Lecture 6 - Perception

The abstract: You have a problem. The goal of vision is to provide you with a description of the external world. However, the input to the visual system is the image that is formed on your retina. It contains information about the world but it is not a simple copy of the world. You need to do a lot of work to get from input to perception. Today, we will talk about several of the tasks that must be accomplished. You need to distinguish between your motion and motion in the world. You need to figure out how isolated bits of the image go together. You need to ignore irrelevant information (e.g. the light sources). You need to infer the most probable world even though there is an infinity of possible worlds that could produce the stimulus on the retina.

Why is it hard to go from the retinal image to visual perception? Three example problems

> The input is a flat (2D) image but the world is 3D

> The input is an interaction of photons and photoreceptors.
> > You want to know the properties of surfaces.

> And the image is sliding around on a moving retina in a moving head

Let start with that

Task 1: Establishing a framework

> DEMO: retinal vs world orientation

> DEMO: moving the eye (efferent copy)

> DEMO: VOR = Vestibular ocular reflex

The Dalmation Dog – You have all the features, now you need a theory

Task 2: What pieces go together?

> Joining line segments

> Dealing with occlusion

> Grouping: (sometimes called "Gestalt grouping rules")

Task 3: Finding Edges

> Why is there a problem?

> Why do you see contours that are not in the image?
> > (Called "Subjective" or "illusory"contours"

> > Conclusion: You infer the presence of contours based on your implicit understanding of the physics of the world

Task 4: "Discounting" the illuminant

> You (mostly) do not care about the light source

> BUT: The retinal stimulus = surface reflectance X illumination

You want to get rid of illumination.

SLIDE: Craik-O'Brien-Cornsweet Illusion

SLIDEs: Adelson's cylinder

SLIDES: Info from shadows (eg. Moony faces)

Task 5: Making the best guesses you can make.

One great example: How can you infer a 3D world from 2D input?

1)

2)

3)

4)

5)

…. And a few others

In summary:

    For all of these topics, the hard work is in the details.
    So, for example, how should you combine information from different sources??

And, if we have time, we will ask that question in the context of motion sickness.
If we don't…ask me sometime.

Questions to answer from Chapter Six:

P205ff - Like the lecture, the book goes into detail about depth cues. You will want to be able to say something intelligent about the difference between binocular and monocular cues (206).

P207 - and you will want to be able to describe a variety of depth cues

P210 - Why doesn't the world appear to jump when you move your eyes?

P213 - What is a template theory of recognition and why is it wrong?

P215 - What is an illusory conjunction and why is it interesting?

P216 - The topics of segmentation, figure/ground, and grouping are all part of the broader problem of parsing the visual stimulus into objects. Be able to say something intelligent on that topic.

P220 - The book is quite taken with network models of recognition. That's nice….

P223 - What the book is calling the "active role of the perceiver" and "the logic of perception" is closely related to "unconscious inference" as discussed in lecture.

P225 - It would not hurt to know the distinction between "magno" and "parvo" cells in the early visual system.

P227 - What is the division between "what" and "where" pathways all about?

P228 - What is the binding problem?

P229 - 232 Covers some of the material from the lecture on attention. I like my story better. But I am biased.

P232ff - What is meant by a perceptual constancy?

P234 - Ah ha, more unconscious inference. How is this related to the material back on p223ff?

P236 - Seeing and knowing….now there is a vexed topic. The attention lecture should have convinced you that a lot of what you think you are SEEING is really something you SAW (or attended to) some moments in the past. Is that the same as KNOWING. Maybe, but if I close my eyes, I still KNOW that my computer is in front of me. That KNOWING is clearly different than what I KNOW about the computer when I have my eyes open but my attention is directed to my copy of Gleitman….Isn't it? Tricky problem.

Something to write about, #6:Bayesian Vision

The formal way talk about making inferences in vision is to talk about making statistical inferences using Bayes' Theorem (or an approximation to it). I am not going to tell you all about Bayes' Theorem because that would take me too long. My current favorite LONG discussion is

An Intuitive Explanation of Bayesian Reasoning
http://yudkowsky.net/bayes/bayes.html
It is in the folder #6 BAYES THEOREM on the webpage BUT the applets didn't come with it. Try them on the web,

Here are four different articles using this approach on different bits of vision. All are in the folder. If you pick this assignment, you job is to explain the Bayesian approach in language that would be accessible to your grandmother, even if she is not a bio-statistician.

Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *J Opt Soc Am A, 14*(7), 1393-1411.

Ernst, M. O., & Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn Sci, 8*(4), 162-169.

Freeman, W. T. (1994). The generic viewpoint assumption in a framework for visual perception. *Nature, 368*(7 April 1994), 542-545.

Kersten, D., & Yuille, A. (2004). Bayesian models of object perception. *Current Opinion in Neurobiology, 13*(2), 150-158.