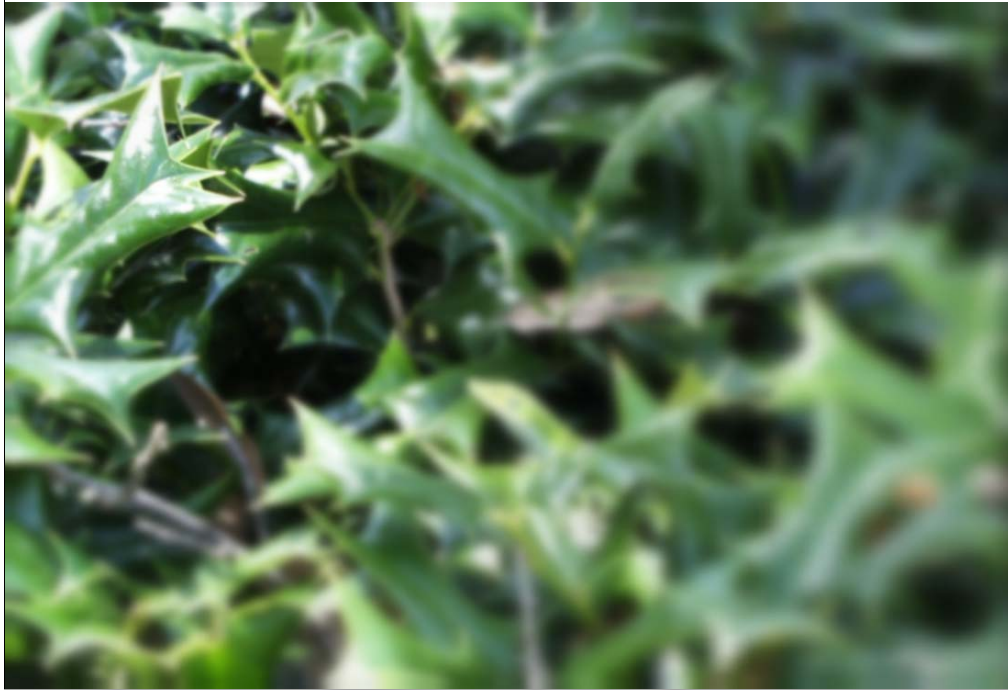


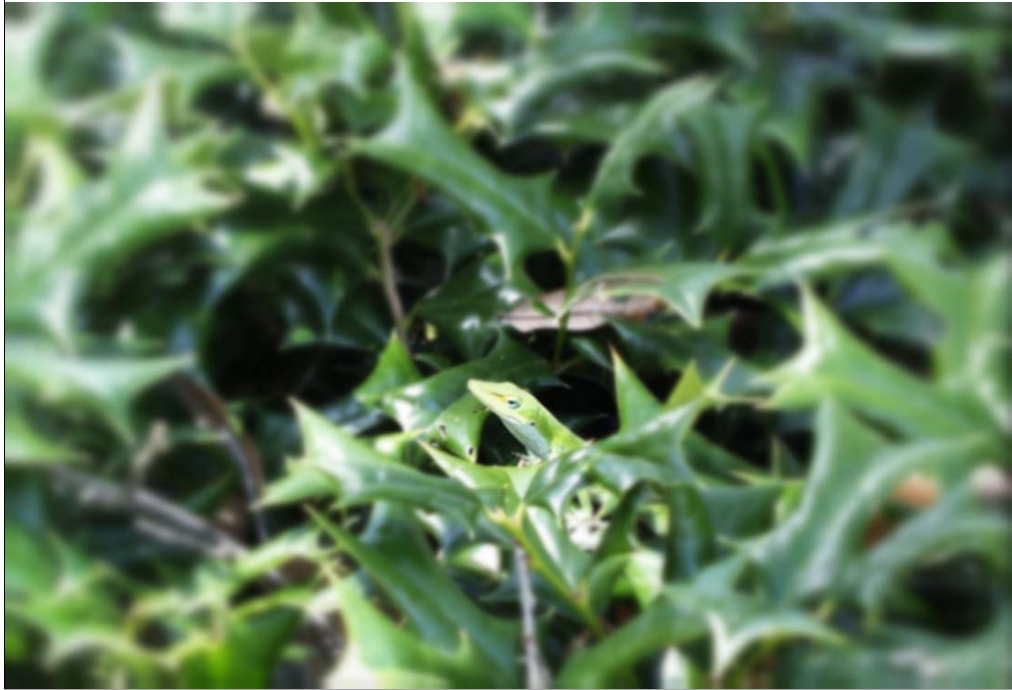
Humans (like many other mammals) encode a large field of view with a variable resolution visual system and then use eye, head and body movements to direct the high resolution region (the fovea) towards potentially relevant locations. Thus, a very fundamental family of attention mechanisms are those that control the selection of eye fixations. There are many natural tasks where selection of eye fixations plays a critical role. Today I want to talk about the visual search task.

Fixation Search



A typical visual search task in Texas is looking for the green anole.

Fixation Search



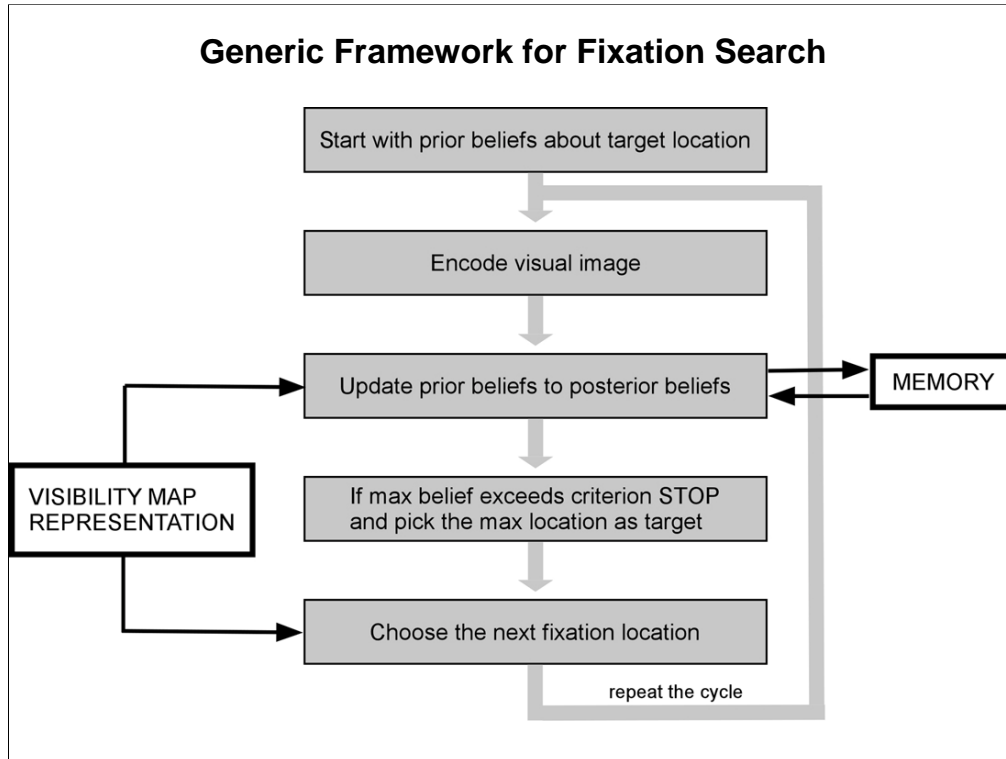
Questions About Fixation Selection in Search

How should the eyes be moved when searching with a foveated visual system?

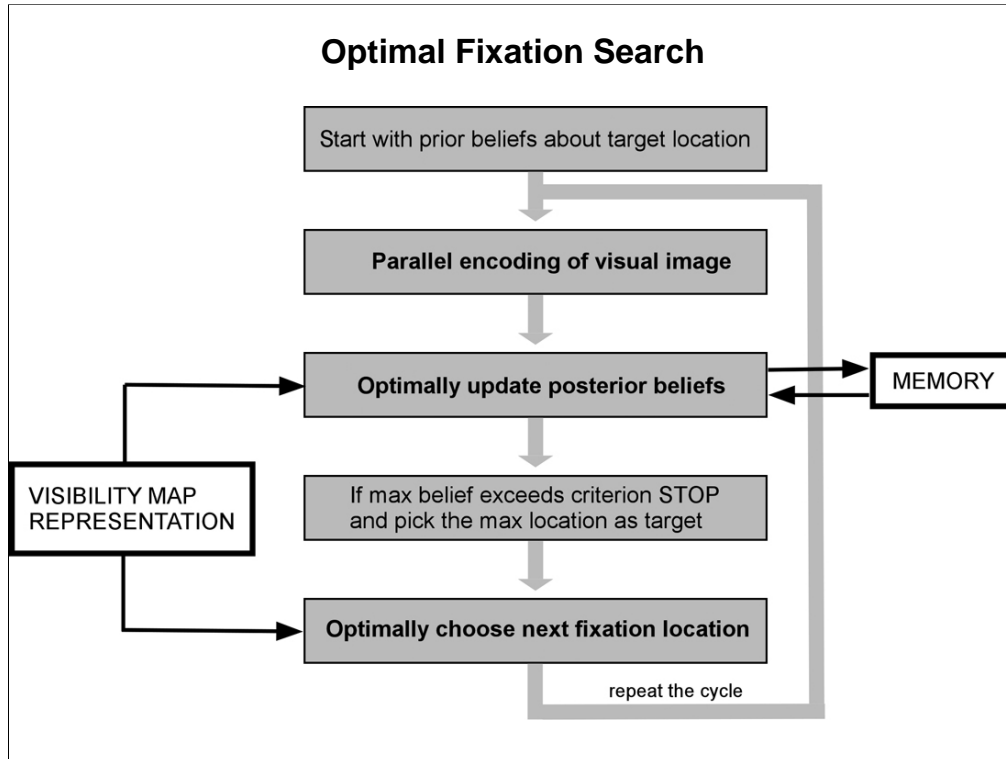
Do humans use rational fixation selection in performing visual search?

If not, what strategies do humans use?

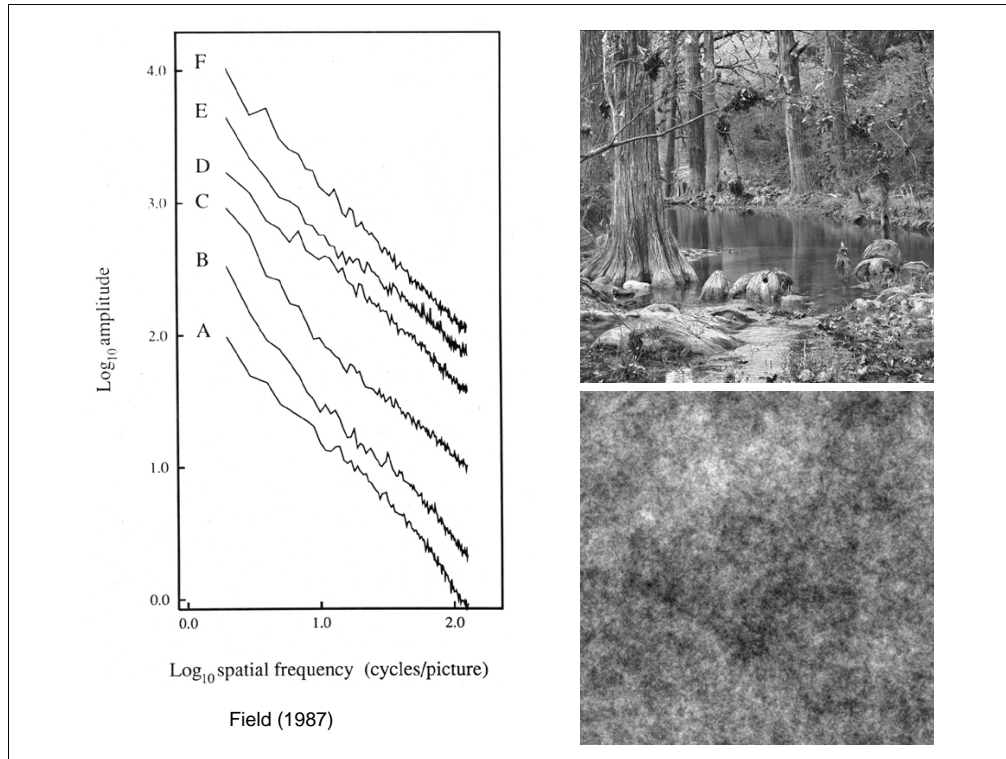
To address these questions we began by deriving and evaluating the ideal Bayesian searcher, and a number of sub-optimal searchers.



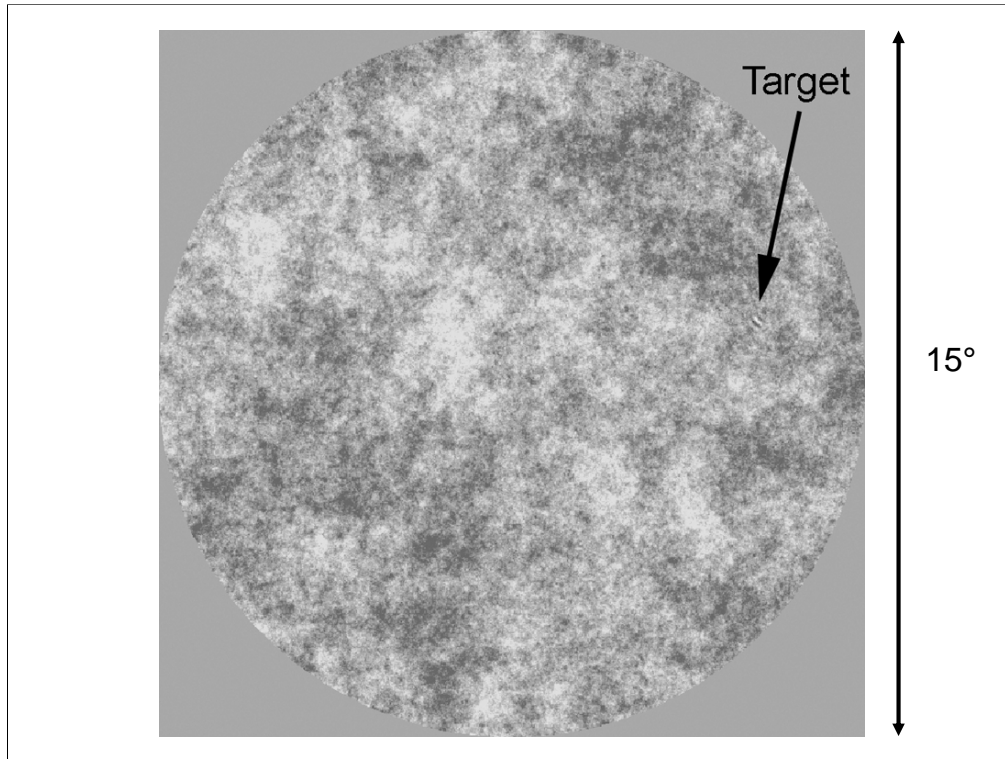
To begin with it is useful to think about what sort of processing is involved in a fixation search task. Here is a generic framework that most, if not all, models of fixation search would fall within. There are many questions about this entire process that need to be addressed. Recently, we have been focusing on how the next fixation is selected.



To make all this concrete we considered a specific search task, where a target is randomly located in a background of naturalistic texture.



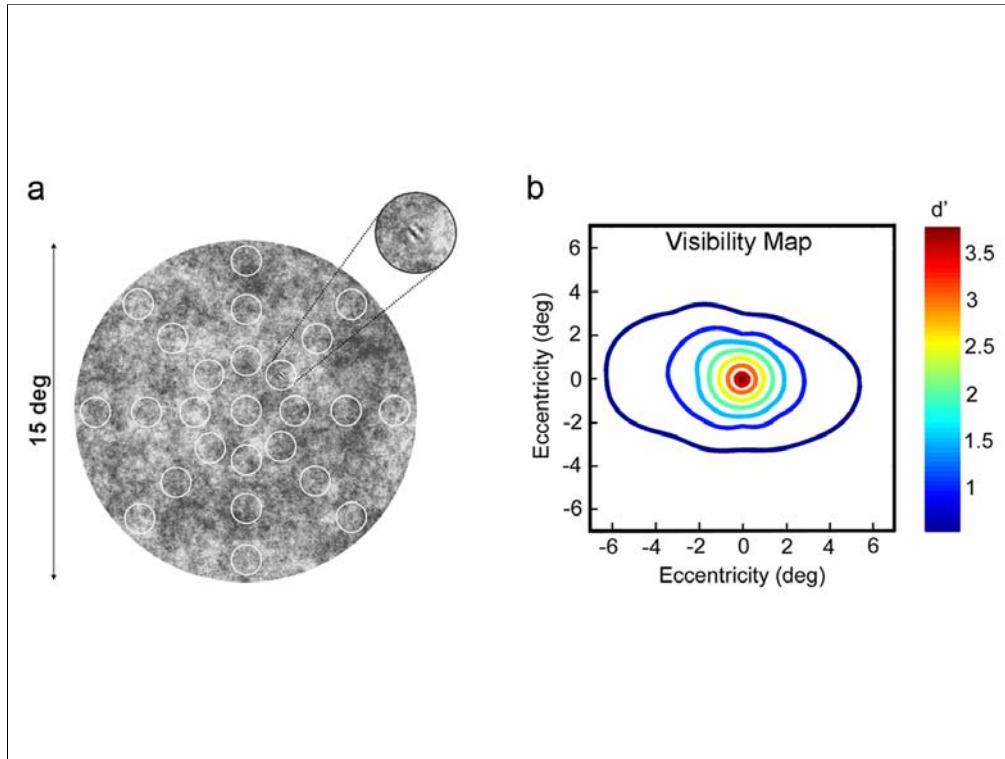
Amplitude spectra of natural images. The two images on the right have the same amplitude spectra.



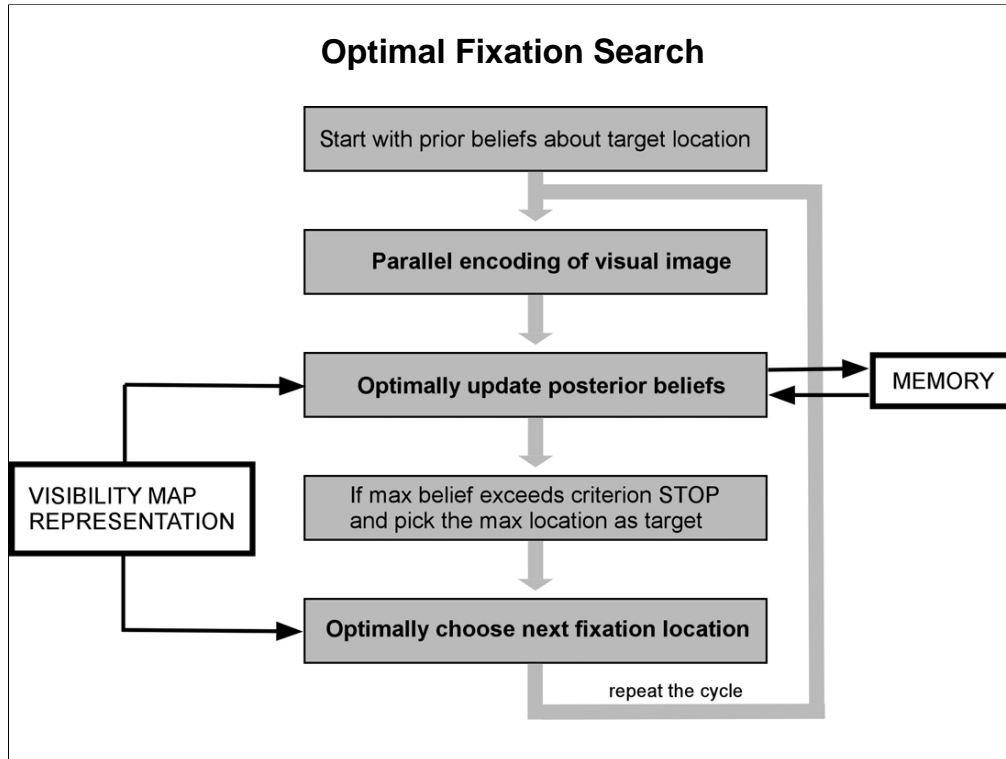
The task is to find the location of the target object as rapidly as possible without making mistakes. Presumably this kind of search is similar to what goes on in many natural search tasks (finding a bird in foliage, a plane in the sky, a life raft in the ocean, an object dropped in grass, or a green anole in a bush).

The specific task was as follows: (1) Fixate a dot in the center of a screen at mean luminance, (2) Press a button when ready, (3) A random time later the search display appears with the target at some randomly selected location, (4) As soon as the subject locates the target he/she presses a button, (5) Then subject fixates the assumed target location and presses the button again. The trial is counted as correct if the second fixation is within a small distance of the target.

We derived the performance of the ideal Bayesian searcher (and various suboptimal searchers) for this task and compared performance to human performance in the same task.



To evaluate the model searchers and compare them to human performance it is necessary to specify or measure the visibility maps for the targets and backgrounds of interest. We measured visibility (detectability) for the 6 cpd target as a function of target contrast and noise background contrast, at the 25 locations indicated by the small circles, in a two interval 2AFC detection task, while monitoring eye position. The stimuli were presented for 250 ms intervals which correspond to the approximate duration of fixations in the search task. From the measurements of detectability for a wide range of eccentricities, target contrasts, and noise contrasts, we can determine maps of detectability (d' maps) for any target contrast and noise contrast. Once we have these maps we can derive the ideal searcher.



The visibility maps describe the parallel encoding of the visual image. How should the posterior beliefs be updated?

**Posterior Probability of Target at Location i after T fixations
(optimal posterior updating)**

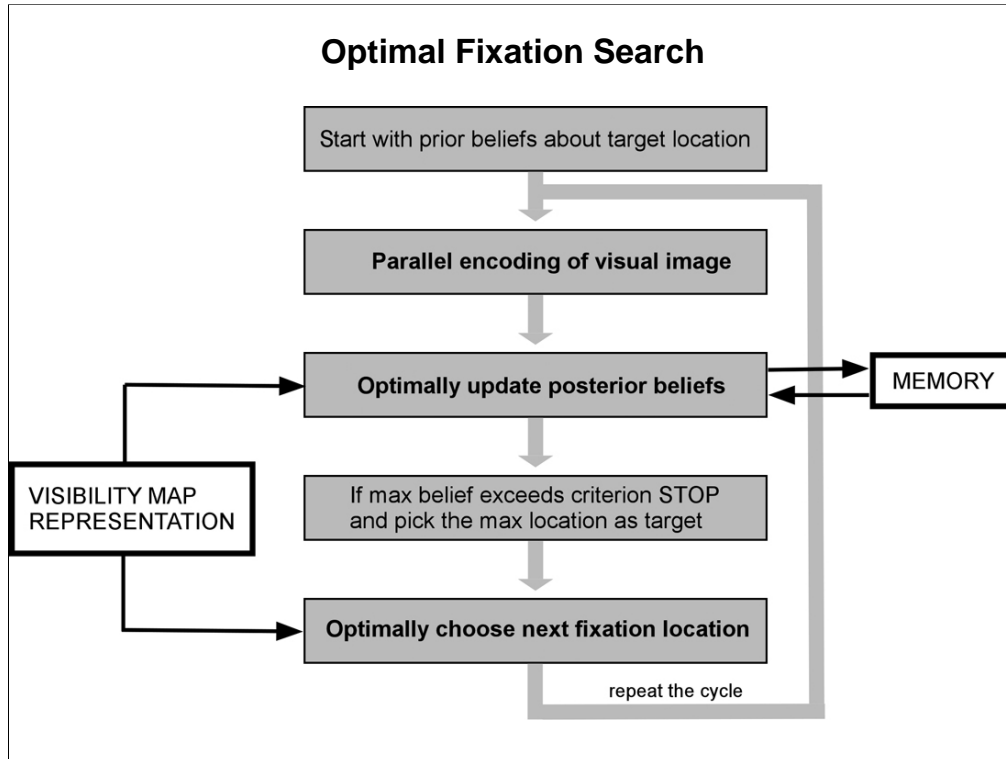
$$p_i(T) = \frac{\text{prior}(i) \exp\left(\sum_{t=1}^T d'_{ik(t)} W_{ik(t)}\right)}{\sum_{j=1}^n \text{prior}(j) \exp\left(\sum_{t=1}^T d'_{jk(t)} W_{jk(t)}\right)}$$

$W_{ik(t)}$ = response at time t when fixation is at $k(t)$

$d'_{ik(t)}$ = visibility at time t when fixation is at $k(t)$

This formula shows how the posterior probabilities at an arbitrary display location are updated during the course of a visual search trial. This formula represents perfect parallel processing, perfect integration of information across fixations, and no forgetting.

The W s are normalized responses from each potential target location (i or j) when the fixation is at display location k . The d' values are the visibilities at each of those locations (which are given by the visibility maps and known to the ideal searcher). The sum in the exponent is over all the fixations since display onset.



How should the next fixation be selected and how do humans select fixations? This is the main question I will consider today.

Fixation Selection Strategies of Model Searchers

Random (Ran) Search:

Fixate by random selection from prior pdf

Tiling (Tile) Search:

Random fixation with optimal inhibition of return

Feature-Based (MAP) Search:

Fixate locations with features most similar to target

Probability Matching (PM) Search:

Fixate by random selection from posterior pdf

Information-Based (Ideal) Search:

Fixate locations most likely to provide the best information about the location of the target

Here are five general classes of fixation search strategies that we have been considering; most are representative of strategies considered in the fixation search literature. Moving from top to bottom the searchers get more sophisticated. All of these are parameter free models.

Properties Of Model Searchers					
	Ran*	Tile	MAP*	PM	Ideal*
Update Posteriors	X	X	X	X	X
Inhibit Return		X	X	X	X
Features			X	X	X
Configuration					X
* will show formulas and results today					

Here are the factors the different searchers take into account. All of them optimally update the posteriors. All but the Ran searcher employ optimal inhibition of return. The MAP and PM also guide fixations based the similarity of the image features to the target. Only the ideal observer takes into account configuration information. I will show you formulas and results for the searchers marked with an asterisk.

Random Fixation Selection
(alias Ran searcher)

$$P[k(T+1)=i] = \textit{prior}(i)$$

Feature-Based Fixation Selection
(alias MAP searcher)

$$k(T+1) = \arg \max_i (p_i(T))$$

Information-Based Fixation Selection
(alias Ideal searcher)

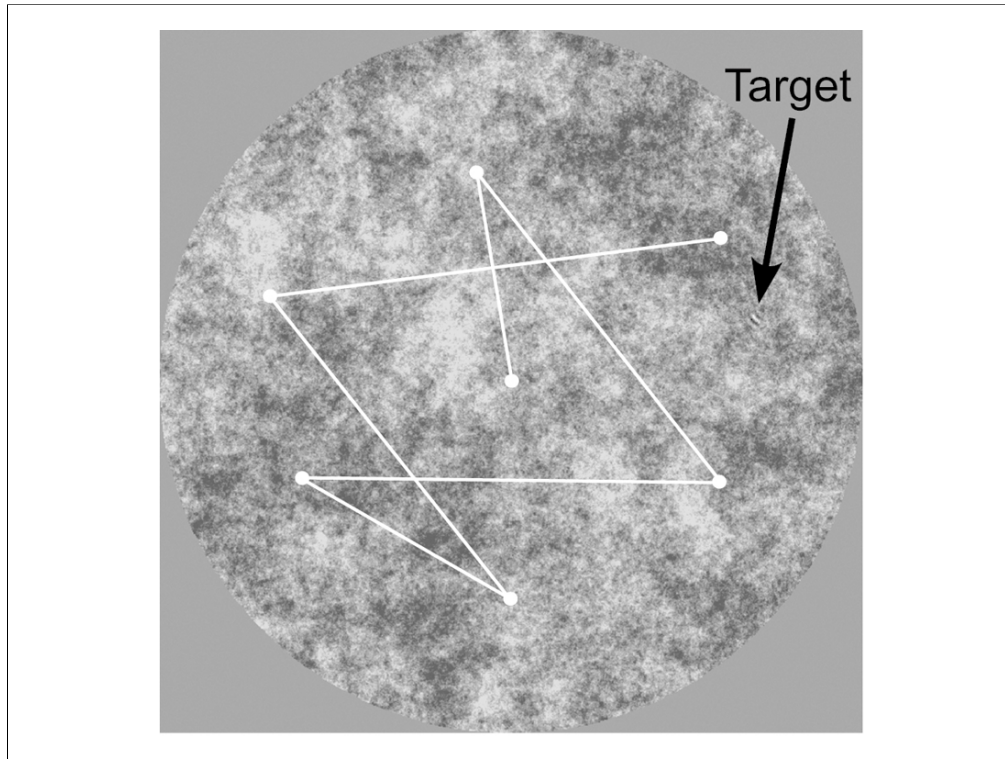
$$k_{opt}(T+1) = \arg \max_{k(T+1)} \left(\sum_{i=1}^n p_i(T) p(C|i, k(T+1)) \right)$$

$$p(C|i, k(T+1)) = \int_{-\infty}^{\infty} \phi(w) \prod_{j \neq i} \Phi \left(\frac{-2 \ln \frac{p_j(T)}{p_i(T)} + d'_{jk(T+1)}{}^2 + 2d'_{ik(T+1)}w + d'_{ik(T+1)}{}^2}{2d'_{jk(T+1)}} \right) dw$$

$\phi(\square) =$ Standard normal density function

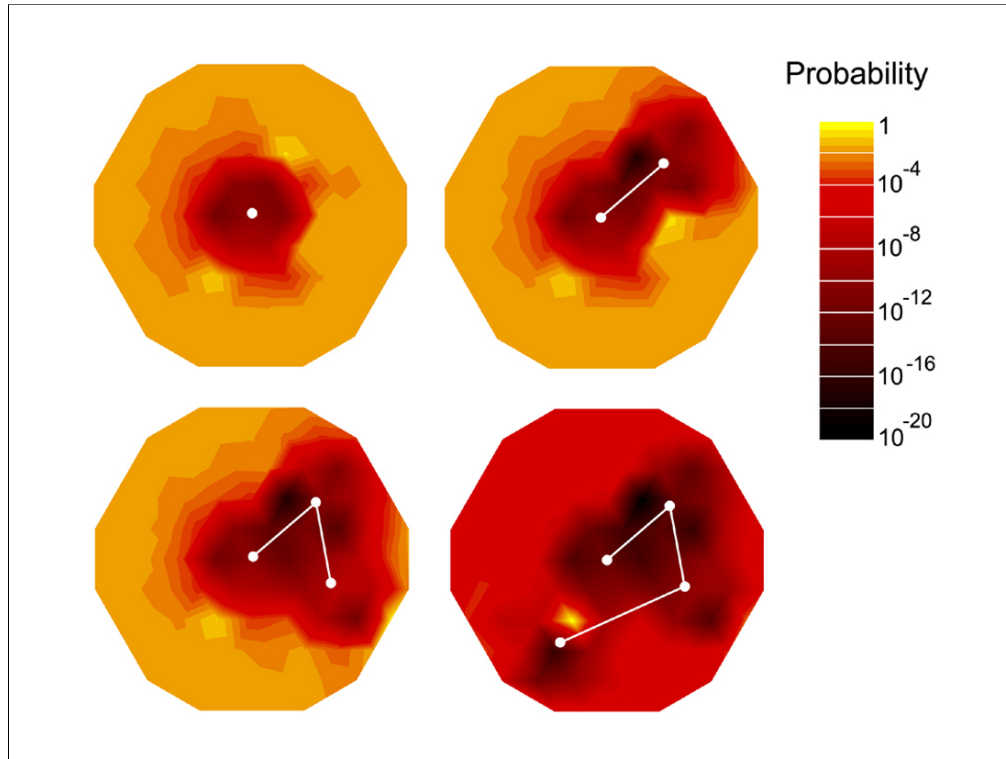
$\Phi(\square) =$ Standard normal integral function

This formula specifies the optimal next fixation, given the current posterior probabilities, and the searcher's visibility map. Specifically, this formula gives the fixation that maximizes the percent correct for localizing the target after that fixation is made.

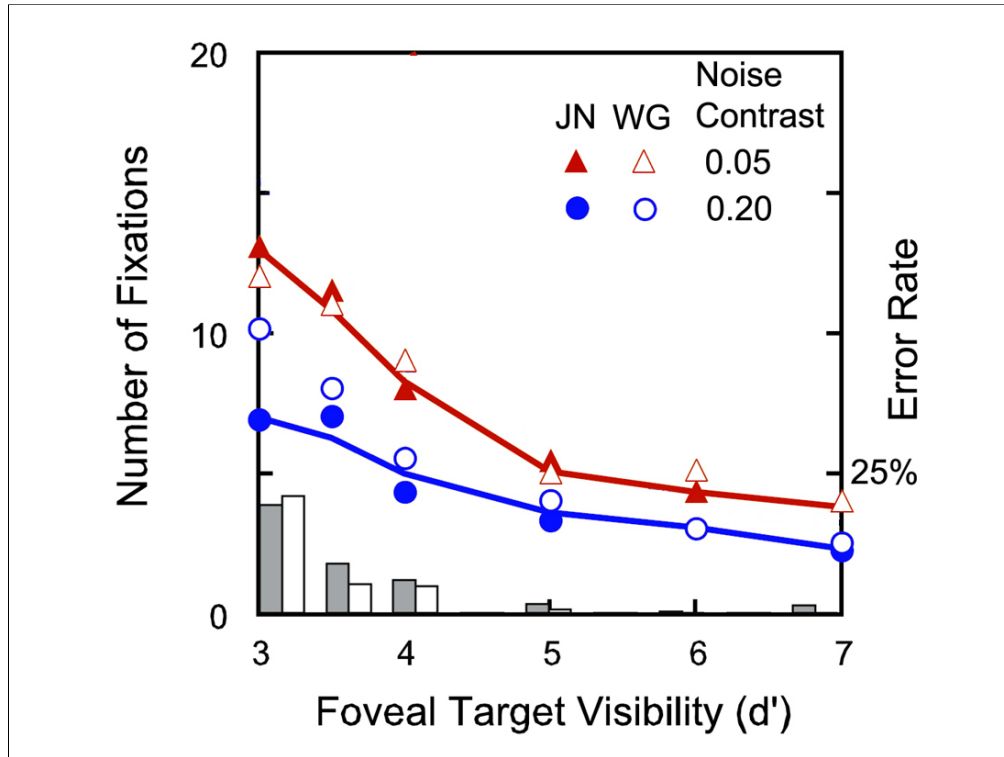


Here is a typical sequence of fixations for the ideal searcher. Notice the haphazard pattern and variable saccade length. The ideal observer found that target in 7 fixations.

To get an idea of what is going on during ideal search it is useful to look at the posterior probability maps.



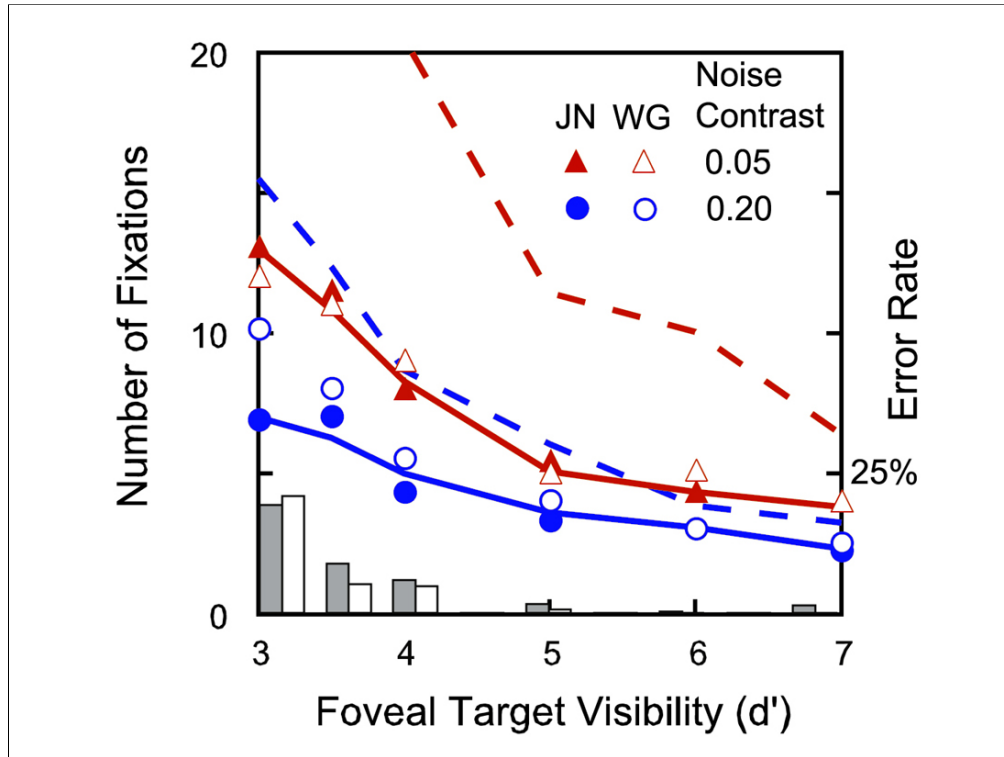
Here are the posterior probability maps of the ideal searcher during a search where the target was found after 4 fixations. There is an elevated posterior probability at the target location (down and left of center), but initially the eye is drawn to other locations. Posterior probabilities are suppressed in the neighborhood of each fixation creating inhibition of return. Notice how the fixations are not always to the location with the highest posterior probability.



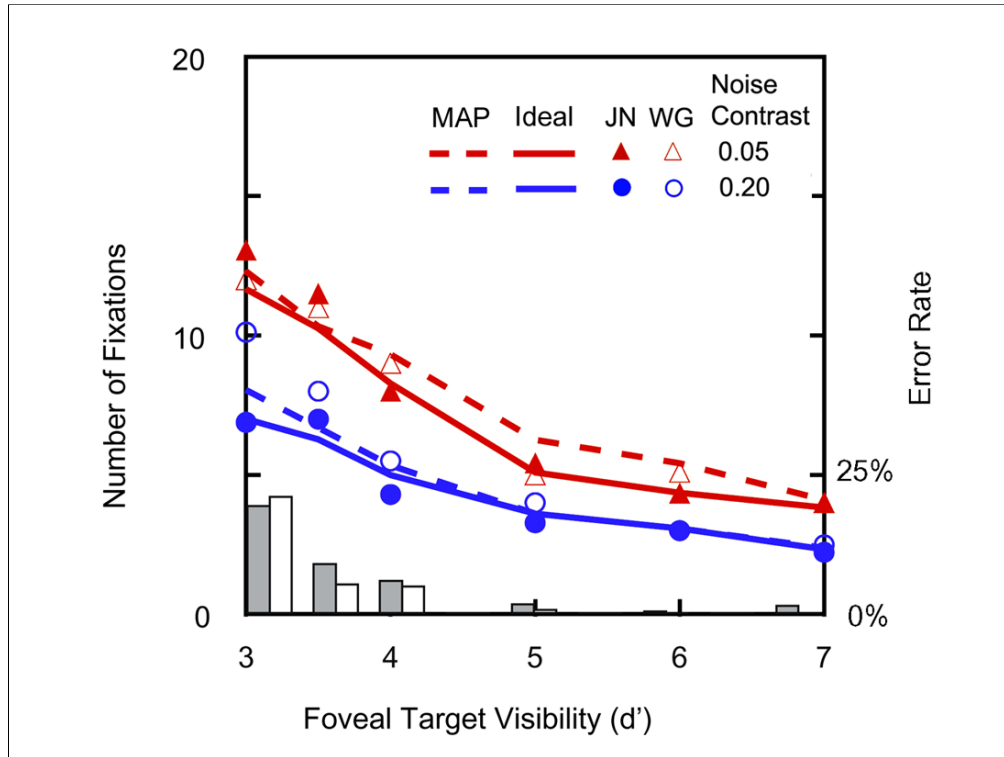
The symbols are the median number of fixations that the two human observers required to locate the target, as a function of the visibility (d') of the target in the fovea, for two levels of background noise contrast.

You can think of the bottom axis as contrast (the greater the d' the greater the contrast). As can be seen, search performance improves as the visibility of the target increases and is better in the high noise condition (this is because the visibility maps are broader in high noise for the same foveal d'). The solid curves show the predictions of the ideal searcher with the same visibility maps as the human observers. (The stopping criterion for the ideal searcher was set so that it has the same error rates as the human observers.) The results imply that humans are remarkably efficient at visual search, at least under these conditions, nearly reaching the performance of the ideal searcher.

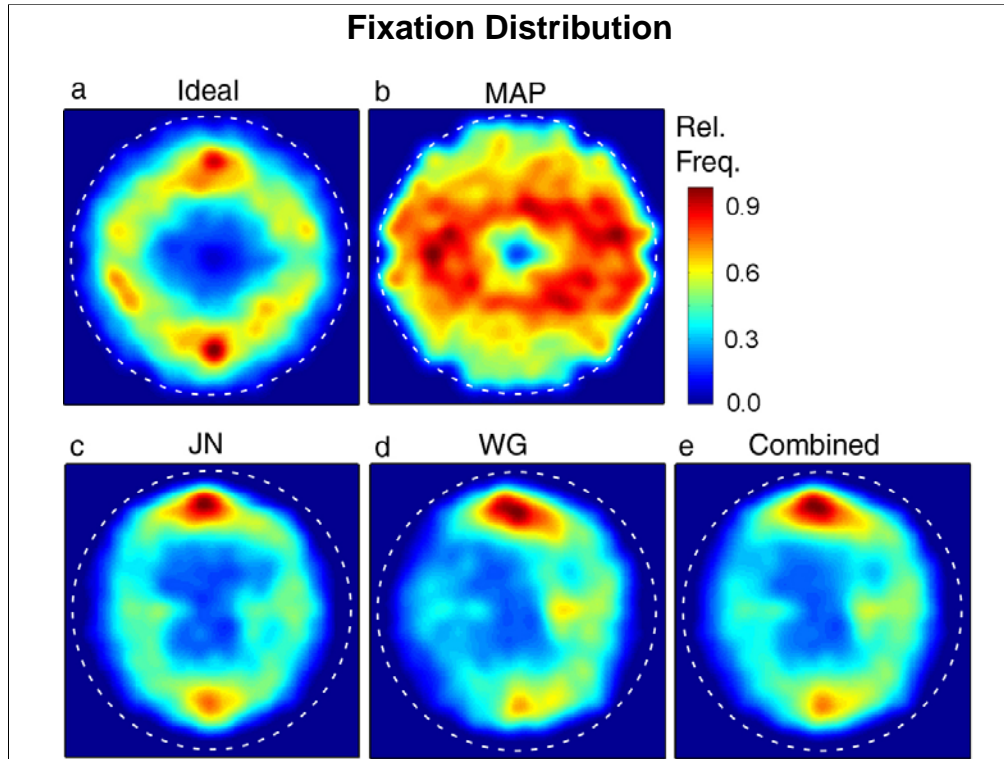
To get a better idea of how well humans are doing it is useful to consider sub-optimal observers.



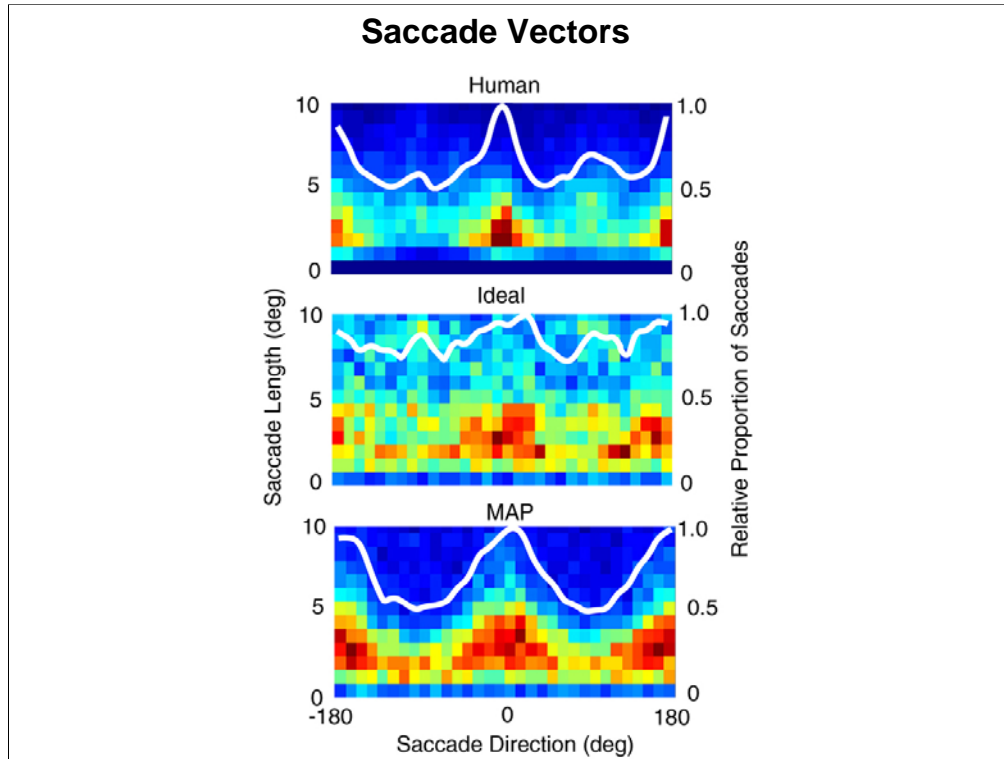
The dashed curves show the performance of a searcher that is ideal in every way except that it makes random fixations. Humans far outperform this searcher. Humans also outperform an enhanced random searcher that has the added feature of inhibition of return. The fact that humans outperform these searchers is a powerful result because it conclusively rejects (for this study) all possible models of visual search where fixation locations are selected at random, with or without replacement.



The dashed curves show the performance of a searcher that is ideal in every way except that it always fixates the location with the greatest posterior probability of being the target location. Under these conditions this MAP searcher performs almost as well as the ideal searcher. Hence the MAP (feature-based) searcher cannot be rejected on the basis of overall performance. To compare models need to look at eye movement statistics.



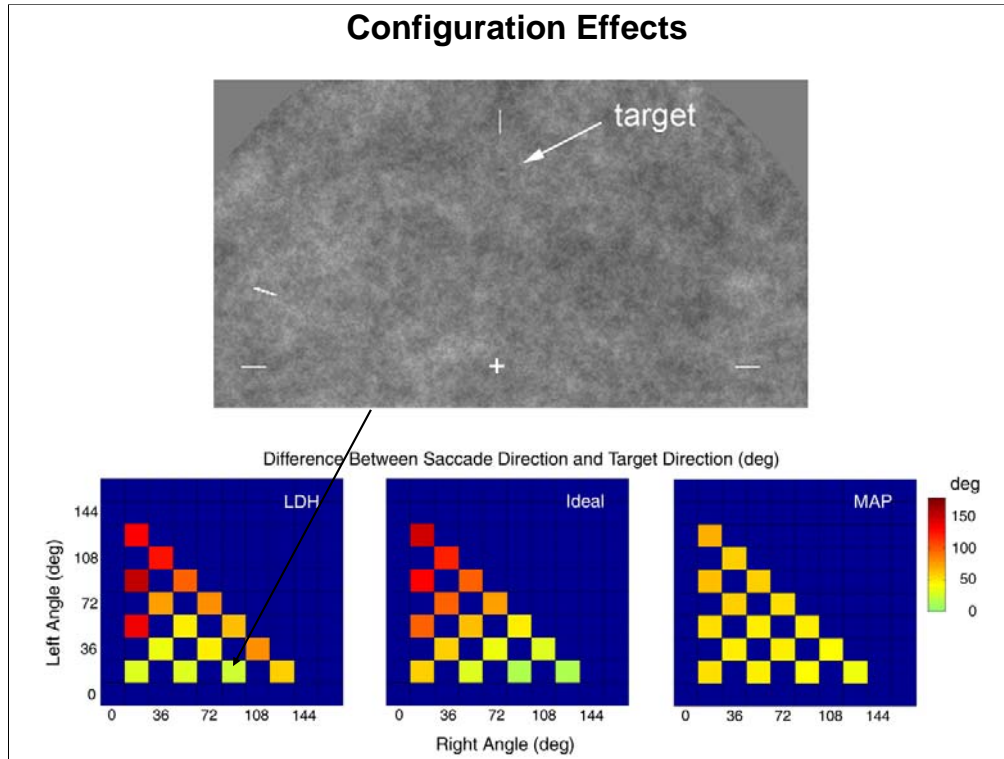
Humans, like the ideal, have a donut shaped distribution with increased density of fixations at top and bottom and with fixations not near the edge of the display. MAP is very different.



Humans: more horizontal saccades, shorter horizontal saccades, secondary peaks in vertical direction

Ideal: same trends but fewer horizontal saccades

MAP: more horizontal saccades, longer horizontal saccades, no secondary peaks



In a current study (with Chris Bradley) we are looking at configuration effects in a single saccade search task. Four locations are cued prior to each trial (16 configurations). After a random time interval the display appears with enough time for one saccade. Here is plotted the mean difference in saccade direction and target direction. There are 4 plots like these, one for every actual target position. The distribution of human saccade directions is similar to ideal and very different from MAP.

Summary

1. Developed a theory of optimal feature-based and information-based searchers
2. Humans are near optimal, ruling out many non-optimal eye movement strategies
3. Eye movement statistics share many properties with optimal information-based searchers
4. The optimal information-based searcher can be approximated with more biologically plausible searchers
5. We are currently testing these models in multiple and single saccade search tasks

Najemnik & Geisler (2005) *Nature*, 434, 387-391

Geisler, Perry & Najemnik (2006) *Journal of Vision*, 6, 858-873

Najemnik & Geisler (2008) *Journal of Vision*, 8, 1-14.

