Introduction

Last lecture we noted that a central challenge to visual perception is dealing with variability in the image—the image of a signal (say your friend’s face) is never quite the same from one time to the next. In this lecture, we’ll start off with what is probably the simplest case of variability in vision—how can one discriminate a slight change in brightness of a spot given that from trial to trial the number of photons landing in the eye is not consistently the same? Or even simpler, what are the limits to just detecting a spot of light? This question was addressed in a classic study by Hecht, Schlaer and Pirenne, in what is probably the best, if not first, study that combined psychophysics, biology, and computational theory in vision. We will take a look at their experiment with the following goals in mind:

- Devise a computational theory of detection and discrimination, called Signal Detection Theory (SDT)
- Quantify limits to the eye’s information capacity
- Learn about the human eye
- Preview of how to generalize SDT to Pattern Theory, the study of perception as inference.

What are the fundamental physical limitations to vision?

Paul Dirac, one of the principal architects of quantum mechanics, was known for being laconic. He made a famous statement about photons that implies the fundamental limits to spatial resolution, and visual light intensity discrimination:

“Each photon interferes only with itself. Interference between photons never occurs.”

These two sentences imply that: 1) Interference or diffraction limits the spatial resolution of any imaging device, including the human eye; 2) Statistical independence of photon emission and absorption limits detection and discrimination of light intensity. In this lecture, we study the second limitation. Within a couple of decades of the establishment of quantum mechanics in the 1920s and the quantum basis of light, it was natural to ask: “What is the least number of photons a human observer can detect?” Hecht, Schlaer, and Pirenne, then in a biophysics lab at Columbia University set out to answer this question which was published in a famous paper in 1942.
numberofphotons[mean_] := Random[PoissonDistribution[mean]]; dotsize = 0.04;

Now let's simulate a trial presentation of a light flash. Select the following cell and evaluate it:

brightmean = 4; numbrightsample = numberofphotons[brightmean]; brightsample = Table[{Random[], Random[]}, {numbrightsample}];
brightg = Graphics[{PointSize[dotsize], Point[brightsample], AspectRatio -> 1, Frame -> False, FrameTicks -> None, Background -> GrayLevel[0.0], PlotRange -> ((-.2, 1.2), (.2, 1.2)), PlotLabel -> "Photon Count: " <> ToString[numbrightsample]; Show[Graphics[brightg], Frame -> False];

Photon Count: 5

Now, select it again, and repeat 10 times. Keep count of how many times you saw ANY dots at all.

Did you always see dots? If not, why not?—we will get into the answers to this question later.

What fraction out of 10 was it? This fraction would correspond to a psychophysically measured dependent variable. We will call it %yes responses. (Later on, when we introduce signal detection theory, we will call it the "hit rate").

So what is threshold?

We see that there is not a specific value of the independent intensity variable (# of quanta) at which the light goes from being invisible to visible. So we arbitrarily decide to define threshold as: "the intensity at which the %yes responses is 60%".

Hecht et al. measured the detectability of flashes of light in a dark room. Before running the experiment, they looked for the optimal stimulus conditions (size, duration, position, wavelength) and state of adaptation of the subjects. The goal was to measure the lowest possible threshold.

Before getting started, let’s get an overview of the visual anatomy to see where the eye is in relationship to the visual pathways.

Human Visual Anatomy: Brief Overview of the "front-end"

Light travels through the eye to the retina where it gets transduced to nerve impulses that leave the eye by the optic nerve to the lateral geniculate nucleus (branches also off to superior colliculus), and from there to the visual cortex.

For the next few lectures, we are going to concentrate on the eye itself. The light enters the eye at the cornea which does the bulk of the refraction, passes through the pupil formed by the iris, through the lens which controls accommodation (if you are young enough), and through the tissue of the retina to the photosensitive parts of the receptor cells called the outer segments.
The experiment of Hecht, Schlaer & Pirenne

Optimal conditions for detection

In order to find the least number of photons that could be seen, it was essential for Hecht, Schlaer, and Pirenne to find the optimal conditions for detection. The needed information about the optimal:

- state of the subject, i.e. adaptation state
- flash placement on the retina
- flash size
- flash duration
- wavelength of the light

Adaptation of the subject

After coming into a dark room after being outside on a bright day, it takes awhile to be able to detect changes in brightness, and your sensitivity gradually improves as you get used to the dark. This is called dark adaptation.

Full dark adaptation can take up to 30 minutes, and involves a change from cone vision (photopic) to rod vision (scotopic). The theory of two receptor systems goes back to the 19th century and is called duplicity theory. It also involves the regeneration of rod photopigment which gets bleached in response to light.

Much psychophysical work has gone into characterizing human dark adaptation. The figures show psychophysical measurements of human threshold to a flash of light as a function of time the subject has been sitting in the dark. Psychophysicists either present their dependent variable as a threshold (e.g. number of photons), or as the reciprocal of threshold called sensitivity.

What is the mechanism of adaptation? At the time of this experiment (1940s), it was thought that the percentage of unbleached rhodopsin molecules directly reflected sensitivity. Rhodopsin has a half-life of 5 minutes, meaning that if all of the molecules are bleached by light, then half of them would be restored in 5 minutes (75% in 10 minutes, etc.). We now know that adaptation involves retinal processes other than just pigment bleaching (see Figure below). Note the hundred-fold drop in threshold (increase in sensitivity) between 15 minutes and 30 minutes, even though 90% of the rhodopsin has already been regenerated.
Hecht, Schlaer, and Pirenne decided that the observer should be completely dark adapted— in other words, the poor subjects had to sit in a dark room for over a half hour before the data collection could begin.

**Flash placement**

The next consideration for the Hecht, Schlaer, and Pirenne experiment is where to put the image of flash on the retina, or in other words: where should the subject look?

There are two spots to avoid. First, one must avoid the blind spot in the eye, the region where visual information leaves the eye en route to the thalamus (and other destinations) of the brain. This is the optic disk, and source of the optic nerve. It consists of axons of the retinal ganglion cells (we will say more about these cells later). If a small light flash is restricted to in this region, it will not be detected. It is curious to note that the blind spot wasn’t discovered until 1668 by Mariotte—despite the fact that it is right before your very eyes! Because the brain completes missing information through the blind spot, you can only detect it in your own eye by carefully looking for it. Mariotte had reason to look for it in his own eye, because anatomists had noticed that when they dissected eyes, there was this curious white spot where the optic nerve joined the eye at about the same location in all the eyes.

**Photoreceptor density vs. eccentricity**

But, perhaps somewhat surprisingly, the observer shouldn’t be looking directly at the flash either. A star is more likely to be seen if you look just to the side of it, rather than directly at it (in which case it would fall on the fovea). This is because the maximum concentration of rods is not at the fovea, where there are virtually no rods, but at about 20 deg of visual angle away from the fovea (see figure).
Hecht et al. decided to place the flash at 20 degrees of visual angle to the left of the fixation point (the blind spot being about 15 degrees to the right of fixation) for right eye viewing. This would avoid the blind spot, and place the flash close to where the density of rod receptors is highest.

By the way, vision scientists measure size in terms of visual angle, rather than meters or feet. This is because visual angle is proportional to size on the retina, and is easily measured if one knows the size and distance of an object. Visual angle \((a)\) is defined in the following way:

\[
\frac{\text{Flash size}}{\text{How large should the spot of light be? Before answering this question, we need to understand something about how size is measured in this kind of psychophysics. We really want to know the size on the retina, not the size of the external stimulus on, say a screen. There isn't an really easy direct way of measuring how big an image is on the retina; however, we can easily measure visual angle—which is proportional to retinal size.}}
\]

Hecht, Schlaer, and Pirenne decided that a spot with diameter of less than 10 min would guarantee that detection would not be limited by the eye's inability to spatially integrate light energy. Note that the degree of spatial integration does depend on where in the retina it is measured, state of adaptation, and duration. Hecht et al. used a size suitable for the conditions of their experiment.

There is little difference in the number of quanta required for detection of spots of light of various sizes as long as their diameters are less than 10 min (1/6 deg). 1000 quanta spread over 1 min is as easily detected as 1000 quanta spread over 10 min; but 1000 quanta spread over 1000 min requires more quanta to be detected. This is called spatial integration.

Hecht, Schlaer, and Pirenne used a 1 millisecond duration for their flash.

Analogous to spatial integration, the eye also shows temporal integration. In this case, there is little difference in the threshold in terms of the number of quanta for a durations under 100 milliseconds. 1000 quanta is detected as easily at 1 ms and at 100 ms.

Hecht, Schlaer, and Pirenne used a 1 millisecond duration for their flash.
What should the wavelength of the light be? Psychophysical measurements again provide the answer.

As you can see from the figure, human thresholds vary with wavelength. Different wavelengths of light vary in the efficiency with which they are transduced, and thus detected. This is a direct consequence of the absorption properties of the photoreceptors. For scotopic (rod) vision, around 500 nm is optimal (Hecht used 510 nm). For photopic vision, 555 nm is optimal.
Summary of experimental conditions

To summarize the conditions chosen by Hecht et al.:

- The subjects were dark adapted and viewed a disk of light against a dark background.
- The flash was placed 20 degrees to the left of the fixation mark (right eye).
- The flash size was 10 minutes of arc.
- The duration of the flash was 1 ms.
- The wavelength was 510 nanometers.

Hecht et al. did not make the measurements necessary to determine these conditions themselves. The data to make the decisions for their experiment were all in the existing literature of the time. If we repeated the experiment today, we could make some minor improvements in their choices, but the results would not be substantially different.

Results of the experiment

- Threshold measured

For the criterion of 60% yes responses, about 90 quanta were required at the cornea.

What is the significance of a threshold of 90 quanta?

Of the 90 quanta, about 3% is reflected from the cornea. Then 50% of the remaining light is actually transmitted through the vitreous fluid of the eye. Finally, 20% of this is lost in the spaces between receptors (not transduced). The result is that they concluded that only about 9 photons are actually required at the receptors to "see" a flash.
Hecht et al. estimated that a single rod can process a single photon. Because the 9 photons are spread over about 500 rods, the chances that 2 photons are absorbed by one rod is very small. (Current estimate of the number of rods in this part of the retina in a 10° diameter spot is closer to 350). In any case:

One photon is sufficient to activate one rod.

Four decades later: “Going inside the box”

Almost 40 years later, the basic results of this experiment still stand. Baylor et al. were able to suck single toad rods into electrodes and measure current responses to single photon events (see figure)
Preview: Computational Theory for Light Discrimination

Why was the subject response curve sigmoidal, rather than a step function? Variability may be due to the statistics of photon emission and absorption, human responses, and neural transmission. Signal detection theory was developed 1940's and 1950's to describe the ability to detect, discrimination, and estimate signals in the presence of variability or "noise".

Let's return to the light detection example, but generalize it to light discrimination.
In[9]:=  

\[
\text{brightmean} = 7; \text{numbrightsample} = \text{numberofphotons[brightmean]}; \\
\text{darkmean} = 5; \text{numdarksample} = \text{numberofphotons[darkmean]}; \\
\text{brightsample} = \text{Table}[[\text{Random[]}, \text{Random[]}], \{\text{numbrightsample}\}]; \\
\text{darksample} = \text{Table}[[\text{Random[]}, \text{Random[]}], \{\text{numdarksample}\}]; \\
\text{brightg} = \text{Graphics}[[\text{PointSize}[0.04], \text{Point} /@ \text{brightsample}], \text{AspectRatio} -> 1, \text{Frame} -> \text{False}, \text{FrameTicks} -> \text{None}, \text{Background} -> \text{GrayLevel}[0.0], \text{PlotRange} -> \{\{-0.2, 1.2\}, \{-0.2, 1.2\}\}]; \\
\text{darkg} = \text{Graphics}[[\text{PointSize}[0.04], \text{Point} /@ \text{darksample}], \text{AspectRatio} -> 1, \text{Frame} -> \text{False}, \text{FrameTicks} -> \text{None}, \text{Background} -> \text{GrayLevel}[0.0], \text{PlotRange} -> \{\{-0.2, 1.2\}, \{-0.2, 1.2\}\}]; \\
\text{Show}[[\text{GraphicsArray}[[\text{brightg}, \text{darkg}]], \text{GraphicsSpacing} -> 0.3, \text{Frame} -> \text{False}];
\]

High Flash: 7  
Low Flash: 5

Histograms for bright and dim photon counts:

Suppose we compile a histogram of photon counts conditional on the switch setting:

Define bright and dim models:

In[15]:=  

\[
\text{brightpdist} = \text{PoissonDistribution}[7]; \\
\text{dimpdist} = \text{PoissonDistribution}[5]; \\
\text{PDF}[\text{dimpdist}, x] 
\]

Out[17]=  

\[
4^x e^{-5} x!
\]

In[18]:=  

\[
\text{sample[ntimes_] := Table[Random[dimpdist], \{ntimes\}]; \\
z = \text{sample[1000]}; \\
\text{domain} = \text{Range}[0, 30]; \\
\text{Freq} = \text{Map}[\text{Count}[z, \#], \text{domain}]; \\
\text{dimp} = \text{ListPlot}[N[\text{Freq}/\text{Length}[z]], \text{PlotJoined} -> \text{True}, \text{PlotStyle} -> \{\text{RGBColor}[1, 0, 0]\}, \text{DisplayFunction} \rightarrow \text{Identity}]; \\
\text{Show}[[\text{dimp}, \text{brightg}], \text{PlotRange} -> \{0, 3\}, \text{DisplayFunction} \rightarrow \$\text{DisplayFunction}];
\]

The point is: photon detection is inherently statistical, and in fact because of the independence of photon absorption, the histogram can be modeled as a Poisson distribution. More on this next time.

Discrimination as inference

How do quantum fluctuations limit discrimination?

Inference is choosing an hypothesis based on data. In statistical inference, the decision is based on a model of the probabilities of the hypotheses, \(H\) and the data, \(k\). What are our hypotheses? Think of the task in terms of signal transmission. The sender wishes to set the intensity of the light switch to one of two positions, corresponding to a bright and dim light. The hypothesis space is simple: \(H = S_1\) (dim), or \(H = S_2\) (bright).

The idea is that we have two hypotheses that each determine a conditional probability distribution.

If we have a high average photon count, and if the means of the two are close, these two distributions would look like this.

Thus, given a measurement of \(k\) photons, there is an inherent ambiguity in determining the cause or signal—\(S_1\) or \(S_2\) Next time, we will study the signal detection theory, and then see how to extend the fundamental principles of SDT to general perceptual inference for general image patterns and visual tasks.
References


© 2000 Daniel Kersten, Computational Vision Lab, Department of Psychology, University of Minnesota.
(http://vision.psych.umn.edu/www/kersten-lab/kersten-lab.html)