

My assignment is to talk about pattern vision and natural scenes

Pattern Vision

The study of pattern vision is concerned with understanding the eye and brain mechanisms underlying detection, discrimination and encoding of spatial patterns.

Before, getting down to details there are a few general points to make, most of which have been presented earlier in the course, either implicitly or explicitly.

Overarching Principle

Vision systems evolve to obtain information about the environment that is relevant for the tasks the organism must perform in order to survive and reproduce.

Corollary: The design of a vision system is constrained by the tasks it performs, by the physical/statistical properties of the environment, and by various biological factors.

Approaches to Understanding Vision

Natural scene statistics and natural tasks

Anatomy

Responses of individual neurons

Responses of neural populations

Perceptual/behavioral performance

Mathematical and computational modeling

Why Measure Natural Scene Statistics?

Characterize natural stimuli

Identify sources of stimulus information available for performing natural tasks

Determine the dynamic range, reliability and utility of the stimulus information

Generate hypotheses for visual mechanisms that might exploit the stimulus information

Design experiments to test for the hypothesized mechanisms



It is useful to distinguish between two general types of natural scene statistics, which can be measured at various levels along the pathway from environment to behavior. Absolute statistics are useful for understanding coding and representation. However, they say nothing about the relationship between the different levels and hence they are not as useful for understanding the information relevant for specific tasks. In most natural tasks the goal is to use the retinal image or some neural representation of the retinal image to make inferences about properties of the physical environment. In other words, the statistics one needs to know is the probability of different states of the environment given the image. These are what I am calling Bayesian statistics.



Most natural tasks involve dealing with one or more of these difficult general problems. Furthermore, the solutions that the visual system can come up with in natural tasks are constrained by various fundamental biological factors.

Approach to Handling Many of the Difficult Problems

- 1. Efficiently encode the attributes of retinal images in small regions with a foveated visual system.
- 2. Combine the measured local attributes into groups, categories or objects using mechanisms based on the physical laws and statistical facts of natural scenes, and on past experience.

The first lecture will say a little about how the initial encoding of retinal images contributes to pattern vision. The second lecture will describe a little about how local image properties are combined (grouped and segmented) and interpreted, with a focus on the measurement of relevant natural scene statistics (because relatively little is known about the relevant brain mechanisms).



The visual pathway from the eye to the primary visual cortex.

Competing Goals for Eye Design

Maximize spatial resolution

Maximize field of view

Minimize neural resources

Solution

High resolution foveal vision

Low resolution peripheral vision

Rapid eye movements

(Gaze contingent display demo.)



Ganglion cell density falls precipitously with eccentricity; faster than the falloff in cone density (based on data from Curcio & Allen 1980).



Ganglion cell receptive fields have a center-surround organization and scale in size so that retinal image is fully covered. As described earlier in the course, this organization helps to reduce the redundancy in the neural representation of natural stimuli, and helps (along with other adaptation mechanisms) to prevent the loss of information by keeping neurons within their operating range.



Retinal ganglion cells respond in a non-linear way with contrast. As described earlier in the course, this may reflect relatively efficient coding that is roughly matched to the typical distribution of contrasts in natural scenes. The nonlinear properties include point nonlinearities and contrast gain control.



What are the consequences of this retinal processing for pattern vision? One way to proceed is to put the pieces together from the anatomical and physiological studies of the macaque retina (and LGN) in order to simulate the responses of the retina in pattern vision experiments. From these simulations one can compute performance in pattern vision experiments assuming the output of the retina is used with perfect efficiency by the brain. This ideal performance (if simulation components are correct) precisely measures the limitations imposed by the eye. Comparison with behavioral performance measures the efficiency of brain mechanisms. This is what some of us call an ideal observer analysis.



A sensible starting point for this analysis is detection tasks in uniform backgrounds. Standard behavioral measurements are contrast thresholds for localized targets.



It is relatively easy to compute the approximate optimal performance given the retinal outputs, if we assume the noise in different neurons is statistically independent (which will not be precisely correct; but appears to be approximately correct).



Contrast sensitivity functions for sine wave gratings are an interesting case because they are known to be similar for humans and macaque monkeys (see DeValois & DeValois, 198x).



Contrast sensitivity data of Robson & Graham (1980) compared with the oneparameter (efficiency parameter) predictions of the retinal ideal observer analysis. (100 ms presentions with a fixed number of cycles)



The predictions can also be plotted as contrast sensitivity functions at various eccentricities.



Contrast threshold as a function of spatial frequency for brief stimuli (e.g., 100 ms) and for long stimuli (e.g., a couple of seconds). Typically the eye jumps from one location to another 3-4 times per second, so the contrast sensitivity function that is most representative of normal conditions does not have much roll-off at the low spatial frequencies. The shape of the CSF also depends on the number of cycles of the grating—an inverted U is more likely with a fixed target area (i.e., variable number of cycles).



The resolution of the human visual system declines rather rapidly with distance from the center of the fovea, as you can see from this figure (explain diagram). The resolution is the period of the cutoff frequency of the CSFs.



What happens when this analysis is applied to detection tasks other than sine wave gratings? Here are targets used in the ModelFest project.



Performance in simple detection experiments on uniform fields seems to be largely explained by retinal factors (uniformly efficient processing by the brain). However, once we leave the realm of simple detection on uniform backgrounds, brain efficiency is no longer approximately constant.



Measuring tuning functions of cortical neurons.



Quantitative characterization of cortical cell response properties using moving (drifting) sine wave gratings.



Summary of cortical neuron responses to sine wave grating stimuli (Geisler & Albrecht 1997).



The behavioral contrast sensitivity function reflects the combined neural contrast sensitivity of many cortical neurons, each tuned to a particular range of frequencies. Similarly (but not shown), behavioral contrast sensitivity as a function of orientation reflects the combined neural contrast sensitivity of many cortical neurons each tuned to a particular range of orientations. This "multiple channels" hypothesis is plausible if the neurophysiological description of primary visual cortex is accurate, but is there any independent behavioral evidence for this hypothesis?



Pattern masking demo. These masking effects cannot be explained by retinal mechanisms but are consistent with what one would expect from properties of neurons in the primary visual cortex.



Vertical masker plus low contrast horizontal target grating.



Vertical masker plus low contrast slightly tilted target grating. The target contrast is the same as in the previous slide.



Vertical masker plus slightly tilted low contrast grating two octaves higher in spatial frequency. The target contrast is the same as in the previous slide.



Cortical neurons respond in a nonlinear fashion; some of this nonlinear behavior they inherit from the retina, some is generated in the cortex. One of their major nonlinear behaviors is consistent with a multiplicative contrast normalization mechanism (Heeger 1991; 1992; Albrecht & Geisler 1991; others); another is consistent with an expansive (accelerating) nonlinear point nonlinearity (Heeger 1991; 1992; Albrecht & Geisler 1991; others).



Typical model of single neuron responses in primary visual cortex. Add to this model selective surround suppression and you have the model that Tony described on Friday. Such models predict some, but not all, pattern masking phenomena. There is a large literature on this that I will not attempt to summarize.



To return to eye design implications. Saccade lengths tend to be larger than decorrelation distances (the distance over which contrast is correlated) in natural scenes. The arrow indicates the average decorrelation distance in natural outdoor scenes. We have measured saccade lengths in search experiments with fairly small displays and even then the mean saccade lengths is greater than 3 deg of visual angle. Fixation lengths are typically 200-300 ms in many tasks. The conclusion is that the time course of the rapid contrast gain control mechanisms (and some luminance gain control mechanisms) must be very fast if it is to be useful under natural conditions. Albrecht, Geisler, Frazor & Crane (2002) find that, in fact, contrast normalization is very fast (and should be clearly distinguished from slower contrast adaptation).