The distribution of visual objects on the retina: connecting eye movements and cone distributions

Alex Lewis
Department of Psychology, University College London, London, UK

Raquel Garcia
Theoretical Physics, Imperial College, London, UK

Li Zhaoping
Department of Psychology, University College London, London, UK

Experimental data on the accuracy and frequency of saccades are incorporated into a model of the visual world and eye movements to determine the spatial distribution of visual objects on the retina. Visual scenes are represented as sequences of discrete small objects whose positions are initially uniformly distributed and then moved toward the center of the retina by eye movements. We then use this model to investigate whether the distribution of cones in the retina maximizes the information transferred about object position. Assuming for simplicity that a single cone is activated by the object, the rate of information transfer is maximized at the receptor stage if the probability that a target lies at a position on the retina is proportional to the local cone density. Although qualitatively it is easy to understand why the cone density is higher at the fovea, by linking the cone density with eye movements through information sampling theory, we provide an explanation for its quantitative variation across the retina. The human cone distribution and the object distribution in our model visual world are shown to have the same general form and are in close agreement between 5- and 30-deg eccentricity.

Keywords: saccades, photoreceptor distributions, sampling theory

Introduction

Imagine a scuba diver. Sea creatures are equally likely to appear anywhere in her visual field. But her eyes tend to align with the interesting fish and corals, bringing their image to the high-resolution region at the center of the retina. Overall, the time she spends watching creatures at different angles with her line-of-sight must decrease with the size of the angle. It must decrease smoothly, because eye movements are subject to errors. But we know that the density of cones on the retina decreases smoothly with eccentricity. Are the two decreases related in a simple quantitative way? Intuitively, it makes sense to have more cones at retinal positions where targets are more likely to lie.

If we restrict our argument to information about positions of attended objects, ignoring the transfer of information about the objects themselves, this intuition can be put formally in the language of sampling theory. The coding of position information by visual receptors is most efficient if the cone density (or possibly the density of retinal ganglion cells) \( D(x) \) at a retinal location \( x \) is proportional to the probability density \( P(x) \) of attended objects at the same visual location (this is shown in Appendix B). In the absence of eye movements, it would be most efficient to have a uniform cone distribution (assuming that objects appear uniformly in the visual field). Instead, when an interesting object attracts our attention, we shift our gaze to fixate it at the center. Eye movements thus allow for a small high-resolution region, which is associated with the fovea. We expect fixation movements primarily to shape the arrangement of cones rather than rods, as rods become saturated and so contribute no information in daylight conditions.

However, we also compare the distribution of ganglion cells, which transmit information from both rods and cones.

Figure 1 shows the change in rod and cone density with eccentricity for a typical human eye (Osterberg, 1935; Curcio, Sloan, Kalina, & Hendrickson, 1990). The cone density has a sharp peak at the foveola, of up to 200,000 cones per square millimetre, and declines very rapidly in the first few degrees. Outside this region, it declines more slowly, to stabilize at 2,000 cones per square millimetre at around 15°. Our hypothesis then is that this arrangement maximizes the information gathered by the retina, given a fixed number of cones. To make this argument quantitative, we must study the accuracy and frequency of eye movements.

In work on robotic vision, models with a distribution of pixels similar to the distribution of cones in the eye are used as a method of data reduction (Bolduc & Levine, 2003).
So we can think of the arrangement of photoreceptors in two complementary ways: either as that which encodes the maximum information with a given number of receptors, or as one that uses the fewest possible receptors to encode the required information.

Our aim here is to compute \( P(x) \), the probability distribution of objects in visual space, from experimental data on the accuracy and frequency of saccades. As we focus on information about positions of objects, this approach is not applicable to the center of the fovea, the function of which is mainly to collect information about objects. We therefore ignore gaze-holding movements, which affect only the center of the fovea. Indeed, within the fovea, we don’t expect the arrangement of cones to match the target distribution determined by eye movements, because other factors become dominant. First of all, close to the foveola, the angular separation of cones approaches the optical limit on the eye resolution, setting a bound to the maximum useful cone density (Wandell, 1995). Second, the size of the peak density region is constrained by factors such as the proximity of capillaries to the cones. Finally, fovea and periphery serve different functions: The fovea is used mainly for object recognition and the periphery for object detection. The distribution of cones optimized for object detection is unlikely to be that for object recognition.

By comparing the calculated distribution \( P(x) \) with the cone distribution, we assess the extent to which information is maximized. Our model gives a quantitative link between two separate bodies of experimental data: cone density measurements and saccade psychophysics (Figure 1).

### Methods: Obtaining the Probability of Target Incidence Across the Retina From Eye Movements

Distinct objects and other luminous stimuli in the visual field of an observer will sometimes induce her to saccade to bring their image closer to the fovea. Because saccades are generally inaccurate, two or more are sometimes required to fixate the target. From the retina’s viewpoint, objects of interest are seen to jump closer to the fovea. We use spherical coordinates \((x, \phi)\) to parameterize points in the visual hemisphere, where \(x\) is the angular distance from the center of the eye and \(\phi\) the angle around the circle at \(x\). Clearly, the centering movements of the eye imply that \(P(x, \phi)\) must decrease with eccentricity \(x\).

What one means by \(P(x, \phi)\) depends on one’s representation of the visual world. Here we adopt a very simple one that we can model easily from available experimental data. So we represent a subject’s visual life as a sequence of point object locations in the visual hemisphere. Note that this does not mean that there is only one object at a time in the scene, but that we pay attention to only one object at a time. Each object location is either a sample from an initial uniform distribution, \(P_0(x, \phi)\), or the result of one or more saccades to that sample point. \(P(x, \phi)\) is the distribution of the object locations both before and after saccades (Figure 2).
Figure 2. In a scene where several objects are present one of them calls the attention of the observer. 1. At the start, the object could be anywhere. This is represented by an initial uniform distribution \( P_U(u) \). 2. The chosen object will with some probability elicit a saccade, which brings it closer to the fovea. 3. This happens to many objects in an observer's visual life, so that the distribution of targets after saccades is concentrated around the center of the eye. 4. Combining the distributions of target before and after saccades, we obtain the average distribution. In the full model, we allow for zero, one, or more than one saccade to be made to each target.

This picture assumes the following additional simplifications. First, targets are depicted as point objects. This is assumed largely because it is difficult to define what a target object is (e.g., whether it is the face, or the eye in the face, or the pupil in the eye). Second, we assume that a new or initial target is equally likely to appear at any point of the visual field. This is adopted due to its simplicity, and due to our lack of knowledge about the initial distribution of visual targets and about how our eyes choose the next target. Third, we assume that our eyes only gather information about the targets between saccades, not during a saccade, because of saccadic suppression (Matin, 1974).

While these strong assumptions represent significant departures from reality, it is appropriate to keep things simple at this early stage of the investigation into the possible link between retinal sampling and eye movements. These simplifications enable us to avoid complexities, which at this point would largely obscure the subject, and thereby to obtain a tentative answer to our main question.

We moreover restrict ourselves to radial features in the visual field by considering only the effect of saccades on \( x \), the target's eccentricity, and ignoring variations with the azimuth \( \phi \). This makes sense because cone density variations and gaze shifts are much more strongly associated with changes in \( x \). In fact, in this work, we consider mainly saccades along the horizontal meridian, \( \phi = 0 \).

Further, we assume, based on experimental evidence, that the post-saccade eccentricity \( x \) of a target that lies initially at \( y \) comes from a probability distribution \( K(x,y) \). We allow for the probability of making a saccade toward an object, \( \alpha'(y) \), to depend both on the eccentricity of the target, \( y \), and the number of saccades previously made toward the target, \( n \). Thus the decision to switch attention to a new target can occur without making a saccade to the previous one detected at \( y \), with probability \( 1 - \alpha'(y) \), or after \( n \) saccades have brought the previous target to eccentricity \( y \), with probability \( 1 - \alpha'(y) \).

In either case, if attention is not switched to a new target, a saccade is made toward the currently attended target. The probabilities \( \alpha'(y) \) are unknown, but they are closely related to the proportion of natural saccades made to a target at \( y \), denoted \( f(y) \), which can be determined from psychophysical experiments, as can \( K(x,y) \) (see the following sections).

Under these assumptions, we find that the probability distribution, as a function of eccentricity, for fixed \( \phi \), can be obtained from \( K(x,y), f(y) \) and \( P_0(x) \) as

\[
P(x) = (1 - \omega) P_0 + \frac{\omega}{2 \pi \sin(\alpha)} \int f(y) K(x,y) dy, \tag{1}
\]

where \( \omega \) is a free parameter representing the probability to saccade to an object. We derive this equation in Appendix A, using the fact that our assumptions above define a stationary Markov process. An alternative (and equivalent) approach is to obtain \( P(x) \) numerically by simulating the saccadic process (see Appendix A).

In the following sections, we explain how we obtained \( K(x,y) \) and \( f(y) \) from experimental data.

**The Accuracy and Precision of Saccades**

If we display a visual stimulus at an angle \( y \) from a subject's fixation axis and ask her to saccade toward it, the probability that the postsaccade eccentricity lies between \( x \) and \( x + dx \) is given by \( K(x,y)dx \). Thus \( K(x,y) \) can be thought of as the probability distribution for the error \( x \) of a gaze shift as a function of initial target eccentricity \( y \).

In many studies of saccades, a subject sits with his head fixed and waits for luminous stimuli to appear randomly at different points, usually in the horizontal plane (Becker, 1991; Frost & Poppel, 1976; Kapoula, 1985; Deubel, 1987). The saccades observed in these conditions have been termed "normal" saccades (Becker, 1991), and are believed to reflect the properties of saccades in more natural conditions.

Experiments measuring the error of normal saccades generally agree on a simple linear relationship between initial target eccentricity \( y \) and the mean error \( \mu \), namely

\[
\mu(y) = a(y - y_0), \text{ with } a \text{ typically in the range } 0.1 - 0.2 \text{ and }
\]
\( y_0 \) in the range \( 5^\circ - 10^\circ \) (Becker, 1991). Saccades to targets at \( y > y_0 \) usually undershoot the target, whereas saccades to targets at \( y < y_0 \) typically overshoot (Figure 1B). The scatter in the error of saccades also increases linearly with increasing eccentricity (van Opstal & van Gisbergen, 1989), so we take the standard deviation for saccade errors \( x \) to be \( \sigma(y) = b + cy \). A linear fit on data (from Frost & Poppel, 1976) (see Table 1) gives the following values for the parameters: \( a = 0.15, y_0 = 8.5^\circ, b = 0.5^\circ, c = 0.1 \).

Table 1. Mean Saccadic Error \( \mu \) and SD \( \sigma \) for Saccades to Targets at Eccentricity \( y \) (all in degrees).

<table>
<thead>
<tr>
<th>( y )</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

We have averaged over the further split given in each case into nasal-fixation (target in relatively temporal position) and temporal-fixation (target in relatively nasal position) conditions.

For each initial eccentricity \( y \), we model \( K(x,y) \), the distribution of saccade errors, as a Gaussian over \( x \) with mean \( \mu(y) \) and standard deviation \( \sigma(y) \).

\[
K(x, y) = \begin{cases} \frac{A(y)}{\sigma(y)} \left( \exp \left( -\frac{(x-\mu(y))^2}{2\sigma(y)^2} \right) + \exp \left( -\frac{(x+\mu(y))^2}{2\sigma(y)^2} \right) \right), & \text{if } x < y \\ 0, & \text{if } x > y \end{cases}
\]

(2)

where \( A(y) \) is the normalization factor that ensures \( \int dx \cdot K(x, y) = 1 \), and the second term is the contribution from overshoots (\( \mu(y) > 0 \) for most \( y \)) under the assumption of no asymmetry between nasal and temporal fixations.

Although we have obtained \( \mu \) and \( \sigma \) from data on eye-only saccades, studies on eye-head gaze shifts show that their accuracy is very much the same as that of eye-only saccades (Stahl, 2001). Therefore \( K(x,y) \) describes the redistribution of target eccentricities under a general saccadic gaze shift, involving both eye and head movements.

Our estimates for \( a \) and \( y_0 \) are derived from data in which \( y = 5^\circ \) is the smallest initial target eccentricity. Below \( 5^\circ \), the simple linear relation between target eccentricity and saccade error ceases to hold. Indeed, saccades with small amplitudes can be extremely accurate (Kowler & Blaser, 1995). A point where the above expression for \( \mu \) manifestly breaks down is \( Y \) with \( \mu(Y) = a(Y-y_0) = -Y \), because for \( y < Y \), the mean error would become greater than the initial target eccentricity. For our values of \( a \) and \( y_0 \), we get \( Y = 1.11^\circ \), so we take \( y' = 1.5^\circ \) as the lower end for the range of validity of our model (the lower limits of our integrals are accordingly set to \( 1.5^\circ \)). This is not a problem because, as discussed in the Introduction, we only expect saccade accuracy to determine the cone distribution at the larger eccentricities.

### Frequency Distribution of Saccades

Our main source for the frequency distribution \( f(y) \) is an experiment in which three subjects were asked to wander freely outdoors while their eye movements were recorded (Bahill, Adler, & Stark, 1975). They found the relative frequency of eye-only saccades to decay exponentially with amplitude, that is \( f_0(y) \propto \exp(-y/7.6^\circ) \). This means a mean saccade amplitude of \( 7.6^\circ \). This value has been confirmed in a more recent experiment (Andrews & Coppola, 1999), in which subjects with their heads fixed viewed a variety of scenes and performed different visual tasks. The mean size of the saccades recorded was very close to \( 7.6^\circ \).

The empirical \( f_0(y) \) differs from \( f(y) \) in two ways. First, saccades with amplitudes greater than \( 10^\circ \) usually involve both eye and head movements, and their total amplitude is therefore larger. So we expect \( f(y) \) to take higher values at large eccentricities. Second, the argument of \( f_0 \) is saccade amplitude, whereas that of \( f \) is initial target eccentricity. We bridge over these differences one at a time: First, we find \( f_0 \), the frequency of eye-head saccades as a function of amplitude, from \( f_0 \) and data on eye-head coordination; then we deconvolve \( f(y) \) with \( K(y; y') \) to obtain \( f(y) \). We assume that \( f(y) \) and \( f_0 \) will have a similar functional form to \( f(y) \), and so we model them as a constant term plus an exponential decay, namely

\[
f(y) = d \cdot e^{-y/\lambda}.
\]

(3)

\[
f_G(y) = d_G + e^{-y/\lambda_G}.
\]

(4)

### General Gaze Shifts

The probability distribution for eye and head components, \( y_E \) and \( y_H \), associated with a gaze-shift \( y'' = y_E + y_H \) can be determined from experiments on combined eye-head visual fixations (Stahl, 1999, 2001). In these experiments, the total gaze shift and the eye movement relative to the head were measured as subjects moved their eyes and head freely to targets at various eccentricities. The head component was found to follow a piecewise linear fit as a function of the total gaze shift. If \( y < B \) then only the eyes move and \( y_H = 0 \), and if \( y > B \), the average head component increases linearly with \( y \).

\[
y_H = \begin{cases} 0 & \text{if } y \leq B \\ D + m y & \text{if } y > B \end{cases}
\]

(5)

The three parameters \( B, D, \) and \( m \) vary considerably among the subjects, and even for left and right fixations by the same subject (for details, see Stahl, 1999, 2001). Because we need to combine these results with the average eye-only saccade frequency from a different experiment, we use the values that give the best fit to the
combined data from five subjects. We get $D \approx 0$, $B = 7.8^\circ$, and $m = 0.84$. We also find that the standard deviation of $y_H$ is $s = 7.6^\circ$ for all $y > B$.

The probability distribution of an eye-only saccade amplitude $y_E = y - y_H$, given a total gaze shift of size $y$ is then

$$ p(y_E \mid y) = \begin{cases} 
\delta(y_E - y) & \text{if } y \leq B \\
N(y)\exp\left[\frac{(y_E - (m-y_H))^2}{2\sigma^2}\right] & \text{if } y > B \text{ and } y_E \leq y \\
0 & \text{if } y_E > y
\end{cases} \tag{6} $$

The normalization factor $N(y)$ is obtained by requiring that $\int_{-\infty}^{B} p(y_E \mid y)dy_E = 1$ for $y > B$. The frequency distribution of total gaze-shifts $f_G(y)$ and that of saccade amplitudes $f_E(y_E)$ are then related by

$$ f_E(y_E) = \int_{0}^{\infty} f_G(y) dy \tag{7} $$

To solve for $f_G$, we could try to discretize the integral and then solve the resulting linear equations through $f_G = p^{-1}f_E/df_E$. The problem is that $p$ is not invertible: all the elements of the left-most columns of $p$ are zero, because for large gaze shifts the corresponding saccades are always accompanied by head movements. Instead we find the parameters in Equation 4, which give the best fit of the function $\int_{0}^{\infty} p(y_E \mid y)f_G(y)dy$ to $f_E(y) = \exp(-y/7.6^\circ)$. This gives $d_0 = 0.017$ and $\lambda_0 = 3.2^\circ$.

**From Saccade Amplitude to Target Eccentricity**

A gaze shift of amplitude $y'$ toward a target at $y$ corresponds to an error of $y - y'$. It follows that

$$ f_G(y') = \int f(y)K(y - y', y)dy \tag{8} $$

The kernel here differs from that in Equation 2 in that we allow for negative $y - y'$ and dismiss the second term in Equation 2. This is needed to distinguish undershoots and overshoots — which have a larger amplitude — to the same target. Equation 8 can be inverted to obtain a numerical solution for $f(y)$. However, the resulting $f(y)$ remains very close to $f_G(y)$, so that $d \approx d_0$ and $\lambda \approx \lambda_0$. (The reason $f_G$ is very close to $f$ is that at large $y$, where saccadic errors are large, $f_G$ is almost constant.)

The above values of $d$ and $\lambda$ were obtained by averaging the results of the eye-head coordination experiments for all five subjects. Because there is considerable variation between subjects, we repeated the computations using the individual subjects’ data for $P(y_E \mid y)$. This gives values of $\lambda$ ranging from $2.0^\circ$ to $7.6^\circ$ and values of $d$ ranging from 0.004 to 0.09. Because these parameters were obtained using the average $f_G(y)$, for three different subjects in another experiment, this does not necessarily mean that the actual $f(y)$ varies between individuals. Indeed the data that show large individual variation in eye-head coordination come from experiments in which all subject made saccades to targets with approximately the same distribution of positions (Stahl 1999, 2001). Thus, it is reasonable to assume, as we have done, that in a natural environment $f(y)$ will not vary widely even if people use different eye and head movement strategies. However, in view of the wide variation in eye-head coordination strategies, our estimate of $f(y)$ cannot be very reliable. To accurately determine $f(y)$, and to check whether it really does vary between individuals or not, we would need to have data on eye-head coordination and on $f_E(y)$ for the same subjects (or ideally, to measure $f(y)$ directly).

**Saccades to Auditory Targets**

Up to now, we have considered only the accuracy of “normal” saccades: saccades to isolated visual targets under neutral laboratory conditions. In other circumstances, saccades can be much more or much less accurate. For example, undershoots can be eliminated by the range effect (Kapoula & Robinson, 1986), and very high accuracy and precision were observed when subjects were instructed to be as accurate as possible (in Kowler & Blaser, 1995). On the other hand, saccades to auditory targets are much less accurate than saccades to visual targets (Yao & Peck, 1997). They report that auditory saccades to targets at up to $10^\circ$ are as accurate as visual saccades, but they become much less accurate for higher target eccentricities, with errors of $10^\circ$ for targets at $30^\circ$.

In principle, we could try to include the different types of saccades by finding a separate kernel $K'(x,y)$ for each type and decomposing $f(y)$ into the corresponding contributions, $f'(y)$. We would then replace $K(x,y)f(y)$ in Equation (1) by the sum $\mu f(x,y)f'(y)$. This would require a lot of experimental work on the accuracy and frequency of the various types of naturally occurring saccades. What we have done in effect is to assume that the accuracy of most natural saccades is well approximated by that of normal saccades.

We have nevertheless computed a modified $P(x)$ by using saccades to auditory targets. These are no doubt less frequent than visual saccades, but they cannot be neglected: We do often turn our heads looking for the source of a sound. In fact, only auditory saccades can possibly occur to targets at eccentricities outside the visual field. It therefore seems plausible that auditory saccades may have a significant effect on the target distribution.

According to Zambarbari, Schmid, Magenes, and Prablanc (1982), for most target eccentricities, saccades to auditory targets have errors with mean 32% and standard deviation 23% of target eccentricity. So we find $K(x,y)$ for auditory saccades by inserting the parameters $a = 0.32$, $x_0 = 0$, $b = 0$, and $c = 0.23$ in Equation (2) for auditory saccades.
The expression for \( f(x) \) we used for visual saccades was based on an experiment in which \( f_s(y) \) was measured for all saccades, visual and auditory. Because we do not know what fraction of natural saccades are auditory, we calculate \( P(x) \) in the two extreme cases, one where all saccades are visual and one where they are all auditory. The true \( P(x) \) will be somewhere between these extremes. This calculation is intended as a way of assessing how sensitive \( P(x) \) is to experimental uncertainty in \( K(x,y) \).

**Results**

Our motivation to determine \( P(x) \) was to compare it with the density distribution of cones on the retina \( D(x) \). For maximal information, the two should be proportional (a proof of this is given in Appendix B). We restrict the comparison to the cone density and gaze shifts along a horizontal axis through the fovea. According to our model

\[
P(x) = (1 - \omega) P_0 + \frac{\omega}{2 \pi \sin(x)} \int f(y) K(x,y) \, dy
\]

where \( P_0 = 1/(2\pi) \) is the initial uniform target distribution, \( f(y) \) is the normalized frequency of saccades to targets at eccentricity \( y \), and \( \omega \) is a free parameter which measures the fraction of targets that elicit saccades.

By \( D(x) \) we mean the number of cones per unit solid angle in the visual field found at eccentricity \( x \). We obtain it from standard cones/mm\(^2\) density measurements using the curves for retinal projection given in Drasdo and Fowler (1974).

A nonlinear fit of \( P \) to \( D \) gives \( \omega = 0.09 \). In Figure 3, we have plotted \( D(x) \) and the \( P(x) \) corresponding to \( \omega = 0.09 \). The two distributions have the same general form. They both show a peak around the center of approximately equal width, and they are in close agreement between 5\(^\circ\) and 30\(^\circ\). However, they behave differently within the fovea and at large eccentricities. For large \( x \), \( P(x) \) is constant, as a consequence of the assumption \( P_0 = \text{constant} \), whereas \( D(x) \) keeps decreasing with \( x \) (this decrease does not show in Figure 1, because there we plotted \( D(x) \) in units of cones/mm\(^2\) rather than cones/solid angle).

Arguably, we should expect the density of retinal ganglion cells (RGC), rather than that of cones, to be proportional to \( P(x) \), because the RGCs determine the rate at which information can be transmitted from the retina. If we take \( D(x) \) to be the density of RGC (from Sjostrand, Olsson, Popovic, & Conradi, 1999), the best fit of \( P \) to \( D \) gives \( \omega = 0.13 \), and the results are similar to those for cones (see Figure 4).

Figure 5 compares the probability distributions computed using the accuracy parameters for either auditory saccades or visual saccades. Because auditory saccades are less accurate, the result is a \( P(x) \) that is smaller at low eccentricities and larger at high eccentricities. The small departure from \( P(x) \) for visual saccades can be explained as follows. The distribution \( f(y) \) peaks at small eccentricities, so there are few saccades to targets at very large eccentricities, which is where the difference between auditory and visual saccades is most pronounced. In reality, only a fraction of saccades will be made to auditory targets, so the departure from \( P(x) \) will be even smaller. We can conclude that if our estimate of \( f(y) \) is accurate, auditory saccades have only a small influence on the cone distribution.
Figure 5. Probability distributions computed with different parameters for saccade accuracy and frequency distribution. 1. Normal visual saccades, targets up to 90°. 2. Less accurate auditory saccades, with the same ω and f(y) as in 1, but with targets up to 180°. Changing the accuracy of saccades has very little effect except at very small eccentricities.

Figure 6 shows the probability distributions derived using the data on eye-head coordination for each subject separately to compute f(y). In this case, the differences between the calculated probability distributions are relatively large, compared to the difference between P(x) and D(x) in Figure 3. These curves provide an estimate of how large is the uncertainty in P(x) due to the limited data available with which to calculate f(y).

Figure 6. Probability distributions computed using saccade frequency distributions f(y) derived from data on eye-head coordination for five subjects (from Stahl 1999, 2001) (two of the curves are very close together), using the same ω. The vertical scale is the same as in Figure 3. There are large differences between the curves, indicating that changes in f(y) could have a significant effect on the results.

Discussion

We wanted to test the hypothesis that saccadic movements govern the arrangement of cones (or retinal ganglion cells — at this stage, our model does not allow us to distinguish which type of cell is more closely related to eye movements) on the retina through a maximal information principle. This would establish a quantitative link between results from the psychophysics of eye movements and the physiology of the retina. Our results confirm that the distribution of cones in the human eye is to some extent adapted to maximize information flow, but, although P(x) and D(x) are roughly proportional across a broad range of eccentricities, the two show important differences at very small and at large eccentricities.

One possible explanation for the discrepancy is that our proposal is essentially right but that our simplistic representation of visual experience leads to the “wrong” P(x). A first step forward in this case would be to do experiments to obtain a better estimate of f(y). Equally important would be to develop a model that allows for complex visual scenes and takes into account their effect on attention.

We have not considered the distribution of rods, which is very different from that of cones or ganglion cells, with a peak at an eccentricity of 15° (Figure 1). A simple possible explanation for this is that the retina is primarily adapted for photopic viewing conditions, and the rod distribution is highly constrained by the space taken up by cones. It is also not clear whether natural saccades in scotopic conditions foveate the target (Doma & Hallet, 1988).

In any event, the current work provides a useful starting point for further research into how attentional and environmental factors contribute to the distribution of objects on the retina and into the question of whether this distribution determines the distribution of cones. Next we discuss possible improvements and extensions of this work.

Insights Into the Large x Discrepancy

Because saccades with errors greater than 30° are extremely rare, the final target distribution P(x) goes like P₀(x) at large eccentricities. Thus, our assumption that the initial target P₀ is independent of eccentricity could be at the root of the large x discrepancy between P(x) and D(x).

The assumption that P₀(x) = constant makes sense if the positions of objects that we pay attention to are unrelated. However, when viewing complex scenes there may be correlations among the positions of some of the objects present (e.g., different targets for saccades may be part of the same large physical object). For this to induce a persistent decline of P₀(x) at large eccentricities, there would have to be correlations at very large scales (e.g., to
have \(P_\theta(60) > P_\theta(80)\), there should be significant correlations between objects 60° apart). Another important factor that is not yet included in our simple model is that the cone density affects the resolution of the eye, it can itself influence the number of interesting objects that are detected, and so \(P_\theta(x)\) might not be independent of the cone distribution.

Although we have been treating \(P_\theta(x)\) as an independent function, according to Equation 13 in Appendix A, \(P_\theta(x)\) is related to the frequency distribution of saccades \(f(x)\) and probabilities for making saccades \(\alpha'(x)\). Indeed, for large \(x\), Equation 13 implies that \(f(x)\) is simply proportional to the product \(P_\theta(x)\ \alpha'(x)\). Because \(f(x)\) is a constant for large \(x\) (Equation 3), our assumption that \(P_\theta(x)\) is independent of \(x\) at large eccentricities is equivalent to assuming that \(\alpha'(x)\), the probability of making a primary saccade to an attended object, is a constant at large eccentricities. Because \(\alpha'\) drops out of Equation 1, the resulting probability distribution \(P(x)\) is independent of that constant. To get \(P(x)\) to fit the data \(D(x)\), we could instead assume that \(P_\theta(x)\) is proportional to the cone density at large eccentricities. We would then have \(\alpha'(x) \propto 1/D(x)\), also at large eccentricities so that the resulting \(f(x)\) is constant there. Our hypothesis that the distribution of cones maximizes information then predicts the form of both the initial distribution of visual objects before saccades \(P_\theta(x)\) and the probability of making at least one saccade to an attended object \(\alpha'(x)\) at large \(x\). However, these predictions are sensitive to the form of \(f(x)\) at large \(x\), which we have estimated from very limited data. At smaller eccentricities, \(P(x)\) depends on all the functions \(\alpha'(x)\), which is why we had to ignore them and use the experimentally observable \(f(x)\).

On the other hand, the large \(x\) departure between \(P(x)\) and \(D(x)\) may indicate that the information collected by the eye is not maximized there. For example, if visual attention declines with \(x\), there may be little to gain by fine-tuning the cone distribution to the target incidence curve there.

### Experimental Parameters in the Model

Our model for \(P(x)\) relies on six parameters derived from available experimental results – four describing the accuracy and precision of saccades, and two describing their frequency. The first four seem well established after many experiments on the errors of normal saccades. However, a good deal of uncertainty clouds the two parameters in the saccade frequency curve. Perhaps the main improvement to our model would come from new experiments to re-evaluate \(f(y)\).

Here we have combined the distribution of eye-only saccades for three subjects (Bahill, Adler, & Stark, 1975) with average values for eye-head coordination from five subjects (Stahl, 1999). Although this is the best we can do without further experiments, it seems hardly justified, because both the parameters characterizing eye-head coordination and the average size of saccades vary noticeably among individuals (Stahl, 1999, 2000; Andrews & Coppola, 1999). As can be seen from Figure 6, a different value for \(f(y)\) could have a significant effect on our results. It would clearly be preferable to have data on both eye-movements and eye-head coordination for the same subjects, or even better, to be able to measure the frequency of gaze shifts involving both eyes and head directly. In other words, the uncertainty in \(f(y)\) can only be resolved by further experiments.

### Radial Asymmetries

Here we have considered only horizontal eye movements and cone distributions. The curve shown in Figure 3 is the average of nasal and temporal cone densities along the horizontal meridian, but the cone distribution is not radially symmetric. There are differences between the nasal and temporal directions, and between the horizontal and vertical directions.

#### Horizontal-Vertical Asymmetry

The density declines more sharply along the vertical meridian than along the horizontal meridian, so that at 3.5°, the density is 20,000 cones/mm² on the horizontal meridian and 16,000 cones/mm² on the vertical meridian (Curcio et al., 1990).

This radial asymmetry has three possible explanations within our model: (a) vertical saccades could be less accurate (radial asymmetry in \(K(\theta, \phi)\)); (b) they could be less frequent (radial asymmetry in \(f(\theta)\)); (c) and the initial probability \(P_\theta(x)\) could be different. Of these, there is some evidence that vertical saccades are less accurate than horizontal saccades (Becker, 1991).

When saccades are less accurate, \(P(x)\) is smaller at small values of \(x\). However, vertical saccades are still much more accurate than auditory saccades, which, as we saw, cause only a very small change in \(P(x)\). Thus the lower accuracy of vertical saccades cannot account for the substantial horizontal-vertical asymmetry of the cone distribution. Once more we are led to conclude that to exploit the predictive power of our model, we need more detailed information about the relative frequencies \(f(\theta)\) of vertical and horizontal gaze-shifts \(\phi\). This said, we also expect a contribution to the asymmetry from \(P_\theta(x)\).

#### Nasal-Temporal Asymmetry

The most prominent feature in the retinal periphery is the “cone streak” (a region of higher cone density extending along the horizontal meridian into the nasal retina). Thus the cone density in the nasal retina, which is close to the density in the temporal retina for small eccentricities, starts to become slightly larger at about 5° and this difference increases up to 40°, where the nasal density is 40-45% higher than the temporal density. An explanation of the cone streak is beyond the scope of our
model, because an object that is in the nasal visual field for one eye is in the temporal field for the other eye. Moreover, the asymmetry persists at eccentricities much larger than the errors of almost all saccades, so it cannot be related to the accuracy of saccades.

**Predictions**

The theory developed in this work could be used to relate a species’ photoreceptor distribution to the environment in which it lives. The species’ habitat would be reflected both in $P_s(x)$ and $f(y)$. For instance, animals that live in open environments often have horizontal streaks of high cone densities in their retinas. Cone distributions are already known for many species, for example, various monkeys, (Packer, Hendrickson, & Curcio, 1989; Wikler, Williams, & Rakic, 1990; Andrade, da Costa, & Hokoc, 2000), squirrels (Kryger, Galli-Resta, Jacobs, & Reese, 1998), and pigs (Chandler, Smith, Samuelson, & Mackay, 1999), but a systematic study of saccade accuracy and frequency across different species is still needed.

The angular width of the high cone density region varies by a factor of 5 across different primate species (Franco, Finlay, Silveira, Yamada, & Crowley, 2000). Now, the angular width of the cone density peak in humans is one of the successful predictions of our model. According to the model, the more accurate saccades are, the narrower this peak will be. (Earlier we saw that for fixed $f(y)$, $P(x)$ is not sensitive to uncertainties in $K(x,y)$. But a different $K(x,y)$ implies a different $f(y)$ [see Equation 11 and Equation 13]). Therefore, we could use the model to predict the accuracy of eye movements in different species: We expect those with a narrower peak to have more accurate eye movements.

It is interesting to note that the foveal cone density in human infants does not reach the adult value for several years (Hendrickson, 1994), and that infants’ saccades are much less accurate than adults’, with corrective saccades of size similar to primary saccades (Salapatek, Aslin, Simonson, & Pulos, 1980).

More speculatively, by developing our theory further, we could learn about the visual world of different species from knowledge of their photoreceptor densities and saccades. Specifically, one could compare $D(x)$ with the $P(x)$ arising from data on eye movements through different models $P_0(x)$ of the visual world. The model that gives the best fit between $P(x)$ and $D(x)$ would most adequately describe the visual life of the species in question. Perhaps our simple model, which is at best partially adequate to describe human vision, would give a better fit between $P(x)$ and $D(x)$ in animals with a more basic visual system.

**Appendix A**

Formally, the chain of visual events described in Methods corresponds to a sequence $(u, n)$, where $t$ labels time step, $u_t = (x_t, \phi_t)$ is the position of the current target in the visual hemisphere, and $n_t$ is the number of saccades made so far to that target. This chain is the outcome of a Markov process, because the probability of occurrence of different states at time $t$ depends only on the state at $t-1$, and not on previous states or on $t$ itself. If $p(u, n; u, n)$ represents the transition probability from “target at $u$ after $n$ saccades” to “target at $u$ after $n$ saccades,” then clearly $p(u, n; u, n) = 0$ unless $n' = n + 1$ or $n' = 0$, corresponding respectively to making a further saccade to the same target or switching attention to a new target.

The idea is to find the probabilities $P(u, n)$ of the different states from $P_0(u)$ and the transition probabilities. Then we can compute $P(u)$ as the sum $P(u) = \Sigma_p P(u, n)$. However, we first reduce the problem to a one-dimensional one, in which we consider only the effect of saccades on $x$. This allows us to replace the two-dimensional probability $P(x, \phi)$, for which $|P(x, \phi)\sin(x)dx = 1$, by a one-dimensional probability density $\Pi(x)$, for which $|\Pi(x)dx = 1$. If $P(x, \phi)$ is independent of $\phi$, we have $\Pi(x) = 2\pi \sin(x)P(x)$. This is a good approximation because a saccade causes little change in $\phi$, and so $P(x, \phi)$ will only vary slowly with $\phi$. Once we have a model for $\Pi(x)$, we can revert to the two-dimensional target distribution by dividing by $\sin(x)$, as is needed to compare with the cone density distribution.

The experimental data on which we base our model suggest the following expression for the transition probabilities between events at $t$ and $t + 1$:

$$p(x, n'; y, n) = \begin{cases} 
\alpha^n(y)K(x, y) & \text{if } n' = n + 1 \\
(1 - \alpha^n(y))P_0(x) & \text{if } n' = 0 \\
0 & \text{otherwise}
\end{cases}$$

Here $\Pi_0(x) = \sin(x)$ is how the uniform distribution, $P_0 = 1/(2\pi)$, looks in the one-dimensional reduction. The function $\alpha^n(y)$ is the probability that an object at eccentricity $y$ elicits a saccade, given that $n$ saccades have previously been made to that target. And $K(x, y)$ is the probability distribution of saccade errors, as described previously.

There is experimental evidence that suggests that $\alpha^n(y)$ decreases with $y$. For instance, when shown two identical targets at different positions, we are more likely to saccade to the one at a lower eccentricity (Findlay, 1980). Another indication comes from recording saccade amplitudes for subjects in a natural environment: The frequency of saccades decreases as a function of amplitude (Bahill, Adler, & Stark, 1975). Indeed $f(y)$ captures the joint effect of the $\alpha$’s on $\Pi$, and so we will not need their explicit form. It would be very difficult to determine the $\alpha$’s from experiments. For one thing, there
are many other variables affecting salience, such as target size and luminosity. Hence \( \alpha(y) \) has to be interpreted as an average over all such other parameters. Over a long enough time, we would expect that objects of all sizes, luminosities, etc., will appear at all distances, and \( \alpha'(y) \) can be interpreted as the proportion of all objects at eccentricity \( y \) that elicited a saccade.

The Markov process defined by the transition probabilities (Equation 10) is ergodic. (A Markov process is said to be ergodic if it is positively recurrent and acyclic, namely if it returns to a given volume of event space within finite time and if the period for return can take any value above a minimal \( t \).) For an ergodic Markov process with transition probabilities \( p(s; s') \), the relative frequency with which a state \( s \) occurs satisfies \( \Pi(s) = \sum_s P(s; s') \Pi(s') \). In our case this means

\[
\pi(x, n) = \int dy \alpha^{-1}(y) K(x, y) \pi(y, n-1)
\]

(11)

\[
\pi(x, 0) = \left[ \sum_n \int dy (1-\alpha^n(y)) \pi(y, n) \right] \Pi_0(x)
\]

(12)

These two expressions can also be reasoned from first principles. The first one says that the probability that a given target is taken to \( x \) after \( n \) saccades is the sum over \( y \) that it was at \( y \) after \( n-1 \) saccades, that a new saccade is made toward it, and that the error is \( x \). The second is the probability that the target at \( t \) is no longer the target at \( t+1 \) times the probability that a new interesting object is found at \( x \). Finally, the probability that the target lies at eccentricity \( x \) is given by

\[
\Pi(x) = \sum_n \pi(x, n)
\]

To compute this sum, we start by eliminating the \( \alpha's \) in favor of \( f(y) \), the proportion of natural saccades made toward targets at eccentricity \( y \). The relative frequency \( f \) primary and secondary saccades to targets at eccentricity \( y \) is given by

\[
\frac{\alpha^0(y) \pi(y, 0)}{\alpha^0(y) \pi(y, 1)}
\]

In general \( f''(y) \propto \alpha''(y) \pi(y, n) \), with the same proportionality factor for every \( n \) and so \( f(x) = \sum f''(x) \) satisfies

\[
\omega f(y) = \sum_{n \geq 0} \alpha''(y) \pi(y, n)
\]

(13)

We take \( f(y) \) to be normalized whereas the integral over \( y \) is the right hand side is less than one because \( \alpha''(y) < 1 \). So \( \omega \), a number between 0 and 1, represents the proportion of objects detected that elicit a saccade. We have no data from which to derive \( \omega \), and so we will treat it as a free parameter of the model. Putting all this together, we get

\[
\Pi(x) = \sum_n \pi(x, n)
\]

\[
= \left[ 1 - \int \alpha^n(y) \pi(y, n) dy \right] \Pi_0(x)
\]

(14)

\[
+ \int \sum_{n \geq 1} \alpha''(y) \pi(y, n-1) \right] K(x, y) dy
\]

\[
= (1 - \omega) \Pi_0(x) + \omega \int f(y) K(x, y) dy
\]

And then \( P(x) = \Pi(x)/(2\pi \sin(x)) \) is as in Equation 1:

\[
P(x) = (1 - \omega) P_0(x) + \frac{\omega}{2\pi \sin(x)} \int f(y) K(x, y) dy
\]

(15)

We can make some analytical statements about the shape of \( P(x) \), which hold irrespective of the specific form of \( K \) and \( f \). Provided that \( K \) has a Taylor expansion, \( K(x, y) = K(0, y) + \text{higher order terms in } x \), with \( K(0, y) \neq 0 \), for small \( x \) we have

\[
P(x) \approx \frac{1 - \omega}{2\pi} + \frac{\omega}{2\pi x} \int_0^\infty K(0, y) f(y) dy \propto x^{-1}
\]

(16)

where we have used the relation \( \sin(x) \approx x + \ldots \) for small \( x \). So for small \( x \), we have \( P(x) \sim a + bx^{-1} \). At large \( x \), we expect \( K(x, y) \approx 0 \) (because eye movements do bring the fovea close to the target), which implies that \( P(x) \approx (1 - \omega) P_0 \) at large \( x \).

**Simulations**

We also carried out simulations of the generation of a distribution of object locations on the retina by saccades. We began with objects with positions chosen at random, corresponding to the uniform initial distribution \( P_0(y) \), and simulated saccades as moving to a new position from the distribution \( K(x, y) \). Given any choice of the \( \alpha's \), the simulation gives the resulting \( f(y) \) and \( \omega \) as well as \( P(x) \). This allows us to compare the simulated distribution to the probability distribution computed using Equation 1 (see Figure 7). Of course, this simulation is mathematically equivalent to a numerical integration of Equation 14, so it does not give us any new information, but it does provide a check of the correctness of the analysis leading to Equation 15.
Appendix B

Here we show that the cone distribution that maximizes information is proportional to the target probability density. If the successive target positions came from independent identical distributions, this would be the standard result of histogram equalization. In our model, the probability of target positions at a given stage depends on the previous position, so there is redundant information in the average probability. We see explicitly that the redundancies drop out to a good approximation so that we retrieve proportionality between $D(x)$ and $P(x)$.

Consider, for simplicity, a one-dimensional array of cones, numbered sequentially starting at the center of vision, so that cone number $\xi$ is at distance $x$. Treating $\xi$ as a continuous variable, we see that it is related to the cone density by

$$\xi(x) = \int_0^x D(x) \, dx.$$ 

Let $\pi_n(\xi)$ be the probability that the target is detected by cone $\xi$, given that $n$ saccades have been made to it. Assuming no noise, the information in the signal from the retina when the target is next detected is the entropy

$$H_n = \int \pi_n(\xi) \log \pi_n(\xi) \, d\xi.$$ 

We use this conditional probability rather than just the probability that a target is detected at $\xi$ to allow for the fact that the visual system already “knows” that it has just made a saccade.

If $n > 0$ some of the information counted in $H_n$ is redundant: Knowing the target position after $n - 1$ saccades and that a saccade will be made to the same target, we can predict the position after the new saccade, even if only imprecisely. Thus the additional information gained by detecting a target after $n$ saccades is only

$$I_n = H_n - I_{n-1}(x, \xi')$$ 

where $I_{n-1}(x, \xi')$ is the information about the position $x$ after the $n$'th saccade contained in the detection of the target by cone $\xi'$ before the $n$'th saccade. We can note that this cannot exceed the information about $x$ contained in the exact position of the target before the saccade, which we call $x'$ [i.e., $I(x, \xi') \leq I(x, x')$]. The equality would be achieved if $\xi'$ gave enough information to determine $x'$ with perfect precision. While this is clearly not the case for the retina, the precision with which eccentricities can be determined by the eye is much greater than the precision of saccades, which have standard deviations of a few degrees. Thus, it is a good approximation to take $I(x, \xi') = I(x, x')$.

The quantity we want to maximize is the average information gained each time a target is detected,

$$I = \sum_n \pi_n I_n$$ 

where $\pi_n = \int \pi(x, n) \, dx$ is the probability that $n$ saccades have been made to the target so far, and $\pi(x, n)$ is the probability that the target is at position $x$ and $n$ saccades have been made. Our aim is to determine the function $\xi(x)$ (or $D(x)$) that maximizes $I$, so if we use the approximation $I(x, \xi') = I(x, x')$, we can ignore the terms $I(x, x')$, which do not depend on $\xi$ or $D(x)$ at all:

$$I \approx \sum_n \pi_n H_n + \text{terms independent of } \xi \text{ and } D$$

The next step is to determine the probabilities $\pi_n(\xi)$. If we consider a small region between $x$ and $x + \Delta x$, the number of cones in that region is $D(x) \Delta x$. The probability that a target is in that region, given that $n$ saccades have been made, is $\Delta x \pi(x, n)/\pi_n$, and so the probability that a particular cone detects the target is

$$\pi_n(\xi) = \frac{\pi(x, n)}{\pi_n D(x)}.$$ 

Combining this with Equations 17 and 20, we obtain

$$I = \sum_n \int \frac{\pi(x, n)}{D(x)} \log \left[ \frac{\pi(x, n)}{D(x)} \right] D(x) \, dx + \ldots$$

$$= -\sum_n \int \pi(x, n) \log D(x) \, dx + \ldots$$

where "..." indicates terms that are independent of $D(x)$.

We have to find the function $D(x)$ that gives the maximum value of this $I$, with the constraint that the total number of cones is fixed, so $\int D(x) \, dx = N_{\text{cone}}$. The optimal $D(x)$ is the function for which the following functional derivative is 0:
where $\lambda$ is a Lagrange multiplier. Thus the optimal distribution of cones is $D(x) \propto \sum_n \pi(x,n) = \Pi(x)$.

\section*{Acknowledgments}

We are very grateful to J. Stahl for providing us with data from his experiments. We also thank the two anonymous reviewers for their very constructive comments. AL and LZ are supported by the Gatsby Charitable Foundation. RG is supported by the Royal Society. Commercial relationships: none.

\section*{References}


