Lecture – Introduction

- Perception Science
  - How do we know what's real
  - How does our nervous system take in what's going on in the world
- Naive realism
  - We sense what's there because it's what is really here
  - Perception is just about figuring out the codes
- We do not see the world, we construct the world
- There is no way our sensor systems can give us unvarnished view of what's out there. What we get is selective, biased and designed to keep us alive
- Judgments about sensation are uncertain
- Signal detection Theory
  - Developed in WII
  - For any given physical signal, or even if there is no signal, perception varies from moment to moment and it is normally distributed like this
  - the selection of a criterion has a relationship to the commission rates of the two types of errors.
    - If one goes down, the other must go up (misses and false alarm)
  - Criterion has nothing do with the shape of underlying distributions for noise and noise + signal, but only do with decision process

Sensation and Perception

- Sensation example: the ability to detect the pressure of a finger and, perhaps, to turn that detection into a private experience
- Perception: the act of giving meaning and/or purpose to detected sensations
- Mental life depends on sensation and perception has deep roots

- Methods of study of sensation and perception
  - Thresholds:
    - Methods of limits: start low and go high, start high and go low, average he results
    - Methods of constant stimuli: present a range of stimuli and just tally the yes's and the no's
    - Method of adjustment: get the participant to control the stimulus themselves
  - Scaling – measure private experience
  - Signal detection theory – measuring difficult decisions
    - how decisions like deciding which deceases one might get can be studied scientifically
  - Sensory Neuroscience
    - black pepper: you feel burn
  - Neuroimaging – an image of the mind
  - Development: the study of the changes over the life span

Thresholds and the Dawn of Psychophysics

- Gustav Fechner
  - Absorbed with the relationship between mind and matter – placed him between adherents of **dualism** and **materialism**
    - Dualism: hold that the mind has an existence separate from the material world of the body
    - Materialism: the mind is not separate
    - Panpsychism: the mind exists as property of all matter – including animals and inanimate things
  - He described his philosophy of panpsychism
  - Thought it should be possible to describe the relation between mind and body using math

- Goal: to describe the relationship between sensation and the energy that gave rise to that sensation
- psychophysics – his theory and his methods
- Absolute threshold: the minimum amount of stimulation necessary for a person to detect a simulus 50% of the time

- Ernst Weber
  - tested the accuracy of sense of touch by using a device
  - measure the smallest distance between two points that was required for a person to feel touch on two points instead of one.
  - Weber's most important finding: judgments of lifted weights
    - the ability of a person to detect the difference between the standard and comparison weights depended greatly on the weight of standard
    - When the standard is relatively light, people were much better at detecting a small difference when lifted a comparison weight
    - When the standard is heavier, people needed a bigger difference before they could detect a change
  - JND(Just noticeable difference); difference threshold
    - smallest change in weight could be detected was close to 1/40 of the standard weight.
    - smallest change in lengths of two lines: 1/100
    - Those ratios: Weber fractions
  - **Weber Rule**: the size of the detectable difference is a constant proportion (K) of the level of the stimulus
  - **Weber's Law**: describe the relationship between stimulus and resulting sensation that says the JND is a constant fraction of the comparison stimulus
  - **Fechner's Law**: $S = klogR$, where S is the psychological sensation, R physical stimulus level
    - Shows that psychological experience of the intensity of light, sound, smell, taste or touch increases less quickly than the actual physical stimulus increases

Psychophysical Methods

- method of constant stimuli: requires creating many stimuli with different intensities in order to find the tiniest intensity that can be detected
  - absolute threshold: the minimum amount of simulation necessary for a person to detect a stimulus 50% of the time
  - inefficient way to conduct an experiment
  - listener's time is spent with stimuli that are clearly well above or below threshold
- method of limits: the particular dimension of a stimulus or the difference between two stimuli is varied incrementally until the participant responds differently
  - tones presents in ascending order: asked to report when they first hear the tone
  - tones presents in descending order: the task is to report when the tone is no longer audible
  - take the average of crossover points (hear tone $\rightarrow$ nothing and vice versa)
- method of adjustment: participant controls the change in stimulus
  - the easiest method to understand
  - not usually used to measure thresholds; good for data in huge changes, not for one with slow changes
- magnitude estimation: participant assigns values according to perceived magnitudes of the stimuli
  - Stevens's power law:

$$S = aI^b$$

  - describes the relationship between stimulus intensity(I) and sensation
    - If stimulus intensity is less than 1, then even huge change in intensity, the sensation will not change a lot. e.g. sweetness, brightness
    - If stimulus intensity is greater than 1, then small changes in intensity will leave a huge change in sensation e.g. apparent length(1.0), electric shock

- Variant in scaling method – cross-modality matching
  - the ability to match the intensities of sensations that come from different sensory modalities
- supertaster: an individual who experience the most intense taste sensations; experience more intense oral burn and oral touch sensations.

Signal Detection Theory

- stimulus to detect is always being detected in the presence of "noise"
- Criterion: an internal threshold that is set by the observer
- Receiver operating characteristic curve
  - When d' increases, the probability of hits and correct rejections increases, the probability of misses and false alarms decreases

Fourier Analysis – mathematical tool

- Complexity of sound – sine wave
  - sine wave – a changing oscillation that repeats across space
  - wavelength – the distance required for one full cycle of oscillation for a sine wave
  - phase – a fraction of the cycle of the sine wave described in degree or radians
  - amplitude = 1 / frequency
  - Any complex sound can be broken down into individual sine wave components through breaking down into combinations of sine waves at many different frequencies with different amplitudes and waves

Sensory Neuroscience and the Biology of Perception

- Doctrine of specific nerve energies (Johannes Muller)
  - Doctrine – we cannot be directly aware of the world itself, only aware of the activity in our nerves
  - which nerve are stimulated is the most important
- Cranial nerve – leading into and out of the skull illustrate the doctrine of specific nerve energies

    - Sensory information: olfactory, optic and vestibulocochlear(both sense of equilibrium and hearing)
    - Dedicated to muscles to move the eyes: oculomotor, trochlear and abducens
  - Polysensory: processing extends beyond primary areas; information from more than one sense is being combined combined in some manner
  - Helmholtz
    - Disliked Muller's beliefs: vitalism (there is a force in life that is distinct from physical entities – violated the physical law of conservation of energy)
    - Believes: all behavior should be explained by only physical forces – obey the purely physical laws
    - Attack Vitalism by claiming that the nerve impulse could never by measured experimentally

Neuronal Connections

  - Ramøn y Cajal
    - suggested neurons do not actually touch one another
    - develop some insights into the organization of neurons in the brain
  - Sir Charles Scot Sherrington
    - Synapse(突触): the tiny gap between the axon(轴突) of one neuron and the dendrite(树突) of the next
  - Otto Loewi
    - Proposed that something chemical, instead of electrical, might be at work at the synapse
  - Neurotransmitter: molecules released by axon
    - drugs: increasing/decreasing the effectiveness of different neurotransmitter

Neural Firing

  - Sir Alan Hodgkin/ Andrew Huxley
    - Isolated a single neuron from the squid and tested how the nerve impulse raveled along the axon

    * Neural firing is actually electrochemical
    * by measuring different aspects of neurons firing – learn about how individual neurons encode and transmit information from sense organs through higher levels of the brain
    * Identify the stimulus makes it fire the most vigorously is one way to investigate what a neuron encodes

Neuroimaging

- EEG:
    * Measures electrical activity through electrodes(电极) on the scalp(头皮).
    * Does not allow to learn what individual neurons are doing or to pinpoint the exact area of neural activity
    * roughly localize whole populations of neuron
    * measure activities with excellent temporal accuracy
    * relatively simple and cheap
- EPR(event-related potential)
    * average all the responses aligned to the moment that the stimulus was present
    * measure of electrical activity from a subpopulation of neurons in response to particular stimuli
- MEG (magnetoencephalograph)
    * measure changes in magnetic activity across populations of many neurons in the brain
    * take advantage of the fact that neurons make small changes in local magnetic fields in addition to small electrical changes
    * expensive and complex
- CT (computed tomography)
    * use X-rays to create images of slices through volumes of material
    * dense tissue absorbs more energy than lighter tissue
    * the detector measures the amount of energy that has been lost on the way through the head

  - MRI (magnetic resonance imaging)
    * use the responses of atoms to strong magnetic fields to form images of structures (can be adapted to measure activity in the brain)
  - fMRI (functional magnetic resonance imaging)
    * measure localized patterns of activity in the brain by measuring changes in the response of oxygenated and deoxygenated blood to strong magnetic fields
    * BOLD: the ratio of oxygenated to deoxygenated hemoglobin
  - PET (positron emission tomography)
    * define locations in the brain where neurons are especially active by measuring the metabolism of brain cells using sate radioactive isotopes
    * the idea is to detect activity in neurons by looking for increased metabolic activity – using oxygen-15
      · the most active in the brain should have the greatest requirement for this oxygen

- Development over the Life Span
  - What comes with the system?
  - What has to be learned?
  - What changes with age?

- Weber's, Fechner's and Stevens's
  - Weber's law involves a clear objective measurement.
  - Fechner's law is a calculation based on some assumptions about how sensation works. Assume all JNDs are perceptually equivalent
  - Stevens's power law: describes rating data quite well, but notice that rating data are qualitatively different from the data that supported Weber's law. We can record the observer's ratings and check if it is reasonable, but we cannot tell if it is correct or wrong

Chapter 2 – Lecture Notes

- **Eyes**
  - Only certain amount of wavelengths is visualize
    - any shorter (UV) and there is less available in the world
    - any longer (microwaves) and too much radiation is absorbed by objects
  - Most visual information does not come from radiant objects but from reflections
  - Light is reflected and transformed by objects
  - The original light sensors were very basic
  - Many eyes can form an image of the external world
    - such eyes require a lens for focusing an image and a receptive sheet for cells for coding the image (the retina)
  - Optics
    - The job of a lens is to focus an image on a plane
    - Lenses vary in power
      - the focal length of a lens is the distance at which parallel rays of light can be focused to a point
      - stronger lens $\rightarrow$ short focal length
      - weaker lens $\rightarrow$ long focal length
    - Measured by diopter = 1/ focal length in meters
    - Human focusing power $\approx$ 40 diopters, added 15-20 diopters in a young person; that is we can focus on 1/60 meters = 16cm from the eye
- **Photoreceptor Transduction**
  - structure: outer segment(magic take places), inner segment, synaptic segment
  - presenting a flash of light to a photoreceptor produces an electrical response
  - Fewer than 10 photons produce a perceptible response, with 1 photon closes about 10^6 million channels
- **The duplex retina**

- Much active processing takes place in the retina.
- It is a pattern recognizing machine and already looking for the places where the most important information is likely to be
- receptors come in two flavours – rods and cones
- Cone: a photoreceptor specialized for daylight vision
  * tight packing in fovea helps with detail vision
  * wavelengths sensitivity helps produce colour vision
  * does not work well in the dark dimilluminations
- Rod: a photoreceptor specialized for night vision
  * Hecht, Schlaer & Pirenne (1942)
  *
- Other cells in the retina
  * Horizontal cells and amacrine cells make lateral interactions possible
  * Ganglion cell is the output cell
- Properties of retinal ganglion cells
  * Cells with centre-surround organization produce the strongest response when the centre is entirely filled with light and the surround is as dark as possible.
  * It is suited to detecting change or contrast
  * Receptive field size differences allow specialization for different scales
- Mach bands phenomena
  * can be explained by center-surround antagonism
- Different numbers of photoreceptors influence single ganglion cells
  * High convergence: Many inputs to one cell
    · sensitivity is gained at the cost of resolution
    · can have a lot of sensitivity, but low resolution
  * Low convergence: One to one
    · resolution is gained at the cost of sensitivity
- Regional specialization: high resolution and high sensitivity

    - visual systems consist of one system(cone) for doing resolution and rod for maintaining sensitivity
    - cones are in the fovea and rods not in the fovea

Reading

- Two ways to conceptualize light:
  - as a **wave**: an oscillation that travels through a medium by transferring energy from one particle or point to another without causing any permanent displacement of the medium
  - as a **photons**: a quantum of visible light or other form of electromagnetic radiation demonstrating both particle and wave properties
- Visible light waves have wavelengths between 400 and 700 nanometers
- **hue** the perceptual attribute of colors that enables them to be classed as similar to red, green or blue, or something in between
- Some of the starlight's photons are absorbed by encounters with dus, vaporized water; and some of light is scattered by there particles.
  - sky is blue → short-wavelength(blue) light is scattered more strongly than other wavelengths
- If the ray of starlight were to strike a light-colored surface, most of the light would be reflected.
- Most of the light striking a dark surface is absorbed. Light that is neither reflected nor absorbed by the surface is transmitted through the surface
- Eyes that Capture light:
  - An eye can form an image fo the outside world, enabling animals that possess eyes to use light to recognized objects. (including if light is present and its direction)
- Eye Structure:
  - Cornea(角膜): the transparent window into the eyeball; has no blood supply

    - cornea can heals within 24 hours even when it is scratched
  - aqueous humor(房水): a fluid derived from blood, fills the space immediately behind the cornea and the lens
  - lens: has no blood supply; the shape of lens is controlled by the ciliary muscle
  - pupil(瞳孔): the dark, circular opening at the center of the iris in the eye, where light enters the eye
  - iris(虹膜): the coloured part of the eye, consisting of a muscular diaphragm surrounding the pupil and regulating the light entering the eye by expanding the contracting the pupil
    - give the eye its distinctive color and controls the size of the pupil
    - the pupil of the iris plays an important role in the image quality
      - low illumination + pupil is large, the depth of focus is reduced, resulting in poor image quality
  - vitreous humour(玻璃体): the transparent fluid that fills the vitreous chamber in the posterior part of the eye
    - longest part of the journey through the eyeball
    - this chamber comprises 80% of the internal volume of the eye
  - retina: detect light and tell the brain about aspects of light that are related to objects in the world
    - transduction: light energy is turned into electrical neural signals

- Focusing light onto the Retina
  - cornea, aqueous humor, lens and vitreous humor must be perfectly matched to the length of the eyeball
  - Cornea is highly curved, it forms the most powerful refractive surface in the eye; aqueous and vitreous humors also help refract light
  - Accommodation: alter the refractive power by changing its shape; the process by which the eye changes its focus
  - Accommodation is accomplished through contraction of the ciliary muscle

    - When the ciliary muscle is relaxed, the zonules are stretched and the lens is relatively flat.
    - In order to look closer, ciliary muscle must contract.
    - Contraction reduces the tension on the zonules and enable the lens to bulge
    - The fatter the lens is, the more power it has, the closer you can focus
  - presbyopia: old sight, the age-related loss of accommodation which makes it difficult to focus on near objects
  - Cataracts: an opacity of the crystalline lens
    - most cataracts are discovered after age 50
  - Emmetropia: the refractive power of the four optical components are perfectly matched to the length of the eyeball
  - Refractive errors: occurs when the eyeball is too long or too short relative to the power of the four optical components
    - too long → myopia (近视)
    - too short → hyperopia (远视)
  - the most powerful refracting surface in the eye is the cornea which contributes about 2/3 of the eye's focusing power
  - astigmatism: a visual defect caused by the unequal curving of one or more of the refractive surface of the eye usually the cornea
  - transduce: to convert from one form energy to another
    - process of seeing begins with the retina, where the light energy from our star is transduced into neural energy that can be interpreted by the brain

- Doctor saw
  - Fundas: the back layer of the retina: what the eye doctor sees through an ophthalmoscope
    - it is the only pace in the body where one can see th arteries and veins directly
    - even with the other eye closed, the visual system "fills it in" with information from the surrounding area
    - fovea is located near the center of the macula

    - photomicropgraph reveals the retina is a layered sheet of clear neurons, about half the thickness of credit card (Rodieck, 1998)
    - the transduction of light energy into neural energy begins in the backmost layer of the retina, which is made up of photoreceptors
    - photoreceptors sense light → stimulate neurons in the intermediate layers, including bipolar cells, horizontal cells, and amacrine cells
    - photoreceptors stay at the back – this arrangement requires light to pass through the ganglion, horizontal and amacrine cells before making contact with the photoreceptors
- Retina contains: rods and cones
- Rods: are absent form the center of fovea and their density increases to a peak at about 20 degrees and then declines again
- Cones: the cones are most concentrated in the center of the fovea, and their density drops off dramatically with retinal eccentricity
- Duplex retina: retina contains rods and cones
- fovea vs periphery
    - photoreceptor type: mostly cones mostly rods
    - bipolar cell type: midget diffuse
    - convergence: low high
    - receptive-field size: small large
    - acuity: high low
    - light sensitivity: low high
- cones:
    - become larger and more spares away from the foveal center
    - require brighter illumination to operate efficiently
    - have one of three different type of photopigments in color that differ in the wavelengths at which they absorb light most efficiently – provide the basis for our color vision
- rods:

    - the small cells that appear outside the fovea
    - function relatively well under conditions of dim illumination
    - have the same type of photopigments; cannot signal differences in color

- Dark and Light Adaptation
  - dark room to bright: pupil constricts to reduce the amount of light arriving at retina
  - Sneezing: current thinking suggests that the photic sneeze reflex is a result of crossed wires in the brain
  - When entering a dark room from bright sunlight, the number of photons of light entering eye might be reduced by a factor of several billion
  - Four ways in which visual system adjusts to change illumination: pupil size, photopigment regeneration, the duplex retina and neural circuitry

- Pupil Size
  - the diameter of the pupil can vary by about a factor of 4, from about 2nm in bright illumination to about 8nm in the dark
  - pupil dilation has a time course of a few seconds

- Photopigment Regeneration
  - in dim lighting conditions, plenty of photopigment is available, and rods and cones absorb and respond to as many photons as they can.
  - As the overall light level increases, the number of photons starts to overwhelm the system: photopigment moleculs cannot be regenerated fast enough to detect all the photons hitting the photoreceptors
  - slow regeneration is a good thing for increasing sensitivity range

- The Duplex Retina
  - rods get overwhelmed when the background light becomes moderately bright, leading to a loss in information quality
  - use rods to see when the light is low

  - use cones to see when the light is too bright
  - cones are sensitive to bright light, they recover sensitivity quickly and then saturate
  - cones are not sensitive to very dim light.
  - rods recover more slowly, but after 20 mins, they are sensitive to dim light

- Neural Circuity
  - ganglion cells are most sensitive to differences in the intensity of the light in the center and in the surround of receptive fields, they are less affected by the average intensity of the light
  - receptive field: the region on the retina where visual stimuli influence the neuron's firing rate
  - ganglion cells will encode the pattern of relatively light and relatively dark areas in the retinal image

- Visual system deals with large variation
  - reduce the scale of the problem by regulating the amount of light entering the eyeball, by using different type of photoreceptors in different situations, and effectively throwing away photons we do not need
  - responding to the contrast between adjacent retinal regions, the ganglion cells do their best to ignore any variation in overall light level is left over

- Retinal information processing
  - Light Transduction by Rod and Cone Photoreceptors
    * outer segment: the part of photoreceptor that contains photopigment
    * inner segment: the part of photoreceptor that lies between the outer segment and the cell nucleus
    * synaptic terminal: the location where axons terminate at the synapse for transmission of information by the release of a chemical transmitter
    * Photoreceptor capture light $\rightarrow$ produce chemical changes $\rightarrow$ send signals by way of the synaptic terminals, specialized structures for contacting other

retinal neurons → synaptic terminal contains connections from the neurons to horizontal and bipolar cells

  * Photoreceptor contains: outer segment, inner segment and synaptic terminal

- Visual pigment molecules
  * made in the inner segment
  * stored in the outer segment
  * consists of a protein (an opsin), the structure of which determines which wavelengths of light the pigment molecule absorbs, and a chromophore (capture light photons)
  * the chromophore is the part of the pigment molecule that determines its color by selectively absorbing specific wavelengths of light
  * the pigment rhodopsin is found in the rods, concentrated mainly in the stack of membranous discs in the outer segment
- Another photoreceptor:
  * adjusting biological rhythms to match the day and night of the external world
  * sensitive to the ambient light level and contain the photopigment melanopsin
  * sned signals to the suprachiasmatic nucleus
- photoactivation: transfers energy to chromophore portion of the visual pigment molecule
  * result in the closing of cell membrane channels that normally allow ions to flow into the rod's outer segment
- hyperpolarization: a change in membrane potential such that the inner membrane surface becomes more negative than the outer membrane surface
- photoreceptors do not respond in an all-or-nothing fashion
  * the pass information on to bipolar cells via graded potentials(an electrical potential that can vary continuously in amplitude)
- Three cones photoopigments are not distributed equally among the cones
  * short wavelength-sensitive cones: constitute only about 5-10% of total cone population, missing from the center of the fovea (the foveal center is dichromatic)

    - long wavelength-sensitive cones twice as many as medium wavelength-sensitive cones

- Lateral Inhibition through Horizontal and Amacrine Cells
  - horizontal cells: perpendicular to the photoreceptors, make contacts between nearby photoreceptors
  - connections between them form of lateral inhibition
    - enables the signals that reach retinal ganglion cells to be based on differences in activation between nearby photoreceptor
    - play important roles in visual perception, as well as in several visual illusions
  - Amacrine cells: run perpendicular to the photoreceptors in the inner layers of the retina
    - receive input from (bipolar cells and other amacrine cells)
    - send signal to (bipolar, amacrine, and retinal ganglion cells)

- Convergences and divergence of information via bipolar cells
  - bipolar cells are the intermediaries – various type of bipolar cells, and their wiring determines the information that is passed from the photoreceptors to the ganglion cells
  - pooling information from many photoreceptors can increase visual sensitivity
  - the high degree of neural convergence in peripheral vision has important consequences for visual acuity(a measure of the finest detail that can be resolved by the eyes), which falls off rapidly with eccentricity
  - in the fovea, midget bipolar cells receive input from single cones and pass this information on to single ganglion cells
  - the low degree of convergence in the fovea ensures high acuity but poor sensitivity to light
  - Each foveal cone contacts two bipolar cells
    - ON bipolar cell: increase in light captured by the cone
    - OFF bipolar cell: hyperpolarize

  - ON and OFF bipolar cells respond differently – express different types of postsynaptic glutamate receptors
    * they provides information about if the retinal illumination increased or decreased

- Communicating to the Brain via Ganglion Cells
  - ganglion cell: a retinal cell that receives visual information from photoreceptors via two intermediate neuron types and transmits information to the brain and midbrain
  - P ganglion cell: a small ganglion cell that receives excitatory input from single midget bipolar cells in the retina
    * contains 70% of the ganglion cells in the human retina
  - M ganglion cell: resembling a little umbrella that receives excitatory input from diffuse bipolar cells and feeds the magnocellular layer of the lateral geniculate nucleus
    * diffuse bipolar cells project to ganglion cells
    * about 8-10 percent of ganglion cells in the human retina are of the M variety
  - The dendrites of both P and M ganglion cells increase in size with retinal eccentricity(离心率).
  - At all eccentricities the P ganglion cells have much smaller dendritic trees than do the M ganglion cells
  - Both M and P ganglion cells together constitute about 80 percent of all ganglion cells
  - Difference between M and P cells:
    * P cells have smaller receptive fields than M cells have
    * M cells listen to more photoreceptors; respond to a larger portion of the visual field
    * M cells are much more sensitive – better detect visual stimuli under low-light conditions
    * P cells provide finer resolution if there is enough light for the P cells to operate

    * Differ in temporal response
        · P cells tend to respond with changes in sustained firing while light shines on excitatory region
        · M cells respond more transiently; M cells respond with a burst of impulses when the spot is turned on, and then it will quickly return to its spontaneous rate
    * M and P ganglion cells signal different information to the brain
        · P cells provide information about the contrast in the retinal image
        · M cells provide information about how the image changes over time
- koniocelluar cell: a neuron located between the magnocellular and parvocellular layers of the lateral geniculate nucleus
- Center-surround Receptive Fields
    * Ganglion cells fire action potentials spontaneously, even in the absence of visual stimulation
    * the receptive field in the region on the retina where visual stimuli influence the neuron's firing rate
        · excitatory, increasing the ganglion's firing rate
        · inhibitory, decreasing the ganglion's firing rate
- Stephen Kuffler:
    * mapped out the receptive fields of individual retinal ganglion cells in the cat, using small spots of light
    * small spot of light on a projector screen, turning on and off while he recorded impulses from a single retinal ganglion cell
    * when the spot was placed on a specific small region of the retina, the ganglion cell increased its firing rate when the light is turned on (this area of retina is called "center" of the ganglion cell's receptive field)
    * when the spot moved to an adjacent area of the retina, the ganglion cell decreased its firing rate
    * ON-center cell: increases firing in response to an increase in light intensity in its receptive-field center

    * OFF-center cell: increases firing in response to a decrease in light intensity in its receptive-field center
    * A small circular area in the center responds to an increase in illumination, and a surrounding ring responds to a decrease in illumination, and size matters
    * The ganglion cell reduces its firing rate when the spot of light begins to encroach on its inhibitory surround.
- center-surround organization two important functional consequences
    * each ganglion cell will respond best to spots of a particular size
        · retinal ganglion cell act as a filter by responding best to stimuli that are just hte right size and less to stimuli that are large or smaller
    * ganglion cells are most sensitive to differences in the intensity of the light in the center and in the surround
        · the contrast – the difference in luminance or brightness between adjacent bits of the scene will be roughly the same regardless of lighting condition
- center-surround antagonism, lateral inhibition *rightarrow* resulting in the illusions of strips and spots

Reading

- Contrast: the difference in luminance between an object and the background, or between lighter and darker parts of the same object
- Rods and cones in the periphery are packed together less tightly and many receptors converge on each ganglion cell → visual acuity is much poorer in the periphery than in the fovea
- Visual acuity in peripheral vision is not uniform
  - Horizontal and vertical asymmetry: visual acuity falls off more rapidly along the vertical midline of the visual field than along the horizontal midline
  - Vertical meridian asymmetry: have better acuity a fixed distance below the midline of the visual field than above
- Difference between central vision and peripheral vision
  - central vision is considerably slower than peripheral vision
  - foveal cones have longer axons than peripheral cones – longer axons transmit slow signals better than fast ones
- Herman Snellen
  - invented the designating visual acuity in 1862
  - $\frac{\text{distance at which a person can just identify the letters}}{\text{distance at which a person with "normal" vision can just identify the letter}}$
  - patient with amblyopia, acuity with Snellen letters may be much more affected by acuity with gratings
- Visual Acuity – four of the most common definitions
  - minimum visible acuity
    - measured by detection of a feature
    - refer to the smallest object that one can detect
    - It is so small for two reasons
      - the optics of the eye spread the image of the thin line, making it much wider on the retina

        - the minimum visible acuity is limited by our ability to discriminate the intensity of the target relative to the background
    - minimum resolvable acuity
        - the smallest angular separation between neighboring objects that one can resolve
        - represents one of the fundamental limits of spatial vision: it is the finest high-contrast detail that can be resolved
    - minimum recognizable acuity
        - refer to the angular size of the smallest feature that one can recognize or identify
        - used by eye doctors
    - minimum discriminable acuity
        - refers to the angular size of the smallest change in a feature that one can discriminate
        - Vernier acuity – smallest misalignment that we can reliable discern
            - vernier scale success: humans are very adept at judging if nearby lines are lined up or not
            - vernier alignment is still widely used in precision machines, and even in the dial switches in modern ovens.

- Acuity for Low-Contrast Stripes
    - spatial frequency: refers to the number of times a pattern, repeats in a given unit of space
    - Schade, Fergus Campbell and Dan Green demonstrated that the human contrast sensitivity function is shaped like an upside-down U
    - CSF (contrast sensitivity function)
        - a function describing how the sensitivity to contrast depends on the spatial frequency of the stimulus
    - There are man factors that influence the exact form of the CSF

- Sine Wave Grating

  - the edge of any object produces a single strip, often blurred by a shadow, in the retinal image

- Retinal Ganglion Cells and Stripes
  - each ganglion cell responds well to certain types of stripes or grating
  - when the spatial frequency of the grating is too low, the ganglion cell responds weakly because part of the fat, bright bar of the grating lands in the inhibitory surround, damping the cell's response
  - when the spatial frequency is too high, the ganglion cell responds weakly because both dark and bright stripes fall within the receptive-field center, washing out the response.
  - with a bright bar filling the center and with dark bars filling the surround, the cell responds vigorously.
    * retinal ganglion cells are "turned" to spatial frequency: each cell acts like a filter
  - Chritina Enroth-cugell and John Robson: first to record the responses of retinal ganglion cells to sine wave gratings.
    * the responses depend on the phase of the grating
  - grating has a light bar filling the receptive-field center and dark bars filling the surround. ON-center ganglion cell increases its firing rate
    * there is no difference between the light intensity in the receptive field's center and its surround

- The Lateral Geniculate Nucleus
  - Lateral geniculate nucleus: A structure in the thalamus, part of the midbrain, that receives input from the retinal ganglion cells and has input and output connections to the visual cortex
  - The neurons in the bottom two layers are physically larger than those in the top four layers
    * bottom two called magnocellular layers

        - receive input from M ganglion cells in the retina
        - pagnocellular pathway responds to large, fast-moving objects
    * top four called parvovellular layers
        - receive input from P ganglion cells
        - process with details of stationary targets
    * Koniocellular cells: A neuron located between the magnocellular and parvovellular layers of the lateral geniculate nucleus
    * The organization of the retinal inputs to the LGNs
        - left LGN receives projections from the left side of the retina in both eyes
        - each layers of the LGN receives input from one or the other eye
        - bottom to top; 1,4,6 of the right LGN receive input from the left(contralateral: opposite side of the body/brain) eye, while layers 2,3,5 get theri input from the right (ipsilateral: same side of the body/brain) eye
  - LGN layer contains highly organized map of a complete half of the visual field
  - topographical mapping: the orderly mapping of the world in the lateral geniculate nucleus and the visual cortex
  - There are many connections between other parts of the brain and the LGN, there are more feedback connection from the visual cortex to the LGN than feed-forward connection from the LGN to the cortex

- The topography of the Human cortex
  - BOLD signals reflect a range of metabolically demanding neural signals

- Some Perceptual Consequences of Cortical Magnification
  - Visual acuity declines in an orderly fashion with eccentricity
  - High resolution requires a great number of resources:
    * a dense array of photoreceptors
    * one-to-one lines from photoreceptors to retinal ganglion cells
    * a large chunk of striate cortex
  - visual acuity is not the major obstacle to reading or object recognition

  - Visual crowding: the deleterious effect of clutter on peripheral object recognition
    * objects that can be easily identified in isolation seem indistinct and jumbled when surrounded by other objects
    * essential bottleneck, setting limits on object perception, eye and hand movements, visual search, reading, and perhaps other functions in peripheral vision
    * impairs ability to recognize and respond appropriately to object in clutter

- Receptive Fields in Striate Cortex
  - Hubel and Wiesel's dismay
    * a cat's cortical cells hardly responded at all to the same spots that made its ganglion cells fire like crazy
    * the response had nothing to do with the spot itself, the cell had been responding to the shadow cast by the edge of the glass slide as it swept across the ophthalmoscope's light pathway
    * discovery: receptive fields of striate cortex neurons are not circular, as they are in the retina and LGN; respond much more vigorously to bars, lines, edges and gratings than to round spots of light
  - Orientation Selectivity
    * an individual neuron will not respond equivalently to any old stripe in its receptive field
    * it responds best when the line or edge is at jus the right orientation
    * orientation tuning: the tendency of neurons in striate cortex to respond optimally to certain orientations and less to others
    * circular receptive fields in the LGN transformed into the elongated receptive fields in striate cortex:
      · Hubel: the concentric LGN cells that feed into a cortical cell are in a row
      · the arrangement of LGN inputs in indeed crucial for establishing the orientation selectivity of striate cortex cells
      · the neural interactions within the cortex play an important role in the dynamics of orientation tuning
  - Other Receptive-Field Properties

    * cortex cells are much more narrowly tunned than retinal ganglion cells
    * cortex cells response well to gratings
    * each striate cortex cell is tuned to a particular spatial frequency which corresponds to a particular line width
    * cortical cells responds well to moving lines, bars, edges and gratings
    * each LGN cell responds to one eye or the other, but never to both eyes
        · striate cortex: can be influenced by input from both eyes
    * ocular dominance: the property of the receptive fields of striate cortex neurons by which they demonstrate a preference, responding somewhat more rapidly when a stimulus is presented in one eye than when it is presented in the other

– Simple and complex Cells
    * simple cells: a cortical neuron whose receptive field has clearly defined excitatory and inhibitory regions
        · "phase-sensitive"
    * A stripe detector responds best to a line of light that has a particular width, surrounded on both sides by darkness
    * complex cells: a cortical neuron whose receptive field does not have clearly defined excitatory and inhibitory regions
        · complex cell is tuned to a particular orientation and spatial frequency and show an ocular preference
        · a complex cell will respond regardless of where the stripe is presented, as long as it is somewhere within the cell's receptive field
        · "phase-insensitive": the complex cell gives a robust response, with little or none of the modulation shown by simple cells
    * the receptive fields of complex cells represent a pooling of the response of several subunits
        · subunits give the complex cell its spatial frequency and orientation tunning
        · complex pooling: makes complex cell insensitive to the precise position of the stimulus within its receptive field

    - substantial evidence suggests that complex cells represent a separate parallel pathway → both simple and complex cells get direct input from LGN neurons
  - Future Complication
    - end stopping: the process by which a cell in the cortex increases its firing rate as the length of a bar increases until the bar fills up its receptive field, and then it decreases its firing rate as the bar is lengthened further
    - End-stopping plays an important role in our ability to detect luminance boundaries and discontinuities
    - Additional idiosyncrasies in the receptive fields of strate cortex neurons
    - Sceniak et al: the cell might respond to a smaller portion of the visual field when the grating stimulus has a high contrast than it will when the difference between light and dark bars is more subtle

- Columns and Hypercolumns
  - neurons with similar orientation preferences are arranged in columns that extend vertically through the cortex
    - column: a vertical arrangement of neurons. Neurons within a single column tend to have similar receptive fields and similar orientation preference
  - neurons share the same eye preference also have a columnar arrangement
  - Hubel and Wiesel
    - 1-mm block of striate cortex contains all the machinery necessary to look after everything the visual cortex is responsible for, in a certain small part of the visual world.
  - Hypercolumn:a 1-mm block of striate cortex containing two sets of columns, each covering every possible orientation, with one set preferring input from the left eye and one set preferring input from the right eye
  - cytochrome oxidase (CO): an enzyme used to reveal the regular array of "CO blobs", which are spaced about 0.5 mm apart in the primary visual cortex
    - CO blob columns have been implicated in processing color, with the interblob regions processing motion and spatial structure

  - the striate cortex is concerned with analyzing the orientation, size, shape, speed and direction of motion of objects in the world
    * it does so using modular groups of neurons (hypercolumns) each of which receives input from the processes a small piece of the visual world
    * the arrangement is like a big bank of filters

- Selective Adaption
  - adaption: a reduction in response caused by prior or continuing stimulation
  - selective adaption provide insights into the properties of cortical neurons
    * 0-degree selective cells elicit the strongest response
    * adaptation results in both a decreases in firing rate and a change in the tuning curve
  - tilt aftereffect: the perceptual illusion of tilt, produced by adaptation to a pattern of a given orientation
    * it strongly supports the idea that the human visual system contains individual neurons selective for different orientations
  - Selection adaptation also provides evidence that the human visual system contains neurons selective for spatial frequency
    * adaptation to the high-contrast top panel is selective, resulting in a loss of sensitivity for spatial frequency for spatial frequencies close to the adapting frequency
    * there is little or no effect on sensitivity to vertical gratings following adaptation to a horizontal grating
  - selective adaption causes the neurons most sensitive to the adapting stimulus to become fatigued

- The site of Selective adaption effect
  - the tilt aftereffect and the decreased contrast sensitivity transfer from one eye to the other
  - the transfer of adaptation from the adapted to the nonadapted eye is known as interocular transfer

  - the transfer of adaptation effect from one eye to the other thus implies that selective adaption occurs in cortical neurons

- Spatial Frequency -Tuned pattern Analyzers in human vision
  - human CSF reflects the sensitivity of multiple individual pattern analyzers
  - spatial-frequency channels: a pattern analyzer, implemented by an ensemble of cortical neurons, in which each set of neurons is tuned to a limited range of spatial frequencies
  - we have fewer neurons tuned to low spatial frequencies in order to compensate for the over-representation of energy in the lower spatial frequencies in natural scenes

- multiple-spatial-freqency model of vision implies that spatial frequencies that stimulate different pattern analyzers will be detected independently, even if the different frequencies are combined in the same image

- visual system does not carry out an actual Fourier analysis, analyzing the world into very narrow bands of spatial frequencies

- visual system filters the image into spatially localized receptive fields that have a limited range of spatial frequencies

- different spatial frequencies emphasize different types of information
  - low frequencies emphasize the broad outlines of the face
  - high frequencies carry information about fine details

- The development of vision
  - Fantz: preferential looking
    * if infants are shown two scenes, they invariably stare at the more complex scene
    * it depends on the willingness of babies to stare at stimuli near threshold level
  - VEPs (visual evoked electrical potentials)

        * measure an entire contrast sensitivity functions in as little as 10 s in a non-verbal infants

- development of the contrast sensitivity function
    - contrast sensitivity reach nearly adult levels as early as about 9 weeks of age
    - the primary postnatal changes in the retina concern differentiation of the macular region
    - foveal receptor density and cone outer segment length both increase, as foveal cones become thinner and more elongated
    - from birth to beyond 4 years of age, cone density increases in the central region, because of both the migration of the receptors and decreases in their dimensions
    - the peripheral retina appears to develop much more rapidly than the fovea
- if two objects appear near each other in the world, they will also be processed by cells near each other on the retina, in the LGN, and in the striate cortex

Reading

- a cell will respond to its preferred stimulus only if that stimulus is presented in a very specific location relative to the point where the observer is fixating its gaze(凝视)
- extrastriate cortex: the region of cortex bordering the primary visual cortex and containing multiple areas involved in visual processing
- border-ownership: when an object is sitting on a background, then edges defining the border between object and background "belong" to the object
- **Where and What Pathways**
  - Where Pathways: heads up into the parietal lobe.
    - visual area in this pathway: it is relating to the location of objects in space and the actions required to interact with (moving the hands, the eyes, and so on).
  - What Pathways: heads down into the temporal lobe.
    - the locus for the explicit acts of object recognition.
  - Lesion: a region of damaged brain. To destroy a section of the brain.
  - Agnosia: a failure to recognize objects in spite of the ability to see them. Agnosia is typically due to brain damage.
  - Inferotemporal cortex: part of the cerebral cortex in the lower portion of the temporal lobe, important in object recognition.
  - Homologous regions: Brain regions that appear to have the same function in different species.
  - Neurons in striate cortex are activated by simple stimuli and respond l=only if their preferred stimuli are presented in very restricted portions of the visual field
  - Cells in the IT cortex were discovered to have receptive fields that could spread over half or more of the monkey's field of view
  - Grandmother recognition is more likely to involve quite a large network of cells with individual cells participating in recognition of more than one stimulus

- IT cortex maintains close connections with parts of the brain involved in memory formation
    - IT cells need to learn their receptive-filed properties
- The visual system has many different problems to solve and it appears to have modules that are specialized for working on different problems
- Human with lesions in the temporal lobe often show symptoms of agnosia, the ability to see without the ability to know what is being seen. e.g. prosopagnosia
- feed-forward process: a process that carries out a computations one neural step after another, without need for feedback from a later stage to an earlier stage
- reverse-hierarchy theory: a theory that fast, feed-forward processes can give you crude information about objects and scenes based on activity in high-level parts of the visual cortex
    - the feed-forward processes give a general, categorical impression of the world but do not become aware of the details until "re-entrant" feedback goes back down the visual pathway
- Electrical activity from the brain can be recorded from electrode placed on the scalp
- there cannot be a lot of feedback from higher visual or memory processes, suggesting that it must be possible to do some rough object recognition on the basis of the first wave of activity as it moves, cell by cell, from retina to strate cortex to extrastriate cortex and beyond

- **mid-level vision**
    - mid-level vision: a loosely defined stage of visual processing that comes after basic features have been extracted from the image and before object recognition and scene understanding
    - find-edges:
        - in the early stages of processing, it is figuring out which edges mark the boundaries of objects and which represent surface features – make the system's best guess about the presence of a contour

    * illusory contours: a contour that is perceived even though nothing changes from one side of it to the other in an image
    * the edges are perceived because they are the best guess about what is happening in the world at that location
    * Occlusion: the visual system seems to come up with the hypothesis that there is another contour occluding the vertical line
    * structuralism: a school of thought that believed that complex objects or perceptions could be understood by analysis of the components
    * Gestalt: in reference to perception, a school of thought stressing that the perceptual whole can be greater than the apparent sum of the parts
    * Gestalt grouping rules: a set of rules describing which elements in an image will appear to group together. The original list was assembled by members of the Gestalt school of thought

- **Rules of evidence**
    * **Structuralist**, Wilhelm Wundt and Edward Bradford Tichener, argued that perceptions are the sum of atoms of sensation – bits of color, orientation
        · In the structuralist view, perception is built up of local sensations the way a crystal might be built up of an array of atoms
    * **Gestalt** held that the perceptual whole is more than the sum of its sensory parts.
        · **Gestalt grouping rules** describe the visual system's interpretation of the raw retinal image
        · the visual system is trying to make sense of the vast and often ambiguous an noisy inputs from the early stage of visual processing
        · good continuation: Gestalt grouping rule stating that two elements will tend to group together if they seem to lie on the same contour
        · closure: the Gestalt principle that holds that a closed contour is preferred to an open contour
        · The visual system has a bias toward closed contours
    * rules are useful parts of that effort because they reflect regularities in the world

- **Texture Segmentation and Grouping**
  - texture segmentation: Carving an image into regions of common texture properties
  - one way that the visual system decides that two regions are different is by looking at the statistics of all the features in one region and determining that those statistics differ from the statistics in the neighboring region
  - Two strongest Principles:
    - similarity: image chunks that are similar to each other will be more likely to group together;
    - Grouping of elements can be based on similarity in a limited number of feature such as color, size, orientation
    - proximity: holds that items near each other are more likely to group together than are items more widely separated
  - Parallelism and symmetry
    - shows that the grouping effects of enclosing or connecting items
- Camouflage: it is used to help us find objects in the world can be exploited to hide them
- Ambiguity and perceptual "committes"
  - When the opinions of different "committee members" collide, the results can be somewhat ambiguous.
  - A consensus view almost always quickly emerges and we settle on a single interpretation of the visual scene
- Committee rules:
  - ambiguous figure: a visual stimulus that gives rise to two or more interpretations of its identity or structure
  - necker cube: an outline that is perceptually bi-stable. Unlike the situation with most stimuli, two interpretations continually battle for perceptual dominance
  - Accidental viewpoint: a viewing position that produces some regularity in the visual image that is not present in the world

    * The visual system can be fooled by an accidental viewpoint
    * visual system makes some assumptions based on an implicit understanding of some aspects of the physics of the world
    * one committee uses the knowledge that opaque objects occlude other objects behind them to generate plausible interpretations of image elements
    * another committee considers all the possibilities and devalues any that involved accidental viewpoints, reducing what is initially a theoretically unsolvable problem to a potentially solvable one
- Figure and ground
    * figure-ground assignment: the process of determining that some regions of an image belong to a foreground object figure and other regions are part of the background
    * all visual stimuli maybe inherently ambiguous but the processes that determine figure and ground almost always manage to come to a single conclusions
    * Surroundedness: if one region is entirely surrounded by another, it is likely that the surrounded region is the figure
        · surroundedness: a rule for figure ground assignment stating that if one region is entirely surrounded by another, it is likely that the surrounded region is the figure
    * Size: the smaller region is likely to be figure
    * Symmetry: a symmetrical region is more likely to be seen as figure
    * parallelism: regions with parallel contours are likely to be seen as figure
    * relative motion: surface details move relative to an edge can also determine which portion of a display is the foreground figure and which is the background
- Complicated business
    * figure-ground processes divide up the image into regions that can be recognized as specific objects
    * object recognition is happening both before and after figure-ground assignment
- Dealing with Occlusion

    * relatability: the degree to which two line segments appear to be part of the same contour
    * heuristic: a mental shortcut
    * The visual system is unwilling to propose elaborate relationship, so it concludes that the lines are not related at all
    * Nonaccidiental feature: a feature of an object that is not dependent on the exact viewing position of the observer
- Parts and Wholes:
    * the global letters interfered with naming of the local letters more than the local letter interfered with recognition of hte global letters
    * global superiority effect: the finding in various experiments that the properties of the whole object take precedence over the properties of parts of the object
        · the first goal is to carve the retinal image into large-scale objects
    * when one blob is pushed into another, a pair of concavities is created in the silhouette of the resulting two-part objects.
- Summarizing Mid-level Vision
    * bring together that which should be brought together
    * split asunder that which should be split asunder
    * use what you know
    * avoid accidents
    * seek consensus and avoid ambiguity
- From Metaphor to Formal Model
    * before you look at anything, how likely is what you are proposing – prior probability
    * how consistent is each hypothesis with observation
    * Bayesian approach:a way of formalizing the idea that our perception is a combination of the current stimulus and our knowledge about the conditions of the world
        · help to reduce to formal, mathematical equations

- Object Recognition
  - In V1, cells respond best to lines and edges in very specific areas of the visual field.
  - In V2, we get early steps from local features to objects
  - functional imaging studies can show us that different regions of the cortex are activated better by some categories of stimuli than by others
  - subtraction method: in functional magnetic imaging, comparison of brain activity measured in two conditions: one with and one without the involvement of the the mental process of interest. The difference between the images for the two conditions may show regions of brain specifically activated by the mental process
  - PPA(parahippocampal place area): a region of extrastriate visual cortex in humans that is specifically and reliably activated more by images of places than by other stimuli
  - FFA(fusiform face area): a region of extrastriate visual cortex in humans that is specifically and reliably activated by human faces
  - EBA(extrastriate body area): a region of extrastriate visual cortex in humans that is specifically and reliably activated by images of the body other than the face
  - decoding: the process of determining the nature of a stimulus from the pattern of responses measured in the brain or, potential, in Artificial system like a computer network
  - pandemonium: define as noise and chaos
  - template: the internal representation of a stimulus that is used to recognize the stimulus in the world. Unlike its use in, for example, making a key, a mental template is not expected to actually look like the stimulus that it matches.
    - key and lock
  - Instead of matching each point in the image to a point in a template, perhaps we perform a more conceptual match
  - Structure description: a description of an object in terms of the nature of its constituent parts and the relationships between those parts

- recognition-by-components models: Biederman's model of object recognition which holds that objects are recognized by the identities and relationships of their components parts
- geon: in Biederman's recognition-by-components model, any of the "geometric ions" out of which perceptual objects are built.
- problems with structure description model:
  - not as viewpoint independent as the model would propose
  - no one has ever come up with a set of geon-like primitives that would work over all objects
- entry-level category: for an object, the label that comes to mind most quickly when we identify it. At the subordinate level, the object might be more specifically named; at the superordinate level, it might be more generally named
- holistic processing: processing based on analysis of the entire object or scene and not on adding together a set of smaller parts or features
- Bruce and young argued that faces conveyed seven distinct types of information
- Damage to specific areas in the temporal lobe of the brain can produce prosopagnosia
  - prosopagnosia: an inability to recognize faces
- congenital prosopagnosia: a form of face blindness apparently present from birth, as opposed to acquired prosopagnosia, which would typically be the result of an injury to the nervous system
- some face-processing ability is present before birth
- there were more fetal head turns toward the upright face
- Once the baby is born, it is much easier to show that very young infants have a preference for faces

Three Steps to Color Perception

- Detection: wavelengths must be detected
- Discrimination: tell the difference between one wavelength and another
- Appearance: assign perceived colors to lights and surfaces in the world.

Color Detection

- short-wavelength cones: the cones that have a peak at about 420 nm. **S-cone**: a cone that is preferentially sensitive to short wavelength; coloquially known as a "blue cone".
- medium-wavelength cones: peaks at about 535nm. **M-cone** a cone that is preferentially sensitive to middle wavelengths; coloquially known as a "green cone".
- long-wavelength cones: peak at about 565nm. **L-cone** a cone that is preferentially sensitive to middle wavelengths; coloquially known as a "red cone".
- **Spectral sensitivities**: the sensitivity of a cell or a device to different wavelengths on the electromagnetic spectrum
- Cones work at photopic(daylight) light levels.
  - **photopic**: referring to light intensities that are bright enough to stimulate the cone receptors and bright enough to "saturate" the rod receptors
- One type of rod photoreceptor is working in scotopic(dimmer) light and has a somewhat different sensitivity profile peaking at about 500nm.
  - **scotopic**: referring to light intensities that are bright enough to stimulate the rod receptors but too dim to stimulate the cone receptors

Color Discrimination

- 400-nm light produces only a small response
- 500-nm light produces a greater response
- 550-nm light produces much greater response

- 600-nm light produces less than the maximal response
- 625-nm light produces a moderate strength
- 650-nm light produces a minimal response

The principle of Univariance

- 450 or 625nm produces a response lower than the peak response obtained at about 535nm
- **principle of univariance**: the fact that an infinite set of different wavelength-intensity combinations can elicit exactly the same response from a single type of photoreceptor. One photoereceptor type cannot make color discriminations based on wavelength
- Univariance explains the lack of color in dimly lit scenes.
- there is only one type of rod photoreceptor.
- All rods contain the same type of photopigment molecule: rhodopsin
- one nightitme color blindness is one hint that color is psychophysical and not physical

The Trichromatic Solution

- detect differences between wavelengths or mixtures of wavelengths because of having more than one kind of cone photoreceptor
- the light produces twice as much M response as S response and twice as much S response as L response
- the ability to discriminate one light from another is defined in our visual system by the relationships among three numbers is the heart of trichromacy or trichromatic theory of color vision
  - **trichromatic theory of color vision** the theory that the color of any light is defined in our visual system by the relationships of three cones. Also called the Young-helmholtz theory

Metamers

- Almost every light and every surface that we see is emitting or reflecting a wide range of wavelengths.
- two lights mixed together, produce a mixture that excites the L- and M-cones equally
- Mixture of different wavelengths that look identical are called metamers
  - **metamers** different mixtures of wavelengths that look identical. More generally, any pair of stimuli that are perceived as identical in spite of physical differences.
- The single wavelength that produces equal M- and L-cone activity will look yellow, and the correct mixture of longer- and shorter-wavelength lights will also look yellow
  - Mixing wavelengths does not change the physical wavelengths
  - For a mixture of a red and green to look perfectly yellow, we would have to have just right red and just the right green
- all the light reaching the retina from one patch in the visual field will be converted into three numbers by the three cone types
- if sufficiently different from the numbers in another patch, be able to discriminate those patches
- if not, patches will be metamers: they will look identical, even if the wavelengths are physically different

History of Trichromatic Theory

- basic theory was established by psychophysical experimentation
- Three-dimensional nature of the experience of color was worked out in the nineteenth century by Thomas Young and subsequently by Hermann von Helmholtz
  - Trichromatic theory is called Yong-Helmholtz theory
- James Clerk Maxwell developed a color-matching technique that was central to Helmholtz's work on this topic

- The observer in a modern version of Maxwell's experiments would try to use different amounts of "primary" colored lights to exactly match another reference color.
- central observation from these experiments was that only THREE mixing lights are needed to match any reference light.
- Yong and Helmholtz deduce that three different color mechanisms must limit the human experience of color.

A brief digression into lights, filters and finger paints

- **additive color mixture**: a mixture of lights. If light A and light B are both reflected from a surface to the eye, in the perception of color the effects of those two lights add together
- **subtractive color mixture**: a mixture of pigments. If pigments A and B mix, some of the light shining on the surface will be subtracted by A, and some by B. Only remainder will contributed to the perception of color.

From Retina to Brain: Repacking the information

- the nervous system will look at differences in the activities of the three cone types to tell the difference between different lights.
- The nervous system computes two differences: (L-M) and ([L+M] - S)
  - comparisons across species suggest that the comparison between S-cones and an LM-cone happened first.
- Create (L-S) and (M-S) signals
  - a single comparison between S and (L+M) can capture almost the same information that would be found in (L-S) and (M-S) signals.
  - combining L and M signals is a good measure of the intensity of hte light
  - Convert three cone signals into three new signals – (L-M), ([L+M] -S) and (L+M) and the visual system does something reasonably close to this.

Cone-opponent Cells in the Retinal and LGN

- **lateral geniculate nucleus (LGN)**: a structure in the thalamus, part of the midbrain, that receives input from the retinal ganglion cells and has input and output connections to the visual cortex
  - many ganglion cells in the retina and the LGN of the thalamus are maximally stimulated by the spot of the light
- **cone-opponent cell**: a cell typefound in the retina, lateral geniculate nucleus, and visual cortex – that in effect subtracts one type of cone input from another
  - (L-M), (M-L), ([M+L] - S), (S - [M+L]) cells
- the cells that were excited by light onset could be thought of as (L+M) cells
- **koniocellular**: referring to cells in the koniocellular layer of the LGN of the thalamus. Konio from the Greek for "dust" referring to the appearance of the cells.
  - the S-cone signals go through the koniocellular layers in the LGN
- **parvocellular**: referring to cells in the parvocellular layer of the LGN of the thalamus. Parvo from the Greek for "small" referring to the appearance of the cells.
  - M- and L-cone opponent signals are mostly found in the parvocellular layers
- L-cones produce red response and M-cones produce green regardless of the wavelength of stimulus
- L-cone could be part of an (L-M) circuit that responds to color and an (L+M) circuit that responds to brightness
- the (L+M) brightness response is just stronger
- It could be the response looks white because some cones just do not contribute to color sensation → explain why spatial resolution is quite bad if use equiluminant
  - **equiluminant**: referring to stimuli that vary in color but not in luminance

ganglion cell help to keep track of day and night

- **circadian:** referring to the biological cycle that recurs approximately every 24 hours, even in the absence of cues to time of day

  - the primary force adjusting internal circadian clock is sunlight.

- ganglion cell receives input form the rods and the cones, but it also contains it own photopigment, melanopsin, so it can detect light even when normal photoreceptors are absent.

  - **melanopsin:** a photopigment, found in a class of photoreceptive retinal ganglion cells.

Color Appearance

- the three cone types and the rods are all active in a middle range of light intensities
- The retina and LGN contain cells that have repackaged the three cones signals into cone-opponent difference signals that constrain our ability to see differences between regions
- **Three numbers, Many colors**
  - the two surfaces must and will appear to be exactly the same color.
  - it is possible to produce bloodred color on the page without mimicking the physical properties of blood
  - ignore lightness, be able to distinguish 26000 colors
  - going beyond the spectrum, having three-dimensional color space analogous to a three-dimensional physical space
    - **color-space**: the three-dimensional space, established because color perception is based on the outputs of three cone types, that describes the set of all colors
  - Brightness is the perceptual consequence of the physical intensity of a light
- **The limits of the Rainbow**
  - the combination of 420nm light with a pure 680-nm light strongly stimulate the L- and S-cones and produce minimal stimulation in the M-cones
  - join the red and blue ends with a set of colors that are called nonspectral hues – hues that can arise only from mixtures of wavelengths

  - not all colors are included in the spectrum's
    - e.g. Brown: is seen when a mixture of wavelengths that would look yellow, greenish yellow, or orange is seen in the company of other, brighter patches of color
    - cannot see an isolated brown light in the dark

- **Opponent Colors**
  - three basic colors: red, green, blue
  - **opponent color theory**: the theory that perception of color is based on the output of three mechanisms, each of them resulting from an opponency between two colors: red-green, blue-yellow, black-white
  - **hue cancellation**: cancel the perceptual of one color by adding its opponent color
  - **Unique hues**: any of four colors that can be described with only a single color term: red, yellow, green, blue. Other colors can also be described as compounds
    - hues that can be described with only a single color term.
  - First, three cones detect light in different ranges of wavelengths. Then opponent processes measure the differences in activity between cone types. Finally, some further transformations are needed to create the color opponency described by Hering.

- **Color in the Visual Cortex**
  - many cells in cortex are interested in color but do not seem to linearly add and subtract inputs from different cone types.
  - there are some evidence that they add and subtract in a nonlinear manner.
  - there were "blobs" in V1 where cells did not seem interested in orientation but seemed very interested in color
  - modern imaging studies show some areas of the human visual cortex that seem particularly interested in color
  - The best evidence of specialized brain areas for color in humans comes from certain cases of achromatopsia

    - **achromatopsia**: an inability to perceive colors that is caused by damage to the central nervous system
  - people with achromatopsia are able to find the boundaries between regions of different colors, but they cannot report what those colors might be

- **Language and Color**
  - **basic color terms**: color words that are single words are used with high frequency, and have meanings that are agreed upon by speakers of a language.
  - **cultural relativism**: in sensation and perception, the idea that basic perceptual experience may be determined in part by the cultural environment
  - each group was free to create its own linguistic map of color space
  - Berlin and Kay's
    - various maps used in different cultures are actually rather similar
    - a big color term is partitioned into two smaller terms
  - new basic terms tend to emerge at the boundary between two existing color term, in the area where neither existing term works well
  - mola: light-warm colors
  - mili: dark-cool colors
  - The less similar to the two test colors are, the easier to pick the color saw before
  - The Dani's perfromance on such tasks reflected the same color boundaries, even when their language did not recognize the distinction between the two colors
  - People were faster to identify the different one and more accurate when the patches were separated by the border between red and brown than when both patches were within the "red" or "brown" category.
  - Color names matter, even when you compare two colors side by side.

- Genetic differences in Color Vision
  - the individual differences described in the previous two sections are either small or in the cases of inverted qualia.

    * **qualia**: in philosophy, a private conscious experience of sensation or perception.
- some of these differences will be due to factor like age, which turns the lens of the eye yellow
- the genes that code for the M- and L-cone photopigments are on the X chromosome
- Males have only one copy of the X chromosome; female have two copies
    * some women can end up with four different cone pigments, and in very rare cases, that produces tetrachromatic
        · **tetrachromatic**: referring to the rare situation where the color of any light is defined by the relationships of four numbers – the outputs of those four receptor types.
- The S-cone photopigment is coded elsewhere, so everyone has two copies. – S-cone color deficiencies are rare.
- color-blindness:
    * One determining factor is the type of cone affected
    * a second factor is the type of defect – either the photopigment for that cone tye is anomalous or the cone type is missing altogether
    * M- and L-cone defects are the most common. most color-blind individual have difficulty discriminating lights in the middle-to-long-wavelength range.
- **deuteranope**: an individual who suffers from color blindness that is due to the absence of M-cones
- **protanope**: an individual who suffers from color blindness that is due to the absence of L-cones
- **tritanope**: an individual who suffers from color blindness that is due to the absence of S-cones
- **color-anomalous**: a better term for the commonly used term color-blind. Most color-blind individuals can still make discriminations based on wavelength. Those discriminations are different from the norm – anomalous

- **cone monochromat**: an individual with only one cone type. Cone monochromats are truly color-blind
- **rod monochromat**: an individual with no cones of any type. In addition to be truly color-blind, rod monochromats are badly visually impaired in bright light.
- Brain lesions can also produce various forms of color agnosia or anomia
- **agnosia**: a failure to recognize objects in spite of the ability to see them. Agnosia is typically due to brain damage
- **anomia**: an inability to name objects in spite of the ability to see and recognize them.

- **color of lights to a world of color**
  - the mix of wavelengths reflected from the grass will be very different if the light source is the sun rather than a dim bulb.
  - In color contrast effect, the color of one region induces the opponent color in a neighboring region
    - **color contrast**: a color perception effect in which the color of one region induces the opponent color in a neighboring region
  - In color assimilation effect, two colors bleed into each other, each taking on some of the chromatic quality of the other
    - **color assimilation**: a color perception effect in which two colors bleed into each other, each taking on some of the chromatic quality of the other
  - **unrelated color**: a color that can be experienced in isolation
  - **related color**: a color such as brown or gray, that is seen only in relation to other colors.

- **Adaption and Afterimages**
  - Color contrast effects show how the spatial relations between colors can influence color appearance.
  - Adapting to a bright light makes a moderate light look darker. Adapting to darkness would make that same moderate light appear brighter

- Adaption can be color-specific, as we see in the phenomenon of negative afterimages
    - **negative afterimages**: an afterimage whose polarity is the opposite of the original stimulus. Light stimuli produce dark negative afterimages. Colors are complementary
- **Adapting stimulus**: a stimulus whose removal produces a change in visual perception or sensitivity
- The L-cones will be more adapted than M- or S-cones, as will the later processes in the retina and brain that were more stimulated by the red spot.
- **neutral point**: the point at which an opponent color mechanism is generating no signal. If red-green and blue-yellow mechanism are at theri neutral points, a stimulus will appear achromatic
- Adaption occurs at multiple sties in the nervous system, though the primary generators are in the retina.

- **Color Constancy**
    - **color constancy**: the tendency of a surface to appear the same color under a fairly wide range of illuminants
    - **illuminant**: the light that illuminates a surface
    - **spectral reflectance function**: the percentage of a particular wavelength that is reflected from a surface
    - **spectral power distribution**: the physical energy in a light as a function of wavelength
    - Sunlight is a yellowish light, richer in middle and long wavelengths; skilight is more bluish, with more short-wavelength energy.

- **Physical Constraints Make Constancy Possible**
    - **Reflectance**: the percentage of light hitting a surface that is reflected and not absorbed into the surface. Typically reflectance is given as a function of wavelength.

  - there are ambiguous stimuli. Even if the two interpretations are very different, one person can see them both.
  - even if the two interpretations are very different, one person can see them both.

- **Color vision Good for**
  - evolutionary theory tells us that color vision must provide an advantage that makes it worth the trouble.
  - Most diurnal animals have two photopigments: roughly an S-cone and LM-cone.

Key Terms

- **realism**: a philosophical position arguing that there is a real world to sense
- **positivism**: a philosophical position arguing that all we really have to go on is the evidence of the senses, so the world might be nothing more than an elaborate hallucination.
- **binocular**: referring to two eyes.
- **probability summation**: the increased detection probability based on the statistical advantage of having two detectors rather one
- **binocular summation**: the combination of signals from both eyes in ways that make performance on many tasks better than with either eye alone.
- **binocular disparity(不一致)**: the differences between the two retinal images of the same scene. Disparity is the basis fo stereopsis, a vivid perception of three-dimensionality of the world that is not available with monocular vision.
- **monocular**: referring to one eye
- **stereopsis(实体影像)**: the ability to use binocular disparity as a cue to depth.
- **monocular depth cue**: a depth cue that is available even when the world is viewed with one eye alone
- **binocular depth cue**: a depth cue that relies on information from both eyes stereopsis is the primary example in humans, but convergence and the ability of two eyes to see more of an object than one eye sees are also binocular depth cues.
- **pictorial depth cue**: a cue to distance or depth used by artists to depict three-dimensional depth in two-dimensional pictures.
- **occlusion**: a cue to relative depth order in which, for example, one object obstructs the view of part of another object.
- **nonmetrical depth cue**: a depth cue that provides information about the depth order but not depth magnitude

- **metrical depth cue**: a depth cue that provides quantitative information about distance in the third dimension
- **projective geometry**:the geometry describes the transformations that occur when the three-dimensional world is projected onto a two-dimensional surface
- **relative size**: a comparison of size between items without knowing the absolute size of either one
- **texture gradient**: a depth cue based on the geometric fact that items of the same size form smaller images when they are farther away. An array of items that change in size smoothly across the image will appear to form a surface tilted in depth.
- **relative height**: as a depth cue, the observation that objects at different distances from the viewer on the ground plane will form images at different heights in the retinal image. Objects farther away will be seen as higher in the image.
- **familiar size**: a depth cue based on knowledge of the typical sizes of objects, such as humans or pennies.
- **relative metrical depth cue**: a depth cue that could specify, for example, that object A is twice as far away as object B without providing information about the absolute distance to either A or B
- **absolute metrical depth cue**: a depth cue that provides quantifiable information about distance in the third dimension.
- **haze or aerial perspective**: a depth cue based on the implicit understanding that light is scattered by the atmosphere. More light is scattered when we look through more atmosphere. Thus, more distance objects are subject to more scatter and appear fainter, bluer, and less distinct.
- **linear perspective**: a depth cue based on the fact that lines that re parallel in the three-dimensional world will appear to converge in a two-dimensional image.
- **vanishing point**: the apparent point at which parallel lines receding in depth converge.

- **anamorphosis or anamorphic projection**: use of the rules of linear perspective to create a two-dimensional image so distorted that it looks correct only when viewed from a special angle or with a mirror that counters the distortion.
- **motion parallax**: an important depth cue that is based on head movement. The geometric information obtained from an eye in two different positions at two different times is similar to the information from two eyes in different positions in the head at the same time.
- **optic flow**: the pattern of apparent motion of objects in a visual scene produced by the relative motion between the observer and the scene
- **accommodation**: the process by which the eye changes its focus
- **convergence**: the ability of the two eyes to turn inward, often used in order to place the two images of a feature in the world on corresponding locations in the two retinal images. Convergence reduces the disparity of that feature to zero.
- **divergence**: the ability of the two eyes to urn outward, often used in order to place the two images of a feature in the world on corresponding locations in the two retinal images. Divergence reduces the disparity of that feature to zero.
- **corresponding retinal points**: two monocular images of an object in the world are said to fall on corresponding points if those points are the same distance from the fovea in both eyes. The two foveas are also corresponding points.
- **Vieth-Muller circle**: the location of objects whose images fall on geometrically corresponding points in the two retinas. If life were simple, this circle would be the horopter, but life is not simple.
- **diplopia**: double vision. If visible in both eyes, stimuli falling outside of Panum's fusional area will appear diplopic.
- **horopter**: the location of objects whose images lie on corresponding points. The surface of zero disparity.
- **crossed disparity**: the sign of disparity created by objects in front of the plane of fixation. The term crossed is used because images of objects located in front of the

horopter appear to be displaced to the left in the right eye and to the right in the left eye

- **uncrossed disparity**: the sign of disparity created by objects behind the plane of fixation. The term uncrossed is used because images of objects located behind the horopter will appear to be displaced to the right in the right eye and to the left in the left eye.
- **stereoblindness**: an inability to make use of binocular disparity as a depth cue. This term is typically used to describe individuals with vision in both eyes. Someone who has lost one eyes is not typically described as "stereoblind"
- **random dot stereogram(RDS)**: a stereogram made of a large number of randomly placed dots. Random dot stereograms contain no monocular cues to depth. stimuli visible stereo.
- **Cyclopean**: referring to stimuli that are defined by binocular disparity alone. Named after the one-eyed Cyclops of Homer's Odyssey
- **correspondence problem**: in reference to binocular vision, the problem of figuring out which bit of the image in the left eye should be matched with which bit in the right eye. The problem is particularly vexing when the images consist of thousands of similar features, like dots in random dot stereograms.
- **uniqueness constraint**: in reference to stereopsis, the observation that a feature in the world is represented exactly once in each retinal image. This constraint simplifies the correspondence problem
- **continuity constraint**: in reference to stereopsis, the observation that, except at the edges of objects, neighboring points in the world lie at similar distances from the viewer. This is one of several constraints that have been proposed as helpful in solving the correspondence problem.
- **bayesian approach**: P(A|o) = P(A) * P(O|A)/ P(O)
- **binocular rivalry**: the competition between the two eyes for control of visual perception, which is evident when completely different stimuli are presented to the two eyes.

- **stereoacuity**: a measure of the smallest binocular disparity that can generate a sensation of depth
- **dichoptic**: referring to the presentation of two different stimuli, one to each eye. Different from binocular presentation, which could involve both eyes looking at a single stimulus.
- **critical period**: a period of time during development when the organism is particularly susceptible to development change. There are critical periods in the development of binocular vision, human language, and so on
- **strabismus**: A misalignment of the two eyes such that a single object in space is imaged on the fovea of one eye and on a nonfoveal area of the other eye
- **esotropia**: strabismus in which one eye deviates inward
- **exotropia**: strabismus in which one eye deviates outward
- **tilt aftereffect**: the perceptual illusion of tilt, produced by adaption to a pattern of a given orientation.
- **suppression**: in reference to vision, th inhibition of an unwanted image. Suppression occurs frequently in people with strabismus.

Reading

- The ability to perceive and interact with the structure of space is one of the fundamental goals of the visual system
- If we want to appreciate the three-dimensional world, we have to reconstruct it from the distorted retinal input
- our visual system is a reconstruction of the world based on two distorted inputs: the two distinct retinal images
- the point is general one: the two retinal images always differ.
  - they differ because your two eyeballs are in slightly different places in your head.
- two eyes:

  - lose one eye, another still works
  - doubling the number of eyes is that they enable you to see more of the world
- visual field is limited to about 190 degrees from left to right, 110 degrees of which is covered by both eyes
- the exact size of your visual field will be limited by the specific anatomy of your cheeks and eyebrows
- overlapping, frontal, binocular visual field give fields give predator animals to spot small, fast-moving objects
  - get the advantage of two detectors looking at the same thing
  - evolution found that overlapping visual fields is useful for binocular disparity
- prey animals use very wide visual field allowing them to monitor the whole scene for predators
- stereopsis is not a necessary condition for depth perception or space perception
- stereopsis does add a richness to perception of the three-dimensional world

Monocular Cues to Three Dimension Space

- it is geometrically impossible for the visual system to create a perfectly faithful reconstruction of the true layout of space, given the distorted, two-dimensional input we receive through each eye
- **Pictorial Depth Cues**
  - imaged is viewed from the correct position, the retinal image formed by the two-dimensional picture will be the smae as the retinal image that would have been formed by the three-dimensional world
- **Occlusion**
  - Occlusion was a cue to the presence of an otherwise invisible edge
  - As a depth cue, occlusion gives information about the relative position of objects

  - Occlusion is present in almost every visual scene and many researchers argue that it is the mot reliable of all the depth cues
  - Occlusion is a nonmetrical depth cue; it just gives us the relative orderings of occluders and occludees: provide depth order; A metrical depth cue is one that does provide information about distance in the third dimension

- **Size and Position Cues**
  - Projective geometry describes how the world is projected onto a surface
  - The smaller image of a more distant rabbit will be projected higher in your visual field
  - Texture fields that provide an impression of three-dimensionality are really combinations of relative size and relative height cues
  - different modules in the visual system perform different tasks
  - if your visual system knew the acutal size of an object and the visual angle of the object's projection on the retina, it could calculate the exact distance from object to eye

- **Aerial perspective**
  - objects farther away are subject to more scatter and appear fainter and less distinct

- **Linear perspective**
  - it is based on the rules that determine how lines in 3d space are projected onto a 2d image
  - these are intended to show the parallel lines on the ground or on the sides of buildings running in depth

- **Seeing depth in pictures**
  - the ability to cope with distortion is limited
  - in anamorphic projection, the rules of linear perspective are pushed to an extreme

  - the projection of 3d into 2d creates a 2d image that is recognizable only from an unusual vantage point → anamorphic art

Triangulation Cues to 3d space

- Motion Cues:
  - motion parallax allows you to see the world from multiple viewpoints at the different times by moving your head
  - The geometric information obtained from an eye in two different positions at two different times is similar to the information from two eyes in different positions in the head at the same time
  - the downside of motion parallax is that it works only if the head moves
  - other motion signals produce information about depth, objects get bigger and smaller as they get closer and farther away
- Accommodation and Convergence
  - We need to point out eyes differently to focus on objects at different distances
  - the visual system takes advantage of both cues for objects closer than 2-3 meters away. Convergence is used more than accommodation
  - humans are not particularly precise about measuring the exact angles

Binocular Vision and Stereopsis

- binocular disparity refers to differences between the images falling on two retinas and stereopsis refers to the impression of 3d of objects "popping out in depth" – that most humans get when they view real-world objects with both eyes.
- the visual system is designed so that the object of our gaze always falls on the fovea, the rays from objects fall on the fovea in each eye
- if the two eyes are looking at one spot, then there will be a surface of zero disparity running through that spot. – surface is horopter
- objects that lie on the horopter are seen as single objects when viewed with both eyes

- objects significantly closer to or farther away from the sruface of zero disparity form images on decidedly noncorresponding points in the two eyes, and we see two of each those objects
- Objects that are close to the horopter but not quite on it can still be seen as single objects
- Panum's area provides a little room for small errors in eye alignment, while still maintaining single vision
- the larger the disparity, the greater the distance in depth of the object from the horopter
- right in left, and left in right – crossed disparity
- when the disparity is not too large, the depth information from stereopsis is quantitative
- Once disparities are larger than the upper disparity limit, there is no longer any useful depth information.
- **Stereoscopes and Stereograms**
  - The stereoscope confirmed that the visual system treats binocular disparity as a depth cue, regardless of whether the disparity is produced by actual or simulated images of a scene
  - when free-fuse, the opposite misalignments become the binocular disparity, and visual system converts that disparity into a perception of depth
  - The depth seen depends on if you converged or diverged your eyes
  - divergence requires focusing on a point beyond the plane of the page so that the image of the left-hand set of squares falls on the left fovea and the image of the right-hand set falls on the right fovea
  - Stereoblind individuals are able to achieve the perception of three sets of squares, but the little white squares will not pop out in depth
  - Stereoblindness is usually a secondary effect of childhood visual disorders such as strabismus, in which the two eyes are misaligned.

- Random Dot Stereograms
  - Bela Julesz thought the conventional wisdom might be backward. The stereopsis might be used to discover objects and surfaces in the world
  - stereopsis might help reveal camouflaged objects.
  - important point of RDS is cannot see the squares in either of the component
  - cannot see the squares using any monocular depth cues
  - Wheatstone showed that his stereoscope that binocular disparity is a necessary condition for stereopsis.
  - Julesz demonstrated with RDSs that disparity is sufficient for stereopsis.
- Using Stereopsis
  - Stereopsis can provide useful information about metric depth only for distance up to 40 meters
  - In people with normal binocular vision, visually guided hand movements are significantly impaired when viewing is restricted to one eye
  - Stereoscopic displays are beginning to be used in radiology, and they can reduce the error rate in these important tasks
  - loss of stereopsis may also result in unstable gait, especially reduced accuracy when a change of terrain occurs, and difficulties for children in playing some sports.
- Stereoscopic Correspondence
  - corresponding problem is the problem of figuring out which bit of the image in the left eye should be matched with which bit in the right eye
  - The retinal images of the circles are labeled to make it clear which image on the left retina corresponds to which image on the right retina, but your visual system has no such labels.
  - Matching thousands of dots for both eyes require a lot of work – simpler if look at a blurred version of the stereogram $\rightarrow$ a few larger blobs.

  - In addition to starting with low-spatial-frequency information two more heuristics for solving the correspondence problem – uniqueness and continuity constraints
    * uniqueness constraint – feature in the world is represented exactly once in each retinal image
    * continuity constraint – neighboring points in the world lie at similar distances from the viewer
  - recent work suggests that identifying correct matches may not be the optimal strategy

- The physiological Basis of Stereopsis and depth perception
  - the most fundamental requirement is that input from the two eyes must converge onto the same cell
  - The convergence does not happen until the primary visual cortex, where most neurons can be influenced by input from both the left and right eyes
  - A binocular neuron has two receptive fields, one in each eye – well suited to the task of matching images in the two eyes
  - binocular neurons respond best when similar images occupy slightly different positions on the retinas of the two eyes
  - Stereopsis can be used both metrically and nonmetrically
    * nonmetrical stereopsis might just tell you that a feature lies in front of or behind the plane of fixation
    * stereopsis can also be used in a very precise, metrical manner
    * Both forms of stereopsis have their uses and functional magnetic resonance imaging suggest that the dorsal where pathway is most interested in metrical stereopsis, while the ventral what pathway makes do with more categorical, near-versus-far information
  - early visual areas, particularly V2, are invovled in computing depth order, based on the contour completion and border ownership process
  - Intermediate visual areas such as V4 encode depth intervals, based on relative disparities and higher cortical areas

Combining Depth Cues

- all we really have is a collection of guesses about possible depth relations between different objects in our visual field
- combining and weighting guesses, the visual system generally arrives at a coherent, and more or less accurate, representation of three dimensional space
- it is much more likely that the points of contact between the images of the two pennies reflect occlusions
- combining depth cues, our choice of the metaphor of a committee is not arbitrary
- the visual system must be estimating how reliable each depth cue might be
- illusion and the construction of space
  - we make a plausible guess about the 3d world that is being represented in the 2d picture
  - the job of the visual system is to use available cues to make an intelligent guess about the world
  - it reflects a more general aspect of the visual system's response to titled lines

Key Words

- **Motion aftereffect(MAE)** the illusion of motion of a stationary object that occurs after prolonged exposure to a moving object
- **interocular transfer** the transfer of an effect from one eye to the other
- **middle temporal area(MT)** an area of the brain thought to be important in the perception of motion. Also called V5 in humans.
- **apparent motion** the illusory impression of smooth motion resulting from the rapid succession.
- **correspondence problem** in reference to motion detection, the problem faced by the motion detection system of knowing which feature in Frame 2 corresponds to a particular feature in Frame 1.
- **aperture problem** The fact that when a moving object is viewed through an aperture, the direction of motion of a local feature or part of the object may be ambiguous.
- **aperture** an opening that allows only a partial view of an object.
- **first-order motion**: the motion of an object that is defined by changes in luminance
- **luminance-defined object**: an object that is delineated by differences in reflected light
- **second-order motion**: the motion of an object that is defined by changes in contrast or texture, but not by luminance
- **texture-defined object/ contrast defined object**: an object that is defined by differences in contrast, or texture, but not by luminance
- **akinetopsia**: a rare neuropsychological disorder in which the affected individual has no perception of motion
- **double dissociation**: the phenomenon in which one of two functions, such as first- and second-order motion, can be damaged without harm to the other, and vice versa.

- **optic array**: the collection of light rays that interact with objects in the world that are in front of a viewer. The term was coined by J.J.Gibson
- **optic flow**: the changing angular positions of points in a perceptive image that we move through the world
- **focus of expansion**: the point in the center of the horizon from which, when we are in motion, all points in the perspective image seem to emanate. The focus of expansion is one aspect of optic flow.
- **time to collision(TTC)**: the time required for a moving object(such as a cricket ball) to hit a stationary object
- **tau**: information in the optic flow that could signal time to collision(TTC) without the necessity of estimating either absolute distance or rates. The ratio of retinal image size at any moment to the rate at which the image is expanding is tau, and TTC is proportional to tau.
- **biological motion**: the pattern of movement of living beings
- **smooth pursuit**: a type of voluntary eye movement in which the eyes move smoothly to follow a moving object.
- **superior colliculus**: a structure in the midbrain that is important in initiating and guiding eye movement
- **microsaccade**: an involuntary, small, jerk-like eye movement.
- **vergence**: a type of eye movement in which the two eyes move in opposite directions; for example, both eyes turn toward the nose or away from the nose
- **saccade**: a type of eye movement, made both voluntarily and involuntarily, in which the eyes rapidly change fixation from one object or location to another
- **reflexive eye movement**: a movement of the eye that is automatic and involuntary
- **optokinetic nystagmus(OKN)**: a reflexive eye movement in which the eyes will involuntarily track a continually moving object.

- **saccadic suppression**: the reduction of visual sensitivity that occurs when we make saccadic eye movements. Saccadic suppression eliminates the smear from retinal image motion during an eye movement
- **efference copy/ corrollary discharge signal**: the phenomenon in which outgoing signals from the motor cortex are copied as they exit the brain and are rerouted to other areas in the sensory cortices
- **comparator**: an area of the visual system that receives one copy of the command issued by the motor system when the eyes move. The comparator compares the image motion signal with the eye motion signal and can compensate for the image changes caused by the eye movement

reading

- **Motion Aftereffects**
  - A phenomenon called the waterfall illusion provides another piece of evidence that there is something special about motion.
  - As color aftereffects are caused by opponent processes for color vision, MAEs are caused by opponent process for motion detection
  - MAE implies an opponent-process system much like the one that plays a role in color vision
  - Neurons sensitive to upward motion fire at about the same rate as neurons sensitive to downward motion, so the signals cancel out and no motion is perceived
  - strong MAE is obtained when one eye is adapted and the other tested means that the effect must be reflecting the activities of neurons in a part of the visual system where information collected from the two eyes is combined.
  - Recent advances in functional imaging techniques may make it possible to locate the site of motion aftereffects even more precisely.
  - The emerging evidence suggests that the MAE in humans is caused by the same brain region shown to be responsible for global-motion detection in monkeys: the middle temporal area of the cortex, an area commonly referred to as MT or V5

  - Direction-selective adaption produced a selective imbalance in the fMRI signal in human MT, providing evidence that MAEs are due to a population imbalance in are MT.

- **Computation of Visual Motion**
  - The mechanism would also be tuned to speed because when the bug moved at just the right speed, the delayed response from receptor A and the direct response from receptor B would occur at the same time and therefore reinforce each other.
  - Cell X fires if and only if both its inputs are firing simultaneously and it passes this message on to the motion detection cell M
  - A more realistic circuit would include additional receptors to detect longer-range motion.
  - The M cell fires continually as the bug moves across the fields of the five receptors at the top of the circuit.

- **Apparent Motion**
  - first demonstrated by Sigmund Exner in 1875.
  - Set up a contraption that would generate electrical sparks separated from each other by a very short distance in space and very short period of time.
  - Have motion detectors for all directions, one detector will sense the diagonal motion
  - a different detector may win this competition when an object is viewed through an aperture than would win if we would see the whole object – aperture problem
  - The motion component parallel to the grating cannot be inferred from the visual input.
  - A variety of contours of different orientations moving at different speeds can cause identical responses in a motion-sensitive neuron in the visual system
  - every v1 cell sees the world through a small aperture

- **Detection of global motion in area MT**

  - information from magnocellular neurons feeds into V1 and is then passed on to the middle temporal area of the cortex, an area commonly referred to as MT in nonhuman primates, and then to the medical superior temporal area, MST. MT and MST are considered to be the hub for motion processing
  - The vast majority of neurons in the MT are selective for motion in one particular direction, but they show little selectivity for form or color.
  - To detect the correlated direction, a neuron must integrate information from many local-motion detectors.
  - Lesion studies have been central to our understanding of the specificity of brain areas.
  - Monkeys showed a strong tendency to report motion in the stimulated neurons' preferred direction, even when the dots they were seeing were actually moving in the opposite direction
  - The MT is critically involved in the processing of global motion.

- **Second-order Motion**
  - In the first order apparent-motion displays, nothing actually moves in second-order motion
  - the visual system includes specialized mechanisms for second-order motion
  - brain damage that impaired the perception of first-order motion but not of second-order motion
  - it turns out that second-order-like motion does occur in the real world, especially when an object is effectively camouflaged

- **Motion induced blindenss**
  - MIB seems to be somewhat related to well-known Troxler effect.
  - This effect can be observed under conditions in which the retinal image is stabilized so that the involuntary eye movements that occur during fixation do not move the target onto new receptive fields

- **Using Motion information**

- when we move through our environment, we experiene patterns of optic flow that our visual system use to determine where we are going
- outflow indicates that you are approaching a particular destination, inflow indicates retrea
- human can estimate their direction of heading to within about 1 or 2 degrees, using as their sole guide the pattern of optic flow simulated by the moving dots, even when the display contained only a very small number of dots
- Warren showed that observers were able to discount these radial components, both when the observers moved their own eyes and when the computer-generated display mimicked the radial flow caused by an eye movement
- the visual system can make use of the copies of eye muscle signals when it is processing optic flow information

- **Avoiding imminent Collision**
  - cricketers are extremely good at judging precisely when a tiny ball, hurtling toward them at speeds approaching 100 miles per hour, is about to collide with their head.
  - humans are far better at judging TTC than would be predicated on the basis of their ability to judge distance
  - there is an alternative source of information in the optic flow that could signal TTC without the need for absolute distances or rates to be estimated.
  - the advantage of using tau to estimate TTC is that it relies solely on information available directly from the retinal image; all you need to do is track the visual angle subtended by the cricket ball as it approaches your eye
  - estimating the time to imminent collision is critical important to animals and humans, and almost every species tested will attempt to avoid a simulated collision
  - certain neurons in the visual systems of pigeons and locusts responds to objects on a collision course with them and can signal a particular time to collision.

- **using motion information to identify objects**
  - motion can help inform us about the nature of object

- biological motion helps us identify both the moving object and its actions
- the mechanisms that analyze biological motion obey different rules for integrating motion over space and time than do mechanisms for other forms of complex motion
- observers use estimates of the center of motion as a cue to the walker's gender
- The amount of body sway might provide an even more salient gender cue

- **Eye movement**
  - to avoid confusion, we will ignore this inconvenient bit of physics for now and pretend that images move on the retina in the same direction that the objects in the world are moving
  - small peripheral targets may fade and disappear during steady fixation – an illusion known as the Troxler effect
  - the microsaccades are directed toward objects we are attending
  - Three types of voluntary eye movement
    - smooth-pursuit eye movement – can help distinguish individuals with schizophrenia from others
    - vergence eye movement – occur when we rotate our eyes inward to focus on a near or far object
    - saccade – a fast jump of the eye that shifts our gaze from one spot to another
  - eyes are more likely to make saccades in response to contours than to board featureless areas of an image
  - reflexive eye movements – when the eyes move to compensate for head and body movement while maintaining fixation on a particular target
  - optokinetic nystagmus – reflexive eye movement in which the eyes will involuntarily track a continually moving object, moving smoothly in one direction in pursuit of the object moving in that same direction and then snap back
  - all of the information processing must take place during the fixation – occurs during only a small fraction of each fixation

- if a word remains on the screen for only 50 ms after it is first fixated, reading proceeds normally.
- When we make a saccade, the viusal system essentially shuts down for the duration of the eye movement
- Saccadic suppression acts mainly to suppress information carried by the magnocellular pathway
- Saccadic suppression is little more than masking and that the magnocellular pathway is not suppressed during saccades.
- One copy goes to the eye muscles, another coly goes to an area of the visual system that has been dubbed the comparator.
- While a lot of work has been dedicated to the analysis of extraretinal signals during smooth pursuit, visual stability during saccades remains a matter of active research, and attention seems to play na important role.
- since the brain generates the neural signals for saccadic eye movements, it can perceptually compensate for them
- receptive fields are generally considered to be fixed in space relative to the fixation point
- the neural representation of the visual field is rigidly translated just prior to the eye movement
- the receptive fields of neurons in the frontal eye fields shift transiently toward the target location
- the receptive field shifts do not predict the retinal displacements produced by saccades, but rather reflect the fact that space is perceived to be compressed just before a saccade.
- neurons in V1 have adult-like sensitivity to motion direction
- sensitivity to global motion appears to dvelop more slowly reaching maturity at about 3-4 years of age, while sensitivity to motion-defined form and biological motion takes even longer.

Key words

- **amplitude**: the magnitude of displacement of a pressure wave. Amplitude is perceived as loudness
- **frequency**: the number of times per second that a pattern of pressure change repeats. Frequency is perceived as pitch
- **hertz**: a unit of measure for frequency; 1 herz equals 1 cycle per second
- **loudness**: the psychological aspect of sound related to perceived intensity
- **pitch**: the psychological aspect of sound related mainly to the fundamental frequency
- **decibel**: a unit of measure of the physical intensity of sound. Decibels define the difference between two sounds as the ratio between two sound pressures.
- **sine wave**: the wave form for which variation as a function of time is a sine function
- **spectrum**: a representation of the relative energy present at each frequency
- **harmonic spectrum**: the spectrum of a complex sound in which energy is at integer multiples of the fundamental frequency
- **fundamental frequency**: the lowest-frequency component of a complex periodic sound.
- **timbre**(音色): the psychological sensation by which a listener can judge that two sounds with the same loudness and pitch are dissimilar. Timbre quality is conveyed by harmonics and other high frequencies.
- **pinna**(耳廓): the outer, funnel-like part of the ear
- **ear canal**(耳道): the canal that conducts sound vibrations from the pinna to the tympanic membrane and prevent damage to the tympanic membrane(膜).
- **tympanic membrane**: the eardrum; a thin sheet of skin at the end of the outer ear canal.
- **outer ear**: the external sound-gathering portion of the ear, consisting of the pinna and the ear canal.

- **middle ear**: an air-filled chamber containing the middle bones, or ossicles. The middle ear conveys and amplifies vibration from the the tympanic membrane to the oval window.
- **ossicle**: any of three tiny bones of the middle ear: malleus, incus and stapes
- **malleus**: the most exterior of the three ossicles. The malleus receives vibration from the tympanic membrane and is attached to the incus.
- **incus**: the middle of the three ossicles, connection the malleus and the staples.
- **stapes**: the most interior of the three ossicles. Connected to the incus on one end, the stapes presses against the oval window of the cochlea on the other end.
- **oval window**: the flexible opening to the cochlea through which the stapes transmits vibration to the fluid inside.
- **inner ear**: a hollow cavity in the temporal bone of hte skull, and the structures within this cavity: the cochlea and the semicircular canals of the vestibular system.
- **tensor tympani**(鼓膜张肌): the muscle attached to the malleus. Tensing the tensor tympani decreases vibration.
- **acoustic reflex**: a reflex that protects the ear from intense sounds, via contraction of the stapedius and tensor tympani muscles.
- **cochlea**(耳蜗): a spiral structure of the inner ear containing the organ of Corti.
- **tympanic canal**: One of three fluid-filled passages in the cochlea. The tympanic canal extends from the round window at the base of the cochlea to the helicotrema at the apex. Also called scala tympani
- **vestibular canal**: One of three fluid-filled passages in the cochlea. The middle canal is sandwiched between the tympanic and vestibular canals and contains the cochlear partition. also called scala media.
- **helicotrema**: the opening that connects the tympanic and vestibular canals at the apex of the cochlea.

- **Reissner's membrane**: a thin sheath of tissue separating the vestibular and middle canals in the cochlea
- **basilar membrane**: a plate of fibers that forms the base of the cochlear partition and separates the middle and tympanic canals in the cochlea.
- **cochlear partition**: the combined basilar membrane, tectorial membrane, and organ of Corti, which are together responsible for the transduction of sound waves into neural signals.
- **round window**: a soft area of tissue at the base of the tympanic canal that releases excess pressure remaining from extremely intense sounds.
- **organ of Corti**: a structure on the basilar membrane of the cochlea that is composed of hair cells and dendrites of auditory nerve fibers.
- **hair cell**: any cell that has sterocilia for transducing mechanical movement in the inner ear into neural activity sent to the brain. Some hair cells also receive inputs from the brain.
- **auditory nerve**: a collection of neurons that convey information from hair cells in the cochlea to the brain stem and from the brain stem to the hair cells.
- **stereocilium**: any of the hairlike extensions on the tips of hair cells in the cochlea that, when flexed, initiate the release of neurotransmitters.
- **tectorial membrane**: a gelatinous structure, attached on one end, that extends into the middle canal of the cochlea, floating above inner hair cells and touching outer hair cells.
- **tip link**: a tiny filament that stretches from the tip of a stereocilium to the side of its neighbor.
- **place code**: tuning of different parts of the cochlea to different frequencies, in which information about the particular frequency of an incoming sound wave is coded by the place along the cochlear partition that has the greatest mechanical displacement.
- **afferent fiber**: a neuron that carries sensory information to the central nervous system

- **efferent fiber**: a neuron that carries information from the central nervous system to the periphery
- **threshold tuning curve**: a graph plotting the thresholds of a neuron in response to sine waves with varying frequencies at the lowest intensity that will give rise to a response.
- **characteristic frequency**: the frequency to which a particular auditory nerve fiber is mot sensitive
- **two-tone suppression**: a decrease in the firing rate of one auditory nerve fiber due to one tone, when a second tone is presented at the same time
- **isointensity curve**: a map plotting the firing rate of an auditory nerve fiber against varying frequencies at varying intensities
- **rate saturation**: the point at which a nerve fiber is firing as rapidly as possible and further stimulation is incapable of increasing the firing rate.
- **rate-intensity function**: a graph plotting the firing rate of an auditory nerve fiber in response to a sound of constant frequency at increasing intensities
- **low-spontaneous fiber**: an auditory nerve fiber that has a low rate of spontaneous firing. Low-spontaneous fibers require relatively intense sound before they will fire at higher rates
- **high-spontaneous fiber**: an auditory nerve fiber that a hight rate of spontaneous firing. High-spontaneous fibers increase their firing rate in response to relatively low levels of sound.
- **mid-spontaneous fiber**: an auditory nerve fiber that has a medium rate of spontaneous firing. The characteristics of mid-spontaneous fibers are intermediate between those of ow- and high-spontaneous fibers.
- **phase locking**: firing of a single neuron at one distinct point in the period of a sound wave at a given frequency.

- **temporal code**: tuning of different part of the cochlea to different frequencies, in which information about the particular frequency of an incoming sound wave is coded by the timing of neural firing as it relates to the period of the sound.
- **volley principle**: the idea that multiple neurons can provide a temporal code for frequency if each neuron fires at a distinct point in the period of a sound wave but does not fire on every period.
- **cochlear nucleus**: the first brain stem nucleus at which afferent auditory nerve fibers synapse.
- **superior olive**: an early brain stem region in the auditory pathway where inputs from both ears converge.
- **inferior colliculus**: a midbrain nucleus in the auditory pathway.
- **medial geniculate nucleus**: the part of the thalamus that relays auditory signals to the temporal cortex and receives input from the auditory cortex.
- **tonotopic organization**: an arrangement in which neurons that respond to different frequencies are organized anatomically in order of frequency
- **primary auditory cortex(A1)**: the first area within the temporal lobes of the brain responsible for processing acoustic information.
- **belt area**: a region of cortex, directly adjacent to the primary auditory cortex(A1), with inputs from A1, where neurons respond to more complex characteristics of sounds.
- **parabelt area**: a region of cortex, lateral and adjacent to the belt area, where neurons respond to more complex characteristics of sounds, as well as to input from other senses.
- **psychoacoustics**: the branch of psychophysics that studies the psychological correlates of the physical dimensions of acoustics in order to understand how the auditory system operates.
- **audibility threshold**: the lowest sound pressure level that can be reliably detected at a given frequency

- **equal-loudness curve**: a graph plotting sound pressure level against the frequency for which is listener perceives constant loudness.
- **temporal integration**: the process by which a sound at a constant level is perceived as being louder when it is of greater duration. The term also applies to perceived brightness, which depends on the duration of light.
- **masking**: using a second sound, frequently noise, to make the detection of another sound more different.
- **white noise**: noise consisting of all audible frequencies in equal amounts. White noise in hearing is analogous to white light in vision, for which all wavelengths are present.
- **critical bandwidth**: the range of frequencies conveyed within a channel in the auditory system.
- **conductive hearing loss**: hearing loss caused by problems with the bones of the middle ear.
- **otitis media**: inflammation of the middle ear, commonly in children as a result of infection
- **otosclerosis**: abnormal growth of the middle-ear bones that causes hearing loss.
- **sensorineural hearing loss**: hearing loss due to defects in the cochlea or auditory nerve.

Key points

- **interaural time difference**: the difference in time between arrivals of sound at one ear versus the other.
- **azimuth**: the angle of a sound source on the horizontal plane relative to a point in the center of the head between the ears. Azimuth is measured in degrees, with 0 degrees being straight ahead. The angle increases clockwise toward the right, with 180 degrees being directly behind.
- **medial superior olive**: a relay station in the brain stem where inputs from both ears contribute to detection of the interaural time difference.

- **interaural level difference**: the difference between levels of sound at one ear versus the other.
- **lateral superior olive**: a relay station in the brain stem where inputs from both ears contribute to detection of the interaural level difference.
- **cone of confusion**: a region of positions in space where all sounds produce the same time and level differences.
- **directional transfer function**: a measure that describes how the pinna, ear canal, ead, and torso change the intensity of sounds with different frequencies that arrive at each ear from different locations in space.
- **inverse-square law**: a principle stating that as distance from a source increases, intensity decreases faster such that decrease in tensity is equal to the distance squared. This general law also applies to optics and other forms of energy.
- **attack**: the part of a sound during which amplitude increases.
- **decay**: the part of a sound during which amplitude decreases.
- **source segregation**: processing an auditory scene consisting of multiple sound sources into separate sound images.
- **auditory stream segregation**: the perceptual organization of a complex acoustic signal into separated auditory events for which each stream is heard as a separate event.
- **similarity**: Gestalt grouping rule stating that the tendency of two sounds to group together will increase as the acoustic similarity between them increases.
- **common fate**: Gestalt grouping rule stating that the tendency of sounds to group together will increase if they begin and/or end at the same time.
- **good continuation**: Gestalt grouping rule stating that sounds will tend to group together as continuous if they seem to share a common path, similar to a shared contour for vision.

- **acoustic startle reflex**: The very rapid motor response to a sudden sound. Very few neurons are involved in the basic startle reflex, which can also be affected by emotional state.

reading

- A complex sound is best described in a spectrum that displays how much energy or amplitude, is present at multiple frequencies.
- The shape of the spectrum (spectral shape) is one of the most important qualities that distinguish different sounds.
- Sounds are first collected from the environment by the pinna
  - the shape of the pinnae plays an important role in our ability to localize sound sources
- the pinna and ear canal make up a division of the auditory system called the outer ear
- The first ossicle, the malleus, is connected to the tympanic membrane on one side and to the second ossicle, the incus, on the other. The incus is connected in turn to the third ossicle, the stapes, which transmits the vibrations of sound waves to the oval window, another membrane, which forms the border between the middle ear and the inner ear.
- ossicles are the smallest bones in the human body
  - the joints between the bones are hinged in ways that make them work like levers a modest amount of energy on one side of the fulcrum (joint) becomes larger on the other
    - increases the amount of pressure change by about a third
  - increase the energy transmitted to the inner ear is by concentrating energy from a larger to a smaller surface area.
- The middle ear has two muscles: the tensor tympani (attached to the malleus) and the stapedius (attached to the stapes). The tensor tympani and the stapedius are the smallest muscles in the body.

- the inner ear is the cochlea (from the Greek kochlos, "snail"), a tiny coiled structure embedded in the temporal bone of the skull
- The cochlea is filled with watery fluids in three parallel canals: the tympanic canal (or scala tympani), the vestibular canal (or scala vestibuli), and the middle canal (or scala media).
- three canals of the cochlea are separated by two membranes
  - Reissner's membrane between the vestibular canal and the middle canal
  - basilar membrane between the middle canal and the tympanic canal
- the cochlea is a closed system, changes in pressure cannot spread out in all directions.
- Listeners can detect differences between onsets of two sounds as small as 1 millisecond; they can detect gaps between sounds as brief as 2-3 ms.
- the stereocilia connected by tip links bend together as a set when deflected by the shearing motion of the tectorial membrane.
- The firing of the auditory nerve fibers finally completes the process of translating sound waves into patterns of neural activity.
- Process:
  - the tympanic membrane vibrates the malleus
  - the malleus vibrates the incus
  - the incus vibrates the stapes
  - pushes and pulls on the oval window
  - the oval window causes pressure bulges to move down the length of the vestibular canal
  - bulges in the vestibular canal displace the middle canal up and down
    - up-and-down motion forces the tectorial membrane to shear across the organ of Corti
  - The pivoting of the stereocilia initiates rapid depolarization

  - spurts of neurotransmitter released into synapses between the hair cells and dendrites of auditory nerve fibers
  - These neurotransmitters initiate action potentials in the auditory nerve fibers that are carried to the brain

- the larger the amplitude, the higher the firing rate of the neurons that communicate with the brain.
- different parts of the cochlear partition are displaced to different degrees by different sound wave frequencies.
- High frequencies cause the largest displacements closer to the oval window, near the base of the cochlea.
- Lower frequencies cause displacements farther away and nearer the apex.
- Cochlear tuning to frequency is caused, in large part, by differences in the structure of the basilar membrane along the length of the cochlea.
- The cochlea narrows from base to apex, but the basilar membrane inside actually widens toward the apex.
- Higher frequencies affect the narrower, stiffer regions of the basilar membrane near the base more, and lower frequencies cause greater displacements in the wider, more flexible regions near the apex.
- Over 90% of the afferent fibers in the auditory nerve fibers that take information TO the brain
- It turns out that most of the nerve fibers that synapse with the outer hair cells are efferent fibers, conveying information FROM the brain.
- Over 90% of the afferent fibers in the auditory nerve fibers that take information to the brain.
- an AN fiber will increase firing to only a very restricted range of frequencies.
- The frequency that increases the neuron's firing rate at the lowest intensity is called the neuron's characteristic frequency.

- The sharp tuning measured from outputs of inner hair cells depends greatly upon the unsung heroes of the basilar membrane, the outer hair cells.
- Outer hair cells make parts of the cochlear partition stiffer in ways that make the responses of inner hair cells more sensitive and more sharply tuned to particular frequencies
- the way the ear transduces acoustic energy at different frequencies into a pattern of neural responses seems fairly straightforward.
- A low-intensity sine wave tone with a certain frequency will cause certain AN fibers to increase their firing rates, while other AN fibers continue to fire at their spontaneous rates
- High-spontaneous AN fibers are somewhat analogous to rods in the retina: they are especially sensitive to low levels of sound, responding at rates above resting level even when decibel levels are quite low.
- AN fibers with relatively high-frequency CFs encode lower-frequency energy in the temporal pattern of their responses
- Some cochlear nucleus neurons use lateral inhibition to sharpen the tuning to one frequency by suppressing nearby frequencies
- some of the neurons that project from the cochlear nuclei to the superior olives cross over to the opposite side of the brain
- Neurons from the cochlear nucleus and superior olive travel up the brain stem to the inferior colliculus
- neurons most responsive to low-frequency energy lie on one edge of each structure, neurons responding to high frequencies lie on the other edge, and neurons responding to other frequencies are spread out in an orderly fashion in between.
- a relatively large proportion of the processing in the auditory system is done before A1
- Auditory capabilities that may be most important to humans, listening to speech and music, are subserved almost entirely by cortical areas.

- the inequality of sound pressure level and loudness: equal-amplitude sounds can be perceived as softer or louder than each other, depending on the frequencies of the sound waves
- the number of photons must be more than doubled to double the perceived brightness of a light
- The loudness of a sound also depends on its duration. Within limits, longer sounds are heard as being louder.
- larger intensities correspond to higher firing rates, and smaller intensities (quiet sounds) correspond to lower firing rates.
- as sounds become more intense, many more AN fibers become excited
- humans are remarkably good at detecting very small differences in frequency
- use masking experiments to investigate frequency selectivity
- a greater proportion of the basilar membrane vibrates in response to low frequencies, and higher-frequency ranges vibrate smaller portions of the membrane
- masking effects are asymmetrical
- the most common, and most serious, form of auditory impairment is sensorineural hearing loss
- Sensorineural loss is caused in two general ways: metabolic and sensory
- Metabolic losses are caused by changes in the fluid environment of the cochlea that decrease the activity of hair cells.
- Sensory losses are caused by injury to hair cells.
- The major cause of sensory hearing loss is damage to the hair cells by excessive exposure to noise
- With fewer inner hair cells, the neuronal firing pattern described by the volley principle for temporal coding of frequency would become more difficult to maintain, because there would be fewer neurons available to take turns firing

- cumulative exposure to even everyday noises present in the environments of industrialized countries can cause hearing loss
- move energy from frequency regions in which hearing is poor (usually high frequencies) into regions where hearing is normal.
- One advantage of the old horns over electronic hearing aids was that they permitted listeners to direct their hearing toward the sound source they were most interested in
- loss of synapses results in a loss of connectivity
- hidden hearing loss might explain some of the difficulties that human listeners have in noisy situations even when their audiograms are normal
- improving the ability of listeners to detect sounds is unlikely to improve the ability of people with hidden hearing loss to use sounds.

Concepts

- even though sound travels fast, the pressure waves do not arrive at both ears at the same time.
- the intensity of a sound is greater at the ear closer to the source.
- the term is used to describe locations on an imaginary circle extending around us in a horizontal plane-front, back, left, and right-is azimuth.
- The medial superior olives (MSOs) are the first places in the auditory system where inputs from both ears converge.
- the brain takes advantage of the traveling wave to use frequency differences to measure time.
- The second cue to sound localization is the interaural level difference (ILD) in sound intensity.
- Neurons that are sensitive to intensity differences between the two ears can be found in the lateral superior olives

- When the sound is more intense at one ear, connections from that ear are better both at exciting LSO neurons on that side and at inhibiting LSO neurons on the other side.
- Cones of confusion are real perceptual phenomena, not just theoretical problems for the auditory system.
- children may update the way they use DTF information during development.
- The simplest cue for judging the distance of a sound source is the relative intensity of the sound.
- The effectiveness of relative intensity decreases quickly as distance increases, because sound intensity decreases according to the inverse-square law
- listeners tend to consistently underestimate the distance to sound sources farther away, and the amount of underestimation grows as distance becomes longer.
- Intensity works best as a distance cue when the sound source or the listener is moving
- A final distance cue stems from the fact that, in most environments, the sound that arrives at the ear is some combination of direct energy and reverberant energy.
- reverberations appear to be important for judging the loudness of sounds
- listeners maintain constant perceptions of loudness across changing distances by scaling direct energy relative to reverberant energy.
- The most impressive demonstration of enhanced auditory perception of the spatial environment by visually impaired listeners is echolocation.
- harmonic sounds are among the most common types of sounds in the environment
- The most straightforward explanations of the missing-fundamental effect involve the temporal code for pitch
- the richness of complex sounds like those in our world depends on more than simple sensations of loudness and pitch
- perception of timbre is related to the relative energies of different acoustic spectral components

- trombone's third (786 -Hz) component is stronger than its fourth (1048 -Hz) component, whereas for the saxophone, the relationship between the energies of these two components is reversed.
- Our auditory systems are very sensitive to attack and decay characteristics
- The amplitude of the resulting sound increases very quickly before more gradually dissipating
- auditory system uses a number of strategies to segregate sound sources
  - spatial separation between sounds
    - if a sound stays in the same place relative to the path of a listener, it will be easier for that sound to be sorted out from other sounds
  - sounds can be segregated on the basis of their spectral or temporal qualities
  - Grouping by timbre is particularly robust because sounds with similar timbres usually arise from the same sound source
- Grouping of sounds with common onsets is consistent with the Gestalt principle of common fate.
- The compelling nature of perceptual restoration suggests that at some point the restored missing sounds are encoded in the brain as if they were actually present in the signal.
- Complex sounds such as music and speech can also be perceptually restored.
- Effects of attending to a particular sound source can be so strong that we completely miss out on hearing other sounds in a kind of inattentional deafness.

Key Words

- **pitch**: the psychological aspect of sound related mainly to perceived frequency
- **octave**: the interval between two sound frequencies having a ratio of 2:1
- **tone height**: a sound quality corresponding to the level of pitch. Tone height is monotonically related to frequency
- **tone chroma**: a sound quality shared by tones that have the same octave interval.
- **chord**: a combination of three or more musical notes with different pitches played simultaneously
- **absolute pitch**: a rare ability whereby some people are able to very accurately name or produce notes without comparison to other notes
- **melody**: a sequence of notes or chords perceived as a single coherent structure
- **tempo**: the perceived speed of the presentation of sounds
- **syncopation**: any deviation from a regular rhythm.
- **vocal tract**: the airway above the larynx used for the production of speech. The vocal tract includes the oral tract and nasal tract.
- **phonation**: the process through which vocal folds are made to vibrate when air pushes out of the lungs.
- **articulation**: the act or manner of producing a speech sound using the vocal tract.
- **formant**: a resonance of the vocal tract. Formants are specified by their center frequency and are denoted by integers that increase with relative frequency.
- **spectrogram**: in reference to sound analysis, a 3d display that plots time on the horizontal axis, frequency on the vertical axis, and amplitude on a color or gray scale.
- **coarticulation**: the phenomenon in speech whereby attributes of successive speech units overlap in articulatory or acoustic patterns.

- **categorical perception**: for speech as well as other complex sounds and images, the phenomenon by which the discrimination of items is no better than the ability to label items.

Concepts

- Humans are capable of producing an incredible range of distinct speech sounds
- One notorious disadvantage of such a low larynx is that humans choke on food more easily than any other animal does.
- The production of speech has three basic components: respiration (lungs), phonation (vocal folds), and articulation
- Speaking fluently requires an impressive degree of coordination among respiration (lungs), phonation (vocal folds), and articulation.
- The rate at which vocal folds vibrate depends on their stiffness and mass.
- The first harmonic corresponds to the actual rate of physical vibration of the vocal folds- the fundamental frequency
- Peaks in the speech spectrum are referred to as formants
- These concentrations in energy occur at different frequencies, depending on the length of the vocal tract.
- One of the most distinctive characteristics of speech sounds is that their spectra change over time
- Vowel sounds are all made with a relatively open vocal tract
- Three articulatory dimensions:
  - place of articulation:
    - at the lips (bilabial speech sounds: 'b', 'p', 'm')
    - at the alveolar ridge just behind the teeth (alveolar speech sounds: 'd', 't', 'n')
    - at the soft palate (velar speech sounds: 'g', 'k', 'ng')

  - manner of articulation
    * total obstructed (stops: b, d, g, p ,t, k)
    * partially obstructed (fricatives: s, z, f, v, th, sh)
    * only slightly obstructed (laterals: l, r; and glides: 'w', 'y')
    * first blocked, and then allowed to sneak through (affricates: ch, j)
    * blocked at first from going through the mouth, but allowed to go through the nasal passage(nasals: n, m, ng)
  - voicing
    * vibrating (voiced consonants, which can be felt by a finger on the throat: b, m ,z, l, r)
    * not vibrating (voiceless consonants: p, s, ch)

- making small, incremental changes to simple acoustic stimuli such as pure tones leads to gradual changes in people's perception of these stimuli.
- the ability to discriminate sounds can be predicted by how listeners label the sounds.
- speech production is at least as complex as speech perception
- Trying to explain speech perception by reference to production is at least as difficult as explaining speech perception on the basis of acoustics alone.
- processes for perceiving speech are unique to humans is that numerous demonstrations have shown that nonhuman animals can learn to respond to speech signals in much the same way that human listeners do
- speech is special because (1) humans have evolved unique anatomical machinery for producing it and (2) we spend a great deal of time practicing the perception of speech
- At the same time, other stimulus differences must be ignored so that multiple instances of the same speech sound or multiple images of the same face can be classified properly
- we don't need individual acoustic invariants to distinguish speech sounds; we just need to be as good at pattern recognition for sounds as we are for visual images.

- Prenatal experience with speech sounds appears to have considerable influence on subsequent perception.
- the longer a person uses only her first language, the longer it takes to learn to produce and perceive sounds from a second language
- Learning also is easier if two new sounds from a new language differ in the same way that two sounds from the first language differ.
- Picking up on distinctions in a second language is easiest if the second language is learned at the same time as the first.
    - The small downside to this strategy is that kids learning multiple languages at the same time usually take a little longer to master each of the languages than do children learning only a single language
- An infant will hear the sounds making up the word pretty in many different contexts
- infants learn to pick words out of the speech stream by accumulating experience with sounds that tend to occur together
- as sounds become more complex, they are processed in more anterior and ventral regions of the superior temporal cortex farther away from A1
- visual cues can have on the way speech sounds are perceived.
- activation of visual areas of the brain is the result of increased experience and ability with lip -reading for these previously deaf individuals.

Key words

- **vestibular organs**: the set of five sense organs – three semicircular canals and two otolith organs – located in each inner that sense head motion and head orientation with respect to gravity
- **spatial orientation**: a sense consisting of three interacting modalities: perception of linear motion, angular motion and tilt
- **vestibular system**: the vestibular organs as well as the vestibular neurons in cranial nerve VII and the central neurons that contribute to the functional roles that the vestibular system participates in.
- **vertigo**: a sensation of rotation or spinning. The term is often used more generally to mean any form of dizziness.
- **vestibulo-ocular reflex(VOR)** A short-latency reflex that helps stabilize vision by counter-rotating the eyes when the vestibular system senses head movement.
- **balance**: the neural processes of postural control by which weight is evenly distributed, enabling us to remain upright and stable..
- **kinesthesia**: perception of the position and movement of our limbs in space.
- **active sensing**:Sensing that includes self-generated probing of the environment. Besides our vestibular sense, other active human senses include vision and touch. Animal active sensing includes the use of echoes by whales and bats, the use of electrical signals by some fishes, and the use of whiskers/ antennae by fishes, insects, and nocturnal rodents.
- **efferent commands**: Information flowing outward from the central nervous system to the periphery. A common example is motor commands that regulate muscle contraction. The copy of such motor commands is often called an efferent copy.
- **afferent signals**: Information flowing inward to the central nervous system from sensors in the periphery. Passive sensing would rely exclusively on such sensory inflow, providing a traditional view of sensation.

- **graviception**: The physiological structures and processes that sense the relative orientation of gravity with respect to the organism.
- **angular motion**: rotational motion like the rotation of a spinning top or swinging saloon doors that rotate bak and forth.
- **linear motion**: translational motion like the predominant movement of a train ar or bobblehead doll
- **tilt**: to attain a sloped position like that of the learning tower of Pisa.
- **transduce**: to convert from one form of energy to another
- **semicircular canal**: any of three toroidal tubes in the vestibular system that sense angular motion.
- **angular acceleration**: The rate of change of angular velocity. Mathematically, the integral of angular acceleration is angular velocity, and the integral of angular velocity is angular displacement. Angular acceleration, angular velocity, and angular displacement all mathematically represent angular motion.
- **otolith organ**: either of two mechanical structures in the vestibular system that sense both linear acceleration and gravity.
- **linear acceleration**: the rate of change of linear velocity. Mathematically, the integral of linear acceleration is linear velocity, and the integral of linear velocity is linear displacement, which is also referred to as "translation." Linear acceleration, linear velocity, and linear displacement all mathematically represent linear motion.
- **gravity**: A force that attracts a body toward the center of the Earth.
- **sensory conflict**: Sensory discrepancies that arise when sensory systems provide conflicting information. For example, vision may indicate that you are stationary while the vestibular system tells you that you are moving (or vice versa).
- **sense of angular motion**:The perceptual modality that senses rotation.
- **sense of linear motion**:The perceptual modality that senses translation.

- **sense of tilt**: The perceptual modality that senses head inclination with respect to gravity.
- **amplitude**:In reference to vestibular sensation, the s ize (increase or decrease) of a head movement (with angular velocity, linear acceleration, ti lt, etc.)
- **direction**: The line one moves along or faces, with reference to the point or region one is moving toward or facing.
- **velocity**: The speed and direction in which something moves. Mathematically, velocity is the integral of acceleration. In words, linear velocity is distance divided by time to traverse that distance; angular velocity is rotation angle divided by time to traverse that angle.
- **acceleration**: A change in velocity. Mathematically, acceleration is the derivative of velocity. In words, linear acceleration indicates a change in linear velocity; angular acceleration indicates a change in angular velocity.
- **hair cell**: Any cell that has stereocilia for transducing mechanical movement in the inner ear into neural activity sent to the bra in; some hair cells also receive inputs from the brain.
- **mechanoreceptor**: A sensory receptor that responds to mechanical stimulation (pressure, vibration, or movement).
- **receptor potential**:A change in voltage across the membrane of a sensory receptor cell (in the vestibular system, a hair cell) in response to stimulation
- **ampulla**: an expansion of each semicircular-canal duct that includes that canal's cupula crista and hair cells, where transduction occurs.
- **crista**: any of the specialized detectors of angular motion located in each semicircular canal in a swelling called the ampulla.
- **oscillatory**: referring to back-and-forth movement that has a constant rhythm
- **sinusoidal**: Referring to any oscillation, such as a sound wave or rotational motion, whose waveform is that of a sine curve. The period of a sinusoidal oscillation is the time

that it takes for one full back-and-forth cycle of the motion to occur. The frequency of a sinusoidal oscillation is defined as the numeral1 divided by the period.

- **Fourier analysis**: A mathematical procedure by which any signal-in this case motion trajectories as a function of time-can be separated into component sine waves at different frequencies. Combining these sine waves will reproduce the original motion trajectory.

- **utricle**: One of the two otolith organs. A sad ike structure that contains the utricular macula. Also called utriculus.

- **saccule**: One of the two otolith organs. A sad ike structure that contains the saccular macula. Also called sacculus.

- **macula**: any of the specialized detectors of linear acceleration and gravity found in each otolith organ.

- **otoconia**: tiny calcium carbonate stones in the ear that provide inertial mass for the otolith organs, enabling them to sense gravity and linear acceleration.

- **velocity storage**: Prolongation of a rotational response by the brain beyond the duration of t he rotational signal provided to the brain by the semicircular canals; typically yielding responses that are nearer the actual rotational motion t han t he signal provided by the canals.

- **dizziness**: A commonly used lay term that nonspecifically indicates any form of perceived spatial disorientation, with or without instability.

- **imbalance**: lack of balance; unsteadiness; nearly falling over.

- **mathematical integration**: Computing an integrai–Dne of the two main operations in calculus (the other, the inverse operation, is differentiation). Velocity is the integral of acceleration. Change of position is the integral of velocity.

- **sensory integration**: The process of combining different sensory signa ls. Typically, combining several signals yields more accurate and/or more precise information than can be obtained from individual sensory signals. This is not the mathematical process of integration learned in calculus (e .g., the integral of acceleration is velocity).

- **vection**: an illusory sense of self-motion caused by moving visual cues when one is not, in fact, actually moving.
- **sensory reafference**: Change in afference caused by self-generated activity. For the vestibular system, vestibular afference evoked by an active self-generated head motion would yield sensory reafference.
- **sensory exafference**: Change in afference caused by external stimuli. For the vestibular system, vestibular afference evoked by passive head motion would yield sensory exafference.
- **balance system**: The sensory systems, neural processes, and muscles that contribute to postural control. Specific components include the vestibular organs, kinesthesis, vestibulo-spinal pathways, skeletal bones, and postural control muscles. Because of the vestibular system's crucial contributions to balance, some even informally refer to the vestibular system as the "balance system" and the vestibular organs as the "balance organs." But the balance system is much more than just the vestibular system, and the vestibular system contributes to much more than just balance.
- **autonomic nervous system**: The part of the nervous system that is responsible for regulating many involuntary actions and that innervates glands, heart, digestive system, etc.

Reading

- The vestibular organs are a set of specialized sense organs located in the inner ear right next to the cochlea
- the vestibular system helps us see clearly by reflexively rotating the eyeballs in the sockets to compensate for head rotation
- The vestibular system provides the sensory foundation for spatial orientation, which includes perception of translation, rotation, and tilt.
- the vestibular system helps stabilize our eyes during head motion
- makes crucial contributions to balance but does not provide the sensory foundation for balance

  - clarity of sight but does not provide the sensory foundation

- vestibular system also helps maintain blood flow to the brain via contributions to cardiac (heart) regulation.
- active sensing balances information derived from efferent commands flowing outward from the brain to the periphery
- the human vestibular organs include the unique feature that the vertical canals are relatively larger than in other species.
- perception of spatial orientation
  - angular motion
  - linear motion
  - tilt
- The semicircular canals sense angular acceleration
- The otolith organs transduce both linear acceleration, which is a change of linear velocity, and gravity.
- Tilt perception is fundamentally different from translation perception
- Separating the otolith measurement of gravity and linear acceleration into an estimate of gravity and an estimate of linear acceleration is not easy and must be important, since the brain expends energy and effort to do so
- the rate of action potentials transmitted by afferent neurons increases or decreases following the hair cell receptor potential
- acceleration in one direction increases the receptor potentials of some hair cells, while acceleration in the opposite direction decreases those receptor potentials.
- When the head rotates, the inertia of the endolymph causes it to lag behind the motion of the head
- The firing rate increases and decreases as the angular velocity of the head increases or decreases. the change in firing rate has the same frequency as the head velocity.

- patterns of oscillatory motions that have a predominant frequency are quite common:
- Fourier analysis tells us that any complex motion can be broken down into some number of single-frequency components
- the sensing of gravity and linear acceleration relies on two structures in each ear called the otolith organs
  - utricle
  - saccule
- Small movements caused by parallel shear forces, due to gravity and/or linear acceleration, deflect the hair cells and produce changes in the firing rate of afferent neurons
- Both the utricle and the saccule include a central band called the striola
- Larger changes in the hair cell receptor potential lead to larger changes in the rate of action potentials sent to the brain via afferent neurons.
- With the head near upright, the utricle will be sensitive primarily to any Earth - horizontal linear acceleration, while the saccule will be sensitive to vertical linear acceleration.
- Different hair cells respond maximally to different movement directions, with the direction of maximal sensitivity varying systematically across the plane of each macula.
- three different techniques are frequently used to investigate spatial orientation perception: thresholds, magnitude estimation, and matching.
- The dizziness and imbalance that we experienced when we stopped rotating were due to an illusion of self-rotation caused by the semicircular-canal response.
- apparent calculation suggests that while otolith organs sense linear acceleration, our brains turn this information into a perception of linear velocity.
- Individuals suffering from severe vestibular damage generally report greater vection than do normal individuals

- Signals from the vestibular afferent neurons do not distinguish between active and passive head movements
- The angular VORis the compensatory eye rotation evoked by the semicircular canals when they sense head rotation.
- The most direct neural path for the VORs consists of an arc of three neurons that yields reflexive eye responses with a latency of less than 10 milliseconds between the start of head motion and the eye movement.
- the VOR response is more accurate than a simple reading of its input signal would provide.
- The vestibular system also makes contributions to responses of the autonomic nervous system.
- the vestibular system contributes to the regulation of blood flow to the brain
- vestibular system is helping to compensate for the applied-movement disturbance.
- vestibular system measures the movement of the head and sends commands to the postural control system that help reduce the amount of body sway.
- The vestibulo-spinal response can be thought of as a whole family of reflexes.
-