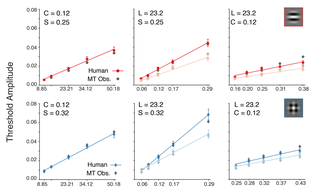
Ideal observer for detection of occluding targets in natural scenes in the fovea and periphery.

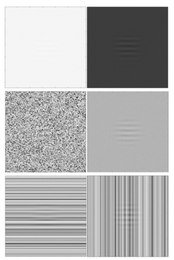
R. Calen Walshe, Stephen Sebastian & Wilson Geisler
Center for Perceptual Systems, The University of Texas at Austin

TEXAS
The University of Texas at Austin

Background and Motivation



Sebastian, Abrams & Geisler (2017)



1. Luminance, contrast and pattern similarity of a background are basic dimensions that mask target detectability.

2. A template matching observer adapted to the statistics of natural scenes predicts foveal thresholds for additive targets very accurately.

Key Questions:

1. When targets occlude the background how does luminance, contrast and pattern similarity impact detectability of occluding targets in natural scenes?

2. Can an extended template matching observer predict detection of occluding targets?

Experiment: Occluding Targets

Statistics of Natural Scenes

$B(x,y)$ Image patch

$env(x,y)$ Target envelope $\sum env(x,y) = 1$

Luminance (%max monitor luminance)

$$L(x,y) = env(x,y) \cdot B(x,y)$$

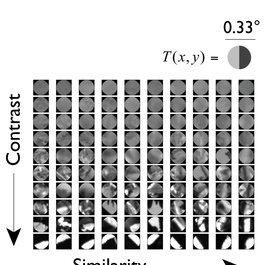
Contrast (RMS contrast)

$$C(x,y) = \left\| \frac{B(x,y) - L(x,y)}{L(x,y)} \sqrt{env(x,y)} \right\|$$

Similarity

Frequency and orientation similarity

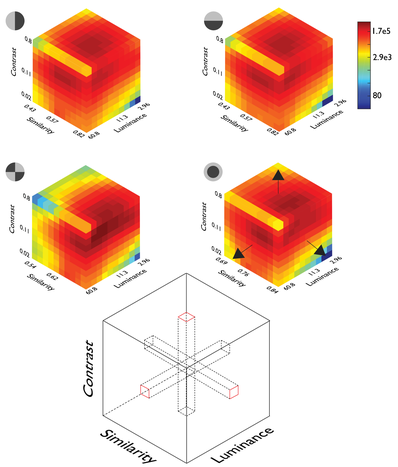
$$S(x,y) = \frac{A_r(u,v) \cdot A_p(u,v)}{\|A_r\| \|A_p\|}$$



Constrained Sampling

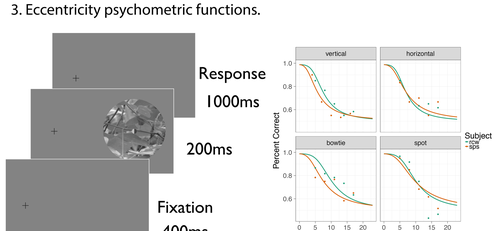
1. Sort patches into discrete bins along luminance, contrast and similarity dimensions.

2. Sample experimental stimuli from the cardinal axes.



Detection Task

1. Select a bin and measure performance across the visual field.
2. Target and background level are fixed. Vary eccentricity.
3. Eccentricity psychometric functions.



Extended Template Matching Observer

1. Luminance edge and pattern information is available for detection of occluding targets.
2. Ganglion cell sampling decreases and receptive field sizes increase as a function of eccentricity
3. An ideal observer should take into account the variability arising from retinal factors and the statistical structure of natural scenes.

Ganglion Cell Encoding

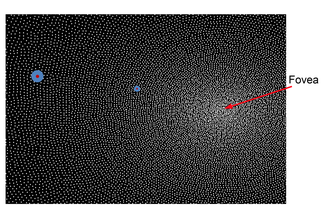
The retinal stimulus I_R is filtered by midget retinal ganglion cell array.

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

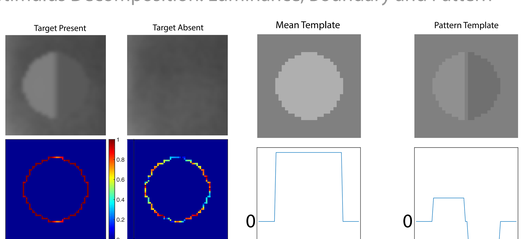
$I(x)$ is the monitor stimulus.

$g(x)$ 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



Stimulus Decomposition: Luminance, Boundary and Pattern



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

Responses to Retinal Images of Natural Scenes

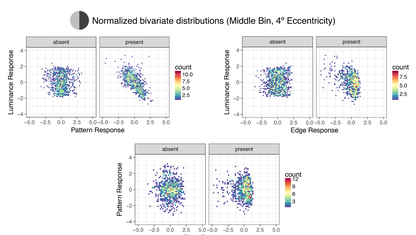
$$R_p = \mathbf{T}_p \cdot \mathbf{I}_R$$

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

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

Gaussian Model

Measure means and covariances of pattern, luminance and edge cues for all background conditions and eccentricities spanning the range of experimental conditions.



1. Minimum Error Rate Classification Rule:

Multivariate normal likelihood function.

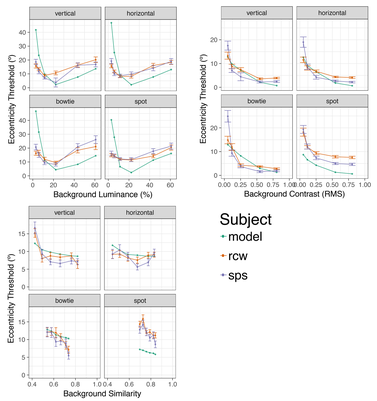
$$X = ln \frac{f(\mathbf{R}|\mu_{present}; \mathbf{\Sigma}_{present})}{f(\mathbf{R}|\mu_{absent}; \mathbf{\Sigma}_{absent})}$$

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

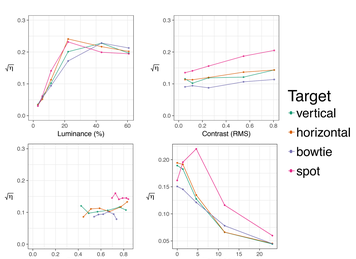
2. Measure Error Rate and Scale by Efficiency

$$\hat{d'}_{ij}^{\,human} = \sqrt{\eta}\, d'_{ij}{}^{\,ideal}$$

Thresholds

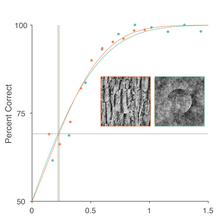


Efficiency



Conclusions and Future Directions

Natural scene statistics and retinal factors alone are a good approximation for human detection of occluding targets in natural scenes.



1. Incorporate biological factors.
   - Feature Normalization
   - Neural Noise
   - Uncertainty
2. Detection of occlusion under conditions of camoflauge.