may be better than detection. *Vision Res.* **21,** 273–278.

Legge G. E. (1984a) Binocular contrast summation I. Detection and discrimination. *Vision Res.* **24,** 373–383.

Legge G. E. (1984b) Binocular contrast summation II. Quadratic summation. *Vision Res.* **24,** 385–394.

MacKay D.M. (1950) Quantal aspects of scientific information. *Phil. Mag.* Ser. 7, Vol. XLC, 289–311.

Mayhew J. E. W. and Frisby J. P. (1978) Suprathreshold contrast perception and complex random textures. *Vision Res.* **18,** 895–897.

Mitchell O. R. (1976) Effect of spatial frequency on the visibility of unstructured patterns. *J. opt. Soc. Am.* **66,** 327–332.

Mostafavi H. and Sakrison D. J. (1976) Structure and properties of a single channel in the human visual system. *Visual Res.* **16,** 957–986.

Nachmias J. and Kocher E. C. (1970) Visual detection and discrimination of luminance increments. *J. opt. Soc. Am.* **60,** 382–389.

Papoulis A. (1965) *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York.

Pelli D. (1981) Effects of visual noise. Ph.D. thesis, Cambridge Univ., England.

Pelli D. (1985) Uncertainty explains many aspects of visual contrast detection and discrimination. *J. opt. Soc. Am.* **A2,** 1508–1532.

Pollehn H. and Roehrig H. (1970) Effect of noise on the MTF of the visual channel. *J. opt. Soc. Am.* **60,** 842–848.

Quick R. F. Jr, Hamerly J. R. and Reichert T. A. (1976) The absence of a measurable "critical band" at low suprathreshold contrasts. *Vision Res.* **16,** 351–355.

Quick R. F., Mullins W. W. and Lucas R. N. (1978) Contrast thresholds of random patterns. *Photo, Sci. & Engng* **22,** 72–75.

Richards W. and Polit A. (1974) Texture matching. *Kybernetick* **16,** 155–162.

Roberts T. A. (1963) Analysis and synthesis of linear and non-linear shift register generators. *Int. Telemetering Conf.* **1,** 390–399. IEE Savoy Place, London.

Rose A. (1942) The relative sensitivities of television pickup tubes, photographic film, and the human eye. *Proc. I.R.E.* **30,** 293–300.

Sachs M. B., Nachmias J. and Robson J. B. (1971) Spatial-frequency channels in human vision. *J. opt. Soc. Am.* **61,** 1176–1186.

Schade O. H. Sr (1956) Optical and photoelectric analog of the eye. *J. opt. Soc. Am.* **46,** 721–739.

Stromeyer C. F. and Julesz B. (1972) Spatial-frequency masking in vision: Critical bands and spread of masking. *J. opt. Soc. Am.* **62,** 1221–1232.

Sturm R. E. and Morgan R. H. (1949) Screen intensification systems and their limitations. *Am. J. Roentg. Radium Ther.* **62,** 617–634.

Tanner W. P. (1961) Physiological implications of psychophysical data. *Ann. N.Y. Acad. Sci.* **89,** 752–765.

Tanner W. P. and Birdsall T. G. (1958) Definitions of d' and η as psychological measures. *J. acoust. Soc. Am.* **30,** 922–928.

Van Meeteren A. and Barlow H. B. (1981) The statistical efficiency for detecting sinusoidal modulation of average dot density in random figures. *Vision Res.* **21,** 765–789.

Van Trees H. L. (1971) *Detection, Estimation, and Modulation Theory. Part III.* Wiley, New York.

Watson A. B. and Pelli D. G. (1983) QUEST: A bayesian adaptive psychometric method. *Percept. Psychophys.* **33,** 113–120.

Watson A. B., Barlow H. B. and Robson J. G. (1983) What does the eye see best? *Nature, Lond.* **302,** 419–422.

APPENDIX 1

The Low-Energy-Coherence (LEC) Approximation

Van Trees (p. 136, part III, 1971) gives an approximate expression for the optimal performance for noise signal detection in white Gaussian noise. We first consider only one frame (e.g. 10 msec), where d'_x is the normal deviate of the proportion correct in a 2AFC procedure. Then d'_x is given by:

$$d'^2_x = 0.5(1/N_0)^2 \iint K^2(x, u)\,dx\,du$$

where $K^2(u, x)$ is the signal covariance function. This approximation is good for S_0/N_0 much less than one. If the signal is Gaussian enveloped and has a flat spectrum inside a passband of width B_x, $K(x, u)$ is given by

$$m(x/s_x)m(u/s_x)S_0 B_x \operatorname{sinc}[\mathrm{B_x(x - u)}]$$

where $m(x/s_x)$ is the Gaussian envelope with space constant s_x. S_0 d'^2_x is given by

$$d'^2_x \simeq 0.5\ (S_0/N_0)^2 B_x^2 \int m^2(x/s_x) \int m^2(u/s_x) \times \operatorname{sinc}^2[B_x(x-u)]\,du\,dx.$$

Now the envelope, $m(x/s_x)$ is given by $\exp\{-(x/s_x)^2\}$. For large $B_x s_x$ ($B_x s_x \gg 2$), sinc^2 can be approximated by a delta function, thus the inner integral is approximately

$$\exp[-(x/s_x)^2].$$

Thus we have

$$d'^2_x \simeq 0.5\ (S_0/N_0)^2 B_x^2 \int m^4(x/s_x)\,dx.$$

Using the fact that

$$\int \exp(-x^2/s_x^2)\cos(bx)\,dx = s_x \exp(-4b^2 s_x^2)$$

we have

$$d'^2_x \simeq 0.25\ \sqrt{\pi}(S_0/N_0)^2 B_x s_x.$$

When observations are made over time, d' is a function of time and is given by

$$d'^2(t) \simeq m^4(t/s_x)d'^2_x$$

and integrating over time

$$d'^2 \simeq (\sqrt{\pi}/2)B_t s_t \sqrt{\pi} d'^2_x.$$

Thus, a factor of $(\sqrt{\pi}/2)s_t B_t$ has to be included to take account "looks" over time. The final expression is:

$$d'^2 \simeq (\pi/8)(S_0/N_0)^2 B_x B_t s_x s_t.$$

APPENDIX 2

A Closed-form Solution and Simulation of the Ideal

The ideal was characterized as follows. The noise signal is assumed to be Gaussian, stationary and spectrally flat (prior to applying the envelope) with a 2-sided spatial bandwidth of B_x. The noise mask is also assumed to be spectrally flat and Gaussian, but with a temporal frequency bandwidth of B_t and a spatial frequency bandwidth greater than the signal.

For the moment assume the signal to be infinite in extent. It can be represented in terms of an infinite, but countable number of uncorrelated values sampled at the Nyquist intervals spaced by $1/B_x$ and $1/B_t$ (Papoulis, 1965). For strictly time or space limited signals, it is no longer true that the sampled values are statistically uncorrelated. The Karhunen-Loeve expansion is used for this case (Papoulis, 1965). However, the assumption is made that neither humans nor the computer simulated ideal benefit very much from the samples outside of plus or minus 4 space (or time) constants from the center of the Gaussian windows. Computer simulations of 8 space constants verified this.

Consider one frame (e.g. 10 msec), temporarily ignoring time. Let S_0 and N_0 be the static spectral densities of the signal and noise respectively. Then the likelihood ratio that the signal is present vs. not present in a given interval is monotonic with the sum of the squared sample contrast function values r_i^2, each weighted in importance by the ratio of the signal spectral density to the signal plus noise spectral density at that point (Van Trees, p. 11 part III, 1971):

$$l_x = \sum r_i^2 \times m_i^2 S_0/(m_i^2 S_0 + N_0)$$

m_i is the envelope value at that point. When the signal is present, r_i is the sum of the Gaussian random variables due to signal and noise at point i. When the signal is absent, r_i is drawn from the noise only distribution. By including time, we just get more uncorrelated samples to help in the decision. The samples are uncorrelated across time because the noise signal samples are "deterministic" across time. Thus the sum is over both space and time and m_{ij} is the value of the envelope at spatial sample point i and temporal sample point j

$$l = \sum r_{ij}^2 \times m_{ij}^2 S_0/(m_{ij}^2 S_0 + N_0).$$

For sufficiently large values of $B_x B_t s_x s_t$, the likelihood ratio is approximately normally distributed. (As a check on this assumption, the simulations described below were done.) This assumption permits the calculation of d'. Analogous to Green (1960), d' is given by

$$d' = \sqrt{2MA}/[\sqrt{(BM^2 + 2CM + 2D)}]$$

where

$$M = S_0/N_0,$$

$$A = \sum m_{ij}^4/(m_{ij}^2 + 1),$$

$$B = \sum m_{ij}^8(m_{ij}^2 + 1)^2,$$

$$C = \sum m_{ij}^6/(m_{ij}^2 + 1)^2$$

and

$$D = \sum m_{ij}^4/(m_{ij}^2 + 1)^2.$$

To simulate the ideal, a value of l was generated for each stimulus presentation of a 2AFC trial. The interval with the largest l was chosen as the interval containing the signal. The QUEST procedure was used to find the contrast threshold. The simulation of the ideal was done in RT-11 BASIC v. 2 in single precision and the values verified with "psychometric function" simulations done with double precision RT-11 FORTRAN v. 4. Both simulations were done on LSI-11/23.

One assumption of this implementation of the ideal is that in the actual experiment, the window is sampled at the same rate as the noise signal. This, in general, is not true because the window was applied after filtering. However, for all but the smallest bandwidths ($B_x < 6$), the LEC (Appendix 1) approximation (which does not share this assumption) gave the same results to within the standard error of the contrast thresholds (less than 8%).

Vision Res. Vol. 27, N
Printed in Great Britai

[1]Neurologis

Abstract—T
approx 1 hr
contrast thr
type of ada

Aftereffects

An economical
suggests that the
that result from
contrast pattern
similar or closel
a common basis
tation and spati
of contrast thres
ent contrast of
the shifts in pe
orientation (Bra
discrepancies be
occasionally not
tenback, 1979;
1981; Magnusse
been suggested
threshold aftere
(Klein *et al.*, 1
O'Connell, 1986
dence for a unit
In a recent
Greenlee, 1985)
threshold elevati
saturation point
and tracked its
conducted a si
aftereffect, and
time-courses of t
ery from contir
these two afterer
The tachistos
suring the tilt a
several previous

*To whom correspo