Week 1

Lecture 1

Introduction

- Perception is **apprehension of the world by means of our senses**
    - It determines what we think is real, and what we think exists (and what we believe to be true)
- Perception **does not need to give us a veridical** (true) **representation of the world out there**
    - Any representation that is sufficient to support our survival is good enough
- *e.g. The black and blue dress*
    - *It's an ambiguous stimulus - we don't know the* ***type, intensity or location of the light source***
    - *We either see a brightly illuminated black and blue dress or a shadowed white and gold dress*
- In a typical analysis of perceptual processing, we consider the stages of perceptual processing:
    - Physical stimulus
    - Receptors
    - Transduction
    - Sensory Coding
    - Perception
- Some textbooks say:
    - **Sensation:** the ability to detect a stimulus, and 'turn it into a private experience' (what?)
    - **Perception:** the act of interpreting and giving meaning to a detected sensation
    - But this is wrong… it's hard to differentiate the two

- Instead, **perceptual experience is hierarchical** in nature
    - In other words, it comprises **perception of local and global images**
        - *(e.g. the dolphins or the lovers vase)*
- But, both local and global structure in an image can be ambiguous and consistent with multiple interpretations
- Even if we see an object and can't recognise it, we may still be able to perceive that it has structure within it - hence **perception occurs independently of recognition**
    - Explicit recognition of something as a familiar object is **not necessary to perceive a structure** in an image
- There are instances where recognition is important, but more frequently in our interactions with the outside world we don't need to know very much more than the structure of an object
    - e.g. you don't care what the make of the car you're about to get hit by is

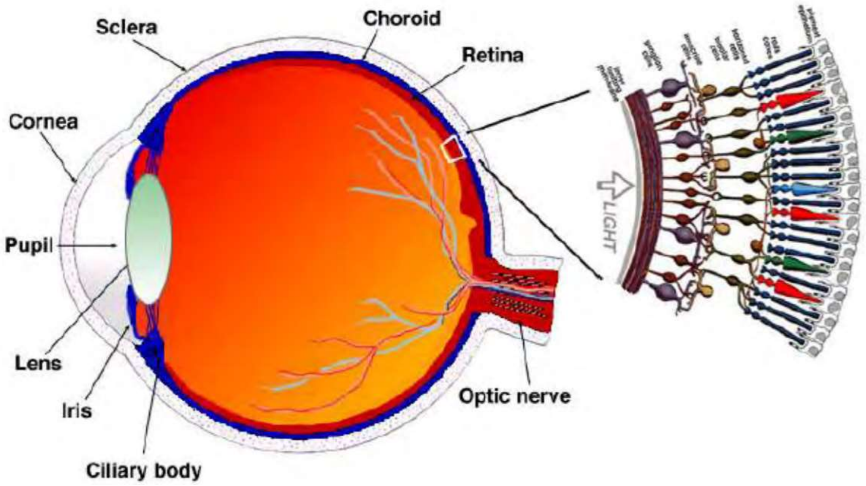
Characteristics of perception

- The **explicit knowledge of an image's structure cannot change our perceptual experience of that image**
    - e.g. think of all those optical illusions that still work, even when you know what the illusion is
- Perception is **immediate, fast, automatic, effortless** and quite sophisticated, even **very early in development**
    - *e.g. the train going through the block on the track surprising the three month old baby - which shouldn't even have the concept of object permanence*
- The simplest forms of perceptual information can convey an enormous wealth of information, and can have **emergent properties** supporting successful interaction with the environment
    - *e.g. just dots on a screen can convey the complicated video of a sad male walking slowly*

- So is perception easy or hard?
  - Even think about those Captcha tests or those "click all the pictures with a flower" things when you sign up for something - for us it's simple, for a computer it's very difficult
  - *"Attempts to construct computer models for the interpretation of arbitrary scenes have resulted in such poor performance, limited range of abilities and inflexibility that,* ***were it not for the human existence proof, we might have been tempted long ago to concluded that high-performance, general-purpose vision is impossible****" - Barrow and Tannenbaum (1978)*
    - i.e. the difficulty in getting computers to recognise letters and objects and things would lead us to believe that it's impossible, if it weren't for the fact that people can do it all the time
- The relationship between physical stimulus and its perceived properties is very complex
- But the question is - why is perception so easy to us humans and other biological organisms?
  - We have dedicated systems in our brain for this - visual cortex, etc.
    - **Modular organisation:** the visual brain consists of many different visual areas, which are functionally specialised to process and perceive different attributes of a scene
  - **Parallel processing:** different processing channels parse incoming signals into parallel streams to provide a compact efficient input for the brain
    - For everything we see, we need to process:
      - Luminance
      - Colour
      - Edges/form
      - Depth
      - Motion
    - Each parallel stream consists of **sequential processing stages** that increase in complexity - so it has a **hierarchical organisation**
    - Eventually though, all these parallel signals need to be elaborated and **integrated back into a unique and coherent percept** - how the brain does this is known as the **binding problem**

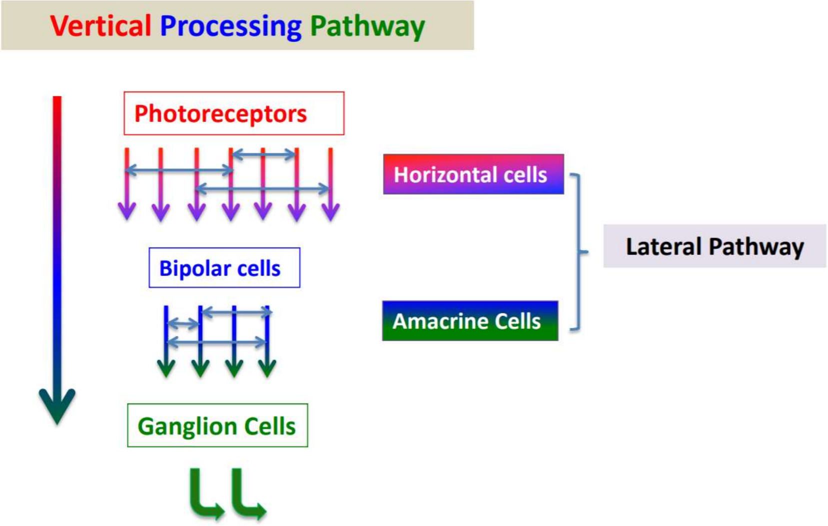
Lecture 2

The first steps in seeing

- The eye looks like so:
  - 

- Images get projected through the pupil on the retina, which has rods and cones which are processed and sent to the brain through the optic nerve

Retinal organisation

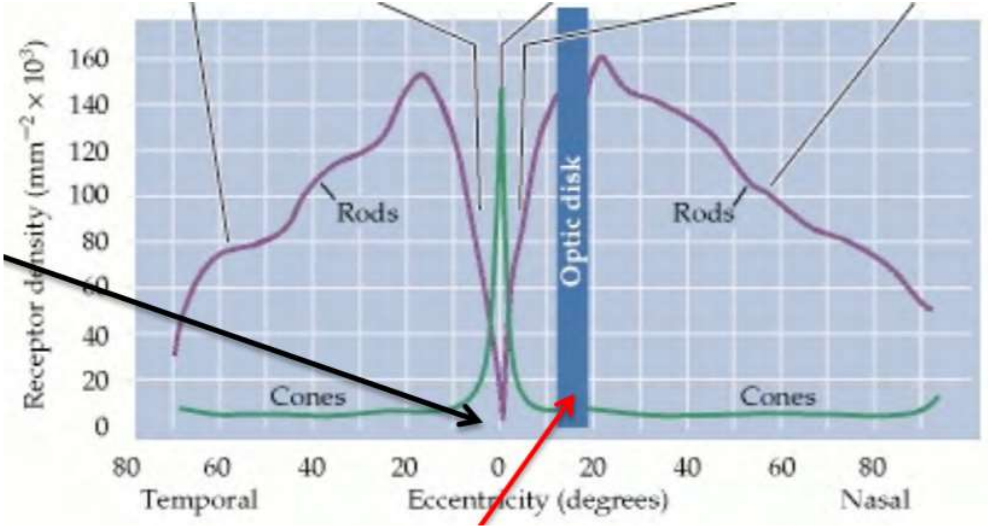
- The **retina** is 0.4mm thick
    - It has six layers; **three 'dark layers' of cell bodies** and **three 'light layers' of axons and synapses**
        - The input processing layer consists of photoreceptors (rods and cones), which are weirdly at the back of the whole structure
        - Then the middle processing layer contains **horizontal cells, bipolar cells and amacrine cells**
        - The output processing layer consists of **ganglion cells which transmit the signals to the brain**
- Schematically speaking, the process looks like this:
    - 

    - We have a **vertical processing pathway**, from photoreceptors down to ganglion cells (and off to the brain)
    - We also have a **lateral pathway,** which includes the horizontal cells and amacrine cells, which actually allow connections between individual cells
        - This is important for determining the structure of the information being sent

The input processing layer

- Photoreceptors consist of **rods and cones**:
    - **Rods** are thin and cylindrical, but **cones** tend to have wider bodies
- The outer segments of these receptors contain **visual pigment molecules**, which is where **visual transduction occurs**
    - **Visual transduction:** conversion of light energy into electrical signals
    - **Chromophores** are the bits of the visual pigments that catch the light
    - **Opsin:** a large protein which determines which light wavelengths the molecule absorbs
        - Cones have **photopsin** and rods have **scotopsin**
- Depending on the light exposure in the environment, the **retina undergoes pigment bleaching**, which results in decreased light sensitivity
    - This is why it's so difficult to see when we go from a bright area into a dark one - the retina is still in a process of pigment bleaching to decrease the amount of light getting in
    - Chromophore and opsin need to **recombine to regain light sensitivity**
        - Cone pigments can regenerate in 6 minutes, and rods take about half an hour
    - This process is called **dark adaptation**, and after about half an hour the retina is back at peak performance
        - At this point, it has the ability to detect an amount of light equal to the energy of a candle 60 kilometres away

Photoreceptors

- We have 120 million photoreceptors, and there are two types: rods and cones
- Rods and cones represent parallel pathways that operate under different light level intensities - this is referred to as the **duplex nature of the retina**
  - **Scotopic system:** our rods only tend to operate in instances of low light intensity, when high absolute sensitivity is required
    - e.g. stars, moonlight, dusk
  - **Photopic system:** our cones will only operate when low sensitivity is needed
    - e.g. daylight, searchlights
  - **Mesopic:** both rods and cones are in operation
    - e.g. office light
- Cones are packed very tightly in the fovea, and are widely spaced in the periphery:
  - 

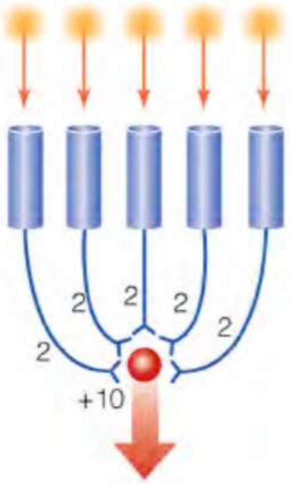
- The **optic disk** refers to the blind spot, the point at which the ganglion cells all reach the optic nerve to be transmitted to the brain - i.e. there are **no photoreceptors**

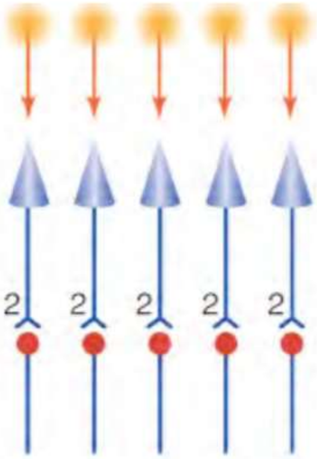
- Retinal diseases affect rods and cones separately:
  - **Macular degeneration:** fovea and small surrounding areas are destroyed - affects cones
    - Creates a central **scotoma** (blind spot) on the retina and is common in older individuals
  - **Retinitis pigmentosa:** genetic disease, progressively kills rods in the periphery and slowly moves in to the centre
    - Literally gives you tunnel vision and is very damaging for night vision

The middle processing layer

- Bipolar cells contact every other neuron type in the retina
  - We have about 10 million
  - They provide the link between the photoreceptors and the highly specialised output of the eye to the brain
- Bipolar cells are functionally specialised, i.e. there are different combinations:
  - Response selectivity to **increases vs. decreases in light**
    - Either **ON-**bipolar cells **or OFF-**bipolar cells
  - **Type of photoreceptors** they contact
    - **Cone** bipolar cells
    - **Rod** bipolar cells
  - **Location of photoreceptors** they receive input from
    - **Centre** bipolar cells (aka **midget** bipolar cells)
    - **Periphery** bipolar cells (aka **difuse** bipolar cells)

  - So, for example, you might have a periphery bipolar cell that connects to rods and responds to increases in light, i.e. a difuse on cell

- In the centre of the retina, **each foveal cone has direct access to two bipolar cells**, an ON and an OFF (midget) bipolar cell - this is known as **divergence**
- In the periphery, depending on the distance from the fovea, **up to fifty cone and rod receptors can converge to a single difuse cell**
  - This is known as **convergence**
- Interestingly, higher convergence, i.e. in the periphery, is associated with **better sensitivity**
  - 



  - *(Rods and cones are blue, bipolar cells are red)*
  - A bipolar cell with a high convergence rate (e.g. on the left) will generate a stronger response to the same level of stimulation, because of **threshold intensities**
    - Every bipolar cell would have a threshold of intensity required before it is activated to send a signal
    - If we assume that the threshold intensity in this example is 4 units, then the individual cells in the right example aren't activated and so don't send
- But, **higher convergence = lower spatial resolution**
  - In the example above, imagine you had light with intensity 4 activating the third and fifth rods/cones in the sequence
  - On the left example, you would get total intensity of 8, but wouldn't know where it came from
  - On the right example, you would get two lots of 4 and so have **greater visual acuity**
- So we have a tradeoff with convergence between sensitivity and resolution

The outer processing layer

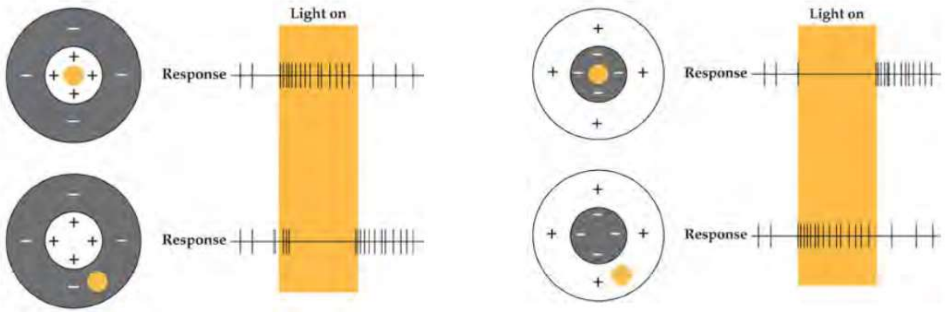
- The **central retinal pathway** uses **parvo** (midget) **ganglion cells**
  - Small cell bodies, short dendritic branching
  - Comprise 80% of all ganglion cells
  - Good chromatic sensitivity and slow conduction rate
  - Good with simulation
  - Used in **analysis of fine spatial detail**, and **colour vision**
- The **peripheral retinal pathway** uses **magno** (parasol) **ganglion cells**
  - Large cell bodies, long dendrites
  - Comprise about 10-15% of ganglion cells
  - Colour insensitive and fast conduction rate
  - Good with transient stimulation
  - Used in **motion perception**

- Retinal receptive fields
    - Generally speaking, each of the neurons in the various layers of the retina sample a specific area in the visual field
    - **Receptive field** refers to the area in space which, when occupied by an appropriate stimulus, affects the firing rate or **activation level** of the neuron
    - Convergence and lateral connections play a crucial role in forming **receptive fields for bipolar and ganglion cells**
        - Turns out they have concentric 'centre' and 'surround' receptive fields

- Kauffler (1953) found two types of ganglion cells which work with these receptive fields:

ON center-OFF surround
(Excitatory-center-inhibitory surround**)**

OFF center- ON surround
(Inhibitory-center-excitatory surround)



- So we see that each of these cells are **excited** by a particular area in the receptive field, and **inhibited** by other areas
- e.g. **ON centre-OFF surround cells** would be excited when there is stimulus in the centre, but are inhibited if stimulus is in the surrounding area, and vice versa for OFF centre-ON surround
- **Lateral inhibition** (aka opponency): stimulation with light excites one region and inhibits another

- For the first type, the ON centre OFF surround, if light covers the whole receptive field the signal would balance out and the firing rate may not be very different from if there was no stimulus - but this would be balanced out with signals from neighbouring cells

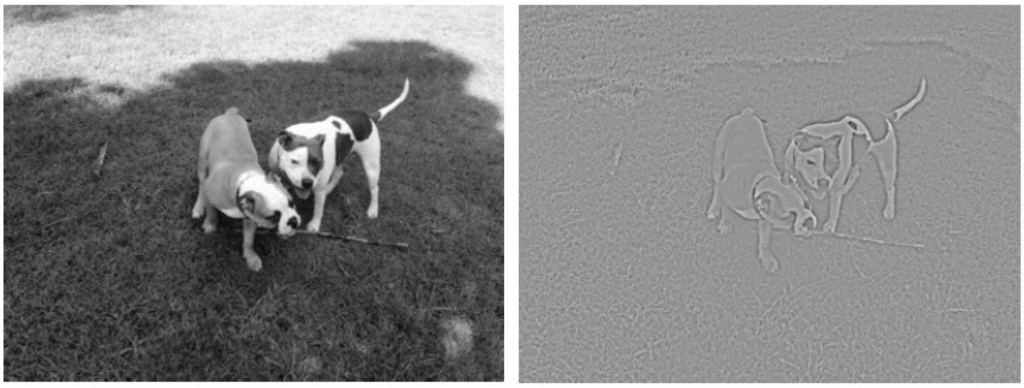
Week 2

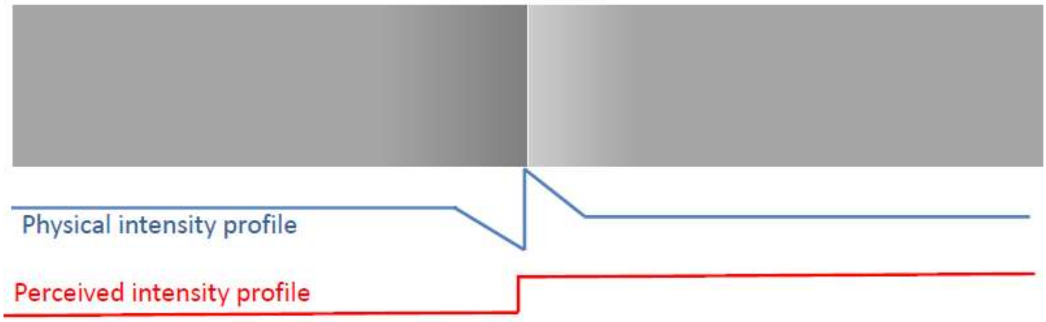
Lecture 1

Lateral inhibition

- The excitation in an ON-center cell would be maximal if the **stimulation only covers the centre and no surround**
    - As the stimulation increases in size and encompasses the whole region, lateral inhibition would cause a decrease in response rate
- With no stimulation or uniform stimulation, signals from neighbouring regions **balance against each other,** and so the neuron doesn't respond
- *Sincich et al (2009): resolving single cone inputs to visual receptive fields*
    - *Showing the firing rate of a neuron in relation to presented stimuli*
- Consequences of lateral inhibition:
    - **Retinal ganglion cells act as a filter**, responding best to stimuli that are just the right type of stimulation
    - Retinal ganglion cells are **most sensitive to differences in intensity between centre and surround**, rather than overall/absolute intensity
- Perceptually, these functional consequences result in
    - **Edge enhancement**
    - Exaggeration of luminance differences (**contrast enhancement**)
        - Mach bands - seeing borders more sharply
        - Simultaneous Brightness Contrast - seeing areas of same luminance differently
        - Hermann Grid - seeing spots at intersections

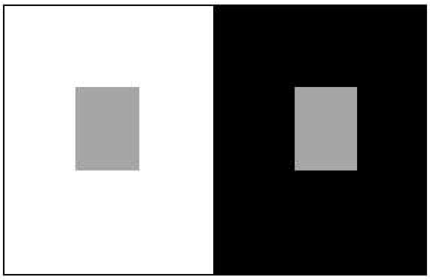
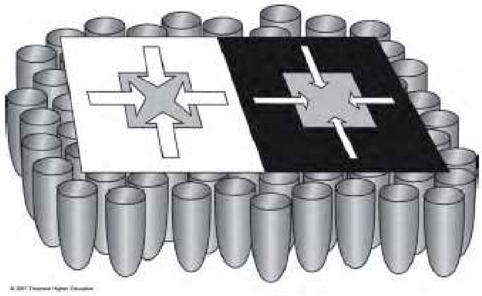
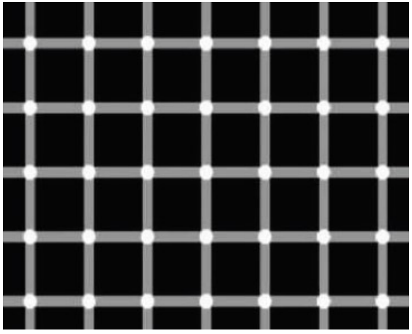
Importance of edges

- Edge enhancement means that **retinal ganglion cells transmit changes in intensities across the image**
    - There is high firing rate on the edges of borders, because it's more likely to have centre without surround (rather than all ON or all OFF)
    - 
        - e.g. on the right, grey refers to low firing rate and white/black refers to high
    - This is beneficial for:
        - **Image compression:** transmitting only information about changes in an image, and reconstructing the rest
        - **Image enhancement:** edge sharpening and de-noising
- We rely on edges for information about uniform regions between them

    - 

        - For example, in the above, apart from the bit in the middle the two sides are the same colour, but our brain applies the edge intensity difference to the entire thing
    - We process surfaces by **generalising the edge contrast to non-edge regions**

Importance of luminance differences

- **Simultaneous contrast:** seeing areas as different brightness due to adjacent areas
    -  
    - We tend to see the grey square on the right as brighter - the left one seems dimmer because it's ON-center OFF-surround are getting a lot more inhibition from the surround side
- Another classic example is the Hermann Grid Illusion
    - 
    - We see illusory spots at the intersections of stripes when we are seeing them in our periphery. But why?
    - The lateral inhibition perspective would say intersections are more inhibited than the single line edges, and hence seem darker
    -