Available online at www.sciencedirect.com

ScienceDirect

Vision Research 47 (2007) 280–288

Vision Research

www.elsevier.com/locate/visres

Limited flexibility in the filter underlying saccadic targeting

Casimir J.H. Ludwig [a,*], Miguel P. Eckstein [b], Brent R. Beutter [c]

[a] Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK
[b] Department of Psychology, University of California, Santa Barbara, CA 93106-9660, USA
[c] Human Performance Division, NASA Ames Research Center, Moffett Field, CA 94035-1000, USA

Received 22 June 2006; received in revised form 19 September 2006

Abstract

The choice of where to look in a visual scene depends on visual processing of information from potential target locations. We examined to what extent the sampling window, or filter, underlying saccadic eye movements is under flexible control and adjusted to the behavioural task demands. Observers performed a contrast discrimination task with systematic variations in the spatial scale and location of the visual signals: small ($\sigma = 0.175°$) or large ($\sigma = 0.8°$) Gaussian signals were presented 4.5°, 6°, or 9° away from central fixation. In experiment 1, we measured the accuracy of the first saccade as a function of target contrast. The efficiency of saccadic targeting decreased with increases in both scale and eccentricity. In experiment 2, the filter underlying saccadic targeting was estimated with the classification image method. We found that the filter (1) had a center-surround organisation, even though the signal was Gaussian; (2) was much too small for the large scale items; (3) remained constant up to the largest measured eccentricity of 9°. The filter underlying the decision of where to look is not fixed, and can be adjusted to the task demands. However, there are clear limits to this flexibility. These limits reflect the coding of visual information by early mechanisms, and the extent to which the neural circuitry involved in programming saccadic eye movements is able to appropriately weigh and combine the outputs from these mechanisms.

Keywords: Saccadic eye movements; Spatial vision; Decision-making; Template-matching; Contrast

1. Introduction

The highest resolution vision is only possible in the central one or two degrees of the visual field. Therefore, in order to explore or interact with the visual environment humans shift their gaze at regular intervals, typically 3 or 4 times every second (Findlay & Gilchrist, 2003). Saccadic eye movements are a critical element of almost any behavioural activity that involves the use of visual information. When an observer decides to look at some part of the visual scene, it is likely that this decision was, at least partly, driven by the visual signals sampled from that region. An important question is "how much" visual information is taken into account. In other words, how does the observer weigh visual signals over space in order to decide where to look?

Such a weighting function is often referred to as a template or filter. Simple perceptual decisions, like detecting whether a signal is present in noise, can be modelled as a process of template-matching (Burgess, Wagner, Jennings, & Barlow, 1981; Lu & Dosher, 1999; Pelli, 1985). A useful tool to estimate the filter used by human observers to perform such a visual task is the so-called 'classification image'. This image is essentially a description of what parts of a visual stimulus the human observer takes into account to make a perceptual decision (Abbey & Eckstein, 2002; Ahumada, 2002; Murray, Bennett, & Sekuler, 2002). The general approach to calculate this description is to present signals in visual noise, and then on each trial relate the observer's decision to the properties of the noise. Work of this type has shown that human observers have

* Corresponding author.
E-mail addresses: c.ludwig@bristol.ac.uk (C.J.H. Ludwig), eckstein@psych.ucsb.edu (M.P. Eckstein), bbeuter@mail.arc.nasa.gov (B.R. Beutter).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.09.009

considerable flexibility in their ability to adapt the filter to the demands of the visual task, ranging from simple contrast detection (Abbey & Eckstein, 2002) to completion of illusory contours (Gold, Murray, Bennett, & Sekuler, 2000). For successful interaction with the visual environment, such flexibility clearly is desirable.

Saccadic eye movements can be regarded as a unique class of perceptual decisions. Saccades are, by definition, directed to visual targets outside the fovea, where sensitivity to fine detail is drastically reduced (Anstis, 1974; Pointer & Hess, 1989; Robson & Graham, 1981). In addition, these movements are generally made in quick succession, and appear to be based on a rather brief temporal integration period (Caspi, Beutter, & Eckstein, 2004; Ludwig, Gilchrist, McSorley, & Baddeley, 2005). These constraints may impose limits on the interaction between the oculomotor and visual systems, which in turn may limit the kinds of weighting that can be achieved. As an extreme example, it may be that the region from which visual signals are sampled is fixed, or perhaps directly dependent on eccentricity in the visual field (Garcia-Perez & SierraVazquez, 1996; Virsu & Rovamo, 1979). In this study we assess (1) the shape of the filter underlying saccadic eye movement decisions in a contrast discrimination task; (2) whether the filter is adjusted according to the task demands (integrate either over small or large regions); and (3) whether the filter depends on eccentricity in the visual field.

2. Methods

2.1. Study outline

Observers were presented with four Gaussian signals embedded in spatially uncorrelated, Gaussian white noise. The contrast of one of the four patterns (the target) was slightly higher than that of the other three (distractors), and observers signalled the location of the target with a manual response. We varied the size of the display items (spatial standard deviation of the pattern, σ, was either 0.175° or 0.8°), and the eccentricity at which they appeared (4°, 6.5°, and 9° of visual angle) in separate blocks. Observers were free to move their eyes and we recorded the landing position of their first saccade after display onset. The first saccade can considered the observers' best guess (decision), at that point in time, of the target location. Because the task is to find the highest contrast pattern, both the stimulus-driven contrast response and the task instructions induce participants to directly aim their first saccade to the target (Ludwig & Gilchrist, 2006). We analysed and modelled the accuracy of these first eye movement decisions.

A first step was to provide a detailed characterisation of saccadic targeting with variations in size and eccentricity of the display items. Here it is important to not just examine accuracy in terms of proportion correctly directed saccades, but to relate this proportion to the amount of information that is available for the task. Thus, in experiment 1 we quantified accuracy of the first saccade relative to an ideal observer in the form of an efficiency measure (Burgess et al., 1981; Eckstein, Beutter, & Stone, 2001). In experiment 2, we assessed whether variations in efficiency could be attributed to changes in the filter. The filter was estimated using reverse correlation. This technique involves extracting the external noise from the location of the first saccade landing position for trials on which the first saccade was directed to one of the distractor items. The extracted noise images are then averaged to compute a classification image.

2.2. Observers

Three observers with normal or corrected-to-normal vision were tested in the two experiments. Author CL is the third observer. The data were collected in multiple sessions over a period of 6 weeks. Each observer completed ~23,000 trials (across the two experiments).

2.3. Stimuli

Stimuli were viewed binocularly on a linearised M17LMAX monchrome monitor (Image Systems, Minnetonka MN) The mean luminance of the display was 31 cd/m^2. Display items were 2-D Gaussians with a standard deviation of either 0.175° or 0.8° (blocked). Dark grey outline circles (radius of 3σ of the large patch) marked the locations where the 4 patterns were presented. They were arranged along the circumference of a circle (i.e. equidistant from the central fixation point). The 4 patterns appeared either in a square configuration or in a diamond configuration. The configuration varied randomly from trial to trial to prevent observers from anticipating the exact locations where items would appear.

The Gaussian patterns were embedded in 0-mean, spatially uncorrelated Gaussian noise with an RMS contrast of 25%. The peak contrast of the three distractors was 12% (pedestal contrast). In experiment 1, the target contrast was varied at 5 levels, resulting in signal-to-noise ratios (SNRs) that ranged from 0 to 5.6 for the small scale patterns, and from 0 to 25.77 for the large patterns. In experiment 2, the SNRs of the small and large items were kept fixed.

2.4. Trial sequence

At the start of a trial, a fixation display appeared. Observers fixated a small black cross in the centre of the screen, and initiated a trial by pressing the space bar. After a variable delay (667–1333 ms), the stimulus appeared, consisting of the circle outline markers, Gaussian patterns, fixation point, and external noise sample. It was presented for 800 ms, followed by a response display containing only the fixation point and the 4 marker circles. Observers moved the mouse cursor to the location where they thought the brightest pattern had appeared and responded by pressing the left button. The background of the fixation and response displays were uniform grey, with the same mean luminance as the noise background in the stimulus.

2.5. Eye movement recording and saccade classification

The position of the left eye was sampled at 250 Hz using an infrared video-based eye tracker (SMI EyeLink). A nine-point grid calibration was performed at the start of each block of 120 (experiment 1) or 96 (experiment 2) trials. Head movements were restricted through a chin rest. Observers viewed the display from a distance of 56 cm.

Eye movement data were analysed off-line. Saccades were detected using velocity and acceleration criteria of 35°/s and 9500°/s^2, respectively. Trials were excluded if the first saccade (i) starting position deviated from the display centre by more than 1°; (ii) latency was less than 80 ms; (iii) amplitude was less than half the distance of the display items. For observers 1–3, respectively, these criteria resulted in the rejection of 6%, 7%, and 5% of the trials. The landing position of the saccade was assigned to a display item if its angle fell within 90° of the direction of that item (i.e. within the correct quadrant). Landing positions were generally tightly clustered around display item locations: 95% of saccades had directional deviations smaller than 14°, 17°, and 26° for observers 1–3, respectively. Assigning the saccade landing position to the nearest item in the display is standard practice in this type of study (Beutter, Eckstein, & Stone, 2003; Ludwig & Gilchrist, 2006).

2.6. Analyses

2.6.1. Psychometric functions

First, the proportion of first saccades directed to the target was mapped onto the signal detection measure of d' according to (Green & Swets, 1966):

$$p(d',N) = \int_{-\infty}^{\infty} g(x-d')\mathrm{d}x\left(\int_{-\infty}^{x} g(y)\mathrm{d}y\right)^{N-1}, \quad (1)$$

where $g(y)$ is the Gaussian probability distribution function, and N is the number of alternatives ($N = 4$). The psychometric function relates d' to SNR. The template-matching model with uncertainty (Beutter et al., 2003; Eckstein, Ahumada, & Watson, 1997; Pelli, 1985) was used to fit these functions. Details of the model can be found in (Beutter et al., 2003); it is only briefly described here.

For a linear observer monitoring the output of N filters centred on the possible target locations, d' is a linear function of SNR: $d' = \alpha \times \mathrm{SNR}$, where α is the slope defined by:

$$\alpha = \frac{r}{\sqrt{1+\frac{\sigma_{\mathrm{int}}^2}{\sigma_{\mathrm{ext}}^2}}}. \quad (2)$$

Here, r is the correlation of the human filter with ideal filter (or signal), and σ_{int}^2 and σ_{ext}^2 are the internal and external noise variances, respectively. The cross-correlation between filter and (noisy) stimulus is performed at each possible target location, the resulting internal responses are corrupted by internal noise, and the observer makes a saccade to whichever patch triggered the largest response (maximum response decision rule).

A nonlinear version of the model incorporates intrinsic uncertainty by having the observer monitor additional, irrelevant filters at each possible target location. Thus, at each location only one filter responds to the target or distractor, and U mechanisms respond to noise. A target-directed saccade is generated if either the correct relevant filter, or an irrelevant filter at the target location triggers a larger response than the $N(U+1)-1$ other filters. Note that for the ideal observer $r = 1$, $\sigma_{\mathrm{int}} = 0$, and $U = 0$. As a result, $\alpha = 1$ and $d' = \mathrm{SNR}$. In other words, the ideal observer's performance is solely limited by the external noise in the stimulus, and is unaffected by our experimental variables of scale and eccentricity.

The model was fit using maximum likelihood (MLE) methods. Estimates of the standard error of the (slope) parameters were generated using bootstrap methods (Efron & Tibshirani, 1993). For each combination of observer, spatial scale and eccentricity, replicate psychometric functions were generated through resampling the binomially distributed proportion correct at each level of contrast. These replicate functions were then fit using the same MLE algorithm that was used to fit the observed psychometric functions. Standard errors were based on bootstrapped distributions of the parameter of interest (1000 replications).

2.6.2. Filter estimates

Classification images were based only on saccades that were erroneously directed to a distractor. For each observer and experimental condition, the classification image was computed by averaging the 87×87 pixel noise patches centred on the location of the fixated distractors. On the basis of the stimuli used (Gaussian patterns) and previous work (Eckstein, Pham, & Shimozaki, 2004) the classification images were parameterised with 2-D Difference-of-Gaussian (DoG) functions. Only the height and width parameters of the DoGs were free to vary; the DC was fixed to 0, and the centres of both Gaussian components were constrained to lie at the centre of the image. Fits were computed using weighted least squares errors (each pixel in the image was weighted by the reciprocal of its squared standard error).

The original noise images that made up an individual classification image, were resampled with replacement and then averaged in order to create a set of 1000 bootstrap classification images for each of the 18 combinations of observer (3), scale (2), and eccentricity (3) (Efron & Tibshirani, 1993). Each of these bootstrapped images was fit with a DoG. These fits allowed us to construct 18 sample distributions of the following parameters of interest: the height of the filter, the magnitude of the inhibitory lobes, the width of the filter, and the overall match between filter and signal. Error bars and null-hypothesis testing were based on these sample distributions. Note that if variability in the set of original noise images is large, this results in broad sample distributions, large error bars, and smaller likelihoods of rejecting the various null-hypotheses.

3. Results

3.1. Experiment 1

Fig. 1 illustrates a set of psychometric functions for one observer, at all combinations of eccentricity and spatial scale. The smooth curves are the template-matching model fits. As described above, the linear model predicts straight psychometric functions. However, at low SNRs psychometric functions frequently display an accelerating nonlinearity (see Fig. 1, top panel). This effect can be modelled as a nonlinearity in the transducer (Eckstein et al., 1997; Lu & Dosher, 1999), or intrinsic uncertainty in the monitoring of filter outputs (Pelli, 1985). We fit a nonlinear model with uncertainty to the data. This model encompasses the linear model when uncertainty, $U = 0$. For all three observers and eccentricities, $U = 1$ in the small scale conditions, and $U = 0$ in the large scale conditions. Bootstrap simulations indicated that the nonlinearity was small, but consistent in the small scale conditions. Most importantly however, the nonlinearity did not vary across observers and eccentricities.

As can be seen in Fig. 1, there were clear effects of eccentricity: stronger signals were required to reach a given performance level as eccentricity increased. These variations are effectively quantified by the parameter α of the model fits, which can be regarded as an approximation of the absolute efficiency underlying saccadic targeting (Burgess et al., 1981): for a linear observer the absolute efficiency is given by α^2. Fig. 2 illustrates α as a function of eccentricity, separately for the two spatial scales (columns) and the three observers. There are three important aspects to these data. First, the values were always less than 1, indicating that human performance was suboptimal. Second, the overall efficiency was much lower when the display items are large compared to the small scale condition. Thus, even

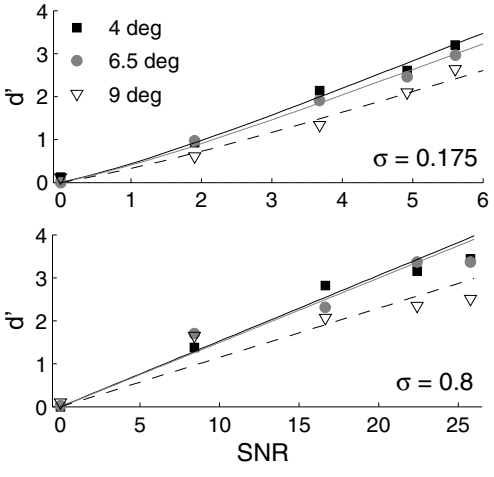

Fig. 1. Psychometric functions relating the accuracy of the first saccade (d') to the signal-to-noise ratio of the target. Curves are fits of a template-matching model (solid black, 4°; solid grey, 6.5°; dashed black, 9°). Data from observer 1.

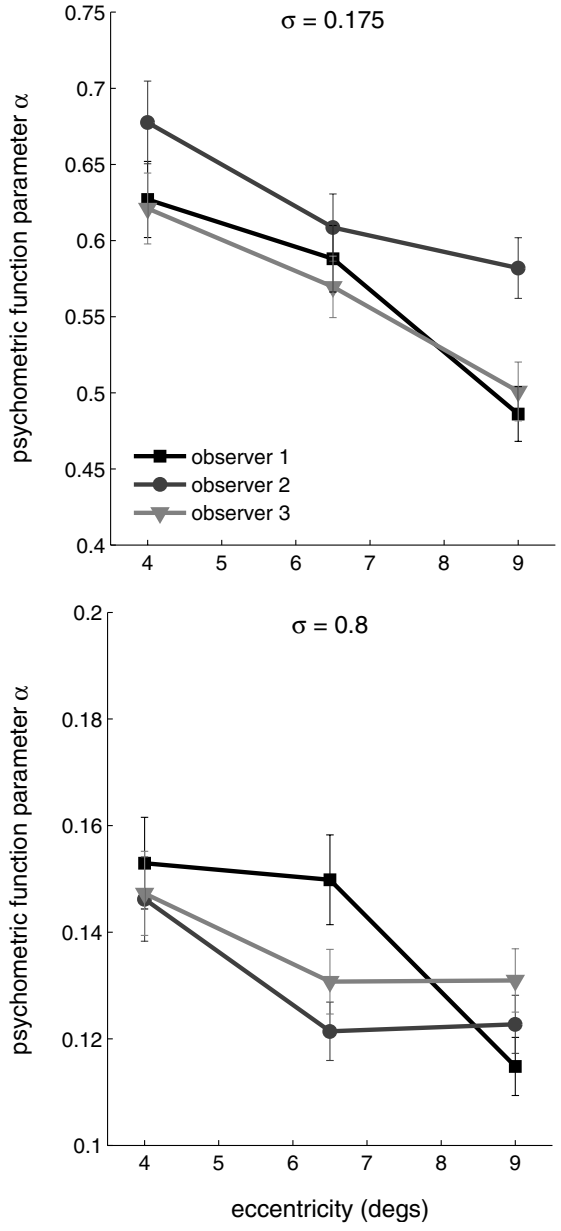

Fig. 2. Parameter α as a function of eccentricity. This measure is derived from the psychometric functions. Error bars indicate bootstrap standard errors.

though performance levels in terms of proportion correct or d' were similar in the two spatial scale conditions (see Fig. 1), humans were much less efficient relative to the ideal observer in the large scale conditions. Third, the effect of eccentricity was very clear and approximately linear for the small scale conditions, but much less pronounced when the display items were large.

We used α as a model-based approximation of the absolute efficiency. A model-free estimate of this quantity can be defined as $d'^2_{\text{human}}/d'^2_{\text{ideal}}$, where $d'_{\text{ideal}} = \text{SNR}$. In order to verify the validity of the approximation, we computed these model-free estimates of efficiency, averaged over SNR. In the small scale conditions, the mean efficiencies across observers were 0.319, 0.263, and 0.196 at eccentricities of 4°, 6.5°, and 9°, respectively. In the large scale conditions, the mean efficiencies were 0.024, 0.020, and 0.019. For the small scale conditions these estimates are consistent with previous published efficiencies for first saccades (Eckstein et al., 2001). The estimates from the large scale conditions are much smaller. Most importantly, the model-free efficiency estimates followed the same pattern as the psychometric function parameters illustrated in Fig. 2.

The main message from experiment 1 then is that human, unlike ideal, performance depended strongly on the spatial scale and eccentricity of the display items. One possible source of these performance variations is the match between filter and signal. In experiment 2, we estimated the filter underlying saccadic targeting in this visual task.

3.2. Experiment 2

In this experiment the SNRs were fixed at 1.02 and 2.34 for the small and large items, respectively. These SNRs were chosen to target values of d' that were just above chance. At these performance levels, one would expect observers' decisions to be at least partly governed by the external noise in the stimulus rather than purely by the signal (Murray et al., 2002). This is important for the classification image method used to derive the filter estimates. As far as the observers were concerned, their task was identical to that of experiment 1. Again, each combination of size and eccentricity was tested separately in different blocks.

Overall accuracy was low, but consistently above chance. Mean d's in the small scale conditions were 0.33, 0.35, and 0.16 at eccentricities of 4°, 6.5°, and 9°, respectively. In the large scale conditions the mean accuracies were 0.27, 0.26, and 0.27 at the three eccentricities. Based on the psychometric functions (Fig. 1) one would not expect large modulations of performance by eccentricity at these low SNRs. Nevertheless, in accordance with the results from experiment 1 there were consistent effects of eccentricity for the small items: For each observer performance was worst at the furthest eccentricity, compared to both the nearest and the middle eccentricity (Mann–Whitney tests with $p < .01$). Any comparison between the two nearest eccentricities was not significant. Accuracy for the large display items was independent of eccentricity. With these levels of accuracy the number of incorrect saccadic decisions ranged from 1773 to 2160 across observers and experimental conditions.

Fig. 3A shows an example of a raw classification image (observer 1, $\sigma = 0.175°$, eccentricity = 4°), median filtered for display purposes to suppress the inevitably noisy appearance of the image. Bright regions in the classification image indicate the pixels that were weighted positively by the observer in making saccadic decisions. Darker regions were weighted negatively. As expected, the bright region in the centre of the classification image in Fig. 3A looks approximately Gaussian.

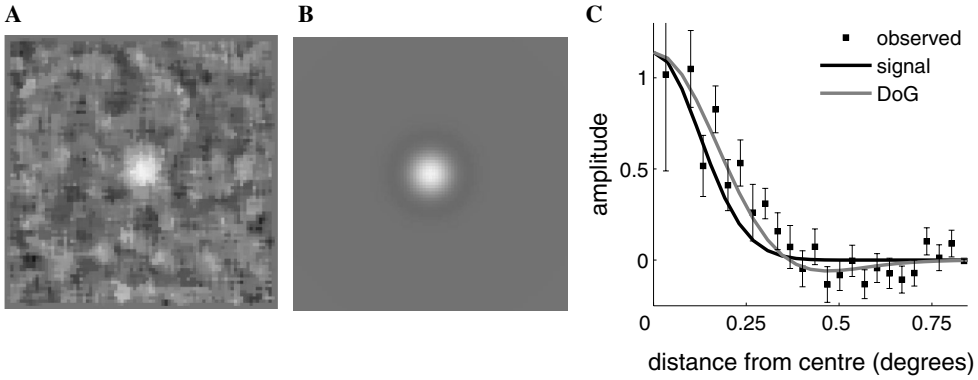

Fig. 3. Example classification image. (A) Raw classification image from observer 1 ($\sigma = 0.175°$, eccentricity = 4°), median filtered. (B) 2D DoG fit to the raw classification image (unfiltered). (C) 1D representation of the DoG fit. Data points are radial averages obtained from the raw classification image (unfiltered). Error bars are ±1 SEM.

Fig. 3B shows the DoG fit to the raw classification image of panel A. Because this function is radially symmetric, a one-dimensional slice is shown in panel C (grey line) along with radial averages derived from the raw classification image (Abbey & Eckstein, 2002). The profile of the signal is shown by the black line. In terms of goodness-of-fit (RMSE = 2.08) the illustrated fit ranked 9th out of a total of 18 classification images (3 observers × 3 eccentricities × 2 spatial scales; RMSE range was 1.98–2.34). The DoG functions provided an effective description of the observed classification images, with only 4 parameters (as opposed to 87^2).

The first parameter of interest was the amplitude of the classification image. Note that if the image had no discernable features (e.g. simply random noise), then on average the amplitude of the DoG should be 0. The bootstrapped sample distributions were used to compute a Z-score associated with this null-hypothesis. The null-hypothesis was easily rejected in each of 18 cases ($p < .001$; Table 1). In other words, all 18 classification images had a clear excitatory centre. A repeated measures ANOVA with scale and eccentricity as factors showed a marginal effect of scale ($F(1,\ 2) = 17.56,\ p = .052$): the small scale classification images tended to have slightly larger amplitudes. There was no effect of eccentricity, nor an interaction between the two factors.

Second, if the filter shape was in actual fact Gaussian, there would be no inhibitory lobes. We computed the maximum magnitude of the inhibitory lobes for each bootstrapped classification image. This allowed us to test the hypothesis that these magnitudes were in fact 0. This analysis suggests that inhibitory lobes were a reliable feature of 17/18 classification images ($p < .05$; Table 2). A repeated measures ANOVA showed no effects of our experimental manipulations on the raw magnitude of the inhibitory lobes.

The third measure of interest was the width of the filter, as an index of the size of the spatial integration window. For each DoG fit, we computed the full-width at half the maximum amplitude (FWHM). These estimates are illustrated in Fig. 4 (top row). The dashed lines indicate the FWHM of the ideal filter (black: $\sigma = 0.175°$; grey: $\sigma = 0.8°$). The solid functions show the width of the DoG filters as a function of eccentricity. Several aspects are noteworthy. The width of the filter estimates in the small scale conditions was consistently smaller than that in the large scale conditions. Thus, observers did adapt their integration window according to the task demands. In addition, the width of the filters, in general, did not match that of the ideal. The pattern of using too large a filter for the small items (as illustrated in Fig. 3C) was highly consistent across observers and eccentricities, albeit small. More dramatic was the mismatch for the large scale items: The human filter was radically smaller than the stimulus. The mismatch in these conditions was highly significant ($p < .001$; Table 3) with only one exception (observer 2, $\sigma = 0.8°$, eccentricity = 4°). For the small spatial scale, the difference between ideal and human filter width was less

Table 1
Z-scores associated with 'amplitude = 0'

Eccentricity	Observer 1		Observer 2		Observer 3	
	$\sigma = 0.175$	$\sigma = 0.8$	$\sigma = 0.175$	$\sigma = 0.8$	$\sigma = 0.175$	$\sigma = 0.8$
4	-9.72^{***}	-5.87^{***}	-8.72^{***}	-3.98^{***}	-6.53^{***}	-8.94^{***}
6.5	-5.83^{***}	-5.56^{***}	-11.08^{***}	-4.90^{***}	-3.88^{***}	-6.99^{***}
9	-6.29^{***}	-4.74^{***}	-9.87^{***}	-7.34^{***}	-5.79^{***}	-7.32^{***}

$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$ (two-tailed).

Table 2
Z-scores associated with 'inhibitory lobe = 0'

Eccentricity	Observer 1		Observer 2		Observer 3	
	$\sigma = 0.175$	$\sigma = 0.8$	$\sigma = 0.175$	$\sigma = 0.8$	$\sigma = 0.175$	$\sigma = 0.8$
4	2.20^{*}	4.80^{***}	2.37^{*}	0.19	2.16^{*}	4.63^{***}
6.5	3.12^{**}	2.51^{*}	2.22^{*}	2.07^{*}	2.36^{*}	3.13^{**}
9	3.68^{***}	3.40^{***}	2.83^{**}	5.80^{***}	4.77^{***}	4.18^{***}

* $p < .05$.
** $p < .01$.
*** $p < .001$ (two-tailed).

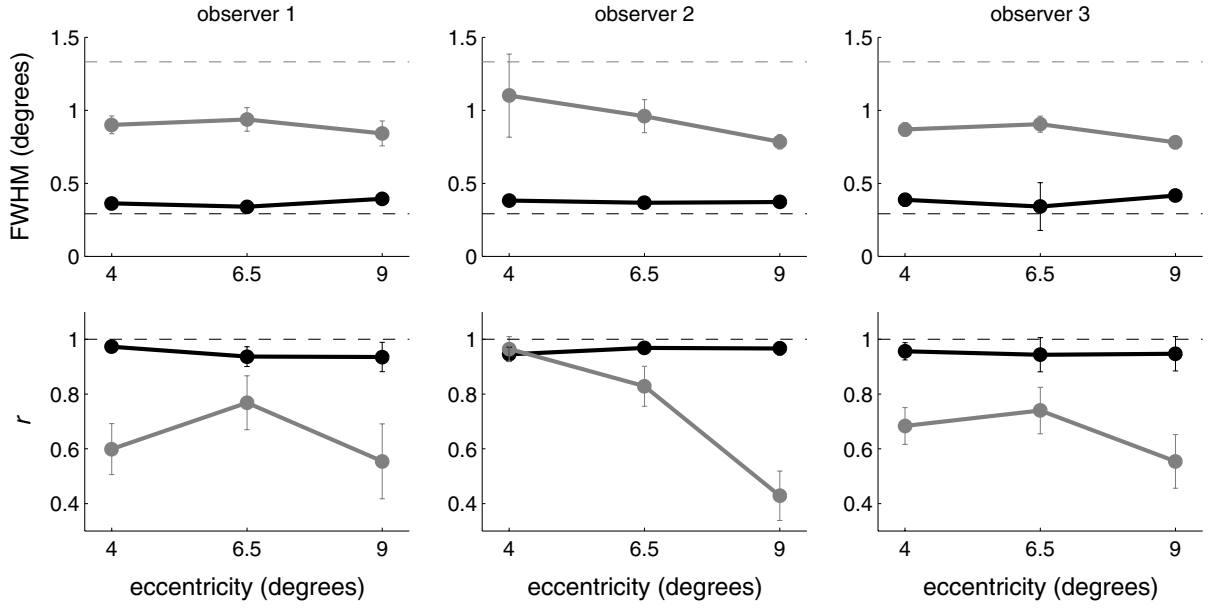

Fig. 4. Filter characteristics as a function of eccentricity and spatial scale. Solid black lines indicate values for $\sigma = 0.175°$; solid grey lines indicate values for $\sigma = 0.8°$. Dashed lines are the relevant parameter values for the small and large signals. Top row, FWHM of the estimated DoG filters. Bottom row, correlation between estimated DoG filters and Gaussian signal. Error bars indicate bootstrap standard errors.

Table 3
Z-scores associated with 'FWHM of the estimated DoG filter = FWHM of the signal'

Eccentricity	Observer 1		Observer 2		Observer 3	
	$\sigma = 0.175$	$\sigma = 0.8$	$\sigma = 0.175$	$\sigma = 0.8$	$\sigma = 0.175$	$\sigma = 0.8$
4	-2.76^{**}	7.12^{***}	-2.62^{**}	0.81	-2.19^{*}	9.91^{***}
6.5	-1.18	4.89^{***}	-3.18^{**}	3.30^{***}	-0.30	7.76^{***}
9	-3.07^{**}	5.75^{***}	-2.87^{**}	11.35^{***}	-3.35^{***}	12.55^{***}

* $p < .05$.
** $p < .01$.
*** $p < .001$ (two-tailed).

striking, but significant in most cases nevertheless. Finally, there were no consistent effects of eccentricity. For observer 2 there appeared to be a slight *decrease* in filter width as eccentricity increases, but the error bars indicate the considerable uncertainty in the width estimate at the nearest eccentricity. Consistent with these observations, only the effect of scale was significant ($F(1, 2) = 304.01$, $p < .05$) in an ANOVA.

Fourth, we estimated the *overall* match between the filter and signal. The analyses reported above already suggest that the human filter is suboptimal due to the presence of inhibition and the mismatch in the width between signal and filter. The overall match is formalised as the correlation between the signal and the estimated filter, the term r in Eq. (2). The bottom row of Fig. 4 shows r as a function of eccentricity and spatial scale. Unsurprisingly, these results are largely consistent with the FWHM analysis. The correlation was close to 1 in the small scale conditions, but there was a substantial mismatch in the large scale conditions. For observer 2 (large spatial scale conditions) there

appeared to be a reliable decrease in the correlation. The analysis of the magnitude of inhibitory lobes showed that this finding was mainly due to stronger inhibition with increasing eccentricity. For the remaining combinations of observers and spatial scales, the correlations were independent of eccentricity. Again, only the effect of scale was significant ($F(1, 2) = 101.85$, $p < .05$).

This drastic mismatch between filter and signal in the large scale conditions is surprising. One may wonder whether a closer match would be found if we analysed the correct, target-directed saccades instead of the erroneous eye movements. Therefore, we computed classification images using the noise patches from the target location on trials in which the first saccade was directed to the target. Because error rates were high, the number of trials available for this analysis was approximately half that for the analysis based on distractor-directed movements. Clearly then, these filter estimates were less stable than the ones derived from the large set of error trials.

We fit these "correct" classification images with a DoG, and computed the correlations with the signal as before. The variability around these estimates was determined by resampling the sets of noise images that made up the correct classification images, and repeating the fitting procedure with these bootstrap classification images (as before with the error trials; see Section 2). We then determined whether the original correlation estimates fell within the 95% confidence interval of this bootstrap distribution. These bootstrap distributions tended to be negatively skewed when the correlation was high (predominantly in the small scale conditions). Therefore, we did not use the kind of Z-approximation used in the statistical tests reported above (Tables 1–3), but instead opted for the nonparametric percentile confidence limits.

As such, 15/18 original correlation estimates were consistent with the estimates based on the correct classification images. Of the remaining three original estimates, two correlations were lower and one was higher than those based on the correct trials. Most importantly, the pattern of correlations was highly similar to that reported above for the error trial analyses: they were high in the small scale conditions (range across observers and eccentricities: .84–1) and much lower in the large scale conditions (.33–.88), without any clear effects of eccentricity. On the whole then, there is no evidence to suggest that the match between the observers' filters and signals was tighter when the first saccade was correctly directed to the target. These analyses support the validity of our original measurements and the conclusions derived from them.

4. Discussion

Deciding to generate a saccadic eye movement to some location in the visual field depends on sampling the visual signals from that location. In the current visual task saccadic eye movement decisions were effectively modelled by assuming the observer cross-correlates the visual input at each item location with a filter, and chooses to look at the location that elicited the strongest internal response. We used the classification image method to estimate the underlying filter, and showed that: (1) the shape of the filter did not match the shape of the expected signal: it was best described by a DoG function, even though the signal was Gaussian; (2) the size of the filter was flexibly adjusted according to the task demands, although the adjustment was far from perfect, with the integration window being much too small in the large scale conditions; (3) despite substantial performance variations with eccentricity we found no evidence for reliable changes in the filter with eccentricity. We will deal with each of these three findings in turn.

The early visual system codes the spatial distribution of light through a bank of relatively narrowband, center-surround mechanisms that are tuned to different sizes (DeValois & DeValois, 1988; Graham, 1989; Wilson & Bergen, 1979). In theory, any arbitrary filter can be constructed by appropriately weighing and combining the outputs of a variety of these mechanisms. Observers were unable to form a precisely matched template to a Gaussian signal: the filters had inhibitory side-bands. When expressed relative to the peak amplitude of the filter, the magnitude of the inhibitory lobes ranged between 1% and 23% (mean across observers and conditions = 10%; the inhibitory lobes illustrated in Fig. 3 had a relative magnitude of 5%). There was a strong trend for the large scale filters to have stronger (relative) inhibition, which clearly contributed to the poor match between filter and signal in these conditions. Thus, although the presence of inhibitory lobes was fairly subtle in some conditions, it was certainly very consistent across observers and conditions.

The center-surround organisation of the filters reported in this study for saccades and previously for perceptual decisions (Abbey & Eckstein, 2002; Eckstein et al., 2004) suggests that under these conditions the saccadic system is driven by signals computed relatively early in the visual processing hierarchy. To some extent this finding may depend on the task given to observers. In contrast discrimination it is conceivable that the eye movement decisions are based on the outputs of early filters. Whether observers are able to adopt a more sophisticated template under different task demands remains to be seen. Rajashekar, Bovik, and Cormack (2006) had observers search for particular geometrical shapes in 1/f noise, and their estimated templates also suggest some flexible adjustment of the filters depending on what the observer is looking for. However, they did not attempt to parameterise their filters to illustrate a link between observers' templates and early visual mechanisms.

The mismatch between the filter and the Gaussian stimulus points to limits in the pooling of outputs from these early visual mechanisms. These limits may arise because the saccadic system only has access to a restricted range of mechanisms, limited temporal integration, or a combination of both. Recent studies using a temporal noise

variant of the contrast discrimination task have demonstrated that only a small portion of the saccade latency period is used for visual integration (Caspi et al., 2004; Ludwig et al., 2005). In particular, it appears that the temporal integration window is restricted to the first ~100 ms. after display onset (Ludwig et al., 2005). It is possible that not all relevant early outputs have become available within this restricted time period. As such, downstream oculomotor structures would miss out on these outputs, preventing a precise match between the functional filter and signal.

Despite the mismatch in shape, observers did scale the size of the integration window with that of the display items. As such, we can reject the hypothesis that the saccadic eye movement system only has access to one class of early visual mechanisms. These findings support the notion that the integration window is under some control, and can be flexibly adjusted according to the task demands (Levi, Klein, & Chen, 2005). Nevertheless, there appear to be clear limits to this flexibility: human observers failed to take optimal advantage of the increase in signal area and their filter was substantially too small for the large scale items. Thus, the reduced efficiency in these conditions is, at least partly, explained by an increased mismatch between the filter and the signal.[1] This result parallels the failure to optimally integrate spatially extended signals in perceptual decisions (Dakin & Bex, 2003).

One possible explanation for this finding is that observers may simply have underestimated the size of the large patterns. With soft-edged patterns the perceived size strongly depends on the peak contrast of the pattern, among other things (Fredericksen, Bex, & Verstraten, 1997). Thus observers may have perceived the low contrast, noise masked patterns of experiment 2 to be smaller (although this would have to hold only for the large scale stimulus), inducing them to integrate over a smaller region. In the absence of perceived size measurements for these stimuli, it is difficult to completely discount this explanation. However, experiment 1 was run partly in order to familiarise the observers with the patterns they had to discriminate. In the context of that experiments observers regularly saw high contrast versions of the Gaussian signal, allowing them to build an accurate representation of the pattern size and shape. Regardless of the role of perceived size, the mismatch in the large scale conditions suggests that the eye movement system did not or could not sample from earlier mechanisms with large enough spatial integration areas (i.e. receptive fields) or from populations of mechanisms with partly overlapping receptive fields that, together, cover the signal area.

Remarkably, the size of the integration window remained relatively constant with eccentricity. It is well-established that the size and spacing of retinal and visual cortical receptive fields increase with eccentricity (DeValois & DeValois, 1988). A functional consequence is that the ability to integrate over small regions of space is reduced outside the fovea (Garcia-Perez & SierraVazquez, 1996; Pointer & Hess, 1989; Robson & Graham, 1981; Virsu & Rovamo, 1979; Wilson & Bergen, 1979). As such, we expected that the size of the integration window would increase with eccentricity, and that this would be a major contributor to the lower efficiency observed further out in the visual field. Although changes in basic visual mechanisms with eccentricity undoubtedly occur, it appears that these changes were not responsible for the performance variations observed in our study.

To conclude, for successful interaction with the visual environment it would be desirable if the saccadic eye movement system could flexibly select and combine outputs from visual mechanisms upstream (e.g. at the level of striate cortex). Only then could visual sampling be perfectly adapted according to the behavioural goals of the observer. We have established that the spatial integration window is indeed under considerable control, and can be flexibly adjusted according to the task demands across a large part of the visual field. Nevertheless, there are clear limits to this flexibility. These limits reflect the coding of visual information by early mechanisms and the extent to which the neural circuitry involved in programming saccadic eye movements is able to appropriately weigh and combine the outputs from these mechanisms.

[1] The estimates of the match between filter and signal, r, can be plugged into Eq. (2) to derive an estimate of the amount of internal noise that is required to completely account for variations in efficiency (quantified by α). Across observers and eccentricities internal-to-external noise ratios were ~4 times larger in the large spatial scale conditions.

Acknowledgments

This work was funded by an EPSRC grant awarded to C.L. (EP/D009898/1) and National Science Foundation grant 0135118 to M.E.

References

Abbey, C. K., & Eckstein, M. P. (2002). Classification image analysis: estimation and statistical inference for two-alternative forced-choice experiments. *Journal of Vision, 2*, 66–78.

Ahumada, A. J. Jr., (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2*, 121–131.

Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research, 14*, 589–592.

Beutter, B. R., Eckstein, M. P., & Stone, L. S. (2003). Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *Journal of the Optical Society of America A, Optics Image Science and Vision, 20*, 1341–1355.

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

Caspi, A., Beutter, B. R., & Eckstein, M. P. (2004). The time course of visual information accrual guiding eye movement decisions. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 13086–13090.

Dakin, S. C., & Bex, P. J. (2003). Natural image statistics mediate brightness 'filling in'. *Proceedings of the Royal Society of London Series B. Biological Sciences, 270*, 2341–2348.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

Eckstein, M. P., Ahumada, A. J., Jr., & Watson, A. B. (1997). Visual signal detection in structured backgrounds II: effects of contrast gain control, background variations, and white noise. *Journal of the Optical Society of America A. Optics Image Science and Vision, 14*, 2406–2419.

Eckstein, M. P., Beutter, B. R., & Stone, L. S. (2001). Quantifying the performance limits of human saccadic targeting during visual search. *Perception, 30*, 1389–1401.

Eckstein, M. P., Pham, B. T., & Shimozaki, S. S. (2004). The footprints of visual attention during search with 100% valid and 100% invalid cues. *Vision Research, 44*, 1193–1207.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. *Monographs on statistics and applied probability* (Vol. 57). New York: Chapman & Hall.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

Fredericksen, R. E., Bex, P. J., & Verstraten, F. A. J. (1997). How big is a Gabor patch, and why should we care? *Journal of the Optical Society of America A. Optics Image Science and Vision, 14*, 1–12.

Garcia-Perez, M. A., & SierraVazquez, V. (1996). Do channels shift their tuning towards lower spatial frequencies in the periphery? *Vision Research, 36*, 3339–3372.

Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology, 10*, 663–666.

Graham, N. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley (pp. 174).

Levi, D. M., Klein, S. A., & Chen, I. N. (2005). What is the signal in noise? *Vision Research, 45*, 1835–1846.

Lu, Z. L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A. Optics Image Science and Vision, 16*, 764–778.

Ludwig, C. J. H., & Gilchrist, I. D. (2006). The relative contributions of luminance contrast and task demands on saccade target selection. *Vision Research, 46*, 2743–2748.

Ludwig, C. J. H., Gilchrist, I. D., McSorley, E., & Baddeley, R. J. (2005). The temporal impulse response underlying saccadic decisions. *Journal of Neuroscience, 25*, 9907–9912.

Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: weighted sums. *Journal of Vision, 2*, 79–104.

Pelli, D. G. (1985). Uncertainty explains many aspects of contrast detection and discrimination. *Journal of the Optical Society of America A. Optics Image Science and Vision, 2*, 1508–1531.

Pointer, J. S., & Hess, R. F. (1989). The contrast sensitivity gradient across the human visual field: with emphasis on the low spatial frequency range. *Vision Research, 29*, 1133–1151.

Rajashekar, U., Bovik, A. C., & Cormack, L. K. (2006). Visual search in noise: revealing the influence of structural cues by gaze-contingent classification image analysis. *Journal of Vision, 6*, 379–386.

Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research, 21*, 409–418.

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research, 37*, 475–494.

Wilson, H. R., & Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research, 19*, 19–32.