

Detection in fixed and random noise in foveal and parafoveal vision explained by template learning

Bettina L. Beard and Albert J. Ahumada, Jr.

NASA Ames Research Center, Human Information Processing Research Branch, Moffett Field, California 94035-1000

Received July 6, 1998; revised manuscript received November 3, 1998; accepted November 4, 1998

Foveal and parafoveal contrast detection thresholds for Gabor and checkerboard targets were measured in white noise by means of a two-interval forced-choice paradigm. Two white-noise conditions were used: fixed and twin. In the fixed noise condition a single noise sample was presented in both intervals of all the trials. In the twin noise condition the same noise sample was used in the two intervals of a trial, but a new sample was generated for each trial. Fixed noise conditions usually resulted in lower thresholds than twin noise. Template learning models are presented that attribute this advantage of fixed over twin noise either to fixed memory templates' reducing uncertainty by incorporation of the noise or to the introduction, by the learning process itself, of more variability in the twin noise condition. Quantitative predictions of the template learning process show that it contributes to the accelerating nonlinear increase in performance with signal amplitude at low signal-to-noise ratios. © 1999 Optical Society of America [S0740-3232(99)00803-0]

OCIS codes: 330.1880, 330.4060, 330.5510, 330.6100, 330.7310.

1. INTRODUCTION

Both image discrimination¹⁻⁷ and template-matching models⁸⁻¹⁰ may be used to predict noise masking in target detection. Image discrimination models measure, or predict, the discriminability between two input images. To predict noise effects with an image discrimination model, a single white-noise sample can be added to both input images while only one image contains the target. A visual system module giving internal image representations processes the two images. The module simulates features of human vision, such as optical blurring, luminance, and contrast effects, and masking by nonlinear transduction within orientation selective channels. Differences between the representations are then summed. Representation components unaffected by the target do not contribute to the aggregated difference image since the noise samples are identical, but the nonlinear processing can allow the noise to reduce the difference in the components responding to the target. Image discrimination models can predict masking effects when a single noise sample (referred to here as the fixed noise condition) is used to mask a target or its absence.^{7,11}

Another condition that permits prediction by image discrimination models is the twin noise condition used by Ahumada and Beard,⁷ where a single noise sample is used for both images within a single two-interval trial but a new random noise sample is used for each trial. Although image discrimination models predict the same average thresholds in twin and fixed conditions and little variation from different white-noise samples, we found a threshold elevation for twin noise conditions relative to fixed noise samples, particularly at higher noise root-mean-square (rms) contrast levels.⁷ Here we present

data that extend this finding to other stimuli and to parafoveal retina.

A threshold elevation for twin noise relative to fixed noise suggests that the observer cannot compare the two eidetic image representations during a trial. One might try to explain the degradation caused by twin noise by including a model of the short-term visual storage system. A simpler, more common, approach is to assume that a sensory representation of each image is compared with an internal memory representation, or template, on each stimulus presentation. Then only the result of the comparison need be remembered. If the template is modified by the current sensory representation, the changing noise samples in the twin condition could add noise to the template as it is being constructed. Fixed noise, however, could contribute to the template construction process.

Template-matching models^{8,9} correlate sensory representations of the images with one or more memory templates. However, when the template models assume a single, unchanging memory template for fixed and twin noise conditions, they do not predict a detection threshold difference. In this paper we introduce template-matching models with adaptable templates. The models are too simple to predict our threshold data, but they do illustrate the possible role of template learning in the lowering of target detection thresholds in fixed noise conditions relative to twin noise conditions.

In contrast to the fixed/twin effect found by Ahumada and Beard,⁷ Watson *et al.*¹² reported no difference between Gabor target thresholds under fixed and twin conditions. Since a Gabor target may be difficult to localize, especially along the grating bars, we compare a Gabor target with a checkerboard target. The reasoning is that

template construction and matching processes should depend on the position alignment between the sensory representation and the memory template. The constantly changing noise in the twin noise condition should increase uncertainty about the target position, while the unchanging noise pixels in the fixed noise should aid in locating the target position. Increasing uncertainty should then dilute the fixed/twin effect. We include two variables that may affect positional uncertainty: (i) Gabor versus checkerboard targets, and (ii) eccentricity.¹³

If template construction is influenced by the noise structure, then following practice on one fixed noise sample with a different fixed noise sample should result in poorer performance (i.e., negative transfer of learning), both from having a less appropriate template and from increased positional uncertainty. To test this, we train observers on one noise sample and observe transfer of training to others.

In the absence of a mask, peripheral target detection thresholds are elevated relative to foveal measurements.^{14,15} Noise mask effects on peripheral target detection are unknown. Here we extend measurements to a parafoveal location. We also investigate whether the cortical magnification scaling factor often used to equate foveal and peripheral detection thresholds¹⁶ equates them in the presence of a white-noise background.

In summary, we

1. attempt to replicate the earlier Ahumada–Beard⁷ finding of a fixed/twin noise threshold difference with two new stimuli,
2. test whether stimulus structure or eccentricity affects the fixed/twin noise threshold difference,
3. measure white-noise effects on object detection in foveal and peripheral vision,
4. determine whether the cortical magnification scaling factor often used to equate foveal and peripheral detection thresholds holds in the presence of a white-noise background,
5. present two template learning models,
6. give reasons to include memory templates and template learning in models of detection and discrimination.

2. METHODS

A. Apparatus

Stimulus images were presented with the green gun of a cathode-ray tube (Sony Trinitron Color Graphics Display Model GDM-20E1) with 640×480 pixel resolution at 60 frames/s. The viewing distance was either 122 cm, which gives a display resolution of 55.3 pixels/deg, or 40.6 cm, which gives 18.4 pixels/deg. Images were 128×128 pixels (2.3 deg of visual angle for the farther distance, and 6.9 deg for the closer distance). The surrounding screen luminance was 19 cd/m^2 . To ensure that luminance was linear with digital value, a lookup table was used. A resistive mixing circuit was used to increase the accuracy of the linearization as described in Ahumada and Beard.⁷

B. Experimental Stimuli

To study noise effects on detection performance two gray-scale digital target images were generated. One image contained a horizontally oriented, odd-symmetric Gabor target with circular Gaussian windowing. The Gabor target had a vertical spatial frequency of 3.7 c/deg, and its $1/e$ spatial half-spread $\{\sigma \text{ in } \exp[-(x/r)^2]\}$ was 0.9 deg. A Gabor stimulus was chosen because Watson *et al.*¹² found no difference between fixed and twin noise thresholds, using a Gabor stimulus. The second digital image contained two light squares and two dark squares arranged as a checkerboard in a 1.8×1.8 deg square, similar in size to the Gabor target.

The upper two panels of Fig. 1 illustrate the targets used in this experiment. The lower two panels show the result of a simulation illustrating the relative spatial uncertainty of the checkerboard and Gabor targets in the presence of noise. These simulation distributions are the best cross-correlating positions of the target with itself in 200 random samples of white noise. Darker pixels represent a higher number of maximum scores. The positional uncertainty is greater for the Gabor target than for the checkerboard and greater for the Gabor target in the direction parallel to the edge rather than perpendicular to it. The signal-to-noise ratio was set so that the ideal observer d' would be 3.5. This value, representing the external as well as the sensory noise, allows for additional decision and memory noise following the signal localization. The simulation therefore supports the possibility that a Gabor stimulus could produce greater spatial uncertainty than a checkerboard stimulus.

Experimental images were constructed by addition of a fraction of the target image to a white-noise image. The white-noise image pixels were independently, identically distributed uniform random variables with a mean of

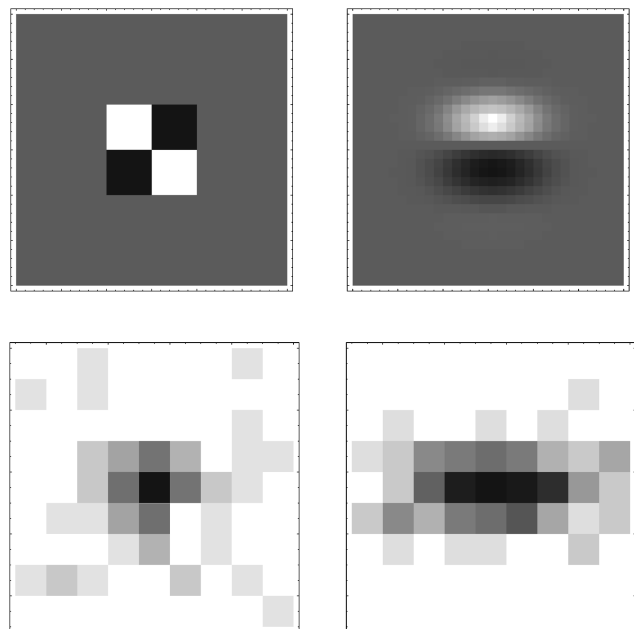


Fig. 1. Top two panels, the checkerboard and Gabor stimuli used in the experiment. Bottom two panels, the 200 best-correlating positions for each target with the (centered) target in noise, showing greater spatial uncertainty for the Gabor target than for a checkerboard target.

zero. One white-noise image was constructed with uniformly distributed amplitudes. It had a peak contrast of 0.33 and an associated rms contrast of 0.19. To get a new sample, a random number from 1 to 128×128 was chosen as the starting position for the upper left-hand corner of the noise, and then subsequent values in the table were copied with wraparound. For fixed noise samples this starting position was unchanged from trial to trial within a block. For twin noise a new starting position was determined for each two-interval trial.

Small location marks were used to increase position knowledge. These marks were 5 pixels wide and were positioned at the same vertical location as the target center just outside the 128×128 pixel image.

C. Noise Type

There were two noise sample randomization conditions: (i) In the fixed noise condition a single fixed noise sample was used throughout a series of trial blocks; (ii) in the twin noise condition a new noise was generated for each trial and was used for both intervals in a two-interval trial.

For each observer the order of the conditions (2 stimuli \times 3 eccentricities) was chosen at random. For observer CM the order of 2 stimuli \times 2 eccentricity conditions was randomized. The noise conditions (fixed versus twin) were always run successively, with their order independently randomized. Limited sets of fixed noise samples were randomly determined for each condition. From this set the experimenter determined the particular fixed noise sample for any set of trials and how many trial block repetitions there were for any single fixed noise sample.

D. Eccentricity

Our earlier measures⁷ were taken in the fovea. Because the periphery codes less phase information,¹³ here we measured detection thresholds in the fovea and at 4 deg eccentricity (inferior visual field) from the same 122-cm viewing distance. Observers also viewed the display from 40.6 cm at 4 deg eccentricity to get an image scaled by a factor of 3. This factor approximates cortical magnification estimates for contrast detection thresholds.¹⁶

E. Procedure

A two-interval forced-choice staircase tracking procedure was used with a stimulus duration of 0.5 s and an interstimulus interval of 1.0 s. The stimulus images appeared in the screen center and were replaced by the background luminance during the interstimulus and intertrial intervals. For one of the intervals, selected at random on each trial, the target image was absent (multiplied by zero). The observer's task was to determine the interval in which the target was presented. Auditory feedback was given if the keyboard response was incorrect. If the observer made an error, the target contrast energy was increased by 1.4 dB. If the observer was correct three trials in a row, the target energy was decreased by 1.4 dB. Conditions (i.e., target, noise type, and eccentricity) were fixed for blocks of 60 trials. Thresholds were determined for each block by maximum-likelihood probit analysis, which estimates the multiplier level leading to 75%

correct.¹⁷ The analysis assumed that the psychometric function is a cumulative normal distribution.

F. Observers

Five observers participated in this experiment: CM, NR, PW, SM, and BLB. Observers NR and BLB had extensive experience detecting an aircraft image in a runway scene masked by fixed noise in a previous experiment.⁷ Observer CM did not gather data in the 4-deg magnified condition.

3. RESULTS

First we present the data averaged across trial blocks for each observer. These results are followed by foveal block-by-block data of three of the five observers, to illustrate practice effects.

In Fig. 2 average target contrast thresholds are plotted for five observers. Thresholds are reported in dBB, a decibel contrast energy scale that is defined as

$$\text{dBB} = 10 \log_{10}(\text{CE}/\text{CE}_0), \quad (1)$$

where CE_0 is $10^{-6} \text{ deg}^2 \text{ s}$, the best contrast threshold reported by Watson *et al.*¹⁸ in their observer HB (H. Barlow).

Contrast energy is determined by

$$\text{CE} = A T \sum c_{ij}^2, \quad (2)$$

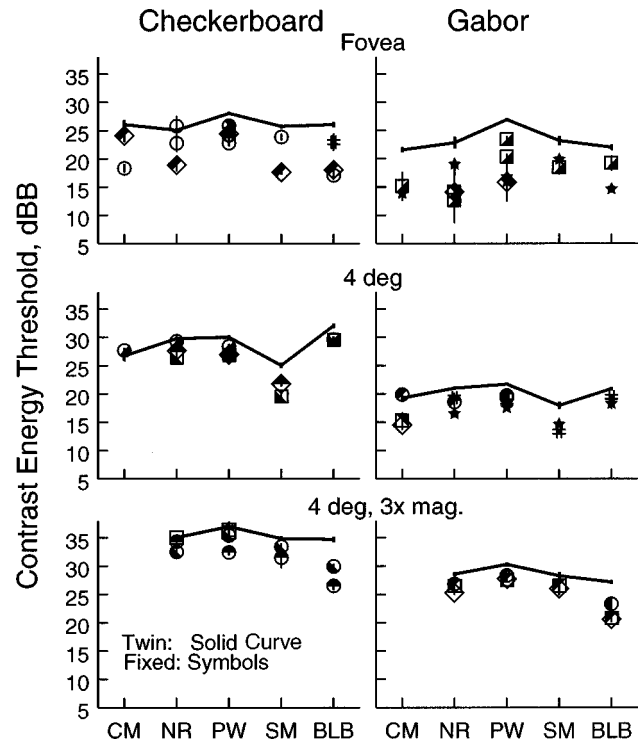


Fig. 2. Average contrast energy thresholds (in dBB) for five observers. Thresholds for the checkerboard target are shown in the left-hand panels; those for Gabor targets are in the right-hand panels. Data collected in the fovea are in the upper panels, data from 4-deg eccentricity (foveal viewing distance) are in the center panels, and data from 4-deg eccentricity from a closer (one third) distance are in the bottom panels. Solid curves represent the twin noise data. Particular symbols represent averaged thresholds for each fixed noise sample.

$$c_{ij} = (a_{ij} - B)/B, \quad (3)$$

where a , the digital amplitude (linear with luminance) of each image pixel is converted to contrast c by B , the digital background level. A is the area of a single pixel in degrees squared, and T is the stimulus duration in seconds. Solid curves connect twin noise averages. Fixed noise averages are shown with different symbols for each different fixed noise sample. Foveal data are presented in the two upper panels, unmagnified parafoveal data are shown in the center panels, and the lower panels show parafoveal data for the image magnified three times. The left-hand panels show data for the checkerboard target, and the right-hand panels show data for the Gabor target.

Figure 2 illustrates several major results. Statistical comparisons were made with 95% confidence intervals based on the variability of the comparison among observers. The major result is the significant performance improvement for fixed over twin. In the fovea, the fixed twin difference (collapsed across target type) averaged 5.5 ± 0.8 dB, while at 4 deg it was 2.6 ± 1.5 dB unmagnified and 3.2 ± 3.1 dB magnified. Excluding observer CM, since she did not make measurements for the 4-deg magnified condition, the foveal effect is significantly larger than the peripheral effects, which do not differ from one another. None of the noise-type (fixed/twin) effects interacted significantly with the target type (checkerboard versus Gabor).

Next we investigate whether there is a significant difference between the checkerboard and the Gabor target thresholds. For the twin noise conditions the Gabor target led to significantly lower contrast energy thresholds than did the checkerboard in the fovea (2.9 ± 1.7), at 4-deg unmagnified (8.6 ± 1.9) and 4-deg magnified (5.5 ± 4.4) conditions. The effects of eccentricity depended on the target type. For the twin conditions, moving the checkerboard target into the periphery led to a 2.5 ± 3.5 dB, nonsignificant performance decrement. To determine whether the magnification factor used equated the 4-deg magnified and foveal thresholds, we transformed the contrast energy (in dBB) scores to contrast threshold units. The graphs of Fig. 2 show contrast energy (in dBB). If the 4-deg magnified curves are shifted down by 9.5 dB [$20 \log_{10}(3)$], the graph can be compared with the others in the contrast domain. The contrast domain comparison shows that the magnification condition led to a threshold that was only 0.4 ± 0.9 dB below that of the fovea. Moving the Gabor target into the periphery led to a significant 3.1 ± 2.4 dB improvement in detection, so no compensatory magnification is required. Again we transformed contrast energy scores to contrast threshold units and found that, for a Gabor target, the magnification led to performance that was 4.7 ± 1.6 dB better than that expected from contrast threshold equality.

Figure 3 illustrates the learning and transfer of training results for the foveal data of observers CM, NR, and BLB. The left-hand panels show data for the checkerboard pattern; the right-hand panels, for the Gabor stimulus. Twin noise data are connected by a solid curve, while data for each fixed noise has a particular

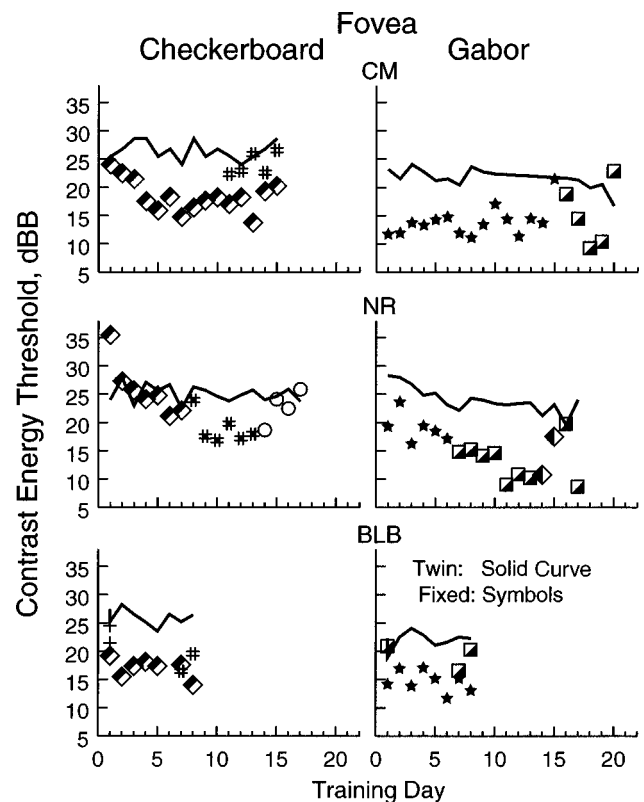


Fig. 3. Foveal contrast energy threshold (in dBB) is plotted as a function of the training day. Each data point is based on 60 two-alternative forced-choice trials. Solid curves represent the twin noise data. The remaining symbols represent fixed noise data. Fixed noise samples have the symbols assigned in Fig. 3. Checkerboard target thresholds are shown in the left-hand panels; Gabor target thresholds in the right-hand panels. The foveal data for three of the five observers are shown.

symbol for that noise sample. The twin noise data show little improvement over sessions or over training day. Regression analyses showed significant initial improvement in twin noise thresholds for observer NR (Gabor) and improvement on the final day for observer CM (Gabor). With fixed noise the more naive observers CM and NR showed strong improvement over as many as 5 days, sometimes in just 1 day (NR, checkerboard, denoted by filled #), and sometimes none at all (CM, Gabor, star). The most experienced observer (BLB) showed no significant learning. There is evidence of idiosyncratic negative transfer of training. When observer CM added the noise (data denoted by #), her performance was poor, while BLB performed well with the same maskers. Also the switch from masking the Gabor by the noised marked by the star to that marked by the square whose lower half is filled appeared to be disruptive for CM and possibly for BLB, but not for NR. Similar, but less significant effects are seen in the parafoveal learning sequences (not shown), consistent with the smaller size of the difference between twin and fixed conditions in the parafovea.

4. DISCUSSION

Ahumada and Beard⁷ found lower contrast detection thresholds for an aircraft target in a single fixed noise as compared with twin random noise. Here we extend this

finding of a fixed/twin threshold difference to multiple samples of fixed noise, to Gabor and checkerboard stimuli, and to parafoveal detection. One explanation is that an internal sensory representation of each input image is compared with templates stored in memory and that observers modify the templates as the experiment progresses. We hypothesize that the fixed/twin effect is the result of template learning. A fixed noise stimulus would be incorporated in the template and would thus help to reduce target position uncertainty. Twin noise samples, which change from trial to trial, would not provide this position information and might contribute noise to the templates.

Because Watson *et al.*¹² did not find a fixed noise advantage over twin noise for a Gabor target, we initially hypothesized that an inherent position uncertainty of the Gabor stimulus may slow learning the fixed noise components, thus removing any potential advantage of the fixed noise display. Our simulation results (see Fig. 1) lent support to this hypothesis. However, our experimental results did show a fixed/twin noise effect. It is likely that differences between the experimental parameters used by Watson *et al.* and those used in our current study can explain this discrepancy. For example, Watson *et al.* used noise pixel sizes that were considerably smaller than those used here, which would make learning the local fixed noise pixels a more difficult task.

We found that, overall, performance was better for a Gabor target than for a checkerboard. Image discrimination models¹⁻⁷ predict this effect as a result of lower sensitivity to the diagonal frequency components of the checkerboard and imperfect summation over different channels. Neither of these features is yet available in template-matching models that have been used to predict target detectability in noise.⁸⁻¹⁰

Since accuracy of positional localization in peripheral vision is poor,¹³ we included conditions with the checkerboard and Gabor targets presented 4 deg parafoveally.

We hypothesized that increased positional uncertainty with eccentricity would decrease the fixed/twin difference. In one condition, the target size was the same as in the fovea (unmagnified). In the other, the target was magnified by a factor of 3 to account for the progressive reduction in cortical area with increasing eccentricity.¹⁶ Confirming our hypothesis, the fixed/twin difference was larger in the fovea.

In the unmagnified 4-deg-eccentricity condition, detection thresholds in noise increased with eccentricity as they have been shown to do in the presence of a pattern mask,¹⁴ whereas, for the Gabor stimulus in noise, thresholds were significantly reduced with eccentricity. We can only speculate that the Gabor stimulus, whose form is based only on low-frequency variations, shows more release from masking by high frequency components.¹⁹

In the magnified eccentricity condition, the checkerboard stimuli again behaved in an expected fashion. The contrast (not contrast energy) thresholds in the periphery were similar to those in the fovea, supporting the use of the scaling factor obtained for detection without noise.¹⁶ This scaling factor was not required for the Gabor stimulus because performance improved without stimulus magnification.

We hypothesize that the fixed/twin effect is the result of template learning. The initial state of the memory templates should depend on the experimenter's description of the experiment, the initial stimulus presentation, prior experience, and other factors and is not addressed here. Our data do provide information about memory template updating. In some cases fixed noise thresholds were substantially below twin noise thresholds on the first trial block, suggesting that template updating can be rapid. Rapid template adjustment may explain the fast perceptual learning described elsewhere.^{20,21} Some fixed noise conditions also show gradual improvement over the course of five to six practice blocks, which may reflect template refinement.

Simulated Detection Model with Learning and Position Uncertainty

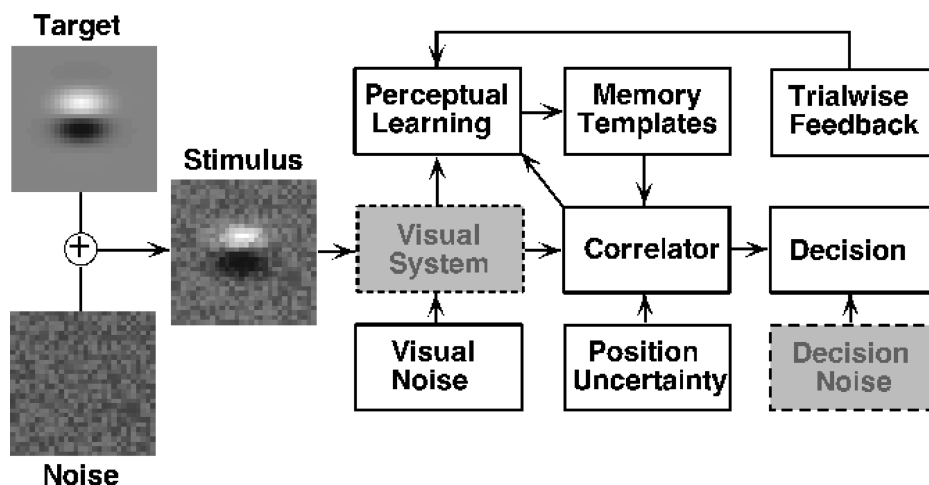


Fig. 4. Schematic of an image detection model with template learning and positional uncertainty. An input image plus added noise enters the visual system, where an internal sensory representation is formed of the stimulus. This noisy sensory representation is correlated with memory templates of the targets over a range of positions, and a decision is made as to which target was present. Based on the trialwise feedback, the template is updated. The visual system module and the decision noise module are in dashed boxes because they are not included in our current models.

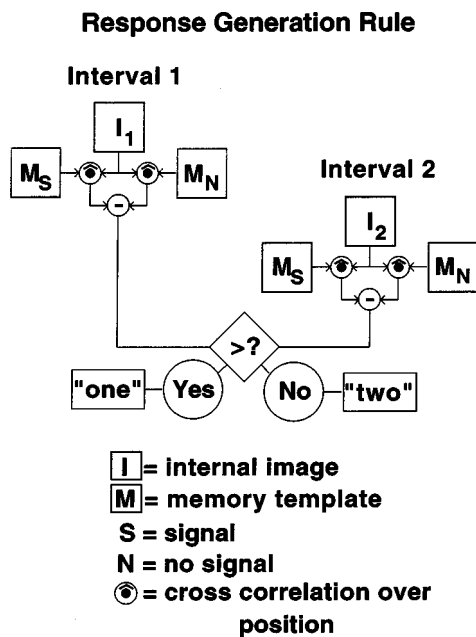


Fig. 5. Template learning model generates a response indicating which input image contains the target. For each image the model forms an internal sensory representation (I) composed of the image plus internal sensory noise. This representation is cross correlated with memory templates for the target (M_S) and no target (M_N). The circled dot symbol indicates the maximum of correlations over a range of positions. The difference of the two maximum correlations represents the signal-presence likelihood for each interval, and a comparison of these likelihoods determines the response.

Data showing positive and negative transfer of training effects are shown in Fig. 3. The data of observer NR show low thresholds for the second noise sample, which appear to be a continuation of the learning curve or a plausible extension of it from the first noise sample. The transfer of improvement suggests that some of the learned relevant features of the target-plus-noise sample are valid in both situations. In other words, the best-matched template to the new sensory representation is similar to the just-refined template and therefore requires little further updating. In other cases there seems to be negative transfer of improvement,²²⁻²⁴ where thresholds are higher for a new noise sample. Negative transfer suggests that observers attend to previously relevant features after these features become irrelevant because the noise sample is changed. A final characteristic that can be taken from these data is that low thresholds observed for almost all fixed noise samples suggest that the observer is able to hold multiple templates for the same task.

A. Template Learning Model with Uncertainty: Response Generation Rule

Figure 4 presents a schematic of a simple image detection model with template learning and positional uncertainty. A model without position uncertainty is presented in Appendix A. Gray-shaded boxes represent functions that have yet to be included in the working mathematical model. An input image plus added noise enters the visual system. An internal sensory representation is formed of the stimulus, which includes internal noise

sources. This noisy sensory representation is correlated with memory templates, a decision is made as to which stimulus was present, feedback is given, and the templates are updated.

In our simplified modeling environment the two 4×4 pixel targets were a simple 2×2 checkerboard and (stretching the term) a Gabor stimulus consisting of a 2×4 light bar above a 2×4 dark bar. The targets were centered in a 6×6 array of white noise. On each simulated two-interval trial there are two input images; one contains the target stimulus plus noise, and the other contains noise alone. The internal sensory representation (I) is composed of the input image plus internal sensory noise (see Fig. 5). For each interval, the internal sensory representation is cross correlated with the 4×4 memory template of the target at all nine possible positions. The largest cross correlation is compared with the largest cross correlation for the no-target template, and the difference is saved for comparison with the same difference from the other interval. The larger of these differences is used to select the interval with the signal.

B. Template Learning Model with Uncertainty: Learning Rule

The memory templates are modified on the basis of feedback indicating the interval containing the signal (see Fig. 6). Each template is replaced by a weighted average of the old template and the internal sensory representation of the appropriate image positioned by the cross correlation. The relative weighting of the average is determined by λ , the learning rate parameter, which is closer to 1 if learning is rapid and closer to 0 if learning is slow. The initial templates in our model were ideal, the

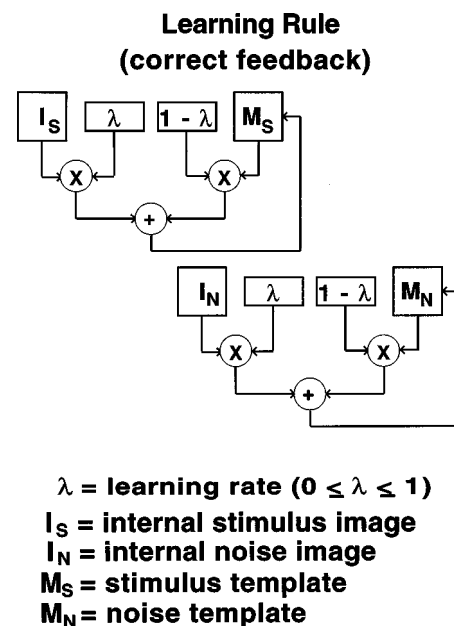


Fig. 6. Learning rule. Correct feedback assigns the internal images to the appropriate memory template, which is assumed to be translated to the best-correlating position. Each template is then replaced by a weighted average of the old template, and the associated internal sensory representation. If λ , the learning rate parameter (a number between 0 and 1), is large, the average contains mostly the current internal image; if it is small, the template is only slightly changed.

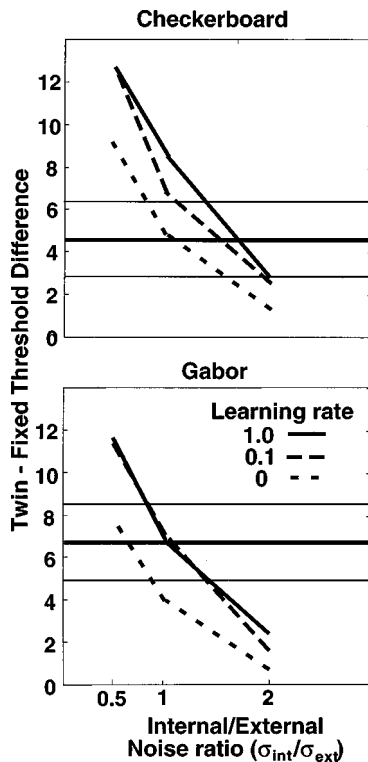


Fig. 7. Model simulations of the fixed/twin effect. Proportion correct scores from six 400-trial repetitions at four signal levels were converted by regression in the d' domain to 79% correct threshold estimates for each of 36 conditions [two noise types (fixed, twin); two target types (Gabor, checkerboard); three internal-to-external noise ratios (0.5, 1, 2); and three learning rates (0, 0.1, 1)]. The graphs show the amount that the twin threshold exceeded the fixed threshold (in dB) as a function of the other variables. The horizontal lines show the average foveal difference for our observers and the 95% confidence interval based on variability over observers. Parameters: $\lambda = 0.0$, 0.1, and internal-to-external noise ratio equals one best fit to the target results.

asymptotic templates being learned with arbitrarily small learning rates. For fixed conditions the initial templates included the external noise sample, while for twin conditions the initial templates were the target and a zero image. For fixed noise the fixed external component of the template noise helps lock in on the signal position. For twin noise, however, the external noise is changing on each trial and effectively adds random noise to the templates whenever the target and no-target templates correlate best at different positions.

To see the role of learning rate and the ratio of internal to external noise on the fixed/twin effect, we estimated the 79% correct threshold level for the two conditions. Each threshold was estimated from six repetitions of 400 trials at four signal levels. Figure 7 shows the difference (twin minus fixed) for the two target patterns, checkerboard and Gabor. The abscissa is the ratio of the internal noise standard deviation to the external noise standard deviation. The parameter is the learning rate. The horizontal lines indicate the average psychophysical values and the associated confidence interval from our data. The largest fixed/twin noise difference occurred for the smaller internal noise level, since the external noise carries the effect. Of the conditions we simulated, the

best matching was for $\lambda = 0.0, 0.1$ and an internal noise level of 1. The fair fit of the model with a learning rate of zero does not mean that learning is not needed to predict the result, since the initial templates were set to reflect learning of the fixed noise sample. In a separate simulation we also estimated the fixed/twin threshold difference with the learning rate set to zero and the initial templates set to the target and zero noise images for both conditions. In this case the sign of the difference actually reversed slightly.

Template-matching models,^{8,9} which correlate sensory representations of the images with one or more memory templates, can predict target detectability in random noise masking conditions, where a different noise sample is added to each input image. If the same template is used in fixed and twin noise conditions, the models do not, however, predict a fixed/twin noise masking difference. Image discrimination models cannot predict either a fixed/twin difference or random noise effects (since the difference calculation would include not only the target but also any changes in the noise). Our template learning model, presented in Section 4, demonstrates that combining template learning with positional uncertainty can explain the fixed/twin noise threshold difference. The model presented in Appendix A shows how template learning alone could explain the fixed/twin effect and how template learning contributes to the accelerating nonlinearity associated with uncertainty and transducer functions.^{25,26}

APPENDIX A: TEMPLATE LEARNING MODEL WITHOUT POSITIONAL UNCERTAINTY

This appendix describes another template learning model for two-alternative forced-choice detection experiments. The main difference between this model and the model presented in Section 4 is that here no position uncertainty is assumed in the model. This simplification allows closed-form expressions for the performance of the model and clarifies the role of the different parameters. If we simply remove position uncertainty from the model presented in Section 4, there is no fixed/twin difference because the same external noise sample is accumulated equally in each template, and they cancel in the response calculation. Here we assume that the observer remembers only the image from the second interval and updates the template that feedback associates with the second interval. The performance formulas show an accelerating nonlinearity of performance with signal level that is generated by the poor quality of the templates at low signal levels.

1. Response Generation Rules

For each stimulus presentation the observer is assumed to have an internal sensory representation vector, I , which is the sum of three vectors: (1) an internal representation of the signal (if present), (2) an internal representation of the external noise, and (3) internal noise:

$$I = S + N_{\text{ext}} + N_{\text{int}}. \quad (\text{A1})$$

The observer is assumed to have internal sensory representations of both stimuli. M_N represents the target-absent sensory representation; M_{SN} , the target-present sensory representation. For a two-alternative forced-choice experiment the observer is assumed to respond "interval 1" if

$$(M_{SN} - M_N) \cdot I_1 > (M_{SN} - M_N) \cdot I_2, \quad (\text{A2})$$

where the dot symbol \cdot indicates inner product. This expression can be rewritten as

$$(M_{SN} - M_N) \cdot (I_1 - I_2) > 0. \quad (\text{A3})$$

2. Template Learning Rule

The template adjustment process replaces the template with an average of itself and the internal sensory representation of the stimulus, by means of a learning rate parameter, λ :

$$M \leftarrow (1 - \lambda)M + \lambda I. \quad (\text{A4})$$

Ideally, λ should be a function of many things, such as the similarity of M and I , the feedback or reinforcement on the trial, and the experience of the observer.²⁷ Template learning processes must be able to proceed without feedback and likely depend on the stimuli in both intervals. It seems reasonable, however, that the strongest influence would be the most recent stimulus (i.e., interval 2) and that feedback does play a role. For simplicity and to accentuate the process that differentiates the fixed and twin noise conditions, we update the template according to the rule (assuming correct feedback)

$$\begin{aligned} M_N &\leftarrow (1 - \lambda)M_N \\ &+ \lambda I_2 \quad \text{if } I_2 \text{ is a target-absent interval,} \end{aligned} \quad (\text{A5})$$

$$\begin{aligned} M_{SN} &\leftarrow (1 - \lambda)M_{SN} \\ &+ \lambda I_2 \quad \text{if } I_2 \text{ is a target-present interval.} \end{aligned} \quad (\text{A6})$$

3. Asymptotic Behavior of the Model

The performance of the model is controlled by the quality of the templates. The updated template is a weighted average of the stimulus plus the internal noise on all the trials on which it was adjusted. If we let M_0 be the initial value of one of the templates and I_i be the image on the i th updated trial, then the template M_n at the end of the n th trial is given by

$$M_n = (1 - \lambda)^n M_0 + \lambda(1 - \lambda)^{(n-1)} I_1 + \dots + I_n \lambda. \quad (\text{A7})$$

Let F be the fixed part of the stimulus on the i th update of this template and R_i be the random part of the stimulus. Then, as n becomes large, the contribution from M_0 becomes negligible, and the template approaches

$$F + \sum R_{n-i} \lambda (1 - \lambda)^i, \quad (\text{A8})$$

where the summation index i ranges from 0 to n .

Since the R_i values are independent, with mean 0 and covariance matrix Σ_R , the asymptotic template distribution has a mean of F and a covariance matrix Σ_M of

$$\Sigma_M = \Sigma_R \lambda^2 / [1 - (1 - \lambda)^2] = \Sigma_R \lambda / (2 - \lambda) = \Sigma_R / n_\lambda. \quad (\text{A9})$$

The quantity $n_\lambda = (2 - \lambda)/\lambda$ can be regarded as the effective number of noise components averaged for a given λ . For $\lambda = 1$, it is 1; for $\lambda = 0.5$, it is 3.

4. Asymptotic Detection Performance of the Model

Let the difference of the templates be $D_M = M_{SN} - M_N$ and the difference of the stimulus representation be $D_S = I_1 - I_2$. Then the response rule of relation (A3) becomes, "respond 'interval 1'" if

$$D_M \cdot D_S > 0. \quad (\text{A10})$$

Because the randomness in D_M comes from past trials, D_M and D_S are independent random vectors. The mean of D_M is S . The mean of D_S is S when the signal is in interval 1 and $-S$ when it is in interval 2. The mean of $D_M \cdot D_S$ is $E = S \cdot S$ when the signal is in the first interval and $-E$ when it is in the second. The performance of the model can be characterized by

$$d = [E - (-E)] / \text{SD}[D_M \cdot D_S] = 2E / \text{SD}[D_M \cdot D_S], \quad (\text{A11})$$

where $\text{SD}[\]$ indicates the standard deviation. If the individual components of the representations are independent and have the same variance, the variance of $D_M \cdot D_S$ can be expressed as

$$\text{SD}[D_M \cdot D_S]^2 = E(\sigma_M^2 + \sigma_S^2) + n\sigma_M^2\sigma_S^2, \quad (\text{A12})$$

where the σ 's are the standard deviations of the individual components and n is their number.

For the twin and the fixed noise cases, the signal difference pixel variance is given by

$$\sigma_S^2 = 2\sigma_{\text{int}}^2. \quad (\text{A13})$$

For the fixed noise case, the template difference pixel variance is given by

$$\sigma_M^2 = 2(\sigma_{\text{int}}^2) / n_\lambda, \quad (\text{A14})$$

while for the twin noise case it is

$$\sigma_M^2 = 2(\sigma_{\text{int}}^2 + \sigma_{\text{ext}}^2) / n_\lambda. \quad (\text{A15})$$

Figure 8 plots the performance of the twin and the fixed models as a function of d' for the ideal observer, limited only by the internal noise:

$$d_{\text{Ideal}} = \text{Sqrt}[2E/\sigma_{\text{int}}^2]. \quad (\text{A16})$$

As the learning parameter λ approaches zero, n_λ becomes infinite, and the performance for both conditions becomes ideal (although it would take infinite time to reach this level). The higher curve below the unity slope line shows predicted model performance in the fixed noise condition with the parameters $n = 16$ and $\lambda = 0.5$. The lower curve shows the twin noise prediction for those same parameters and $\sigma_{\text{ext}} = 2\sigma_{\text{int}}$.

Figure 8 illustrates another feature of the template learning model. In addition to predicting a fixed/twin threshold difference, it also quantitatively predicts an accelerating nonlinearity of performance with signal level generated by the poor quality of the templates at low signal levels.²⁸ Tanner inferred this effect when he failed to find a close correspondence between observer behavior

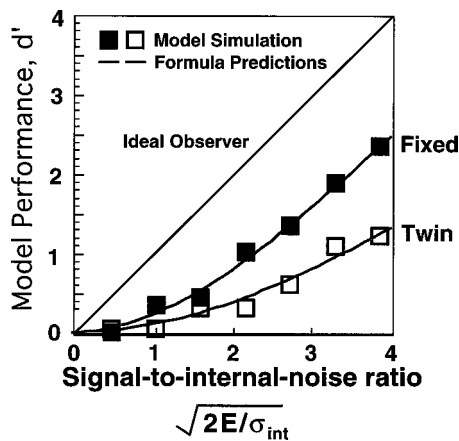


Fig. 8. Mathematical formula predictions and simulation results for the template learning model with no positional uncertainty. Performance in twin and fixed noise conditions is plotted as a function of the signal-to-noise level. The ideal observer with internal noise would perform along the main diagonal. The simulation proportions of correct responses were transformed to d' by means of the cumulative Gaussian distribution, which seems to fit well. Note that the model predicts an accelerating nonlinearity in the absence of uncertainty or a transducer function.

and a stimulus uncertainty model as signal strength increased.²⁸ Template learning now joins signal uncertainty²⁵ and nonlinear²⁶ transduction as a quantified process for explaining nonlinear performance for low signal-to-noise ratios.

ACKNOWLEDGMENTS

This work was supported in part by NASA Space grant 199-06-39, NASA Aeronautics program 505-64-53, and NASA cooperative agreement NCC2-327 with the San Jose State University Foundation. Portions of these data were presented at the Optical Society of America Annual Meeting, Long Beach, California, October 12–17, 1997; the Association for Research in Vision and Ophthalmology Annual Meeting, Ft. Lauderdale, Florida, May 10–15, 1998 and the Optical Society of America Annual Meeting, Baltimore, Maryland, October 4–9, 1998.

REFERENCES

1. A. B. Watson, "Efficiency of a model human image code," *J. Opt. Soc. Am. A* **4**, 2401–2417 (1987).
2. S. Daly, "The visible difference predictor: an algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*, A. B. Watson, ed. (MIT, Cambridge, Mass., 1993).
3. J. Lubin, "The use of psychophysical data and models in the analysis of display system performance," in *Digital Images and Human Vision*, A. B. Watson, ed. (MIT, Cambridge, Mass., 1993).
4. P. C. Teo and D. J. Heeger, "Perceptual image distortion," in *Human Vision, Visual Processing, and Digital Display*, B. Rogowitz and J. Allebach, eds., Proc. SPIE **2179**, 127–141 (1994).
5. H. R. Wilson, "Quantitative models for pattern detection and discrimination," in *Vision Models for Target Detection and Recognition*, E. Peli, ed. (World Scientific, Teaneck, N.J., 1995).
6. A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," *J. Opt. Soc. Am. A* **14**, 2379–2391 (1997).
7. A. J. Ahumada and B. L. Beard, "Image discrimination models predict detection in fixed but not random noise," *J. Opt. Soc. Am. A* **14**, 2471–2476 (1997).
8. H. Barrett, "Evaluation of image quality through linear discriminate models," *Proc. Soc. Inf. Disp.* **23**, 871–873 (1992).
9. A. Burgess, "Statistically defined backgrounds: performance of a modified nonprewhitening observer," *J. Opt. Soc. Am. A* **11**, 1237–1242 (1994).
10. M. Eckstein, A. B. Watson, and A. J. Ahumada, "Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise," *J. Opt. Soc. Am. A* **14**, 2406–2419 (1997).
11. A. M. Rohaly, A. J. Ahumada, Jr., and A. B. Watson, "Object detection in natural backgrounds predicted by discrimination performance and models," *Vision Res.* **37**, 3225–3235 (1997).
12. A. B. Watson, R. Borthwick, and M. Taylor, "Image quality and entropy masking," in *Human Vision, Visual Processing, and Digital Display*, B. Rogowitz, ed., Proc. SPIE **3016**, 2–12 (1997).
13. D. M. Levi and S. A. Klein, "Limitations on position coding imposed by undersampling and univariance," *Vision Res.* **36**, 2111–2120 (1996).
14. G. E. Legge and D. Kersten, "Contrast discrimination in peripheral vision," *J. Opt. Soc. Am. A* **4**, 1594–1597 (1987).
15. A. B. Watson, "Estimation of local spatial scale," *J. Opt. Soc. Am. A* **4**, 1579–1582 (1987).
16. J. Rovamo and V. Virsu, "An estimation and application of the human cortical magnification factor," *Exp. Brain Res.* **37**, 495–510 (1979).
17. D. J. Finney, *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve*, Cambridge U. Press (Cambridge, UK, 1947).
18. A. B. Watson, H. B. Barlow, and J. G. Robson, "What does the eye see best?" *Nature (London)* **302**, 419–422 (1983).
19. L. D. Harmon and B. Julesz, "Masking in visual recognition: effects of two-dimensional filtered noise," *Science* **180**, 1194–1197 (1973).
20. M. Fahle, S. Edelman, and T. Poggio, "Fast perceptual learning in hyperacuity," *Vision Res.* **35**, 3003–3013 (1995).
21. B. L. Beard, D. M. Levi, and L. N. Reich, "Perceptual learning in parafoveal vision," *Vision Res.* **35**, 1679–1690 (1995).
22. C. E. Osgood, "The similarity paradox in human learning: a resolution," *Psychol. Rev.* **56**, 132–143 (1949).
23. M. Fahle, "Human pattern recognition: parallel processing and perceptual learning," *Perception* **23**, 411–427 (1994).
24. B. L. Beard and A. J. Ahumada, "Tuning function changes after practice on a parafoveal vernier acuity task," invited presentation at the Special Symposium on Hyperacuity at the European Conference on Visual Perception, Teubingen, Germany, September 9–13, 1996.
25. D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," *J. Opt. Soc. Am. A* **2**, 1508–1532 (1985).
26. G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Am. A* **70**, 1458–1471 (1980).
27. C. V. Jakowatz, R. L. Shuey, and G. M. White, "Adaptive waveform recognition," presented at the Symposium on Information Theory, Royal Institute, London, August 29–September 2, 1961.
28. W. P. Tanner, "Physiological implications of psychophysical data," *Ann. (N.Y.) Acad. Sci.* **89**, 752–765 (1961).