Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research xxx (2004) xxx–xxx
Vision Research
www.elsevier.com/locate/visres

The footprints of visual attention during search with 100% valid and 100% invalid cues

Miguel P. Eckstein *, Binh T. Pham, Steven S. Shimozaki

Vision and Image Understanding Laboratory, Department of Psychology, University of California, Santa Barbara, CA 93106, USA

Received 26 August 2003; received in revised form 9 October 2003

Abstract

Human performance during visual search typically improves when spatial cues indicate the possible target locations. In many instances, the performance improvement is quantitatively predicted by a Bayesian or quasi-Bayesian observer in which visual attention simply selects the information at the cued locations without changing the quality of processing or sensitivity and ignores the information at the uncued locations. Aside from the general good agreement between the effect of the cue on model and human performance, there has been little independent confirmation that humans are effectively selecting the relevant information. In this study, we used the classification image technique to assess the effectiveness of spatial cues in the attentional selection of relevant locations and suppression of irrelevant locations indicated by spatial cues. Observers searched for a bright target among dimmer distractors that might appear (with 50% probability) in one of eight locations in visual white noise. The possible target location was indicated using a 100% valid box cue or seven 100% invalid box cues in which the only potential target locations was uncued. For both conditions, we found statistically significant perceptual templates shaped as differences of Gaussians at the relevant locations with no perceptual templates at the irrelevant locations. We did not find statistical significant differences between the shapes of the inferred perceptual templates for the 100% valid and 100% invalid cues conditions. The results confirm the idea that during search visual attention allows the observer to effectively select relevant information and ignore irrelevant information. The results for the 100% invalid cues condition suggests that the selection process is not drawn automatically to the cue but can be under the observers' voluntary control.

1. Introduction

A common finding in visual search is that human performance, measured either with response times or accuracy, improves when a spatial cue indicates the probable location of the target (Baldassi & Verghese, 2002; Eckstein, 1998; Eckstein, Thomas, Palmer, & Shimozaki, 2000; Foley & Schwarz, 1998; Palmer, 1994; Palmer, Ames, & Lindsey, 1993; Posner, 1980; Solomon, Lavie, & Morgan, 1997; Verghese, 2001; Verghese & Stone, 1995). This result has been interpreted by some as suggesting that the cue allows the observer to allocate attentional resources to a single location rather than distribute them across many locations and therefore enhances processing at that cued (attended) location (e.g., Bashinski & Bacharach, 1980; Downing, 1988; Hawkins et al., 1990; Luck, Hillyard, Mouloua, & Hawkins, 1996). However, another hypothesis is that when the target is presented among visually similar distractors, the presence of a 100% valid cue can benefit performance by allowing the observer to ignore potentially confusable distractors. This benefit due to selection is expected even without considering limited attentional resources. This concept has been formalized by the theory of signal detection (Green & Swets, 1966). In this theory, each element (target and distractors) elicits an internal response within the observer that is subject to noise. A distractor might be confused for the target because it occasionally elicits a stronger response than the target. In this context, attention improves performance by allowing the observer to select responses from the relevant cued location and to ignore noisy responses arising from irrelevant noise locations that would otherwise bring additional unnecessary variability into the decision.

But how much improvement is expected in human performance based on these principles? One sensible

* Corresponding author.
E-mail address: eckstein@psych.ucsb.edu (M.P. Eckstein).
URL: http://www.psych.ucsb.edu/~eckstein/lab/vp.html.

0042-6989/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.visres.2003.10.026

starting point is to compare humans to an observer that optimally uses the cue to perform the task: the Ideal Bayesian observer. The optimal observer uses all prior information about the data, including that provided by the cue, to calculate the posterior probability of the various hypotheses considered (e.g., signal present vs. signal absent for a yes/no task; interval/location 1 vs. interval/location 2 for a 2 interval/alternative forced choice) and chooses the hypothesis with the highest posterior probability. The posterior probability is calculated as the product of the likelihood of the data given the hypothesis and the prior probability. In some cases (e.g., yes/no task in one of M locations), the Bayesian observer decision rules are non-linear, cannot be calculated using closed form expressions and require Monte Carlo computer simulations. Because running these simulations used to be time consuming given the available computer power, historically investigators have used other models that can be computed from closed form expressions, and are approximations to the optimal observer. One commonly used model bases its decision on the maximum response among the considered responses (max model). In many instances, the ideal Bayesian observer results in performance improvements of similar magnitude to the max model. Many studies have used the set-size effects predicted by the maximum model to compare to human performance. In many situations, the set-size effects in human observers are comparable to that expected from this attentional selection model (Baldassi & Verghese, 2002; Eckstein, 1998; Palmer et al., 1993; Solomon et al., 1997; Verghese, 2001). In other instances where the tasks involve more complex judgments (Poder, 1999), memory, or a rapid temporal sequence of two possibly conflicting cues (Dosher & Lu, 2000a; Lu & Dosher, 2000), the set-size effects present in humans were larger than predicted by these models (Carrasco, Williams, & Yeshurun, 2002; Dosher & Lu, 2000b; Poder, 1999).

A fundamental assumption of both the optimal Bayesian observer and maximum response observer is that for 100% valid cues, humans can perfectly select the information at the relevant cued location and ignore the information at the irrelevant location. However, there has been little confirmation of this selection aside from the comparisons between the performance improvements due to the presence of the cue shown by humans and those expected from the models. In addition, in some instances (Cameron, Carrasco, Tai, & Eckstein, 2004; Foley & Schwarz, 1998), set-size effects were smaller than those predicted by the model, which is sometimes interpreted as suggesting that the observer cannot optimally use the cue to select the relevant information and therefore does not obtain the full benefit in performance of ignoring irrelevant information. Yet, there is no independent method based on the behavioral data to directly assess any ineffective selection in these circumstances.

Our goal is to use a technique introduced recently to visual psychophysics known as classification images to directly estimate perceptual templates at each of the search locations and assess whether observers can efficiently select relevant information. We have previously used the technique to assess the effect of attention with a partially valid cue, a paradigm used prominently by Posner (Eckstein, Shimozaki, & Abbey, 2002; Posner, 1980). Here, we use the technique to investigate attentional selection with 100% valid cues during visual search. Because many previous studies have suggested that cues appearing at the "to be attended" location (peripheral cues) automatically attract attention (Luck & Thomas, 1999; Turatto et al., 2000), we investigate attentional selection under two different conditions. In the first condition, a 100% valid cue indicates the only potential location of the target among eight locations (Fig. 1, left image). In the second condition, we use 100% invalid cues that appear at the seven locations that

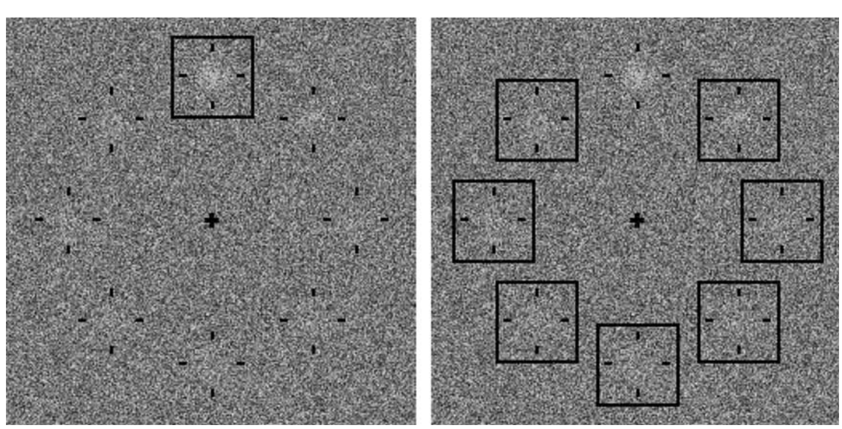

Fig. 1. Left image: 100% valid cue condition. The task is to detect a contrast increment that can only (with 50% probability) appear at the cued location. The cue (black box) n is randomly chosen among the 8 locations. Right image: 100% invalid cues condition. The task remains the same as the 100% invalid cue condition but the target can only appear at the uncued location. Fiduciary crosses are used to indicate the precise location of the target within each of the 8 locations.

will not contain the target (Fig. 1, right image). In this latter condition, the location not containing the cue is the only potential location of the target.

1.1. What is a classification image?

The classification image technique enables the researcher to directly estimate how the observers weight visual information in the image to reach a perceptual decision. It is the psychophysical analog of the reverse correlation technique used to estimate receptive fields in physiology (Sutter, 1975). Ahumada and Lovell (1971) first applied a similar technique (multiple linear regression) to audition. In 1996, it was applied to visual perception (Ahumada, 1996; Beard & Ahumada, 1998) and since then has been applied to study a variety of perceptual problems (illusory contours, Gold, Murray, Bennett, & Sekuler, 2000; stereo, Neri, Parker, & Blakemore, 1999; attention, Eckstein et al., 2002).

For luminance defined signals, the investigator adds random spatially uncorrelated luminance noise to the image. In a yes/no task where the signal is present p% of the times, the investigator records the noisy stimuli presented in the trials corresponding to each of the different human observer decisions: signal present trials in which the observer correctly responded 'signal present' (hit trials), signal present trial in which the observer incorrectly responded 'signal absent' (incorrect rejection or miss trials), signal absent trials in which the observer correctly responded 'signal absent' (correct rejection trials), and signal absent trials in which the observer incorrectly responded 'signal present' (false alarm trials).

The concept of the classification image technique is best exemplified with the false alarm trials. In these trials, noise samples that did not contain the signal led the observer to respond that the signal was present. Therefore, the random luminance variations on that trial must have contained some luminance pattern that the observer identified as the signal. Thus, the sample mean of all the noise images from the false alarm trials will result in average luminance pattern that led the observer to respond that the target was present when it was not. For simple tasks in which the observer is well approximated by a linear operation (the correlation or dot product of a template and the image), it can be shown that the sample mean of the noisy images will accurately estimate the observer's perceptual template, or the set of spatial weights the human observer uses to integrate the information in the image to reach the decision (Abbey & Eckstein, 2002; Ahumada, 2002; Murray, Bennett, & Sekuler, 2002; Solomon, 2002). Methods have been derived to estimate the perceptual templates in a statistically efficient manner for a variety of tasks (Abbey & Eckstein, 2002; Ahumada, 2002; Murray et al., 2002; Solomon, 2002).

In this paper, we are interested in estimating the perceptual templates for all locations during visual search. If the observers are using the information at only the relevant location, then we should obtain a perceptual template at the relevant location, but should obtain classification images that are not different than zero at the irrelevant locations. In addition, if observers are automatically drawn to attend to cued locations, then we should observe in the 100% invalid cues conditions that the classification images at the cued irrelevant locations are significantly different than zero.

2. Methods

2.1. Psychophysical task

The observers' task was to decide whether a contrast increment (5.71%; SNR = 3.41) was present (yes/no) in one of eight Gaussian pedestals (contrast = 9.3%; SNR = 5.1, st. dev. = 6 pixels = 0.226°). The eight Gaussian pedestals were located equidistant along the circumference of a circle with a radius of 3.384° (90 pixels). The eight possible target locations were indicated with fiduciary crosses that were placed 1.2° away from the center of each of the possible target locations. On each trial, the location was chosen randomly from the eight possible locations. White Gaussian luminance noise with a contrast of 19.53% was added to each image. Every trial/image in the study had an independent sample of noise. The viewing distance was 50 cm. The signal was present on 50% of the trials. There were two experimental conditions: (1) 100% valid simultaneous cue condition where a square box (side = 1.955°) appeared around the only possible location of the target. The location was chosen on each trial randomly from the eight possible locations. (2) 100% invalid simultaneous cues condition where 7 square box cues (side = 1.955°) appeared around seven locations indicating that these were not the target locations. Three naïve yet trained observers participated in the study. The observers participated in 110 sessions of 100 trials per condition resulting in 11,000 total trials per condition. An accidental loss of data left observer AK with 9100 total trials of data. Stimuli were presented on an Image Systems monochrome monitor (Image Systems Corp., Minnetonka, MN 55343). Each pixel subtended a visual angle of 0.0376°. The relationship between digital gray level and luminance was linearized using a Dome Imaging Systems board and a luminance calibration system. The mean luminance was 24.8 cd/m^2.

2.2. Procedure

Observers started the trial with a key press. A fixation image was presented for 1 s. Observers were instructed to fixate a central cross at all times. The stimulus was then displayed for 200 ms. The short presentation of the stimulus (200 ms) was chosen to preclude observers from executing a saccadic eye movement to fixate the cued location. A white noise mask with higher RMS (root mean square) contrast than the stimulus display was then presented for 100 ms (same mean background luminance 24.8 cd/m^2). The observers then pressed one of two yes/no boxes on a computer screen to select their decision (signal present or signal absent). Feedback about the correct decision was provided.

2.3. Human and model performance

Performance for human observers for each condition was measured in terms of the proportion of signal present trials in which the observer correctly responded (hit rate) and the proportion of signal absent trials in which the observer incorrectly responded 'signal present' (false alarm rate). We obtained an index of detectability (d') and criterion (λ) as given by the basic signal detection relationship (Green & Swets, 1966):

$$\lambda = -z(\mathrm{FA}) \tag{1}$$

$$d' = z(\mathrm{HR}) - z(\mathrm{FA}) \tag{2}$$

where z is the inverse cumulative Gaussian function, HR is the hit rate and FA is the false alarm rate.

2.4. Classification images

Classification images were obtained by computing the sample mean of the 20×20 pixel noise images presented in the signal absent trials in which the human incorrectly responded 'signal present' (false alarm trials). The actual number of images depended on the false alarm rate of each individual observer. They were: 1725 (BS, 100% valid cue); 1424 (BS, 100% invalid cue); 1919 (AK, 100% valid cue); 1277 (AK, 100% invalid cue); 1188 (TB, 100% valid cue); 1210 (TB, 100% invalid cue). Radial averages across all angles were computed for each of the noise images. A sample mean and a standard deviation of each element of the radial averages were computed, as well as the sample covariance between each element. The classification images for irrelevant locations were calculated for each irrelevant location relative to the randomly assigned relevant location (counterclockwise from the relevant location). In addition, a second set of classification images for irrelevant locations were calculated by averaging noise images for each absolute location (e.g., all noise images at 3 o'clock or all noise images at 6 o'clock).

2.5. Statistical inference for classification images

In order to test for statistically significant classification images we have previously used the The Hotelling T^2 statistic, a multivariate generalization of the univariate t. Here, we used one sample Hotelling T^2 statistics to test whether the radial averages of the classification images were significantly different from zero. The Hotelling T^2 statistic is:

$$T^2 = N[\mathbf{x} - \mathbf{x}_0]^t \mathbf{K}^{-1} [\mathbf{x} - \mathbf{x}_0] \tag{3}$$

where N is the number of observations (false alarm trials for our case), $\mathbf{x}$ is a vector containing the observed radial average of the classification image and $\mathbf{x}_0$ is either a population or a hypothesized radial average classification image. $\mathbf{K}^{-1}$ is the inverse of the covariance matrix that contains the sample variance of each of the elements of the radial average classification images, and the sample covariance between them. To test for significance, the T^2 statistic can be transformed to an F statistic using the following relationship:

$$F = \frac{N - p}{p(N-1)} T^2 \tag{4}$$

where p is the number of dependent variables (number of vector elements in the radial average of the classification images), and N is as defined previously. The obtained F statistic can be compared to an F_{critical} with p degrees of freedom for the numerator and $N - p$ degrees of freedom for the denominator.

To compare the two sample classification images at the relevant locations across the two conditions (100% valid cue vs. 100% invalid cues condition) we used the independent two-sample T^2, which is given by:

$$T^2 = \frac{N_1 N_2}{(N_1 + N_2)} [\mathbf{x}_2 - \mathbf{x}_1]^t \mathbf{K}^{-1} [\mathbf{x}_2 - \mathbf{x}_1] \tag{5}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are vectors containing the observed radial averages of the two classification images, N_1 and N_2 refer to the number of observations for the two classification images. For the two-sample test a pooled covariance $\mathbf{K}$ is computed combining the sum of square deviations and sum of squared products from both samples. To test for significance, the two-sample T^2 statistic can be transformed to an F statistic using the following relationship:

$$F = \frac{N_1 + N_2 - p - 1}{p(N_1 + N_2 - 2)} T^2 \tag{6}$$

where p, N_1 and N_2 are defined before. The obtained F statistic can be compared to an F_{critical} with p degrees of freedom for the numerator and $N_1 + N_2 - p - 1$ degrees of freedom for the denominator.

2.6. Estimating human performance from the estimated templates and internal noise

From the classification images one can estimate the shape of the perceptual templates used by the human observer. In addition, Ahumada (2002) has shown that the internal noise of the observer can be estimated from the amplitude of the classification image. Given the estimated perceptual template and internal noise, we can generate a prediction of human performance by assuming that the observer makes a decision by correlating the template with the data at the relevant location. The scalar response is then perturbed with internal noise (estimated from the classification images) and the observer compares the response to a criterion (λ) to make a decision. If the response exceeds the criterion, the observer says: "yes, target present". Otherwise, the observer says: "no, target absent". It can be shown that for a given estimated template and internal noise, performance (d') is given by (Burgess, Wagner, Jennings, & Barlow, 1981):

$$d'_{\mathrm{t}} = \frac{r}{\sqrt{1 + \frac{\sigma^2_{\mathrm{int}}}{\sigma^2_{\mathrm{ext}}}}}\mathrm{SNR} \tag{7}$$

where d'_{t} is the index of detectability of the template with the internal noise, SNR is the signal to noise ratio, σ^2_{ext} is the external noise standard deviation, σ^2_{int} is the estimated internal noise standard deviation and r is the correlation between the estimated template and the ideal template given by:

$$r = \frac{\sum\sum w_{\mathrm{t}} w_{\mathrm{i}}}{\sqrt{\sum\sum w^2_{\mathrm{t}}}\sqrt{\sum\sum w^2_{\mathrm{i}}}} \tag{8}$$

where w_{t} are the weights for the estimated template and w_{i} are the weights for the ideal observer template.

3. Results

3.1. Human performance for 100% valid cue vs. 100% invalid cues

Fig. 2 shows the hit rates (top left) and false alarm rates (bottom left) for the 100% valid cue and 100% invalid cues conditions for the three human observers. For observer AK the hit rate was significantly larger for the 100% valid cue condition while for observer TB it was significantly larger for the 100% invalid cues condition. We found no significant difference in the hit rates of observer BS. False alarm rates were significantly larger for the invalid cues condition for observers BS and AK. Fig. 2 also shows the index of detectability (d', top right) and the criterion (λ, bottom right) for the three observers and two conditions. For all observers, the index of detectability for the invalid cues condition was larger (7–15%) than the valid cue condition but did not reach statistical significance for observer AK. On the other hand, the criterion was significantly larger ($p < 0.05$) for the invalid cues condition for two of three observers (BS and AK).

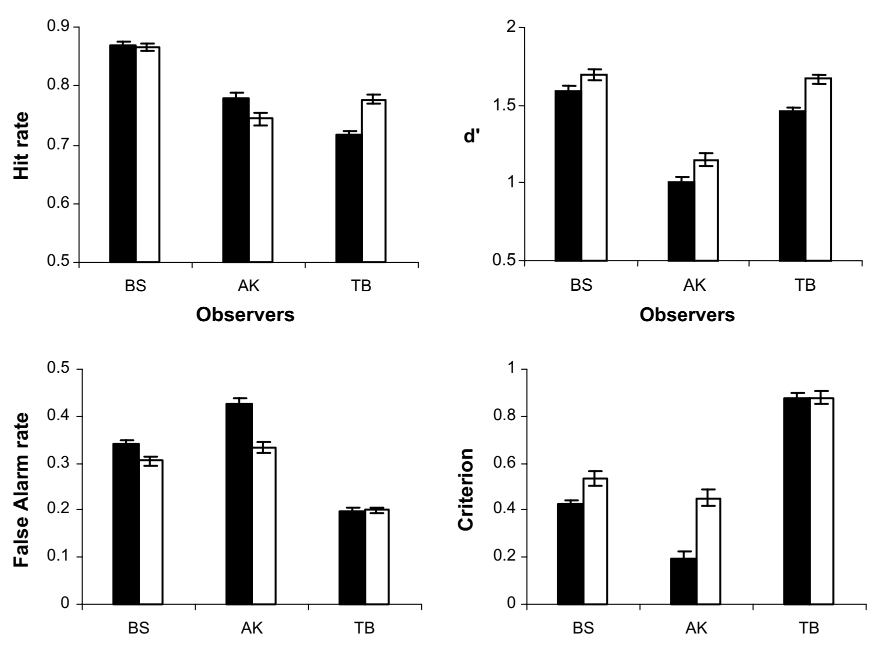

Fig. 2. Top left: hit rate for observers BS, AK and TB for 100% valid cue conditions (black column) and 100% invalid cues condition (white column). Bottom left: false alarm rate for observers for both cue conditions. Top right: index of detectability (d'). Bottom right: criterion (λ).

3.2. Human classification images for 100% valid cue and 100% invalid cues

Fig. 3 shows the classification images for BS, AK, and TB for the relevant location (cued location in the 100% valid cue condition and uncued location in the 100% invalid cue condition) and irrelevant locations (uncued locations for the 100% valid cue condition and cued locations for the 100% invalid cues conditions). Fig. 4 shows the position of the irrelevant locations relative to the relevant possible target location. Overall, the classification images show a general similarity across observers: a bright blob in the classification images from the relevant locations (for both the 100% valid and 100% invalid cues conditions) and virtually nothing at the irrelevant locations. The classification images of the

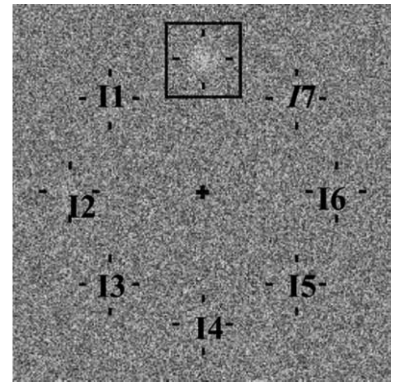

Fig. 4. Classification images for the (rotated) irrelevant locations defined relative to the relevant location. Irrelevant location 1 was always one position counterclockwise from the relevant location. The other irrelevant locations were defined accordingly.

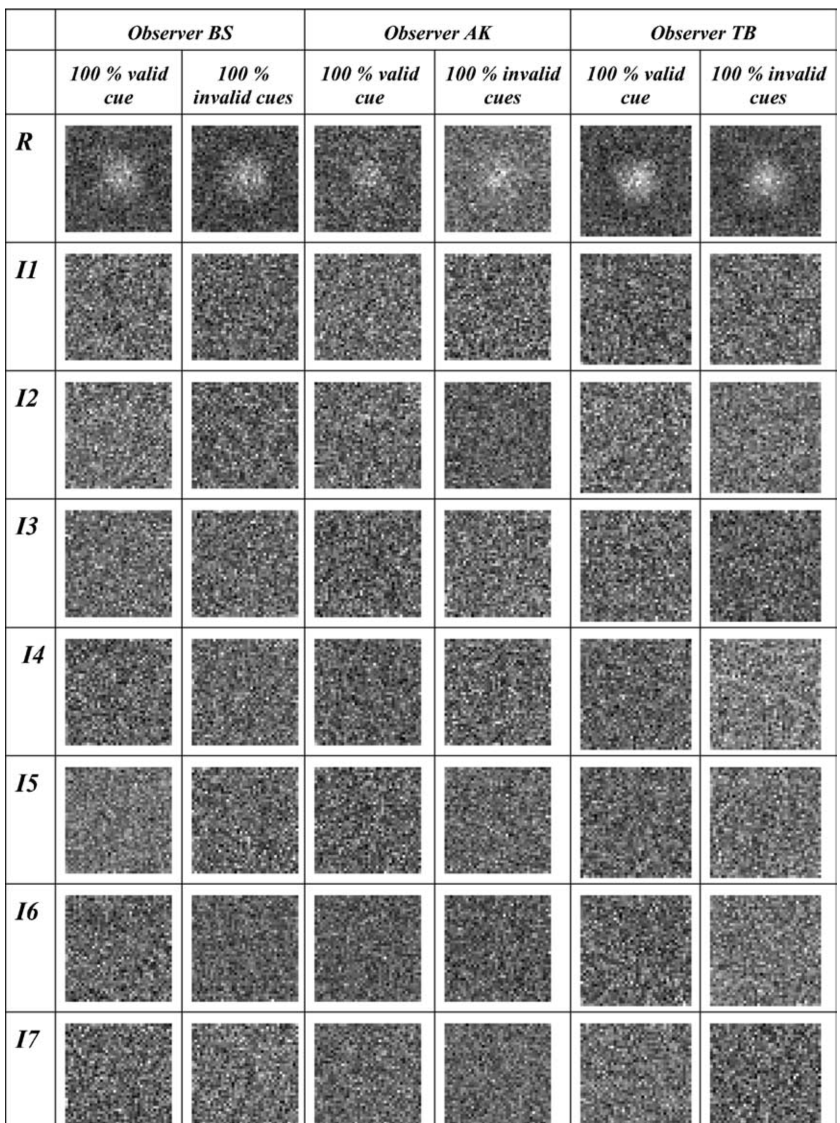

Fig. 3. Classification image for relevant locations (R) and irrelevant locations (I) for both the 100% valid and 100% invalid cue conditions. All classification images calculated relative to the position of the relevant possible target location.

relevant location for observer BS and TB seem to have a higher signal to noise ratio than those of AK.

In addition, classification images were calculated for each irrelevant location based on their absolute locations rather than positions relative to the relevant location. Although not shown, these resulted in similar classification images to those for irrelevant locations shown in Fig. 3.

3.3. Radial averages

In order to analyze the data more clearly, we calculated radial averages across all angles for each noise image from each false alarm trial. Sample mean radial averages were then calculated for the relevant and irrelevant locations for both conditions (100% valid cue and 100% invalid cue), as well as sample variance and covariance among the elements of the radial average vectors. Figs. 5 (BS), 6 (AK) and 7 (TB) show the radial averages for the three observers at the relevant location and a representative irrelevant location for both conditions. The perceptual template of the optimal observer is shown with dotted lines. Table 1 summarizes the F-values and p-values derived from the one-sample Hotelling T^2 statistic. Radial averages of the classification images were significantly different from a hypothesized null classification image (vector of zeros) at the relevant locations for the 100% valid cue and invalid cue conditions ($p < 0.003$). The radial average at every irrelevant location (uncued for the 100% valid cue condition and cued locations for the 100% invalid cues condition) was not statistically significant from the null vector (all zeros). Fig. 8 shows for each observer and cue condition a single radial profile of classification images averaged across all seven irrelevant locations. These did not reach statistical significance (see table in Appendix A for F-values and p-values). Fig. 9 shows representative radial averages of classification images for irrelevant locations calculated based on absolute location rather than position relative to the relevant possible target location. Representative radial averages are for the top location (i.e., 12 o'clock). Radial averages for all locations were not significantly different from a hypothesized null classification image (vector of zeros; see Appendix A for F and p-values).

The radial averages of the classification images at the relevant locations for the 100% valid cue and 100% invalid cue conditions were similar for each observer (with

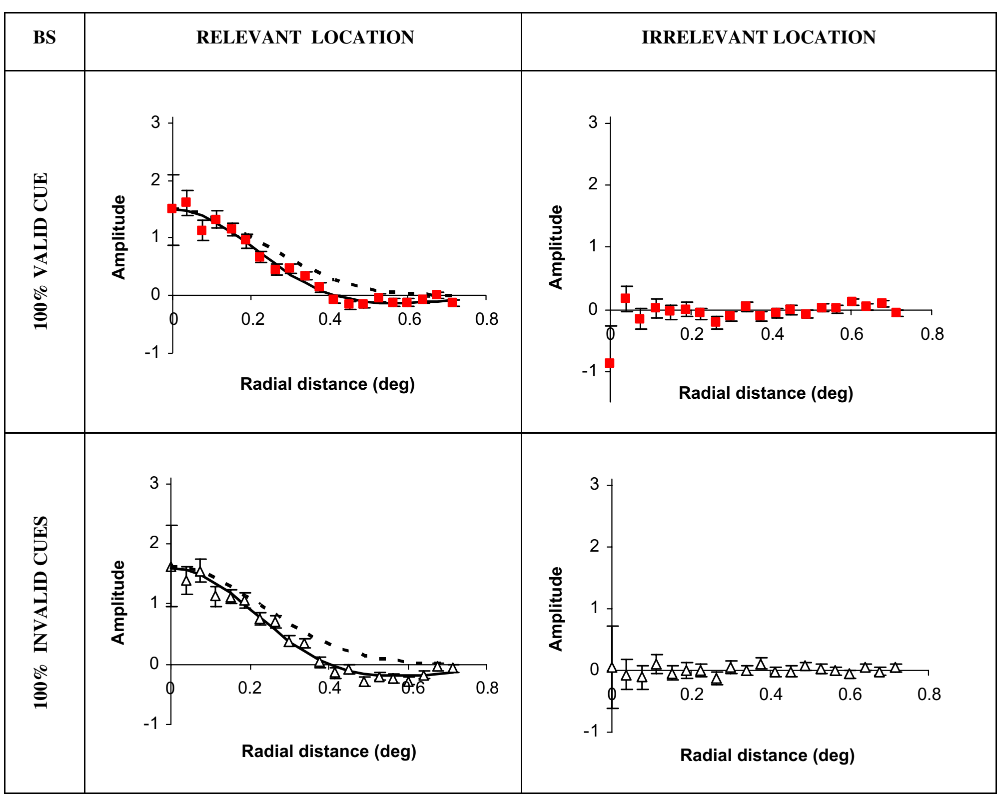

Fig. 5. Radial averages (spatial domain) of the classification images for observer BS. Top left: relevant location in the 100% valid cue condition (cued location). Bottom left: relevant location in the 100% invalid cues condition (uncued location). Top right: one irrelevant location (neighboring location) for the 100% valid cue condition (uncued location). Bottom right: one representative irrelevant location (neighboring location) for the 100% invalid cue condition. Black solid lines are the best-fit difference of Gaussian to the data. Dotted lines correspond to the radial profile of the optimal template: a Gaussian that matches the signal.

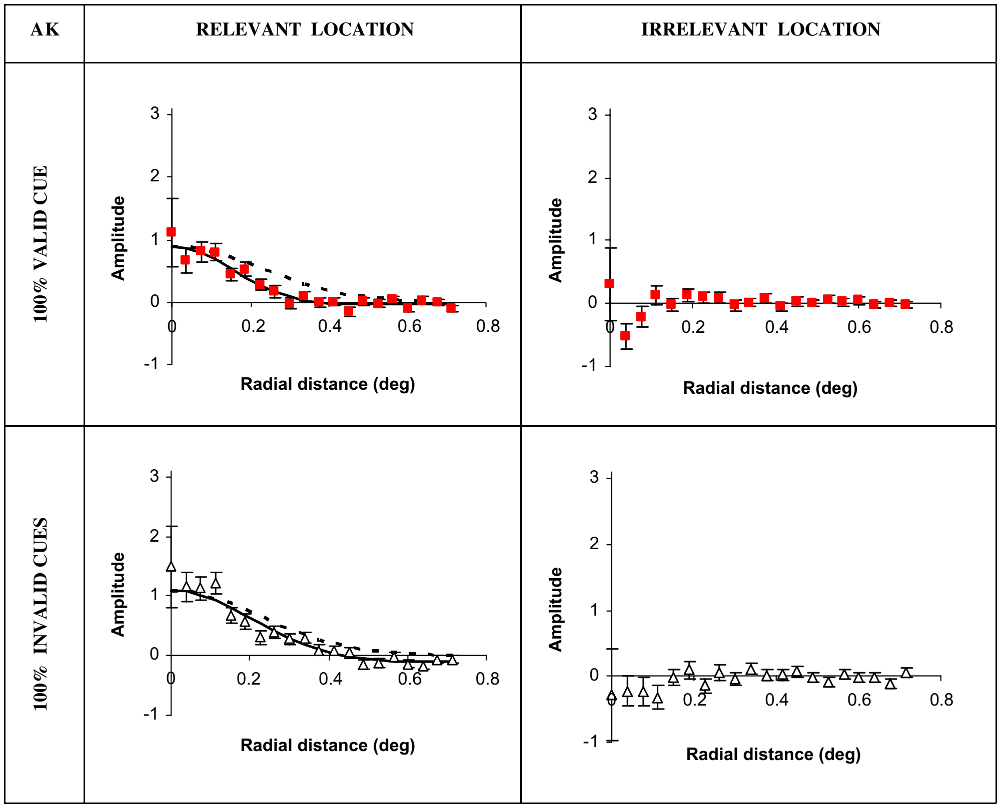

Fig. 6. Radial averages of the classification images for observer AK. Top left: relevant location in the 100% valid cue condition (cued location). Bottom left: relevant location in the 100% invalid cues condition (uncued location). Top right: one irrelevant location (neighboring location) for the 100% valid cue condition (uncued location). Bottom right: one representative irrelevant location (neighboring location) for the 100% invalid cue condition. Black solid lines are the best-fit difference of Gaussian to the data. Dotted lines correspond to the radial profile of the optimal template: a Gaussian that matches the signal.

the exception of higher amplitude for the 100% invalid cue condition for AK). Two-sample Hotelling T^2 statistic showed that the differences between the classification images at the relevant locations for both conditions were not statistically significant for all observers; however, that of AK was closer to reaching significance ($p = 0.016$ for AK).

Radial averages of classification images were fit with difference of Gaussians (DOG) functions with four fitting parameters: one amplitude for each of the two Gaussians (K_1 and K_2) and one standard deviation for each of the two Gaussians (σ_1 and σ_2). The DOG is given by:

$$\mathrm{DOG}(x,y) = K_1 \exp(-x^2/2\sigma_1^2) - K_2 \exp(-x^2/2\sigma_2^2) \quad (9)$$

Table 2 shows the χ^2 best fit parameters for the radial averages for the relevant locations for both human observers in both the 100% valid cue and 100% invalid cues condition. The table also includes a χ^2 goodness of fit for each of the fits. In addition, classification images were fit with a single Gaussian that lacks inhibitory surround (see Table 2) with two fitting parameters (K and σ). For all observers and both conditions the DOG functions resulted in a better fit than the single Gaussian functions. We used the χ^2_r goodness of fit measure to test whether the models could be rejected (other methods are available to compare models of classification images using maximum likelihood, i.e., Solomon, 2002).

The single Gaussian model (with no inhibitory surround) could be rejected for two of three observers for both cue conditions (BS and TB, see Table 2). The DOG model could not be rejected for any of the three observers and cue conditions.

To compare the shape of the estimated perceptual templates for the two conditions we computed their correlation (Table 3, left column). The correlations were high confirming that the shapes of the perceptual templates did not differ at a relevant cued location (100% valid cue condition) vs. a relevant uncued location (100% invalid cues condition).

3.4. *Predicting human performance from the human classification images*

Table 4 shows the values of internal noise (expressed in terms of units of the external noise standard deviation) inferred from the amplitude of the classification images (Ahumada, 2002). From the correlation between the

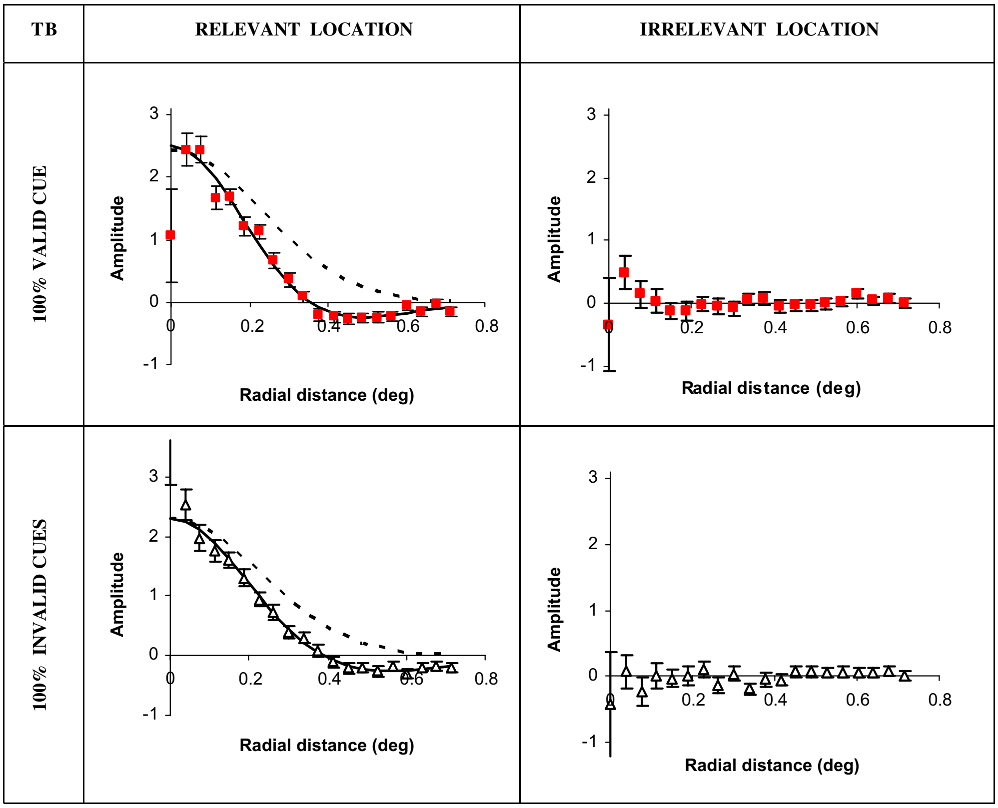

Fig. 7. Radial averages of the classification images for observer TB. Top left: relevant location in the 100% valid cue condition (cued location). Bottom left: relevant location in the 100% invalid cues condition (uncued location). Top right: one irrelevant location (neighboring location) for the 100% valid cue condition (uncued location). Bottom right: one representative irrelevant location (neighboring location) for the 100% invalid cue condition. Black solid lines are the best-fit difference of Gaussian to the data. Dotted lines correspond to the radial profile of the optimal template: a Gaussian that matches the signal.

Table 1
F-values and probabilities obtained from one-sample Hotelling T^2 statistic for the radial averages of each of the classification images (Eqs. (1)–(3))

	Relevant	Irrelevant 1	Irrelevant 2	Irrelevant 3	Irrelevant 4	Irrelevant 5	Irrelevant 6	Irrelevant 7
BS (panel A)								
100% Valid cue	$\boldsymbol{F = 26.172}$	$F = 1.27$	$F = 0.907$	$F = 0.673$	$F = 0.694$	$F = 0.859$	$F = 0.933$	$F = 1.014$
	$\boldsymbol{p < 0.001}$	$p = 0.229$	$p = 0.57S$	$p = 0.856$	$p = 0.837$	$p = 0.641$	$p = 0.544$	$p = 0.441$
100% Invalid cues	$\boldsymbol{F = 26.088}$	$F = 0.378$	$F = 1.881$	$F = 0.907$	$F = 0.634$	$F = 1.055$	$F = 1.116$	$F = 1.391$
	$\boldsymbol{p < 0.001}$	$p = 0.994$	$p = 0.011$	$p = 0.578$	$p = 0.889$	$p = 0.392$	$p = 0.325$	$p = 0.116$
AK (panel B)								
100% Valid cue	$\boldsymbol{F = 7.135}$	$F = 0.874$	$F = 0.647$	$F = 1.972$	$F = 0.794$	$F = 1.029$	$F = 0.957$	$F = 0.657$
	$\boldsymbol{p < 0.001}$	$p = 0.621$	$p = 0.879$	$p = 0.006$	$p = 0.723$	$p = 0.423$	$p = 0.513$	$p = 0.87$
100% Invalid cues	$\boldsymbol{F = 10.77}$	$F = 0.79$	$F = 1.014$	$F = 1.158$	$F = 0.995$	$F = 1.402$	$F = 1.071$	$F = 0.633$
	$\boldsymbol{p < 0.001}$	$p = 0.728$	$p = 0.441$	$p = 0.283$	$p = 0.466$	$p = 0.111$	$p = 0.375$	$p = 0.890$
TB (panel C)								
100% Valid cue	$\boldsymbol{F = 36.4586}$	$F = 0.6739$	$F = 1.7874$	$F = 1.7973$	$F = 1.2694$	$F = 1.0447$	$F = 0.7504$	$F = 0.7791$
	$\boldsymbol{p < 0.001}$	$p = 0.8554$	$p = 0.0177$	$p = 0.0168$	$p = 0.19$	$p = 0.405$	$p = 0.7748$	$p = 0.7411$
100% Invalid cues	$\boldsymbol{F = 35.219}$	$F = 0.68$	$F = 1.221$	$F = 0.9585$	$F = 2.018$	$F = 1.6952$	$F = 0.7418$	$F = 0.8131$
	$\boldsymbol{P < 0.001}$	$p = 0.8487$	$p = 0.2275$	$p = 0.5114$	$p = 0.0052$	$p = 0.0285$	$p = 0.7845$	$p = 0.6994$

Panel A observer BS. Panel B observer AK. Panel C observer TB. The p-value to reach significance with Bonferroni correction for 16 tests was 0.003. F-values that reached significance are in *bold*.

estimated human templates and optimal template (Table 3) and the internal noise inferred from the amplitude of the classification images (Table 4) we generated predictions for human performance (d'), as given by Eq. (7). Fig. 10 compares the measured performance (d') with that predicted from the estimated perceptual templates and internal noise for each observer. The left panel shows the results for the 100% valid cue condition while the

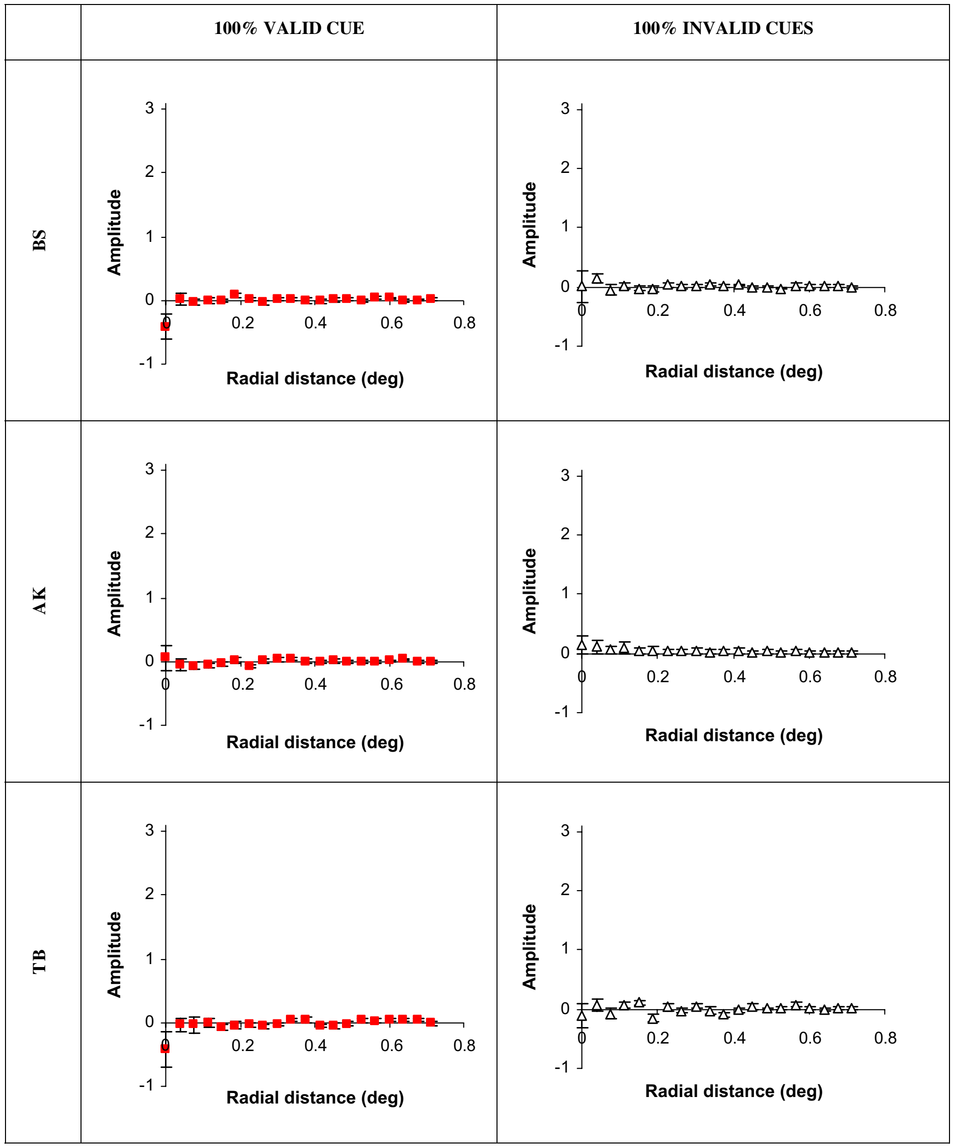

Fig. 8. Radial averages of the classification images for the three observers averaged across all seven irrelevant locations. Left column: 100% valid cue condition. Right column: 100% invalid cues condition.

right panel shows the results for the 100% invalid cues condition. Overall, we found a good agreement between the predicted and measured human performance, with the exception of a slight underestimate of AK's performance in the 100% valid cue condition and that of BS for the 100% invalid cue condition. The discrepancies were on the order of 10–15% in d' units.

4. Discussion

4.1. Attentional selection of information: automatic or voluntary?

The main purpose of the current study was to assess the effectiveness of attentional selection of relevant information during visual search. Our results (Figs. 2–4) clearly argue that for the present task visual attention allows the observers to effectively select information at the relevant locations and ignore information that at irrelevant locations. This finding confirms the underlying assumption in most of the signal detection models of visual search (Bayesian and quasi-Bayesian).

Many studies have drawn a distinction between automatic vs. voluntary processes of attention (Turatto et al., 2000; Yantis & Jonides, 1990). Some have argued that peripheral cues (i.e., cues presented at the "to be attended location") capture attention automatically, given that, even in the absence of predictive value, often times cues facilitate performance (e.g., Luck & Thomas, 1999). If peripheral cues automatically captured attention, we would have observed effective selection in the

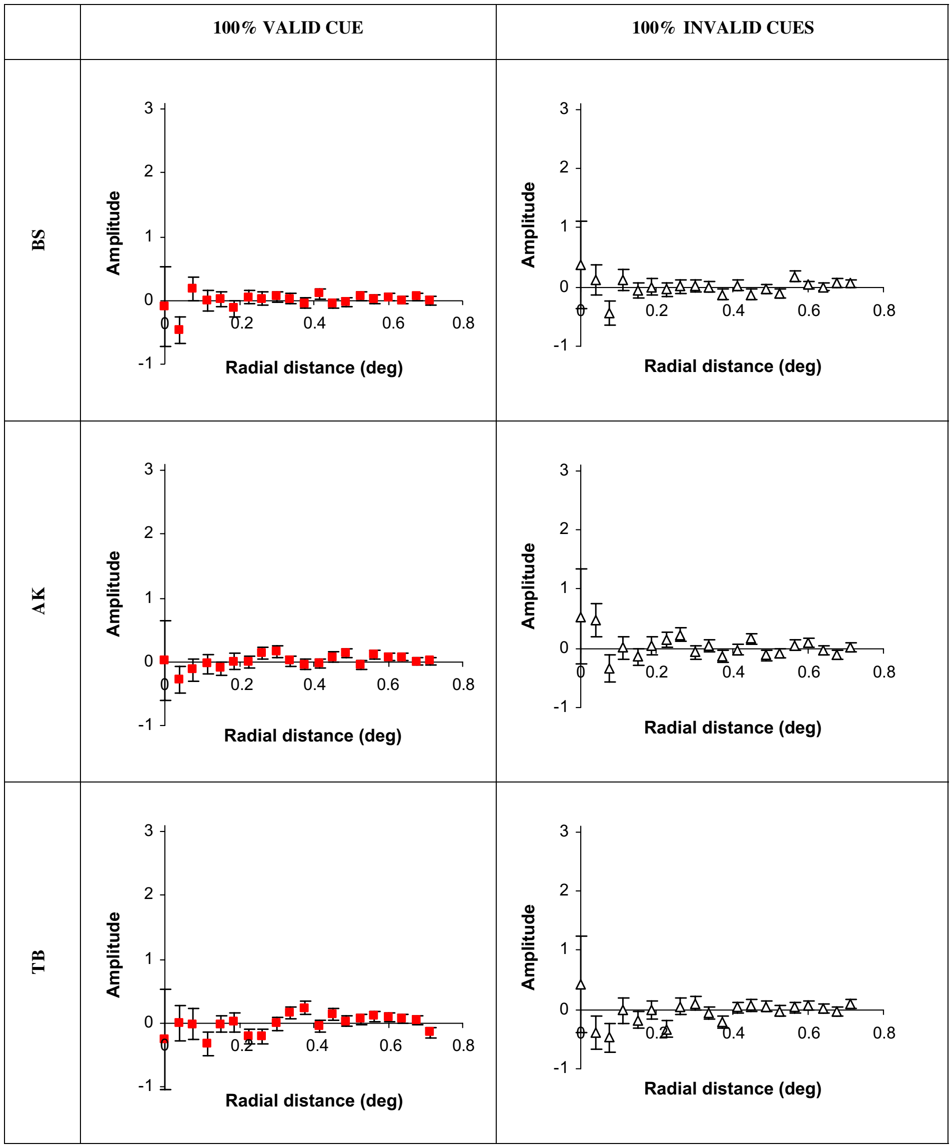

Fig. 9. Radial averages of the classification images for a representative irrelevant location calculated based on absolute location rather than position relative to the relevant possible target location. Shown radial averages for the top location, 12 o'clock. Left column: 100% valid cue condition. Right column: 100% invalid cues condition.

Table 2
Best fit parameters of the DOG functions (left) and single Gaussian functions (right) to the radial averages of the classification plots for the relevant locations for both conditions and the three observers

Observer	Difference of Gaussians						Gaussian			
	K_1	K_2	σ_1	σ_2	χ^2	χ^2_r	K	σ	χ^2	χ^2_r
Perceptual template at relevant location with 100% valid cue										
BS	3.9	2.4	6.4	7.8	19.72	1.23	1.65	4.4	50.9	**2.83***
AK	1.0	0.1	4.2	10.8	17.43	1.089	0.9	3.8	20.9	1.16
TB	4.9	2.4	5.4	7.4	18.75	1.17	2.6	4.0	87.9	**4.89***
Perceptual template at relevant location with 100% invalid cues										
BS	4.1	2.5	6.6	8.4	19.55	1.22	1.7	4.4	94.2	**5.23***
AK	1.3	0.2	5.6	16.4	16.69	1.043	1.2	4.4	29.77	1.65
TB	3.0	0.7	5.4	11.8	8.66	0.6	3.0	5.4	85.5	**4.75***

Goodness of fit $\chi^2_r = \chi^2/\mathrm{df}$ where df are the degrees of freedom defined as the number of data points minus the number of fitting parameters.

* $P(\chi^2_r \geqslant \chi_o)$ the probability of obtaining a χ^2_r that is larger or equal to the observed χ_o given that the model follows the data is less than 5%. Model can be rejected.

Table 3
Correlation between estimated 2-D human perceptual templates for the two conditions obtained from the best fit DOG (first column). Second and third column is the correlation between the estimated 2-D perceptual templates and the template of the optimal observer (a Gaussian)

Correlation	100% valid cue vs. 100% invalid cue	100% valid cue vs. optimal template	100% invalid cue vs. optimal template
BS	0.993	0.904	0.854
AK	0.951	0.851	0.898
TB	0.982	0.814	0.841

Table 4
Internal noise for both conditions expressed in units of the external noise standard deviation

Internal noise	100% valid cue	100% invalid cues
BS	1.70	1.64
AK	3.44	2.44
TB	1.52	1.48

100% valid cue but would not have observed it in the 100% invalid cues condition. Our results for the 100% invalid cues condition, however, show that observers selected the information at the uncued location and ignored the information at the cued locations. Furthermore, the selection process was as good as in the 100% cued condition. The optimality of the human perceptual template at the relevant location, measured by its correlation with the optimal template (Table 3), was very similar for both conditions. For two observers (BS and TB) there were small but significant superiority of performances for the 100% invalid cues condition. We speculate that the somewhat lower performance in the 100% valid cue condition might be due to lateral masking from the presence of the box cue (Foley, 1994). An additional interesting effect of the presence of the cue at the relevant location is that for two observers the 100% valid cue led to an increase in the observers' propensity to say "yes signal present" (false alarm rate).

Together, the results suggest that with 100% invalid cues, attentional selection was not automatic but under voluntary control. So, are these results in conflict with claims that cues automatically capture attention? No, it is very likely that in the presence of non-predictive cues in a laboratory setting, observers are still biased to select information at a cued location. This is perhaps due to the fact that in the real world abrupt onsets and salient cues often are informative. However, when the cost associated with attending to the cues is increased (100% invalid), our results show that the observer can voluntarily ignore the information at the cued locations (i.e., attention is not captured).

4.2. Behavioral attentional selection vs. neural attentional selection

By correlating the noise in the stimuli with the observers' decision we have found that observers effectively select relevant information across space. But how does this technique relate to methods assessing attentional selection in different brain areas by measuring physiological correlates, such as functional magnetic resonance imaging (fMRI) response or single cell recording? The classification technique probes the selection of visual information of the whole system with respect to behavior, while the physiological measures assess the attentional selection of some specific stage of visual processing (e.g., V1, MT, V4). However, the measured attentional modulation or selection of a specific cell or visual area does not necessarily have to agree with that manifested in the organism's behavior. There might be little spatial selection of attended stimuli in a given brain area, but that, might be followed by more selection at a later stage prior to behavior. For example, Saenz, Buracas, and Boynton (2003) have shown that attending to motion direction during a speed discrimination task of a stimulus in the left visual field increased the bold response in area MT to a motion stimulus moving in the same direction but presented in the task-

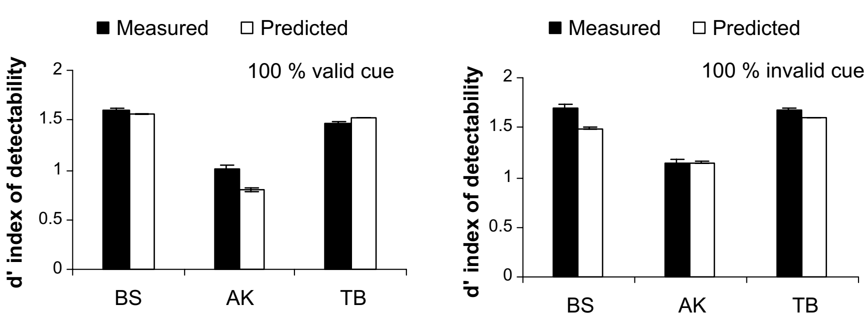

Fig. 10. Measured human performance vs. predicted performance from the estimated perceptual templates and internal noise from the classification images. Left panel: 100% valid cue condition. Right panel: 100% invalid cues condition.

irrelevant right visual field. That is, even if the observer is engaged in a speed discrimination task on the left visual field, MT activity at the right visual field is increased when the directions of motions are the same. This suggests that in this case, MT is not showing much of spatial attentional selection but rather is showing feature selection (i.e. motion direction). However, the observers' performance in the speed discrimination task was not affected by the direction of motion of the stimulus presented in right visual field, suggesting that from a behavioral point of view there is some spatial selection at some stage after MT. For the primary visual cortex V1, a study by Brefczynski and DeYoe (1999) measured the effect of spatial attentional selection on fMRI activity, and found that the cortical topography of attentionally driven signals to be similar to that in which the cued target is presented alone. This result suggests a high degree of spatial selectivity in V1, more similar to that measured behaviorally in the current study.

With respect to single cell recording, many studies have shown that firing rate is modulated by the animal's attentional state. A classic study has shown that neuronal responses to a distractor within V4 receptive fields are greatly attenuated when the monkey's attention is directed to the target (Moran & Desimone, 1985) but is not eliminated altogether. This and more recent studies show that typically the degree of attentional selectivity at a single cell level (e.g., McAdams & Maunsell, 2000; Treue & Maunsell, 1999) is less pronounced than the selection we have measured behaviorally. This difference might be due to the nature of the stimuli and the spatial distance between elements. However, it might be due to the fact that attentional selection at a single cell level is less pronounced than that manifested behaviorally.

4.3. Human vs. optimal perceptual templates

Comparison of the radial averages of the human classification images to the optimal perceptual template (Figs. 5–7) shows that for two observers the human perceptual templates tend to be narrower in the spatial domain than the optimal Gaussian filter, and also have an inhibitory surround. We consistently have observed this pattern of results in the detection and contrast discrimination tasks with Gaussian signals (e.g., Abbey, Eckstein, & Bochud, 1999; Eckstein et al., 2002). Humans have an inability to match the optimal profile when the signal is a Gaussian (see fits in Table 2). The inhibitory surround can be explained in terms of the decreased contrast sensitivity of the human visual system to low frequencies (i.e., the contrast sensitivity function) and/or a center-surround mechanism of many of the cells in the early visual system. The amplitude of the classification images (and radial average) was larger for BS and TB than AK. This is partly explained by the fact that BS and TB participated in 15% more trials than AK, and might be also attributed to AK's larger internal noise. We were able to predict human performance well from the estimated templates and internal noise (Fig. 8), suggesting that modeling the observers as a linear observer that is limited by the suboptimality of a template and internal noise is a good approximation for the present contrast discrimination task.

5. Conclusions

We have applied the classification image technique to assess the effectiveness of attentional selection in the presence of a 100% valid cue and 100% invalid cues during visual search. Our results show that observers can effectively select information at the relevant location and ignore information at the irrelevant locations for both cue conditions. The selection process is under the observers' voluntary control. This finding is consistent with signal detection models of visual search that assume that attentional selection is one main process by which cues benefit search performance.

Acknowledgements

This work was supported by NSF (BCS-0135118), NASA (NAG9-1329) and NIH-HL 53455. The authors thank Aaron Klafter, Briana Schuetz and Tiara Burks for participating as observers in the study. Some of the results in this paper were previously presented at the International Workshop on Visual Attention, San Miniato, Italy, 2003. Commercial relationships: none.

Appendix A

F-values and probabilities obtained from one-sample Hotelling T^2 statistic for the radial averages of each of the classification images calculated based on absolute locations rather than locations relative to the relevant location (Eqs. (1)–(3)). First column also shows the F- and p-value for the classification images averaged across all eight irrelevant locations. In the following table, panel A observer BS, panel B observer AK, panel C observer TB. The p-value to reach significance with Bonferroni correction for 16 tests was 0.003. None of the F-values reached statistical significance

	Average	1	2	3	4	5	6	7	8
BS (panel A)									
100% valid cue	$F = 0.76$	$F = 0.692$	$F = 1.12$	$F = 0.664$	$F = 0.418$	$F = 0.803$	$F = 0.992$	$F = 1.33$	$F = 0.432$
	$p = 0.76$	$p = 0.837$	$p = 0.322$	$p = 0.864$	$p = 0.989$	$p = 0.713$	$p = 0.469$	$p = 0.149$	$p = 0.986$
100% invalid cues	$F = 0.61$	$F = 0.95$	$F = 0.97$	$F = 0.909$	$F = 0.679$	$F = 1.25$	$F = 0.664$	$F = 0.546$	$F = 1.32$
	$p = 0.91$	$p = 0.522$	$p = 0.496$	$p = 0.575$	$p = 0.85$	$p = 0.204$	$p = 0.864$	$p = 0.948$	$p = 0.016$
AK (panel B)									
100% valid cue	$F = 0.77$	$F = 0.844$	$F = 0.992$	$F = 1.03$	$F = 0.841$	$F = 1.14$	$F = 1.06$	$F = 1.24$	$F = 1.12$
	$p = 0.75$	$p = 0.66$	$p = 0.468$	$p = 0.426$	$p = 0.664$	$p = 0.30$	$p = 0.384$	$p = 0.208$	$p = 0.316$
100% invalid cues	$F = 1.18$	$F = 1.16$	$F = 1.56$	$F = 1.18$	$F = 0.579$	$F = 0.601$	$F = 1.36$	$F = 0.87$	$F = 0.981$
	$p = 0.26$	$p = 0.284$	$p = 0.056$	$p = 0.261$	$p = 0.929$	$p = 0.914$	$p = 0.134$	$p = 0.626$	$p = 0.483$
TB (panel C)									
100% valid cue	$F = 1.27$	$F = 0.667$	$F = 1.16$	$F = 1.35$	$F = 0.94$	$F = 0.81$	$F = 1.32$	$F = 1.05$	$F = 0.546$
	$p = 0.18$	$p = 0.861$	$p = 0.286$	$p = 0.138$	$p = 0.534$	$p = 0.702$	$p = 0.158$	$p = 0.402$	$p = 0.947$
100% invalid cues	$F = 1.74$	$F = 1.28$	$F = 0.984$	$F = 1.11$	$F = 0.731$	$F = 0.866$	$F = 0.863$	$F = 1.72$	$F = 1.25$
	$p = 0.02$	$p = 0.186$	$p = 0.479$	$p = 0.337$	$p = 0.796$	$p = 0.632$	$p = 0.636$	$p = 0.026$	$p = 0.204$

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*(1), 66–78.

Abbey, C. K., Eckstein, M. P., & Bochud, F. O. (1999). Estimation of human-observer templates for 2 alternative forced choice tasks. *Proceedings of SPIE, 3663*, 284–295.

Ahumada, A., Jr. (1996). Perceptual classification images from Vernier acuity masked by noise. *Perception*, 26, ECVP Supplement, 18 (Abstract).

Ahumada, A., Jr. (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2*(1), 121–131.

Ahumada, A. J., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America, 49*, 1751–1756.

Baldassi, S., & Verghese, P. (2002). Comparing integration rules in visual search. *Journal of Vision, 2*(8), 559–570.

Bashinski, H. S., & Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception and Psychophysics, 28*(3), 241–248.

Beard, B. L., & Ahumada, A., A. (1998). Technique to extract relevant image features for visual tasksB. E. Rogowitz, & T. N. Pappas (Eds.), *Proceedings of SPIE, 3299*, 79–85.

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the spotlight of visual attention. *Nature Neuroscience, 2*, 370–374.

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

Cameron, L., Carrasco, M., Tai, J., & Eckstein, M. P. (2004). Distracters impair performance in discrimination, detection, localization tasks, but transient covert attention reduces this effect. *Spatial Vision*, in press.

Carrasco, M., Williams, P. E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: support for signal enhancement. *Journal of Vision, 2*(6), 467–479.

Dosher, B. A., & Lu, Z. L. (2000a). Noise exclusion in spatial attention. *Psychological Science, 11*, 139–146.

Dosher, B. A., & Lu, Z. L. (2000b). Mechanisms of perceptual attention in precueing of location. *Vision Research, 40*, 1269–1292.

Downing, C. J. (1988). Expectancy and visual spatial attention: effects on perceptual quality. *Journal of Experimental Psychology—Human Perception and Performance, 14*, 188–202.

Eckstein, M. P. (1998). The lower efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 9*, 111–118.

Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision, 2*(1), 25–45.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception and Psychophysics, 62*, 425–451.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of Optical Society of America A, 11*(6), 1710–1719.

Foley, J. M., & Schwarz, W. (1998). Spatial attention: Effect of position uncertainty and number of distractor patterns on the threshold-versus-contrast function for contrast discrimination. *Journal of the Optical Society of America, 15*(5), 1036–1047.

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

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.

Hawkins, H. L., Hillyard, S. A., Luck, S. J., Mouloua, M., Downing, C. J., & Woodward, D. P. (1990). Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception and Performance, 16*(4), 802–811.

Lu, Z. L., & Dosher, B. A. (2000). Spatial attention: different mechanisms for central and peripheral temporal precues? *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1534–1548.

Luck, S. J., Hillyard, S. A., Mouloua, M., & Hawkins, H. L. (1996). Mechanisms of visual–spatial attention: Resource allocation or uncertainty reduction? *Journal of Experimental Psychology: Human Perception and Performance, 22*, 725–737.

Luck, S. J., & Thomas, S. J. (1999). What variety of attention is automatically captured by peripheral cues? *Perception and Psychophysics, 61*(7), 1424–1435.

McAdams, C. J., & Maunsell, J. H. (2000). Attention to both space and feature modulates neuronal responses in macaque area V4. *Journal of Neurophysiology, 83*(3), 1751–1755.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*(4715), 782–784.

Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: weighted sums. *Journal of Vision, 2*(1), 79–104.

Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature, 401*, 695–698.

Palmer, J. (1994). Set-size effects in visual search: the effect of attention is independent of the stimulus for simple tasks. *Vision Research, 34*(13), 1703–1721.
Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19*(1), 108–130.
Poder, E. (1999). Search for feature and for relative position: measurement of capacity limitations. *Vision Research, 39*, 1321–1327.
Posner, M. I. (1980). Orienting of attention. *Journal of Experimental Psychology, 32*, 3–25.
Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision, 2*(1), 105–120.
Solomon, J. A., Lavie, N., & Morgan, M. J. (1997). Contrast discrimination function: spatial cuing effects. *Journal of Optical Society of America A, 14*(9), 2443–2448.
Saenz, M., Buracas, G. T., & Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision Research, 43*, 629–637.
Sutter, E. E. (1975). A revised conception of visual receptive fields based on pseudorandom spatiotemporal pattern stimuli. In *Proceedings of the First Symposium on Testing and Identification of Nonlinear Systems*, Pasadena, California.
Treue, S., & Maunsell, J. H. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *Journal of Neuroscience, 19*(17), 7591–7602.
Turatto, M., Benso, F., Facoetti, A., Galfano, G., Mascetti, G. G., & Umilta, C. (2000). Automatic and voluntary focusing of attention. *Perception and Psychophysics, 62*(5), 935–952.
Verghese, P. (2001). Visual search and attention: a signal detection theory approach. *Neuron, 31*, 523–535.
Verghese, P., & Stone, L. S. (1995). Combining speed information across space. *Vision Research, 35*(20), 2811–2823.
Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 121–134.