Statistics of boundary, luminance, and pattern information predict occluding target detection in natural backgrounds

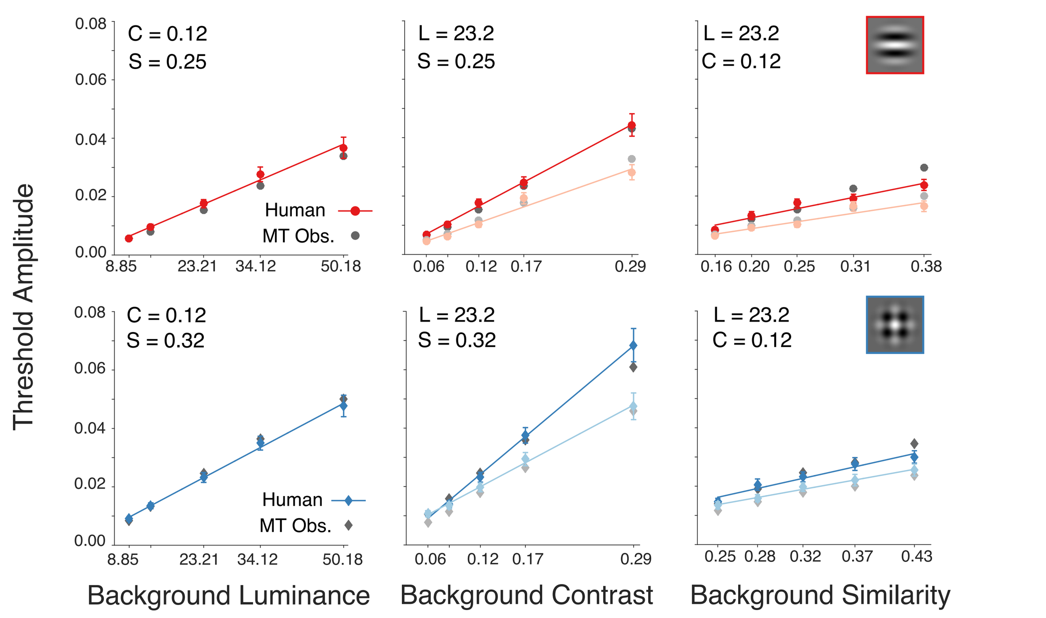
R. Calen Walshe, Stephen Sebastian & Wilson Geisler

Center for Perceptual Systems, The University of Texas at Austin

TEXAS
The University of Texas at Austin

Background and Motivation

1. Masking laws are well characterized for additive targets.
2. Luminance, contrast and similarity identified as fundamental stimulus dimensions.
3. Well developed ideal observer models for additive targets in artificial and natural backgrounds.
4. Currently very little known about occluding target detection.



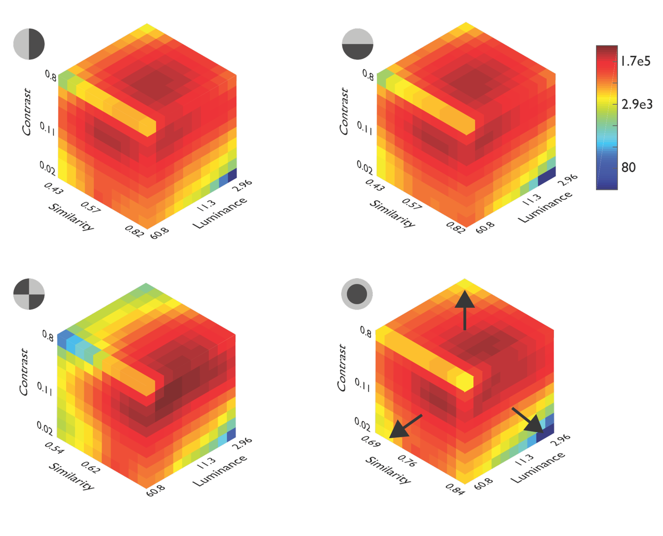
Goals

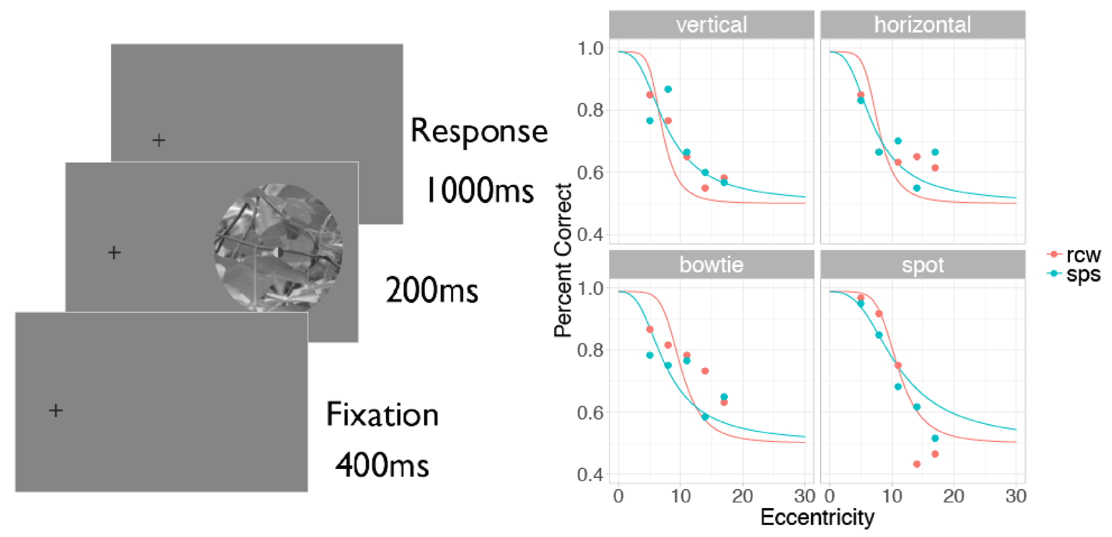
1. Measure masking laws for human detection of occluding targets.
2. Develop an ideal observer and compare with human thresholds.

Masking Experiment

Constrained Scene Sampling

1. 1200 images of the Austin area.
2. Extract millions of 21 pixel patches.
3. Measure luminance, contrast and similarity of each patch. Place in bins.
4. Select a bin and measure performance across the visual field.





Approximately Ideal Observer

Stimulus Encoding

Ganglion Cell Sampling

The retinal stimulus (I_R) is filtered by midget retinal ganglion cell array.

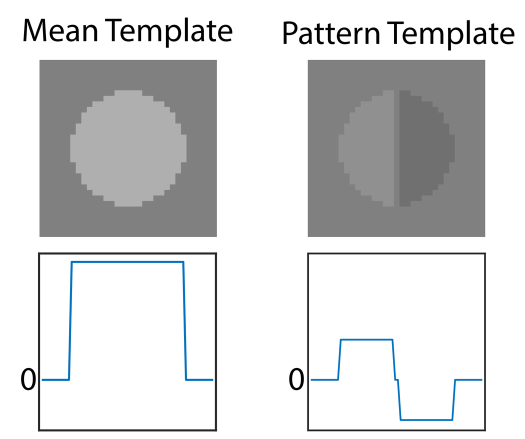
$$I_R(\mathbf{x}) = sample_e\left[I(\mathbf{x}) * g(\mathbf{x}) * f_e(\mathbf{x})\right],$$

$I(x)$ is the monitor stimulus.
$g(x)$ is the optical point spread function of the eye (4mm pupil).
$f(x)$ is a Gaussian kernel with σ_e matched to the average radius of midget receptive fields at eccentricity e :

Target Template

Target template is the sum of the mean and pattern target signal.

$$\mathbf{T} = \mathbf{T}_m + \mathbf{T}_p$$



Note. Target is blurred and downsampled to match the eccentricity condition.

Apply Pattern Template

$$R_p = \mathbf{T}_p \cdot \mathbf{I}_R$$

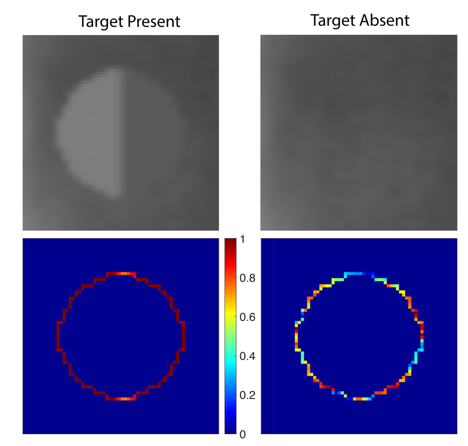
Apply Mean Template

$$R_m = \mathbf{T}_m \cdot \mathbf{I}_R$$

Edge Response to stimulus

$$R_e = \sum_{x \in boundary} \left| \frac{\nabla I(x)}{\nabla ||I(x)||} \cdot \frac{N(x)}{||N(x)||} \right|$$

$N(x)$ is the boundary normal vector.



Noisy stimulus encoding

The ideal stimulus responses are degraded with Gaussian noise.

$$\begin{aligned} R_{p'} &= R_p + \mathcal{N}\left(0, k\left(e;\theta\right) R_p\right) \\ R_{l'} &= R_l + \mathcal{N}\left(0, k\left(e;\theta\right) R_m\right) \\ R_{e'} &= R_p + \mathcal{N}\left(0, k\left(e;\theta\right) R_e\right) \end{aligned}$$

The noise is dependent on eccentricity.

Optimal Response Decoding

Measure the mean and covariance between cues for all stimulus conditions including present/absent.

Respond with stimulus category that is most likely given the observed responses.

Maximum Likelihood

1. Measure mean and covariance matrix for edge, luminance and pattern responses in all experimental conditions:

$$\mu = \left(\mu_e, \mu_l, \mu_p\right)$$

$$\mathbf{\Sigma} = \begin{pmatrix} \mathrm{Var}(R_{e'}) & \mathrm{cov}\left(R_{e'}, R_{l'}\right) & \mathrm{cov}\left(R_{e'}, R_{p'}\right) \\ \mathrm{cov}\left(R_{e'}, R_{l'}\right) & \mathrm{Var}(R_{l'}) & \mathrm{cov}\left(R_{l'}, R_{p'}\right) \\ \mathrm{cov}\left(R_{e'}, R_{p'}\right) & \mathrm{cov}\left(R_{l'}, R_{p'}\right) & \mathrm{Var}(R_{p'}) \end{pmatrix}$$

2. Minimum error rate classification rule:

Multivariate normal likelihood function.

$$X = ln\, \frac{f\left(\mathbf{R}|\mu_{present}; \mathbf{\Sigma}_{present}\right)}{f\left(\mathbf{R}|\mu_{absent}; \mathbf{\Sigma}_{absent}\right)}$$

If $X \geq 0$ then respond present else respond absent.

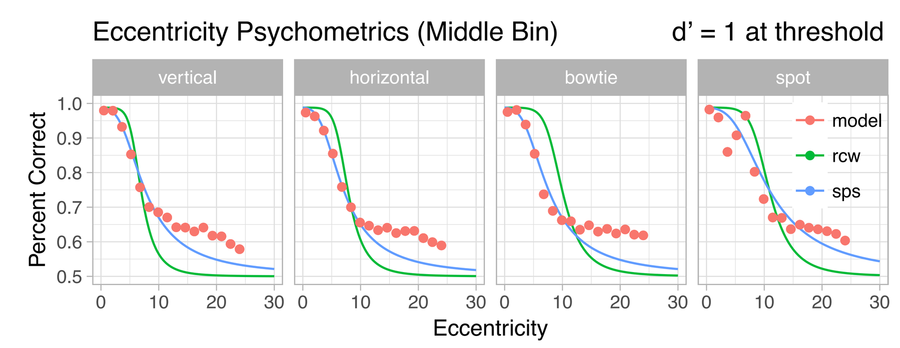
Model Fitting

Select $\hat{\theta}$ that maximizes the likelihood of the ideal observer given the measured human data.

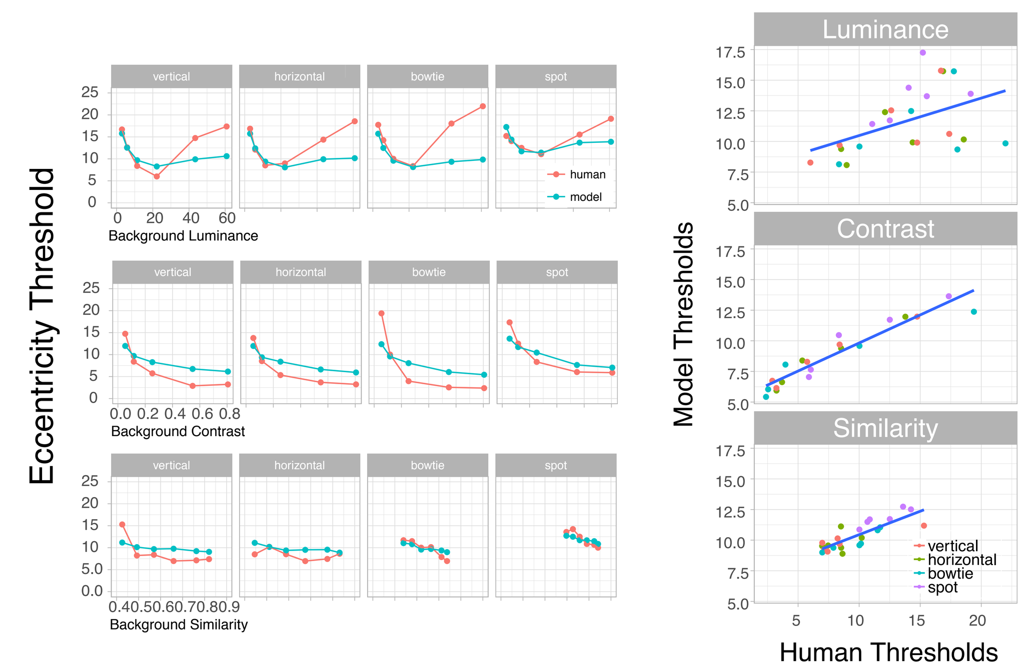
$$\hat{\theta} = \underset{\theta}{\arg\max}\, \hat{\ell}(\theta\, ; x_1, \ldots, x_n) \qquad (1)$$

Results

Fits of the model to human psychometric functions for the middle L,C,S bin:



Eccentricity Threshold Functions:



The ideal observer does a reasonable job to tracking human thresholds.
Early sensory limitations and the statistics of natural scenes can partially explain occluding target detection.