Computational Vision & Imaging Introduction and Overview

Lecture 1.1: Introduction

Hamid Dehghani

Office: CS 241

Schedule

- 2 Lectures / week
  - 3 pm Mondays
  - 2 pm Fridays
- 1 Lab/Tutorial / week
  - Varies, and groups will be assigned
- Modules webpages
  - Canvas

Schedule: Computational Vision & Imaging

Week	Week Commencing	Lecturer	Content	Lab Session
1	15/01/2024	Hamid Dehghani	1.1 Introduction to Computer Vision, Human Vision 1.2 Edge Detection and Noise Filtering	Python Tutorial
2	22/01/2024	Hamid Dehghani	2.1 Advanced Edge Detection, Scale Invariant Feature Transform 2.2 Motion	Edge Detection & Noise Filtering
3	29/01/2024	Hamid Dehghani	3.1 Image Registration 3.2 Image Segmentation	Image Registration
4	05/02/2024	Hamid Dehghani	4.3 Hough Transform 4.4 Dimensionality Reduction & Face Recognition	Hough Transform & Face Recognition
5	12/02/2024	Ales Leonardis	5.1-2 Bridging classical and DL-based CV: Case I: Edge detection, noise filtering	
6	19/02/2024	Ales Leonardis	6.1-2 Bridging classical and DL-based CV: Case II: Dimension reduction, (face) recognition	
7	26/02/2024	Jianbo Jiao	Deep Learning	
8	04/03/2024	Jianbo Jiao	Deep Learning	Catch-up
9	11/03/2024	Hamid Dehghani	Imaging in Medicine	Assignment
10	18/03/2024	Hamid Dehghani	Assignments Help	Assignment

Schedule: Robot Vision

Week	Week Commencing	Lecturer	Content	Lab Session
1	15/01/2024	Hamid Dehghani	1.1 Introduction to Computer Vision, Human Vision 1.2 Edge Detection and Noise Filtering	Python Tutorial
2	22/01/2024	Hamid Dehghani	2.1 Advanced Edge Detection, Scale Invariant Feature Transform 2.2 Motion	Edge Detection & Noise Filtering
3	29/01/2024	Hamid Dehghani	3.1 Image Registration 3.2 Image Segmentation	Image Registration
4	05/02/2024	Hamid Dehghani	4.3 Hough Transform 4.4 Dimensionality Reduction & Face Recognition	Hough Transform & Face Recognition
5	12/02/2024	Ales Leonardis	5.1-2 Bridging classical and DL-based CV: Case I: Edge detection, noise filtering	
6	19/02/2024	Ales Leonardis	6.1-2 Bridging classical and DL-based CV: Case II: Dimension reduction, (face) recognition	
7	26/02/2024	Jianbo Jiao	Deep Learning	
8	04/03/2024	Jianbo Jiao	Deep Learning	
9	11/03/2024	Hyung Jin Chang	Local Features (SIFT, SURF, HoG, etc) Multiple-View Geometry Indexing instances Image categorisation Object Recognition	
10	18/03/2024	Hyung Jin Chang		

Assessment: Computational Vision & Imaging

- Exam:
  - 50%
  - Date - TBA
- Assignment
  - 50 %
  - 3 page experimental write-up of a specified projects relating to your Module (Individual Effort)
  - Submission Date: Preliminary: 18th March @ 12.00

Labs

- Submit lab reports for feedback
- Does not count towards final Module mark

- All labs with be python based
- See canvas pages for self tutorial in week 1
  - Lab support and exercises from week 2

Advance Notice

- Attendance will be monitored
  - Lectures and Labs

Module Description

- Learning Outcomes:
  - Make informed choices about which sort of algorithms to apply to solve specific problems
  - Use standard vision libraries or software to construct working vision systems
  - Apply algorithms to simplified problems by hand
  - Discuss the advantages and drawbacks of different methods, explaining their working

What is Computational Vision?

- First consider ‘Visual Perception’
  - to know what is where, by looking.
  - vision is the process of discovering from images what is present in the world, and where it is.

The acquisition of knowledge

about

objects and events in the environment

through

information processing

of

light emitted or reflected from objects

What is Computational Vision?

- To make computers "*See*"
- "Automatic inference" of "properties" of "the world" from "images"

Automatic inference	•Inference without (or minimal) human intervention
The world	•The real unconstrained 3D physical world •Constrained/Engineered environments
Image	•2D projection of the electromagnetic signal provided by the world
Properties	•Geometric: shape, size, location, distance, •Material : color, texture, reflectivity, transparency •Temporal: direction of motion (in 3D), speed, events •Illumination: light source specification, light source color •Symbolic: objects' class, object's ID

What is Computational Vision?

Is it easy?

- All people can “see” equally well
- Babies can “see”
- Really primitive animals can “see”
- We “see” effortlessly (at least it feels this way)
- Vision is immediate
- Vision appears to be flawless

Computational Vision is challenging

- Vision needs to reverse the imaging process which is a many-to-one mapping (…recover lost information).
- Vision needs to cope with an inherently imperfect imaging process (…recover lost information)
- Vision needs to cope with discretized images of a practically continuous world (…recover lost information).
- The mere complexity of the task is enormous!
- Huge portion of our brain is dedicated to visual perception.

Approaching the problem computationally

- Constrain/simplify the world
- Constrain/simplify the task (i.e., the desired output)
- Devise universal guiding assumptions or heuristics
- Incorporate explicit knowledge
- Use experience (learning)

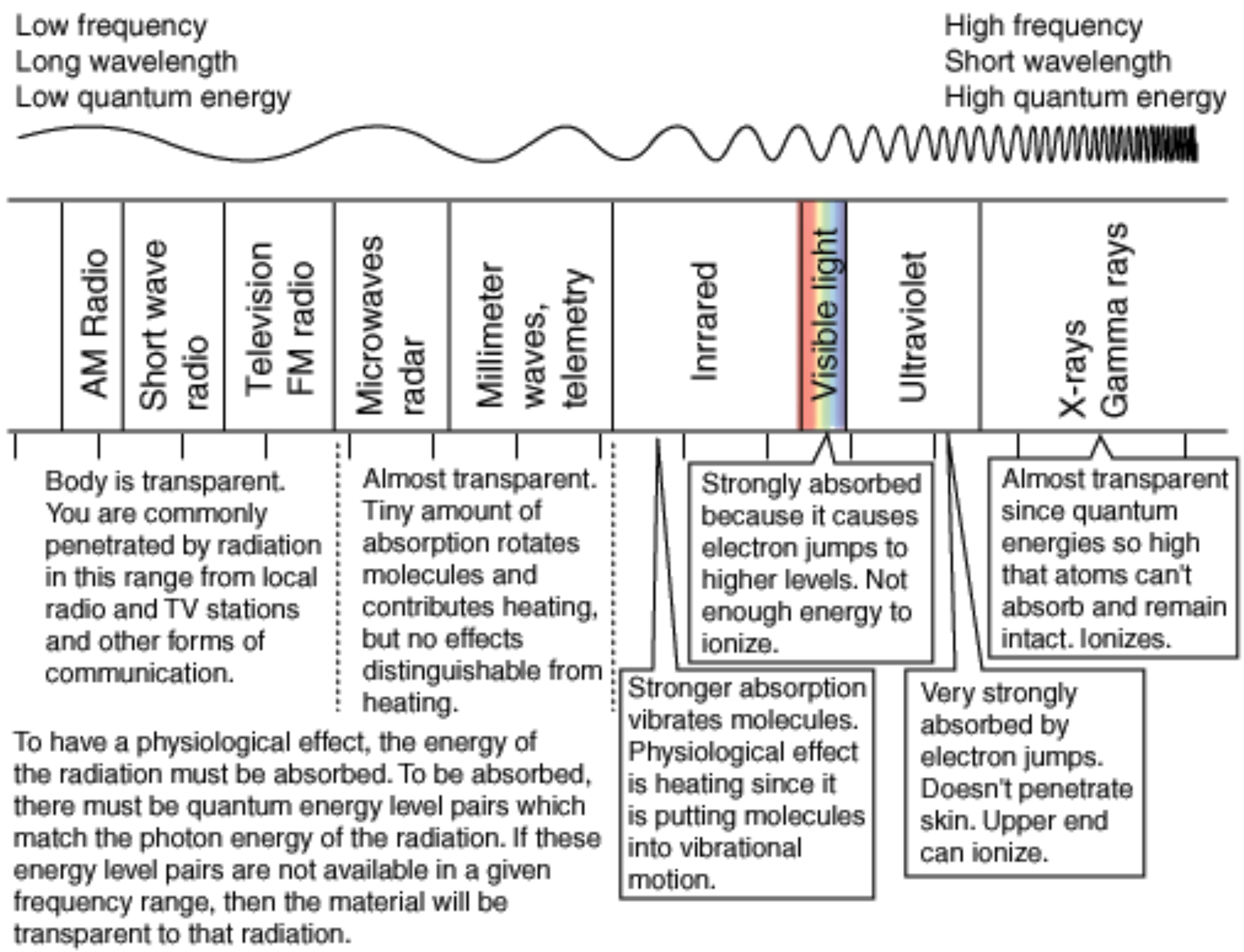
Applications

- Automated navigation with obstacle avoidance
- Object/target detection and recognition
- Place/scene recognition
- Manufacturing and assembly
- Document processing
- Quality control
- Biomedical applications
- Accessibility tools
- Human computer interfaces

Biological Vision

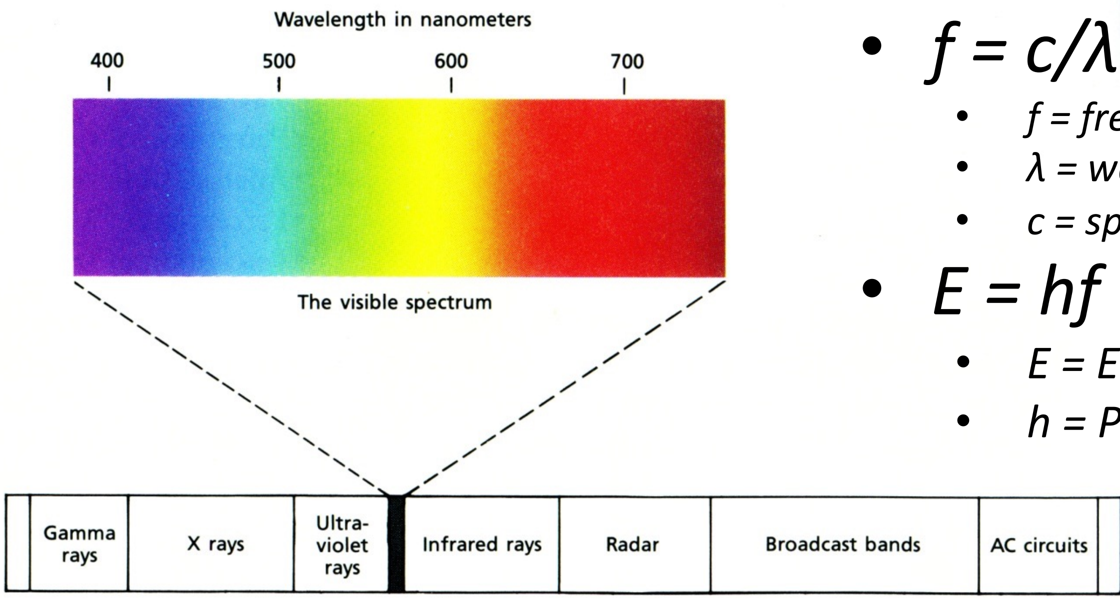
- Light and image formation
- Retinal Processing
- Visual Pathway

Electromagnetic Spectrum



Visible Light

- Humans perceive electromagnetic radiation with wavelengths 380-760nm (1 nm = 10^{-9} m)



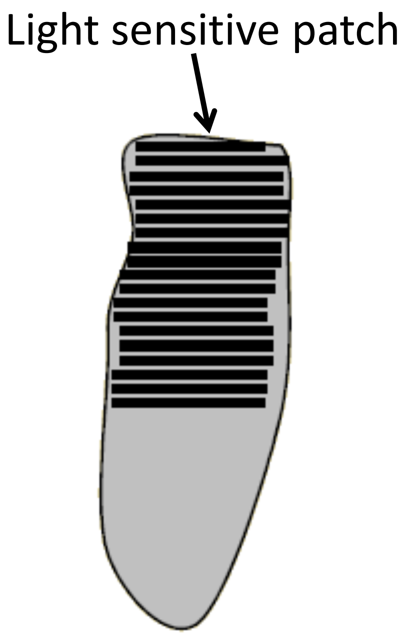
- $f = c/\lambda$
  - f = frequency (Hz)
  - λ = wavelength (m)
  - c = speed of light ($2.998 \times 10^{8}\ ms^{-1}$)
- $E = hf$
  - E = Energy (J)
  - h = Plank's constant (6.623×10^{34} Js)

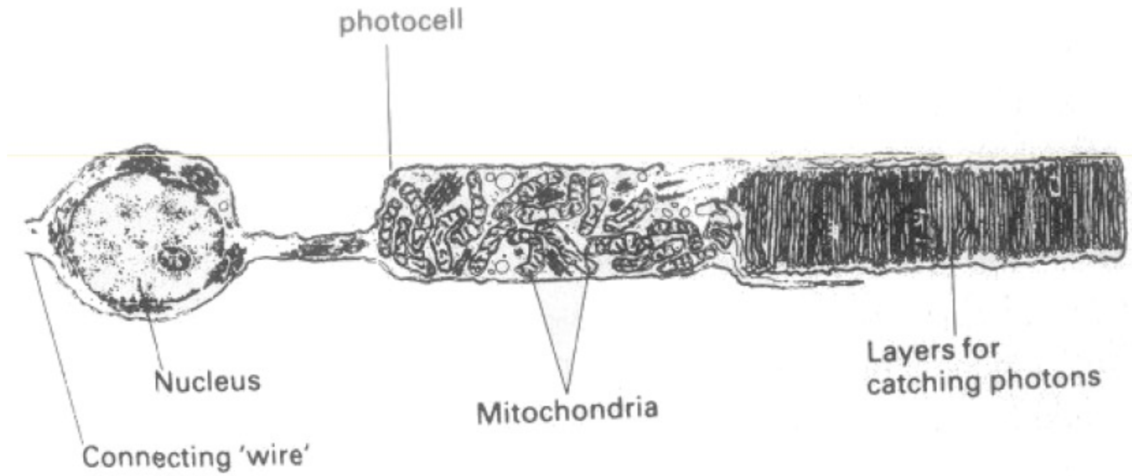
Light Capturing Devices

- In the beginning: Formation of photopigments (>3BYA)
  - Molecules in which light triggers a physical or chemical change.
  - Captured photons lead to release of energy (of different forms)
  - Released energy is used for
    - Building food (photosynthesis)
    - Behavioral reaction (nerve reaction)

Light Capturing Devices

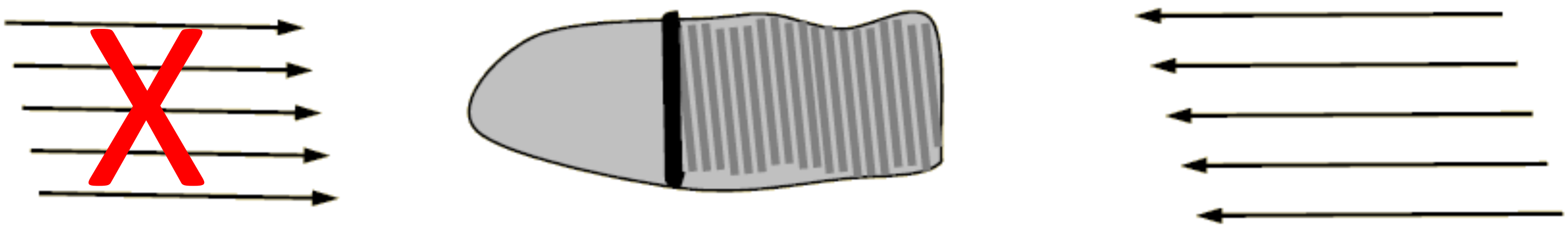
- Photocells



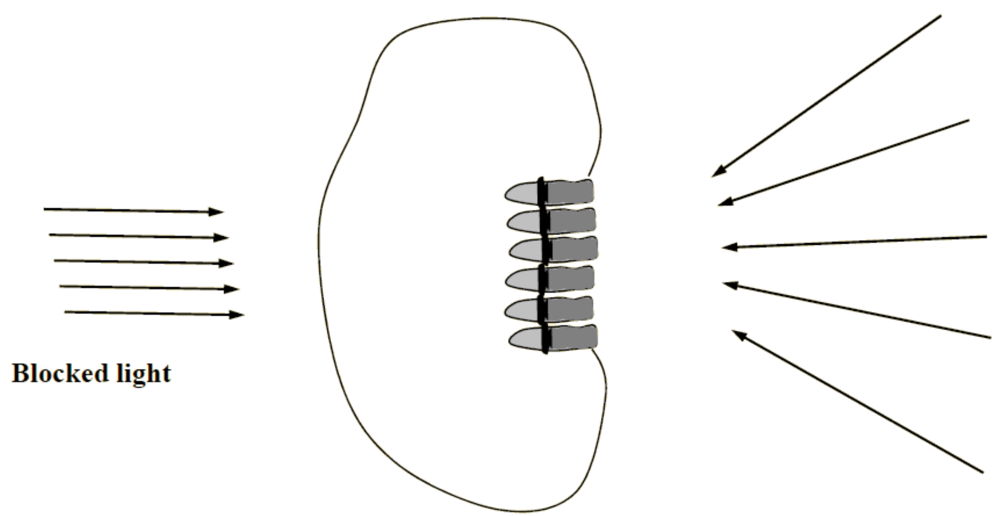


Evolution of eyes

- Single cell – 1D capture of light

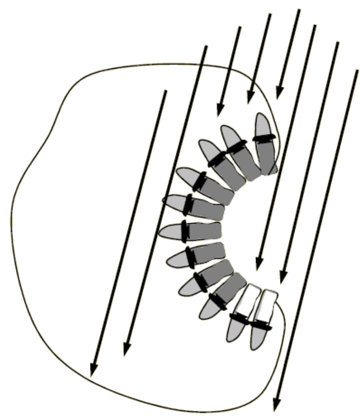


- Multiple cell – Better direction resolution

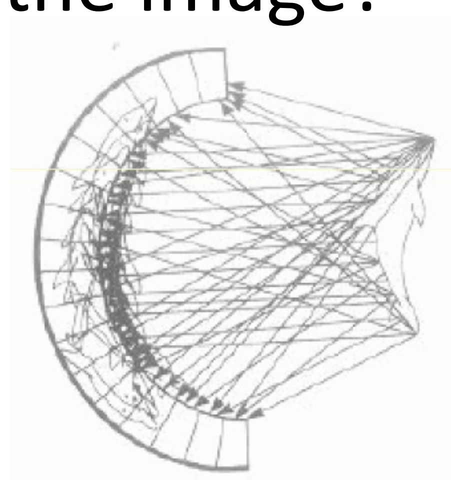


Evolution of eyes

- Multiple cell – Better direction resolution

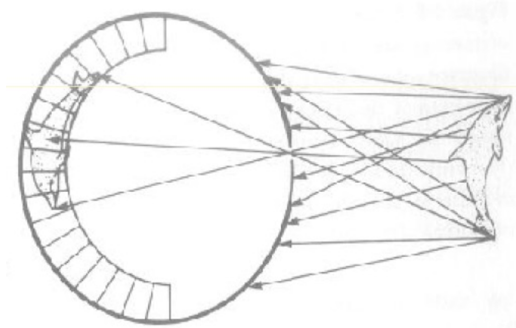


- But...where is the image?



Evolution of eyes

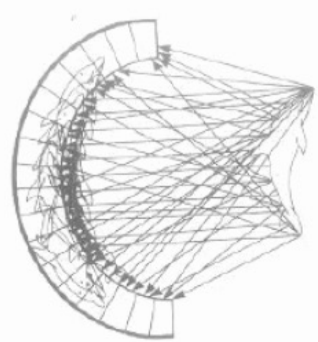
- A pinhole camera



- Dilemma:

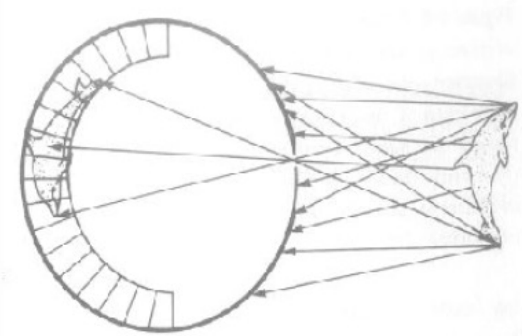
Wide aperture:

- **Bright images**
- **Fuzzy images**



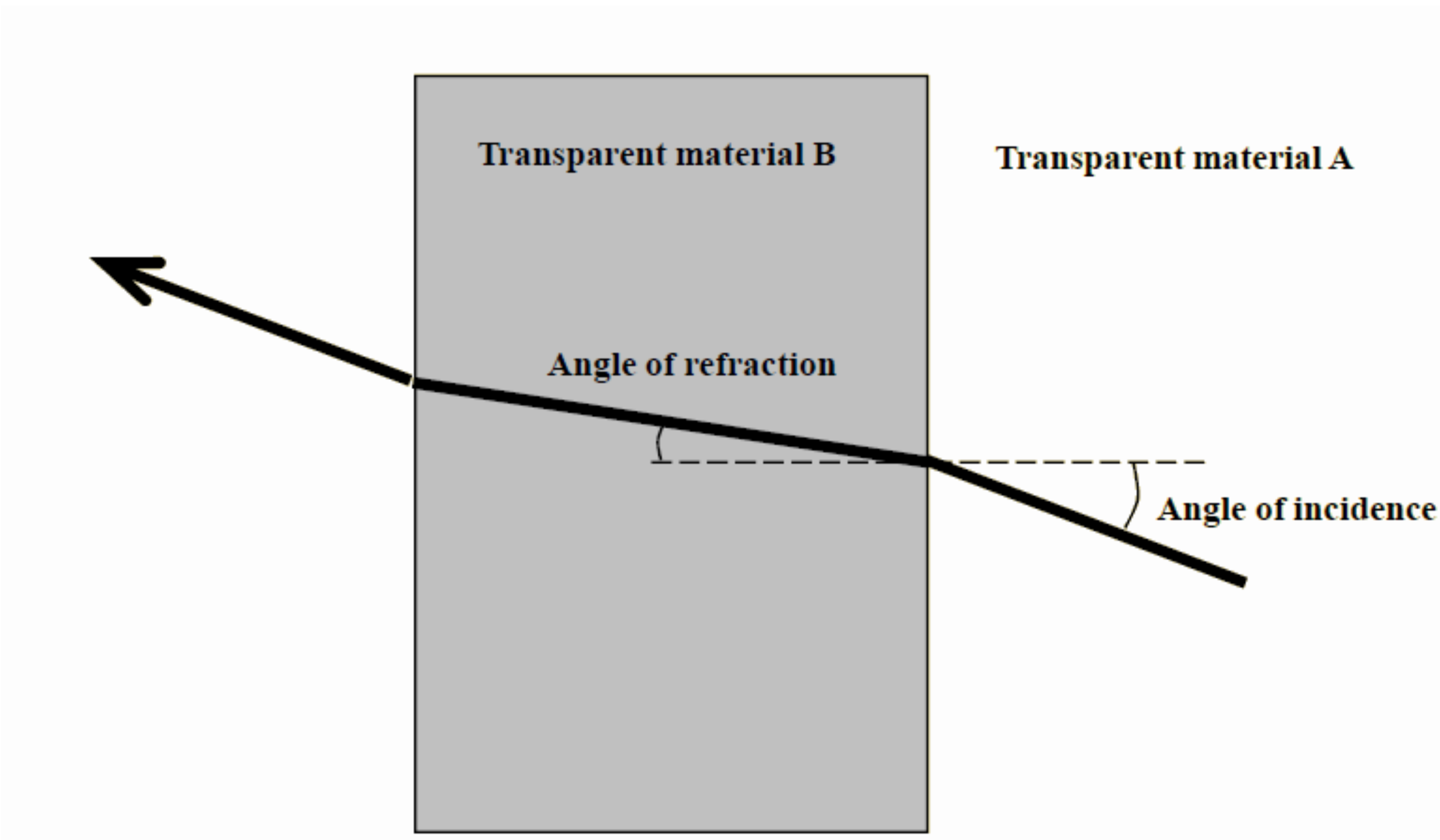
Pinhole aperture:

- **Dim images**
- **Sharp images**

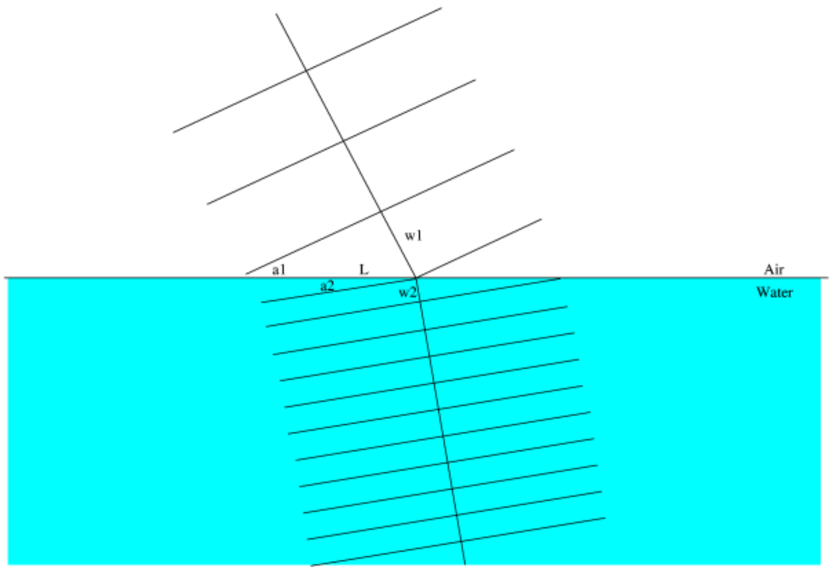


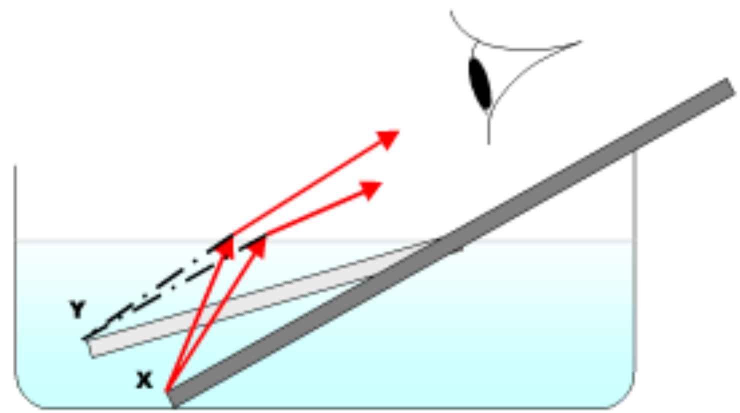
Evolution of eyes

- Solution: Use of light refraction and hence lenses



Refraction (Snell’s Law)

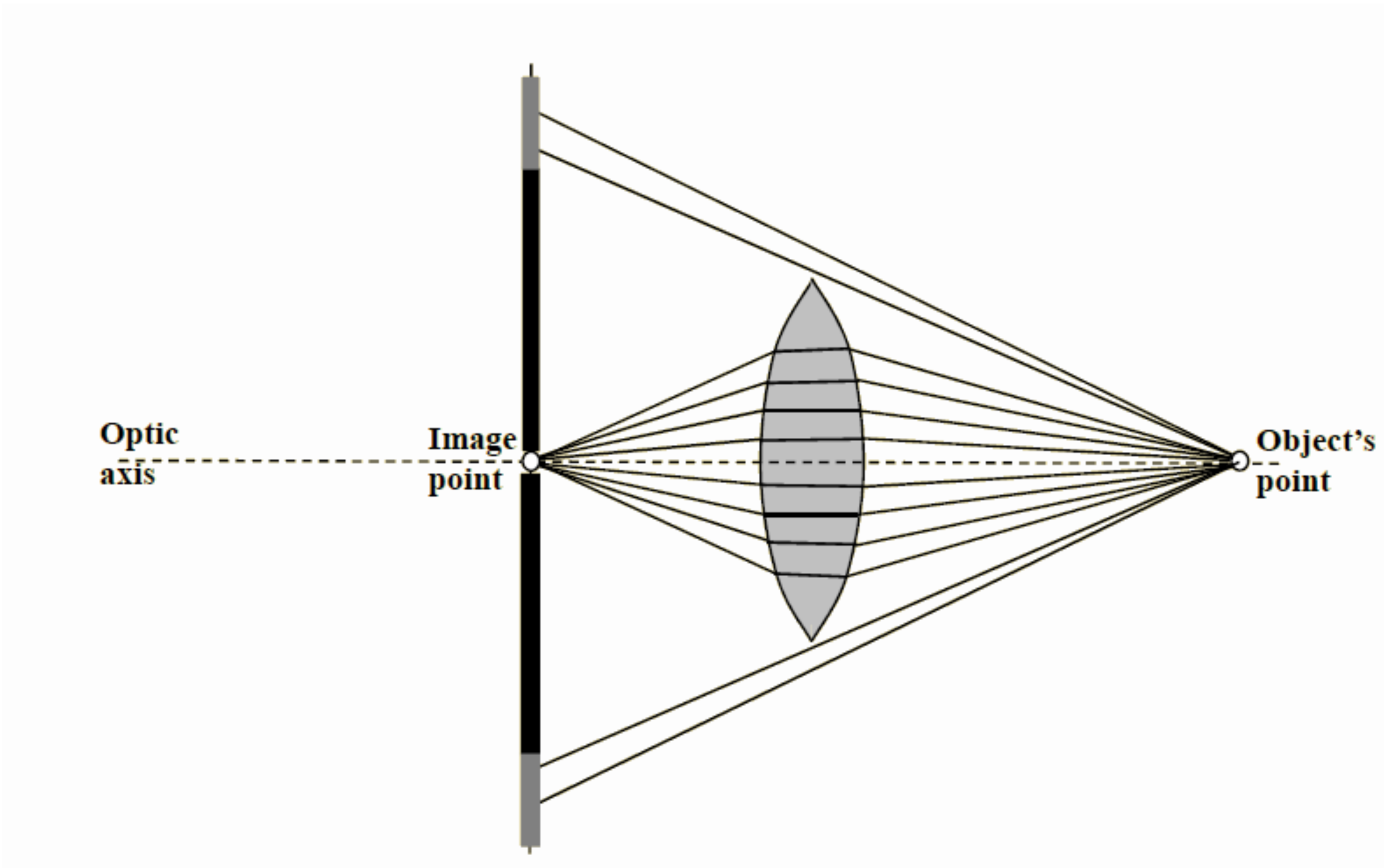




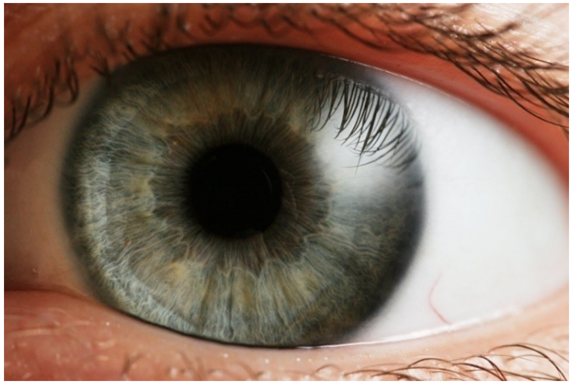
Wave crests can't be created or destroyed at the interface, so to make the waves match up, the light has to change <u>direction</u>.

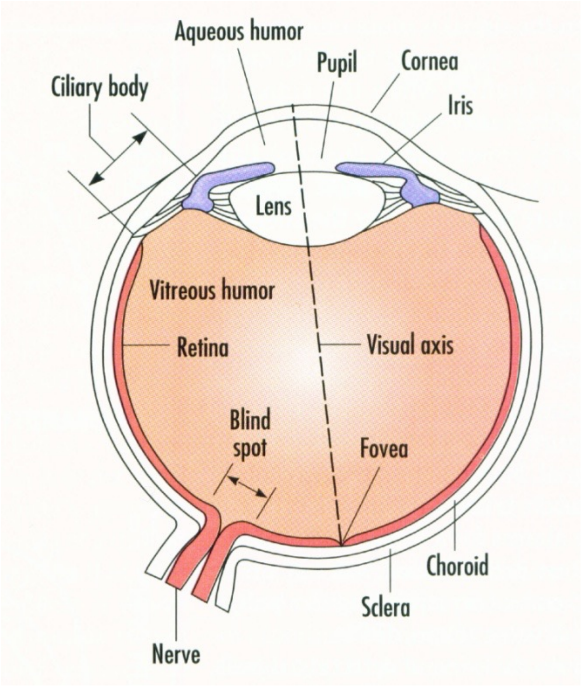
Evolution of eyes

- Formation of lens

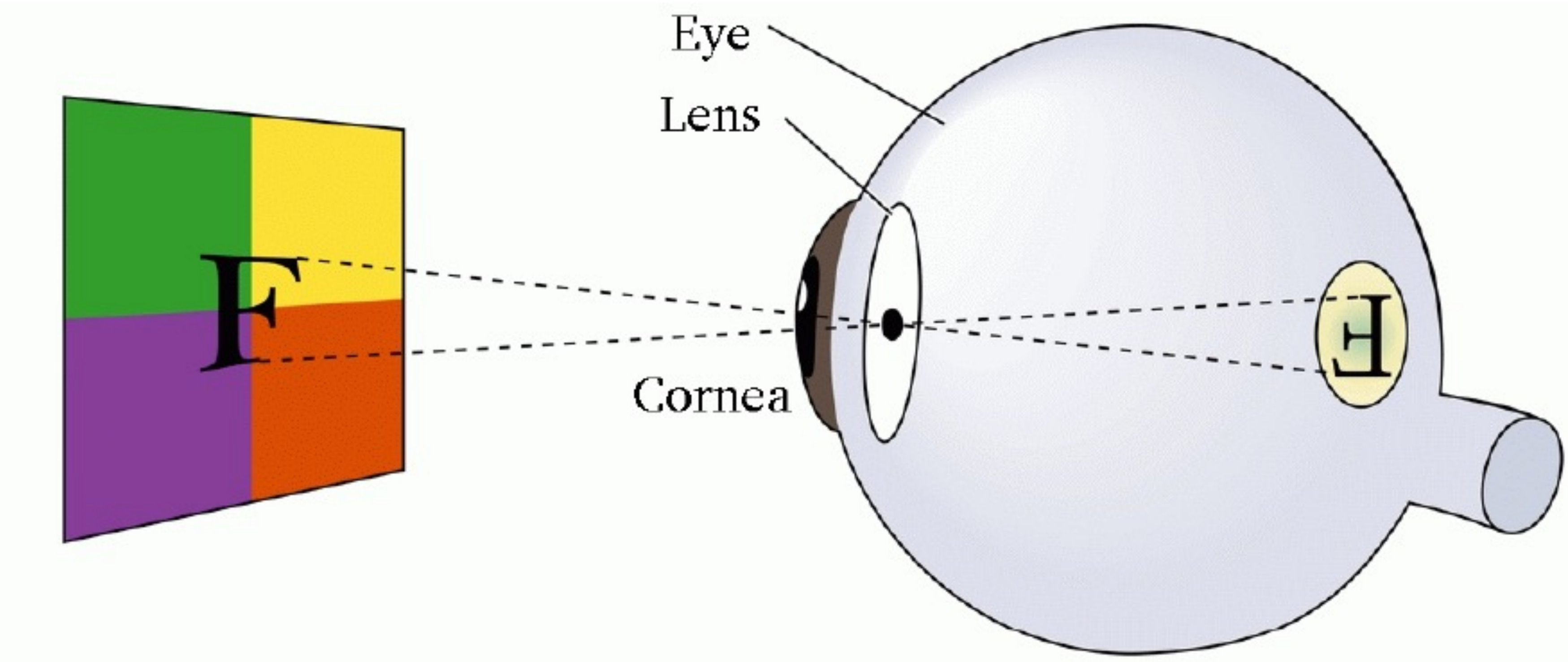


The Human Eye



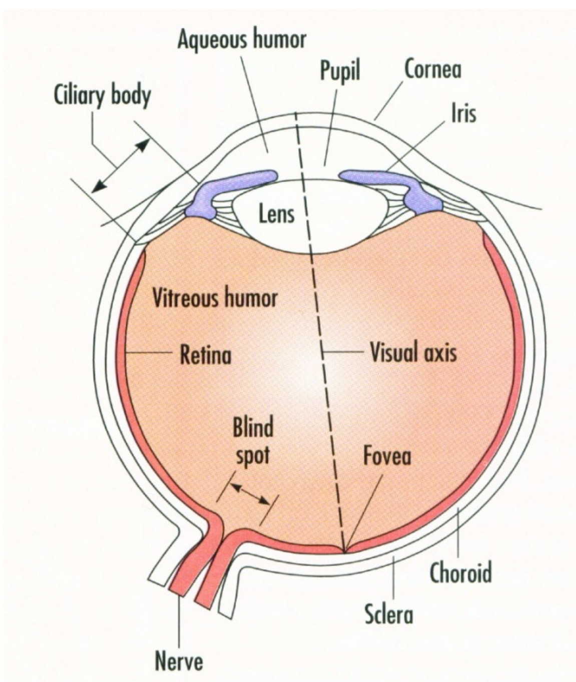


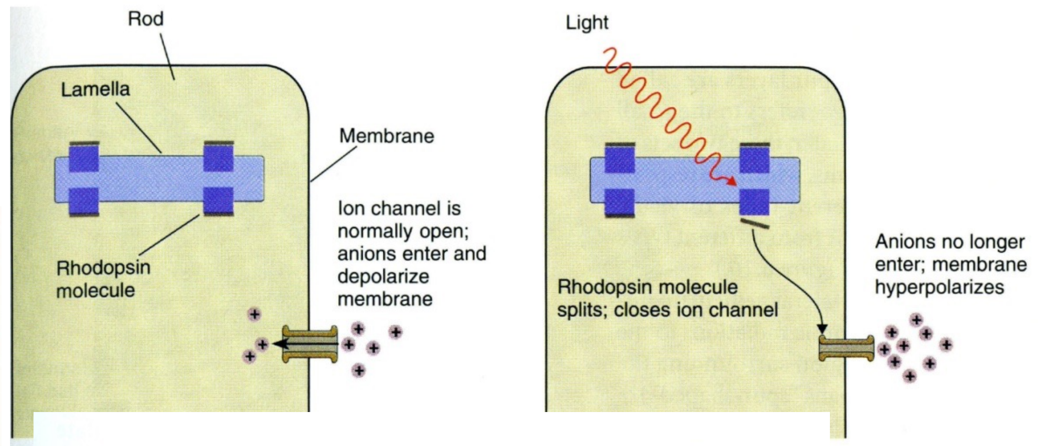
Pinhole Camera Model

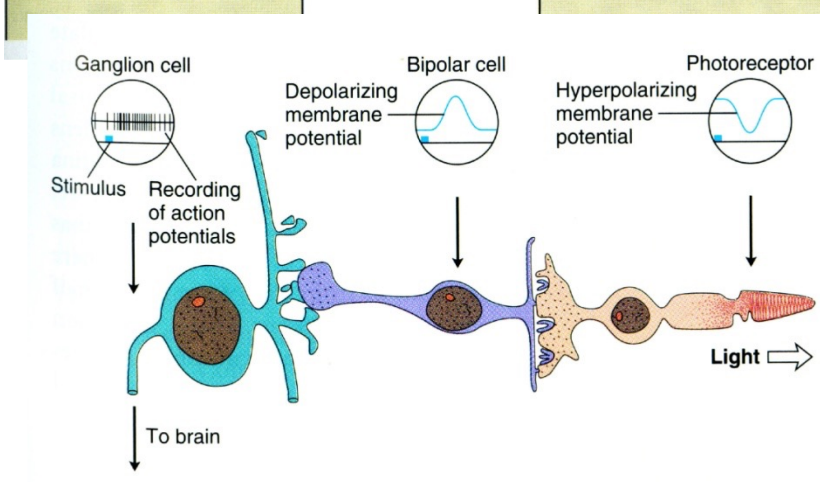


Retina Processing

- The retina contains cells that respond to light:
  - Two types of photorecepetors
    - Rods (~120 million)
    - Cones (~6 million)



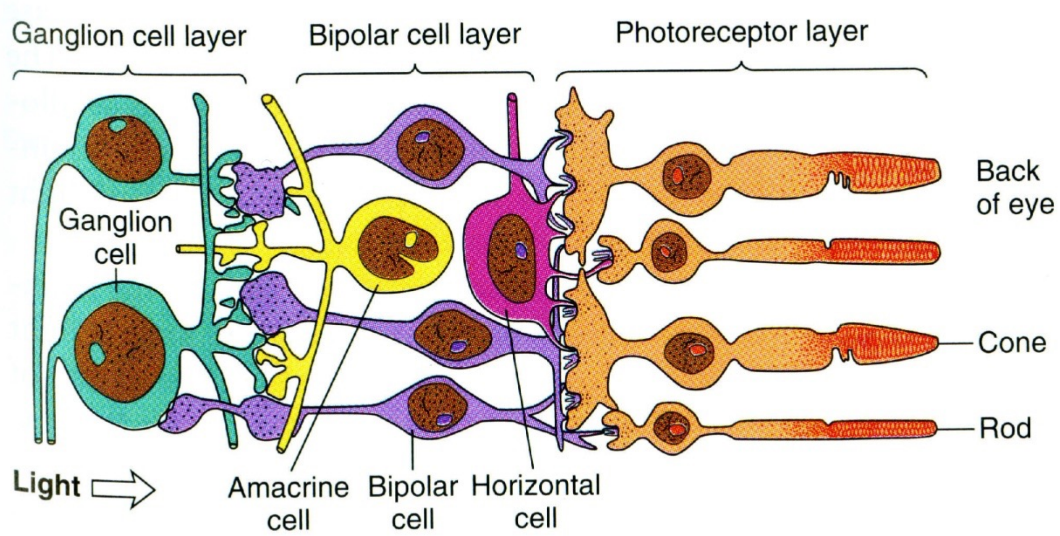




Blind Spot Test

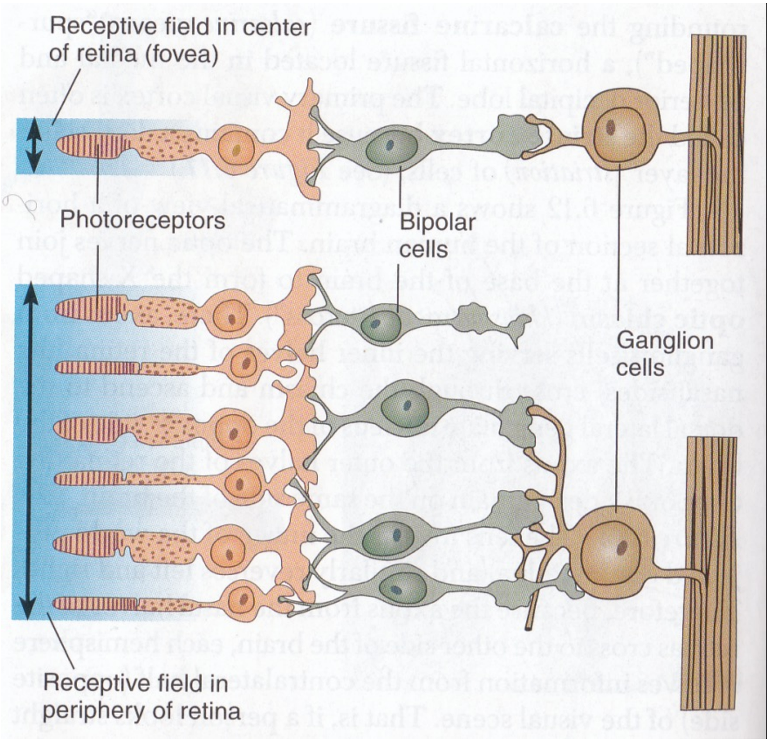
Rods and Cones

- Rods
  - ~120 m
  - Extremely sensitive photosensor
  - Respond to a single photon
  - Poor spatial resolution as they converge to same neuron within retina
- Cones
  - ~6 m
  - Active at higher light levels
  - Higher resolution as Signal processed by several neurons



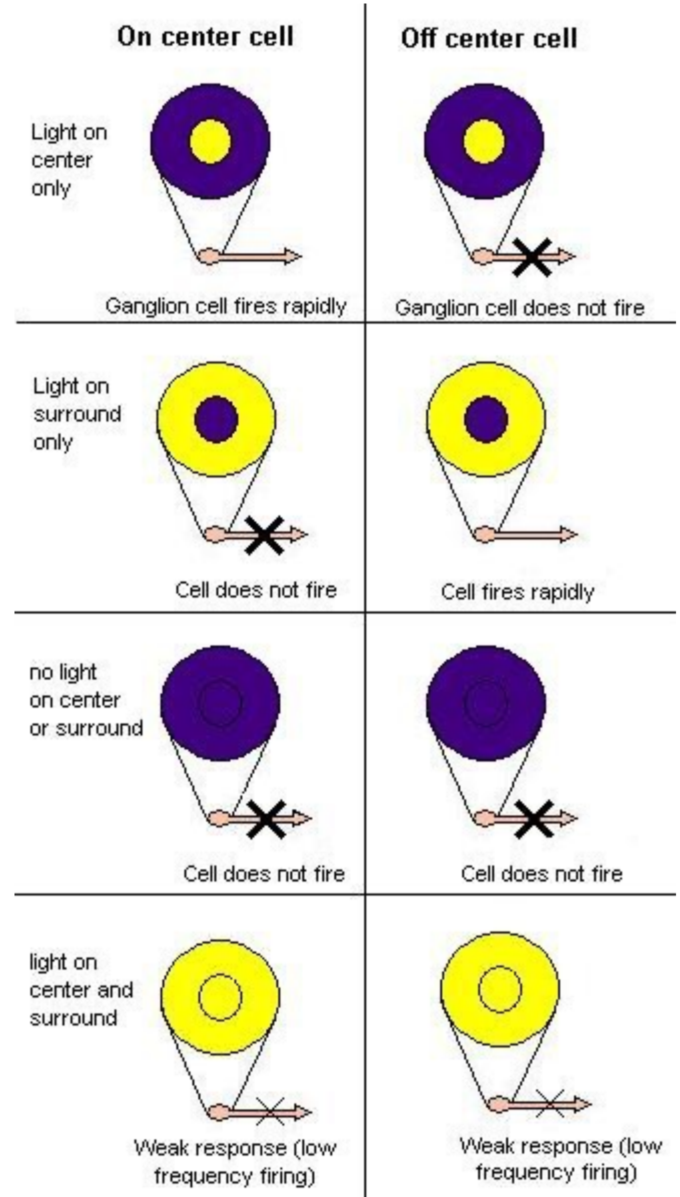
Receptive Field

- Receptive field is the area on which light must fall for neuron to be stimulated.
- Note difference between centre and periphery of field
- As early as 1938, frogs were seen to have different types of ganglion cells
- Using Cats, (electrical recordings from ganglion cells) it was seen that receptive field contains a circular centre surrounded by a ring



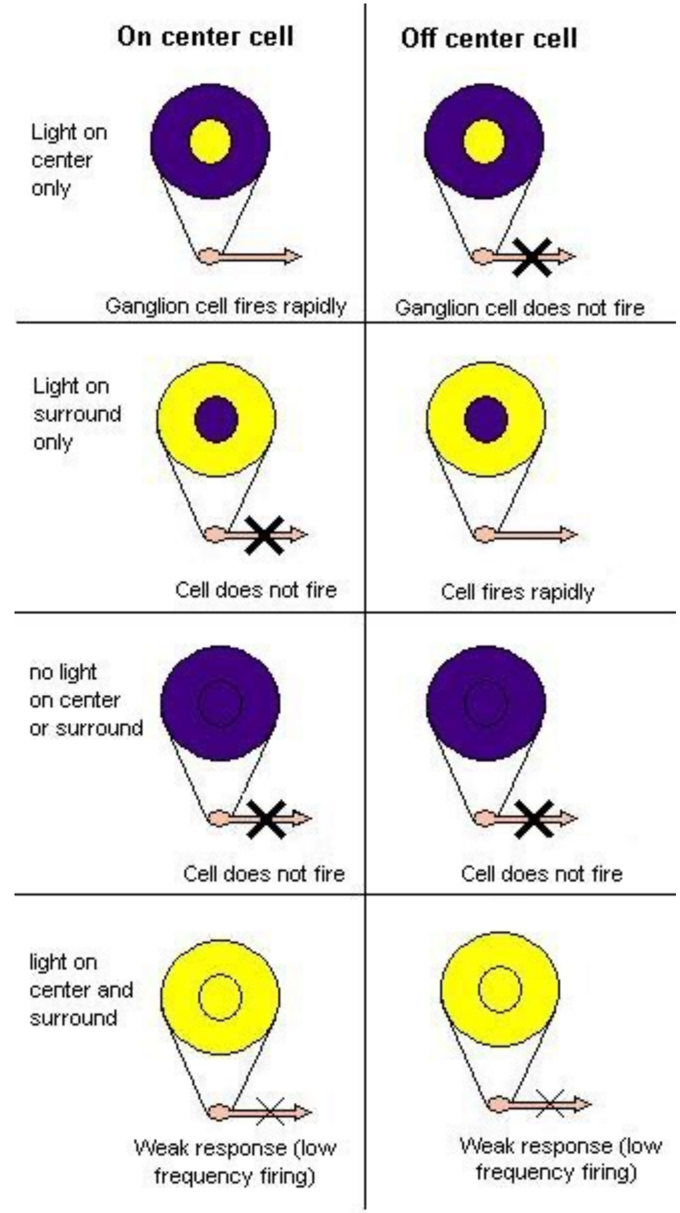
Ganglion cells

- Two types: "on-center" and "off-center".
  - On-center: stimulated when the center of its receptive field is exposed to light, and is inhibited when the surround is exposed to light.
  - Off-center cells have just the opposite reaction



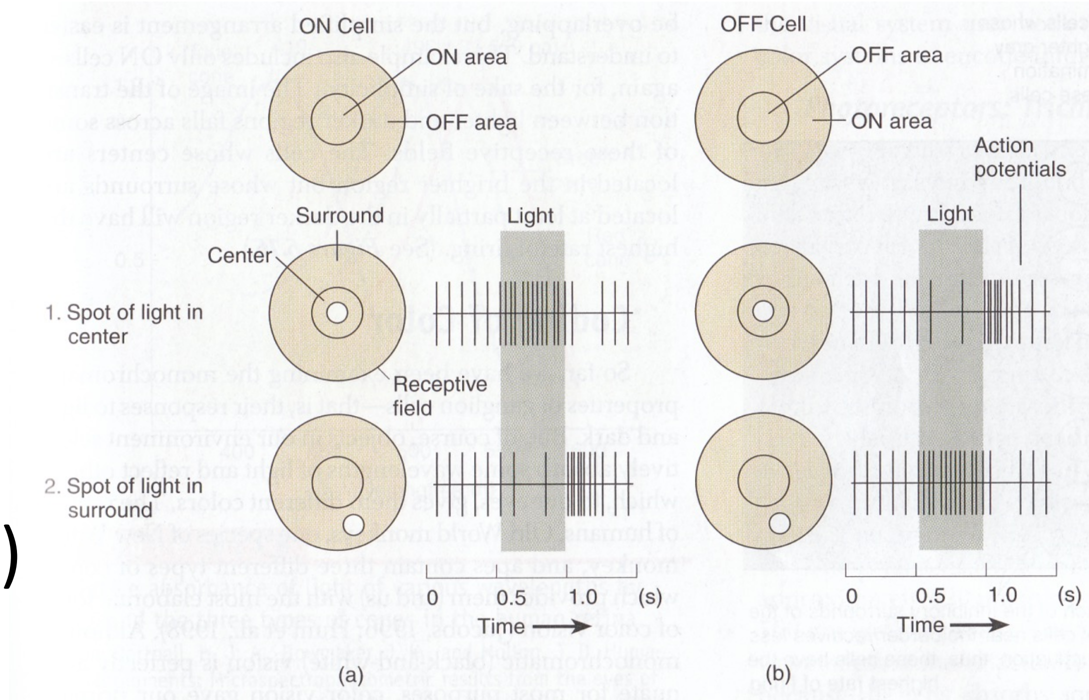
Ganglion cells

- Allows ganglion cells to transmit information not merely about whether photoreceptor cells are firing (Photoreceptors do not actually fire action potentials), but also about the differences in firing rates of cells in the center and surround.
- Allows transmission of information about contrast.
- The size of the receptive field governs the spatial frequency of the information

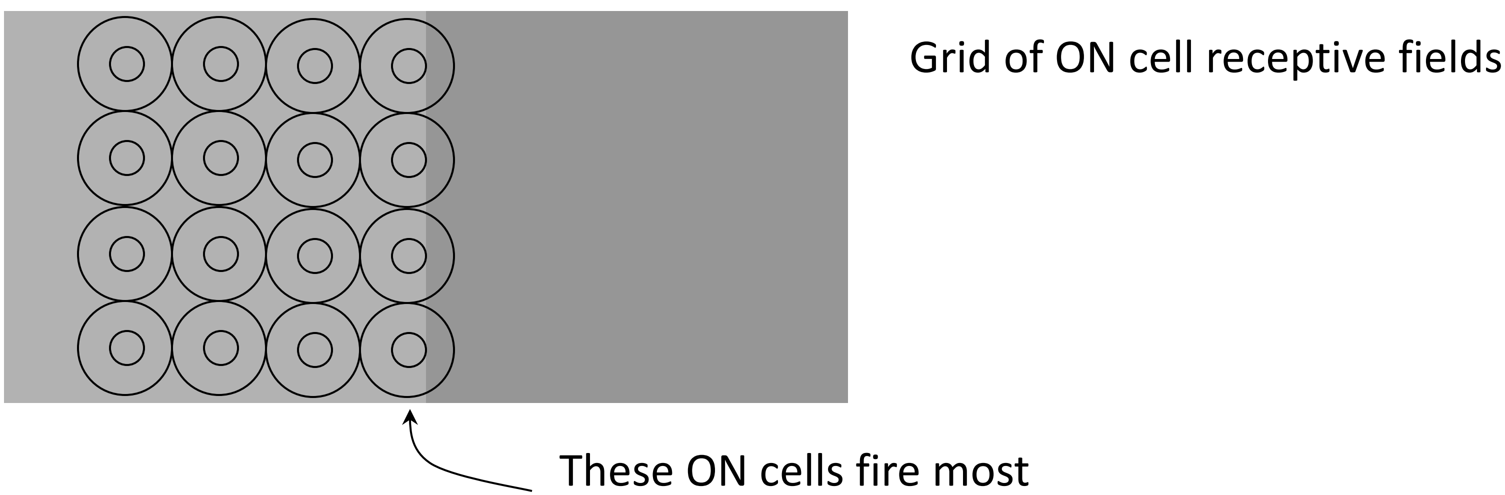


ON/OFF Cells

- Note
  - area of stimulation / inhibition
  - Rate of signal firing (rebound)

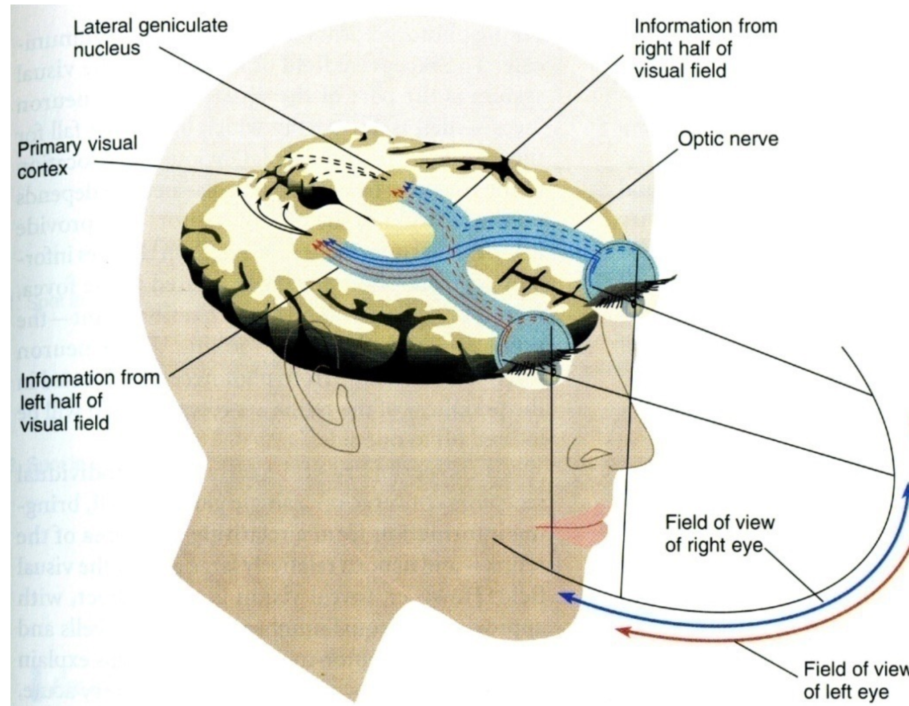


Enhancement of Contrast



Visual Pathway

- Vision generated by photoreceptors in the retina
- The information leaves the eye by way of the optic nerve
- There is a partial crossing of axons at the optic chiasm.
- After the chiasm, the axons are called the optic tract.
- The optic tract wraps around the midbrain to get to the lateral geniculate nucleus (LGN)
- The LGN axons fan out through the deep white matter of the brain and ultimately travel to primary visual cortex, at the back of the brain.



Summary

- Module Outline
- Uses of Computational Vision
- Image formation
- Very early visual processing
- Human vision
- Edge detection

Reading

- Vicki Bruce, Visual Perception, Chapters 1 - 3
- Neil Carlson, Physiology of Behavior, Chapter 3, “Vision”