

Factors limiting suprathreshold vision measured by a flash–sound simultaneity paradigm

Velitchko Manahilov ^{a,*}, Kellyanne Findlay ^a, William A. Simpson ^b

^a *Vision Sciences Department, Glasgow Caledonian University, Cowcaddens Road, Glasgow G4 0BA, UK*

^b *Simulation & Modelling Section DRDC Toronto, 1133 Sheppard Avenue West, P.O. Box 2000, Toronto, Ont., Canada M3M 3B9*

Received 22 October 2003; received in revised form 21 April 2004

Abstract

Internal noise and sampling efficiency are the main factors which limit visual performance. In a previous study [Vis. Res. 43 (2003) 1103] we compared the variance of human reaction time to that of an ideal observer and found that the sampling efficiency to suprathreshold stimuli was much lower than that obtained in detection experiments. In order to bypass the effects of the motor system on visual performance, we used a flash–sound simultaneity paradigm. We found that the sampling efficiency for 0.4- and 4-c/deg near-threshold Gabor patches is higher only by a factor of 2.5 than that to above-threshold patterns. The signal-dependent multiplicative internal noise was similar to the additive internal noise at lower signal contrast levels and exceeded it at higher signal contrast levels. The results show that real observers' performance for detecting suprathreshold stimuli can be accounted for by a model taking into account the non-linear visual–signal transduction and multiplicative components of the internal noise induced by the signal and external noise. In addition, this model assumes that performance depends on the response duration, rather than signal duration. The results imply that the multiplicative internal noise induced by high contrast visual signals determines performance for suprathreshold visual detection.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Vision; Detection; Efficiency; Noise; Arrival time

1. Introduction

Our ability to extract information from incomplete and noisy sensory messages is limited by two main factors: the level of internal noise and sampling efficiency. Any real system, including the visual system, suffers from internal noise which is due to random variations in the level of background neural activity at different visual stages. Additionally, humans do not use all available a priori information for the visual signal which results in a sub-optimal sampling efficiency. Inefficiency can be characterised by measuring human performance and comparing it to the best possible

performance of an ideal observer (Barlow, 1978; Burgess, Wagner, Jennings, & Barlow, 1981; Legge, Kersten, & Burgess, 1987; Nagaraja, 1964; Pelli, 1990). The sampling efficiency and internal noise of a real observer can be estimated by measuring how the signal contrast energy required for an observer to maintain a given performance level depends on the amount of external visual noise added to the signal (Pelli, 1990). This approach has been used successfully to measure internal noise and efficiency for visual detection of near-threshold stimuli (Burgess et al., 1981; Legge et al., 1987; Nagaraja, 1964; Pelli, 1990). In many daily tasks, however, humans have to respond to high contrast visual signals whose signal-to-noise ratio is high and the available signal detection theory methods are not applicable to measure human detection performance.

* Corresponding author. Tel.: +44-141-331-8204; fax: +44-141-331-3387.

E-mail address: vma@gcal.ac.uk (V. Manahilov).

Recently, we reported a method for assessment of the visuomotor sampling efficiency and internal noise in the detection of suprathreshold stimuli which is based on comparing the variance of human reaction times (RTs) to that of an ideal observer (Simpson, Findlay, & Manahilov, 2003). We found that the internal noise was similar to that measured in detection experiments and remained constant as the signal contrast increased. The sampling efficiency was much lower compared to previous detection experiments. We suggested that the discrepancy between our and other findings might be explained by the idea that a large part of the variability in RT arises from the motor system rather than the visual system. Instead of using RTs, we will use a flash–sound simultaneity paradigm. A sound is presented at a fixed time, and a flash is displayed at one of several times either before or after the sound. The observer judges which occurred first. Just as in the case of simple RT, the task requires an estimation of the flash’s time of arrival. As we will show in the next section, the precision of the simultaneity judgements can be used to determine the internal noise and sampling efficiency of suprathreshold visual detection.

Another possible reason for the low sampling efficiencies we found previously may have been our use of a simplified linear amplifier model which does not take into account multiplicative components of the internal noise induced by the signal and the external noise (Burgess & Colborne, 1988; Tolhurst, Movshon, & Thompson, 1981). Moreover, RT distributions are usually skewed and their approximation with a Gaussian function is not always correct.

Here we develop and test a model for estimation of arrival time of a visual signal which is based on the perceptual template model for near-threshold visual detection (Lu & Doshier, 1999).

2. Models

We propose a model for estimation of arrival time of a visual signal. An ideal observer is presented with two signals: an auditory signal (a brief loud tone) and a visual signal (a pattern flash). The auditory signal is always presented at a fixed time (τ_a) in the absence of auditory noise. The visual signal, embedded in dynamic Gaussian white visual noise, is delivered at a random time (τ_v). The observer’s task is to say which signal came first. This requires that the observer compare the times of arrival of the two signals. The ideal observer knows the time of arrival of the auditory signal exactly, since it always arrives at the same time, and because it is not embedded in any auditory noise. The time of arrival of the visual signal, however, must be estimated. The optimal algorithm to do this is to cross-correlate the received noisy visual waveform with the expected visual signal at every

possible time of arrival. The time at which the peak of this cross-correlation function occurs is the estimated time of arrival of the visual signal (Simpson et al., 2003; Woodward, 1980, pp. 104–105).

Due to the visual noise, the estimated time of arrival of the visual signal has a normal distribution with mean equal to the actual time of arrival, and with variance (Simpson et al., 2003; Woodward, 1980, pp. 104–105):

$$\sigma^2 = \frac{D^2 \sigma_N^2}{E}, \quad (1)$$

where D is the signal duration, σ_N^2 is the external noise variance and E is the signal energy (integrated squared contrast signal waveform).

As we noted, the ideal observer knows the time of arrival of the auditory signal (τ_a) exactly, since it always occurs at the same time and is not embedded in auditory noise. On the other hand, the estimated time of arrival of the visual signal ($\hat{\tau}_v$) is distributed normally with mean equal to the actual time of arrival, and with variance σ^2 . Therefore, the discriminability of the times of arrival of a fixed auditory signal and a randomly arriving visual signal could be expressed as

$$d'^2 = \frac{(\hat{\tau}_v - \tau_a)^2}{\sigma^2}. \quad (2)$$

So far we have modelled performance of an ideal observer. Barlow (1978), Legge et al. (1987) and Pelli (1990) developed a linear amplifier model (LAM), which assumes that real observers do not use all the energy in the delivered signal, that is they have sub-optimal sampling efficiency. Real observers also appear to add extra internal noise to the stimulus. On the other hand, brief stimuli evoke prolonged visual responses (Rashbass, 1970; Watson, 1986). Thus, to model performance of real observers we will modify Eq. (1) including these factors:

$$\sigma^2 = \frac{D_R^2 (\sigma_N^2 + \sigma_{\text{add}}^2)}{kE}, \quad (3)$$

where k is the sampling efficiency, σ_{add}^2 is the variance of the additive internal noise and D_R is the response duration.

Lu and Doshier (1999) have postulated that even more parameters need to be considered: the exponent of the power transducer function (γ) and the parameter (m) determining the amounts of the signal energy [$m(kE)^\gamma$] and the external noise variance ($m\sigma_N^{2\gamma}$) which act as multiplicative internal noise. They proposed a perceptual template model (PTM) for detection of visual stimuli which predicts that the detectability index (d') could be expressed as

$$d'^2 = \frac{(kE_{\text{thr}})^\gamma}{\sigma_N^{2\gamma} + m\sigma_N^{2\gamma} + m(kE_{\text{thr}})^\gamma + \sigma_{\text{add}}^2}. \quad (4)$$

In the present study we will describe threshold energy for detection of Gabor patches as a function of external noise variance by rearranging Eq. (4) as follows:

$$E_{\text{thr}} = \frac{d'^{2/\gamma} [\sigma_N^{2\gamma} + m\sigma_N^{2\gamma} + \sigma_{\text{add}}^2]^{1/\gamma}}{k(1 - md^2)^{1/\gamma}} \quad (5)$$

Applying this perceptual template model to supra-threshold visual detection, the variance of the real observers' estimations of arrival time of a visual signal could be described by taking into account the power transducer function (γ) and the multiplicative internal noise components due to the signal energy [$m(kE)^\gamma$] and the external noise variance ($m\sigma_N^{2\gamma}$). Therefore, Eq. (3) could be modified as follows:

$$\sigma^2 = \frac{D_R^2 [\sigma_N^{2\gamma} + m\sigma_N^{2\gamma} + m(kE)^\gamma + \sigma_{\text{add}}^2]}{(kE)^\gamma} \quad (6)$$

Note that Eq. (6) takes account of the increased variance of real observers' performance relative to the ideal observer due to various visual stages. It is likely true that similar stages take place in auditory transduction. However, since the auditory signal always was delivered at the same time, and since it was not embedded in auditory noise, it seems plausible that the noise associated with the representation of the auditory signal relative to that of the visual signal would be negligible.

3. Methods

The stimuli were generated on a RGB monitor having a frame rate of 120 Hz and a spatial resolution of 640×480 pixels. A custom video summation device (Pelli & Zhang, 1991) was used to produce 256 grey levels with a 12-bit precision. The mean luminance was 30 cd/m². The viewing distance was 171 cm.

The visual stimuli were horizontal Gabor patches with a spatial frequency of 0.4 and 4 c/deg and a spatial spread (1 SD) of 1.5°. The stimuli were embedded in dynamic 2D Gaussian white noise whose Gaussian distribution was clipped at 2.5 standard deviations. The noise pixel size was 2.6'×2.6'. The threshold contrast for detection of these stimuli was measured using a two interval forced choice method and a staircase procedure (Manahilov, Calvert, & Simpson, 2003).

In the main experiment, observers were presented with a tone of 8.3-ms duration and 3-kHz frequency using two speakers which were close to the stimulation screen. The sound-pressure level at the observer location was 85 dB. Each trial started with a warning signal (fixation mark) followed by dynamic white Gaussian visual noise of 1.5-s duration. The auditory signal was presented 1 s after the fixation mark. The visual signal was displayed randomly at various flash-sound signal onset asynchronies (SOAs): -250, -166, -83, 0, 83, 166, 250 ms. The signal duration was one screen frame

(8.3 ms) in which only the visual signal was presented. Temporal integration within the visual system combined the noiseless visual signal with noise from frames before and after it. The contrast of the Gabor patches was approximately 5, 7 and 9 times above the threshold contrast for detection of each stimulus in the absence of visual noise. Observers were asked to press one of two keys: "Flash was first" or "Flash was second". For each condition at least 50 trials were presented in two different experimental sessions. One of the authors and a naïve observer took part in the experiments. They had normal vision and viewed the stimuli binocularly. Both observers were experienced in other psychophysical experiments and were given a few training sessions.

4. Results

Fig. 1 shows the threshold energy for detection of Gabor patches of 0.4 c/deg (circles) and 4 c/deg (squares) for two observers as a function of external noise variance. The lines in Fig. 1 denote the predictions of the PTM (Eq. (5)) fitted to the data points of each observer and spatial frequency by the least-square method. At threshold, $d' = 1.14$ (proportion correct responses 0.79 in a three-down one-up 2IFC staircase). The signal energy was computed numerically for the actual stimuli by squaring and summing the contrast of each pixel. The energy values could be converted in units of $\mu\text{deg}^2\text{s}$ by a multiplicative factor of 2.314 ($1000000 \times 0.0083 \times 0.0166^2$). The results show that for both Gabor patches the additive internal noise, the multiplicative coefficient and the exponent of the power transducer function are similar (Table 1). The sampling efficiency for 0.4-c/deg

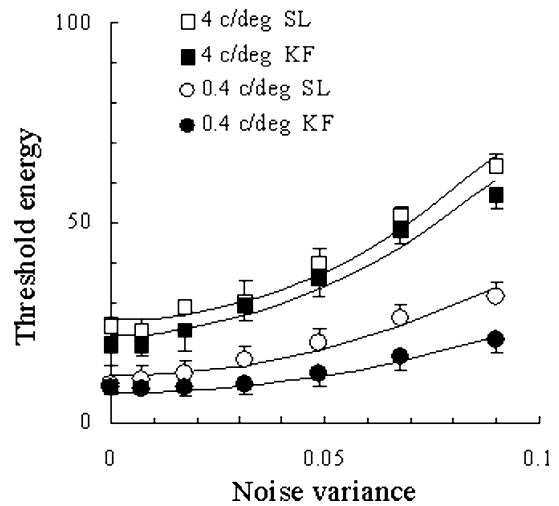


Fig. 1. Threshold energy for detection of 0.4- and 4-c/deg Gabor patches as a function of external noise variance. Vertical bars show 95% confidence interval. Data for two observers.

Table 1

Best-fitting values of the parameters of the PTM (Eq. (5)) used to fit the threshold data shown in Fig. 1

Spatial frequency	k		σ_{add}^2		m		γ		R^2	
	KF	SL	KF	SL	KF	SL	KF	SL	KF	SL
0.4 c/deg	0.0021	0.0014	0.0060	0.0067	0.60	0.55	2.07	2.00	0.95	0.95
4 c/deg	0.0005	0.0003	0.0065	0.0050	0.30	0.25	1.96	2.10	0.95	0.96

R^2 —the proportion of the variance accounted for by the fit, adjusted by the number of free parameters.

Gabor patches was about 4–5 times higher than that for 4-c/deg Gabor patches.

The proportion of flash-first responses/SOA functions were fitted by a cumulative Gaussian function using a bootstrap procedure with 1000 iterations (Foster & Bischof, 1991). The standard deviation of the Gaussian functions increased as external noise variance increased which is illustrated by the effect of external noise level on the steepness of the response proportion/SOA functions shown in Fig. 2.

The external noise also shifted these functions towards negative SOAs. This means that as the external noise increased, the visual signal was perceived to occur earlier. It seems that degrading the signal-to-noise ratio of the visual stimulus has the paradoxical effect of

spreading up the response proportion/SOA functions relative to the auditory signal.

Fig. 3 shows the variance of flash-first responses (symbols) for both observers as a function of external noise variance. The data for each observer and spatial frequency were fitted with a set of nested linear amplifier models that systematically varied the parameters [sampling efficiency (k), additive internal noise (σ_{add}) and response duration (D_R)] of Eq. (3). When the data were fitted with a model, in which a common set of free parameters was used for the three signal contrast levels, the R^2 values (see Appendix A) were 0.902 (KF) and 0.737 (SL) for 0.4 c/deg; 0.808 (KF) and 0.803 (SL) for 4 c/deg. The proportion of variance accounted for by the fit was higher when different free parameters were

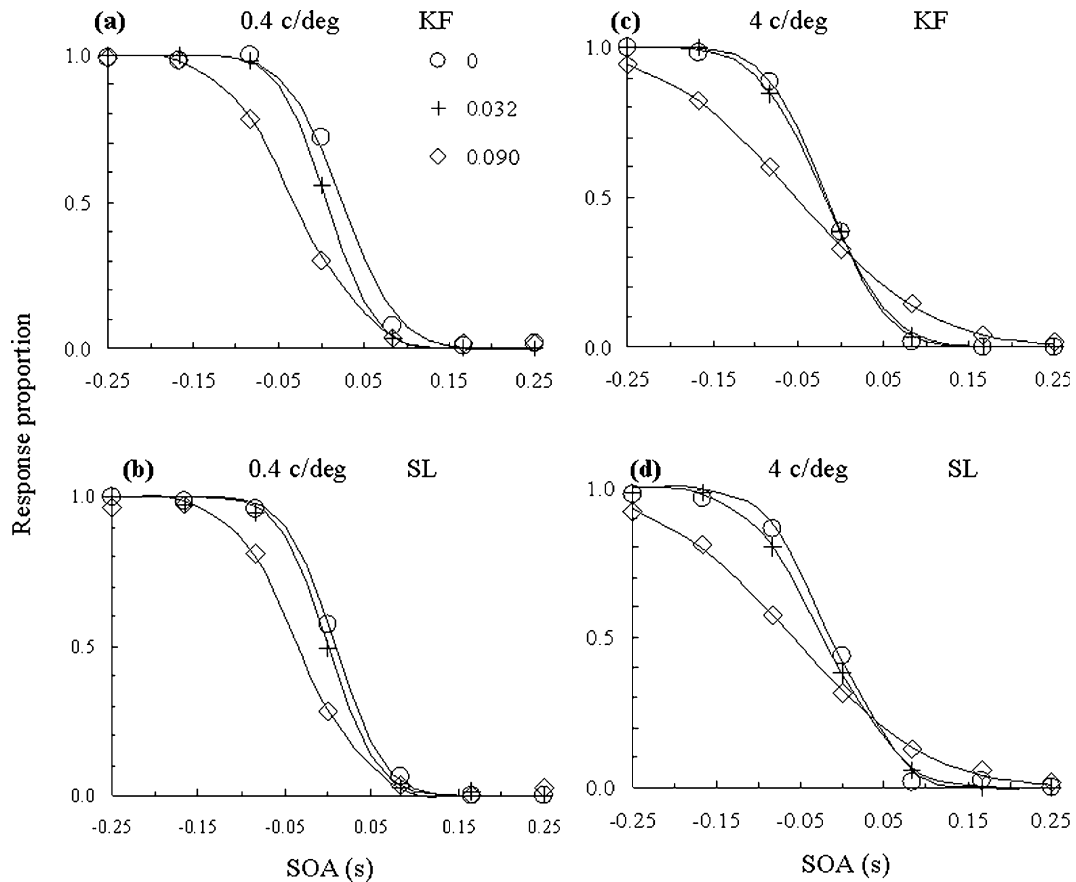


Fig. 2. Flash-first response proportion as a function of SOA (symbols) for various levels of external noise variance as given in insets. Lines represent fitted cumulative Gaussian functions to the data points. Data for detection of 0.4-c/deg Gabor patches of 0.28 contrast [(a) and (b)] and 4 c/deg-Gabor patches of 0.35 contrast [(c) and (d)]. Data for two observers.

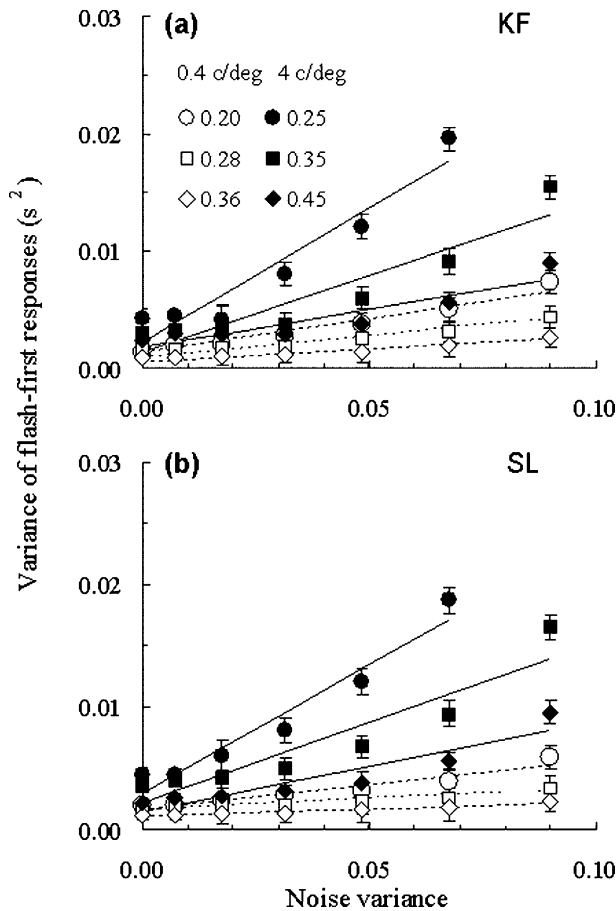


Fig. 3. Variance of flash-first responses as a function of external noise variance. Data for two observers obtained with Gabor patches of 0.4 (empty symbols) and 4 c/deg (filled symbols). Lines show the predictions of the LAM-based $3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$ model whose parameters are shown in Tables 2 and 3. Vertical bars show 95% confidence interval.

allocated to each signal-contrast function. For example, a model with different parameters k and common

parameters σ_{add} and D_{R} ($3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$) increased the R^2 values to: 0.937 (KF) and 0.887 (SL) for 0.4 c/deg; 0.873 (KF) and 0.887 (SL) for 4 c/deg; a model with different parameters k , σ_{add} and D_{R} ($3k - 3\sigma_{\text{add}} - 3D_{\text{R}}$) yielded R^2 values of: 0.947 (KF) and 0.909 (SL) for 0.4 c/deg; 0.845 (KF) and 0.863 (SL) for 4 c/deg.

We compared the quality of fit of these models using the Akaike’s method (Burnham & Anderson, 2002, pp. 60–85; Appendix A). We found that the $3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$ model produced the smallest value of the corrected Akaike’s Information Criterion (AIC_c): -318 (KF) and -321 (SL) for 0.4 c/deg; -246 (KF) and -250 (SL) for 4 c/deg (Tables 2 and 3). The differences between the AIC_c values yielded by the other models and those of the $3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$ model were in the range of 6–25 (evidence ratio >20 ; Akaike’s weight $>95\%$). According to the Akaike’s method, these findings indicate that the $3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$ model would be 20 times more likely of being correct than the other models. It should be noted that the Akaike’s method does not compute a probability value and does not make conclusions about statistical significance. The Akaike’s Information Criterion determines how well the data supports each model, which model is more likely to be correct and quantify how much more likely.

The data were also fitted by a set of nested perceptual template models (Eq. (6)). The simplest model having a common set of free parameters for the three signal contrast levels yielded R^2 values of 0.958 (KF) and 0.939 (SL) for 0.4 c/deg; 0.918 (KF) and 0.880 (SL) for 4 c/deg. The R^2 values increased for models with more free parameters and they were in the range of: 0.973–0.989 (KF) and 0.959–0.983 (SL) for 0.4 c/deg; 0.948–0.990 (KF) and 0.945–0.990 (SL) for 4 c/deg.

The smallest AIC_c value was found with the $3k - 1\sigma_{\text{add}} - 1D_{\text{R}} - 1m - 1\gamma$ model: -337 (KF) and -340 (SL) for 0.4 c/deg; -300 (KF) and -294 (SL)

Table 2

Best-fitting values of the parameters of the LAM-based $3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$ model (Eq. (3)) and the PTM-based $3k - 1\sigma_{\text{add}} - 1D_{\text{R}} - 1m - 1\gamma$ model (Eq. (6)) used to fit the time-of-arrival data for Gabor patches of 0.4 c/deg. R^2 values and model comparison based on the corrected Akaike’s Information Criterion (AIC_c)

Parameter	Contrast	KF		SL	
		$3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$	$3k - 1\sigma_{\text{add}} - 1D_{\text{R}} - 1m - 1\gamma$	$3k - 1\sigma_{\text{add}} - 1D_{\text{R}}$	$3k - 1\sigma_{\text{add}} - 1D_{\text{R}} - 1m - 1\gamma$
$k \times 10^{-4}$	0.20	0.48	7.9	0.45	8.0
	0.28	0.38	6.3	0.34	5.5
	0.36	0.37	5.3	0.30	4.9
σ_{add}^2		0.024	0.0026	0.058	0.0022
$D_{\text{R}}(s)$		0.024	0.123	0.018	0.108
m		–	0.05	–	0.09
γ		–	1.74	–	2.08
R^2		0.937	0.981	0.887	0.964
AIC_c		-318	-337	-321	-340
ΔAIC_c			19		19
Akaike’s weight (%)			99.992		99.992
Evidence ratio			13360		13360

Table 3

Best-fitting values of the parameters of the LAM-based $3k - 1\sigma_{\text{add}} - 1D_R$ model (Eq. (3)) and the PTM-based $3k - 1\sigma_{\text{add}} - 1D_R - 1m - 1\gamma$ model (Eq. (6)) used to fit the time-of-arrival data for Gabor patches of 4 c/deg. R^2 values and model comparison based on the corrected Akaike's Information Criterion (AIC_c)

Parameter	Contrast	KF		SL	
		$3k - 1\sigma_{\text{add}} - 1D_R$	$3k - 1\sigma_{\text{add}} - 1D_R - 1m - 1\gamma$	$3k - 1\sigma_{\text{add}} - 1D_R$	$3k - 1\sigma_{\text{add}} - 1D_R - 1m - 1\gamma$
$k \times 10^{-4}$	0.20	0.39	2.0	0.33	2.4
	0.28	0.32	1.5	0.24	1.6
	0.36	0.30	1.3	0.20	1.4
σ_{add}^2		0.011	0.00046	0.015	0.00068
$D_R(s)$		0.052	0.112	0.046	0.131
m		–	0.16	–	0.10
γ		–	2.20	–	2.21
R^2		0.873	0.994	0.887	0.991
AIC_c		–246	–300	–250	–294
ΔAIC_c			54		44
Akaike's weight (%)			>99.999		>99.999
Evidence ratio			5.3×10^{11}		3.6×10^9

for 4 c/deg (Tables 2 and 3). The differences between the AIC_c values of the other perceptual template models and those of the $3k - 1\sigma_{\text{add}} - 1D_R - 1m - 1\gamma$ model were in the range of 6–39 (evidence ratio >20; Akaike's weight >95%). These findings suggest that the $3k - 1\sigma_{\text{add}} - 1D_R - 1m - 1\gamma$ model is more than twenty times more likely to be correct than the other perceptual template models.

Figs. 3 and 4 illustrate the predictions of the best-fitting LAM and PTM, respectively. For both observers, the AIC_c values of the $3k - 1\sigma_{\text{add}} - 1D_R - 1m - 1\gamma$ model were smaller than those of the $3k - 1\sigma_{\text{add}} - 1D_R$ model. The AIC_c differences were 19 for 0.4 c/deg and 44–54 for 4 c/deg (Akaike's weight >99.999% and evidence ratio >13360). Therefore, one may conclude that the PTM-based model provides better description of our data than the LAM-based model.

The sampling efficiency for both stimuli decreased as signal contrast increased. The sampling efficiencies for 0.4-c/deg Gabor patches were higher by a factor of 3.8 than those for 4-c/deg Gabor patches. The response duration averaged across observers and stimuli was 0.118 ± 0.016 s. The multiplicative component of the internal noise $[m(kE)^2]$ increased with signal contrast (Fig. 5, circles and squares). This induced internal noise was higher for 0.4-c/deg Gabor patches than for 4-c/deg Gabor patches. At low contrast levels, the multiplicative internal noise was comparable to the additive internal noise (horizontal lines) and exceeded the additive internal noise at higher contrast levels.

The best-fitting values of the additive internal noise calculated in the threshold experiment were higher than those obtained in above-threshold experiments by a factor of: 2.6 for 0.4 c/deg and 10.5 for 4 c/deg. The sampling efficiencies for both spatial frequencies measured

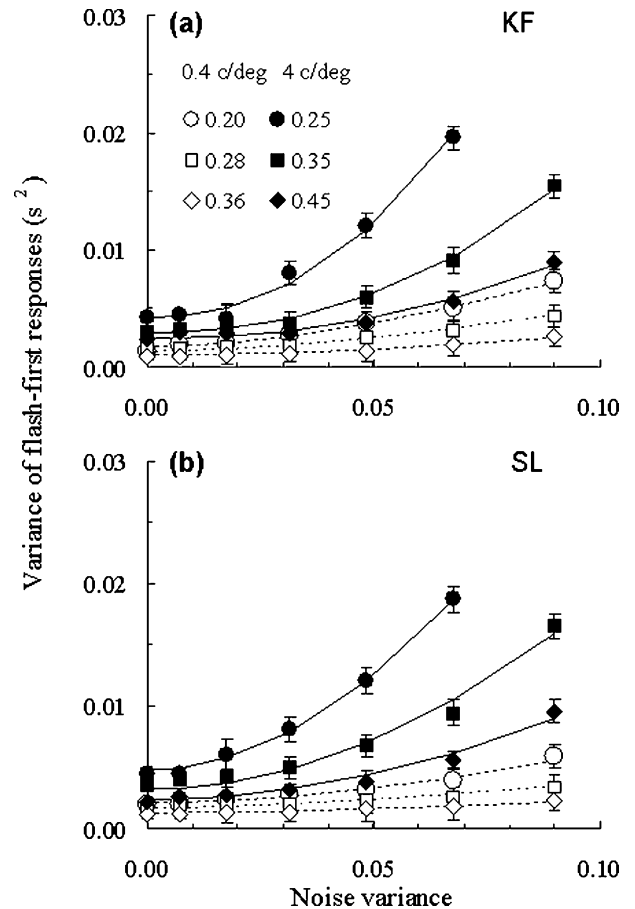


Fig. 4. Variance of flash-first responses as a function of external noise variance. Lines illustrate the predictions of the PTM-based $3k - 1\sigma_{\text{add}} - 1D_R - 1m - 1\gamma$ model whose parameters are shown in Tables 2 and 3. The other designations are as in Fig. 3.

in the threshold experiment were higher by a factor of 2.5 than those obtained by the flash-sound paradigm.

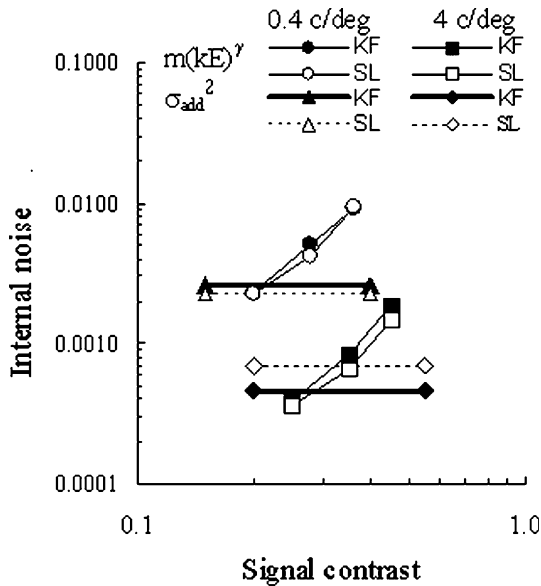


Fig. 5. Internal noise as a function of signal contrast. Circles and squares show the multiplicative signal-dependent internal noise for 0.4- and 4-c/deg Gabor patches, respectively. Horizontal lines illustrate the additive internal noise. Data for two observers.

5. Discussion

The present study has shown that variance of visual performance in a flash–sound temporal order task depends on variance of external visual noise. A perceptual template model for estimation of arrival time of a visual signal provided the best fit of the data. We found that above-threshold Gabor patches of low spatial frequency were detected with higher sampling efficiency than Gabor patches of higher spatial frequency. A similar result was obtained with threshold stimuli. Our stimulus of 0.4 c/deg had a width (cycles per 2 SD of the Gaussian envelope) of 1.2 cycles, while the width of the 4-c/deg Gabor patch was 12 cycles. Cortical neurones which may act as matched filters have localised receptive fields whose width is about 1–2 cycles (De Valois, Albrecht, & Thorell, 1982). Given the similarity between Gabor patches of lower spatial frequency and the receptive fields of cortical neurones sensitive to low spatial frequencies, one might suggest that observers would detect more efficiently lower spatial frequency than higher spatial frequencies in line with our findings.

The present results have shown that the multiplicative component of the internal noise due to the signal increases as signal contrast increases (Fig. 5). This is in line with electrophysiological data which have shown that the variance of V1 cells' responses increases with signal contrast (Tolhurst et al., 1981). We have found that the signal-dependent multiplicative internal noise at low contrast levels is much the same as the additive internal noise, but it exceeds the additive internal noise at higher contrast levels. This implies that the multipli-

cative internal noise which is induced by the signal is an important factor limiting the response variance. In addition, the internal noise has a component which is proportional to the external noise variance. This result is in accordance with the psychophysical findings showing that the total internal noise is higher at higher levels of external noise (Burgess & Colborne, 1988).

Our real-observer model for estimation of arrival time of a visual signal assumes that performance depends on the response duration (Eq. (6)). The best-fitting values of the response duration (D_R) averaged across observers were 0.115 s for 0.4-c/deg Gabor patches and 0.121 s for 4-c/deg Gabor patches (Tables 2 and 3). Previous study (Manahilov et al., 2003) showed that the half-amplitude width of the positive lobe of the impulse responses to luminance gratings of 0.5, 2 and 7 c/deg were in the range of 0.035–0.044 s. The total duration of the temporal impulse responses was in the range of 0.10–0.15 s. Similar results were reported by others (Gorea & Tyler, 1986; Georgeson, 1987). Therefore, the response durations obtained in the present study are in accordance with psychophysical estimations of impulse response duration.

In a previous study (Simpson et al., 2003) using reaction times to suprathreshold stimuli of 0.4 c/deg, we found that the sampling efficiency was much lower (3.3×10^{-6} – 8.2×10^{-6}) compared to that obtained in detection experiments. In these experiments, the sampling efficiency was estimated by a model, similar to the LAM-based model (Eq. (3)) in which the duration parameter was assumed to be equal to the duration of the signal (0.0083 s). In the present experiments, the estimated sampling efficiency was about 100 times higher than in the RT experiments. The fit of the present data with a perceptual template model whose duration parameter was set to 0.0083 s yielded sampling efficiency which was only 7 times higher than in the RT experiments. Therefore, the use of signal duration instead of the response duration seems to be the main reason for the estimated low levels of sampling efficiency in the RT experiments.

In the present study the observers estimated the arrival time of a visual signal using a flash–sound temporal order task. We assumed that since the auditory signal always was delivered at the same time, and it was not embedded in auditory noise, the noise associated with the representation of the auditory signal relative to that of the visual signal would be negligible. If this assumption were not true, then the additive internal noise would have an additional component due to auditory additive noise. Therefore, the additive internal noise estimated in the threshold experiments would be lower than that obtained by the flash–sound paradigm. The results, however, have shown the opposite effect: the best-fitting values of the additive internal noise calculated in the threshold experiment were higher than those

obtained in above-threshold experiments by a factor of 2.6 for 0.4 c/deg and by a factor of 10.5 for 4 c/deg.

In summary, applying a flash–sound paradigm and a perceptual template model for estimation of arrival time of a visual signal, we were able to estimate factors which limit performance for detecting suprathreshold stimuli: sampling efficiency, additive internal noise and multiplicative internal noises due to the visual signal and external noise. The sampling efficiencies for 0.4- and 4-c/deg Gabor patches measured in the threshold experiment were higher by a factor of 2.5 than those obtained by the flash–sound paradigm. The additive internal noise calculated in the threshold experiment was higher than those estimated in the suprathreshold experiments by a factor of 2.6 for 0.4 c/deg and by a factor of 10.5 for 4 c/deg. The signal-dependent multiplicative internal noise was similar to the additive internal noise at lower signal contrast levels and exceeded it at higher signal contrast levels.

The present study has shown that real observers' performance for detecting suprathreshold stimuli can be accounted for by a model which takes into account the non-linear visual–signal transduction and multiplicative components of the internal noise induced by the signal and external noise. In addition, this model assumes that performance depends on the response duration, rather than signal duration. The results suggest that the multiplicative internal noise due to high contrast visual signals determines performance for suprathreshold visual detection.

Acknowledgments

This research was supported by a grant from the Engineering and Physical Sciences Research Council. We thank the anonymous reviewers for their valuable comments and suggestions.

Appendix A. The goodness of fit of model predictions to the data was estimated by an R^2 statistic which is the proportion of the variance accounted for by the fit, adjusted by the number of free parameters (Judd & McClelland, 1989). The R^2 value was calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n \frac{(\alpha_i - \alpha_{i\text{est}})^2}{n-k}}{\sum_{i=1}^n \frac{(\alpha_i - \alpha_{\text{ave}})^2}{n-1}}, \quad (\text{A.1})$$

where α_i represent the observed data values, $\alpha_{i\text{est}}$ denotes the model calculations, k is the number of free parameters, n is the number of data points and α_{ave} is the mean value of the experimental data.

The quality of fit of several nested to the data obtained in the flash–sound simultaneity experiment was assessed by the corrected Akaike's Information Criterion (AIC_c) which was calculated using the following equation (Burnham & Anderson, 2002, pp. 60–85):

$$\text{AIC}_c = n \ln \left(\frac{\sum_{i=1}^n (\alpha_i - \alpha_{i\text{est}})^2}{n} \right) + 2K + \frac{2K(K+1)}{n-K-1}, \quad (\text{A.2})$$

where α_i are the data values, $\alpha_{i\text{est}}$ are the model calculations, n is the number of data points and K is the number of free parameters plus one.

The AIC_c approach is based on information theory, and does not use the traditional “hypothesis testing” statistical paradigm, rather it determines how well the data supports each model. The model with the smallest AIC_c value is most likely to be correct.

If A_a and A_b are the AIC_c values for models a and b , respectively, and $A_a < A_b$ ($\Delta = A_b - A_a > 0$), the Akaike's weight:

$$\text{Akaike's weight} = \frac{e^{-0.5\Delta}}{1 + e^{-0.5\Delta}} \quad (\text{A.3})$$

shows the probability that model a is correct.

The evidence ratio defined as follows:

$$\text{Evidence ratio} = \frac{1}{e^{-0.5\Delta}} \quad (\text{A.4})$$

shows how many times more likely model a is compared to model b .

References

- Barlow, H. B. (1978). The efficiency of detecting changes of density in random dot patterns. *Vision Research*, 18, 637–650.
- Burgess, A. E., & Colborne, B. (1988). Visual signal detection. IV. Observer inconsistency. *Journal of the Optical Society of America A*, 5, 617–627.
- Burgess, A. E., Wagner, R. F., Jennings, R. J., & Barlow, H. B. (1981). Efficiency of human visual signal discrimination. *Science*, 214, 93–94.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference—A practical information-theoretic approach*. Springer.
- De Valois, R. E., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22, 545–559.
- Foster, D. H., & Bischof, W. F. (1991). Thresholds from psychometric functions: Superiority of bootstrap to incremental and probit variance estimators. *Psychological Bulletin*, 109, 152–159.
- Georgeson, M. (1987). Temporal properties of spatial contrast vision. *Vision Research*, 27, 765–780.
- Gorea, A., & Tyler, C. W. (1986). New look at Bloch's law for contrast. *Journal of the Optical Society of America A*, 3, 52–61.
- Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model comparison approach*. New York: Harcourt Brace Jovanovich.

- Legge, G., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America A*, 4, 391–406.
- Lu, Z. L., & Doshier, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A*, 16, 764–778.
- Manahilov, V., Calvert, J., & Simpson, W. A. (2003). Temporal properties of the visual responses to luminance and contrast modulated noise. *Vision Research*, 43, 1855–1867.
- Nagaraja, N. S. (1964). Effect of luminance noise on contrast thresholds. *Journal of the Optical Society of America*, 54, 950–955.
- Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 3–24). New York: Cambridge University Press.
- Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, 31, 1337–1350.
- Rashbass, C. (1970). The visibility of transient changes of luminance. *Journal of Physiology*, 210, 165–186.
- Simpson, W. A., Findlay, K., & Manahilov, V. (2003). Efficiency and internal noise for detection of suprathreshold patterns measured using simple reaction time. *Vision Research*, 43, 1103–1109.
- Tolhurst, D. J., Movshon, J. A., & Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortex neurones on stimulus contrast. *Experimental Brain Research*, 41, 414–419.
- Watson, A. B. (1986). Temporal sensitivity. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance I: Sensory processes and perception* (pp. 6-1-6.41). New York: John Wiley.
- Woodward, P. M. (1980). *Probability and information theory, with applications to radar*. New York: Pergamon.