INTRODUCTION

Ole-Johan Skrede

23.01.2017

INF2310 - Digital Image Processing

Department of Informatics
The Faculty of Mathematics and Natural Sciences
University of Oslo

After original slides by Fritz Albregtsen

- **Practical information**
- **Course content overview**
- **Motivation and applications**
- **Imaging systems**

PRACTICAL INFORMATION

Lecturers

Are Charles Jensen `are@ifi.uio.no`

Ole-Johan Skrede `olejohas@ifi.uio.no`

Teaching Assistant

Kristine Baluka Hein `krisbhei@ifi.uio.no`

- Main textbook: *Digital Image Processing, 3rd edition, R. C. Gonzalez and R. E. Woods, 2008*
- In addition, some extra material may be published on the course website `http://www.uio.no/studier/emner/matnat/ifi/INF2310/v17/`
- A complete list of this year's curriculum will be available near the closing of the course.
- Additionally, I will post the content from my part of the course at `https://ojskrede.github.io/inf2310/`.

WEB RESOURCES

As mentioned, the course website is at
`http://www.uio.no/studier/emner/matnat/ifi/INF2310/v17/`
There, you will find

- Lecture slides
- Weekly exersices
- Solution to weekly exersices
- Mandatory assignments
- Course material
- Schedule
- Exam information
- Additional messages

Lectures	Wednesday 12:15 - 14:00 at room *Java* (2423). Presentation of weekly content.
Group lectures	Friday 12:15 - 14:00 at room *Fortress* (3468). Assistance with weekly exercises.

Mandatory assignments In order to qualify for the exam you have to pass *both* mandatory assignments. They will consist of practical problems, and will involve some programming (choose whatever language you want, but python or matlab is recommended).

Midterm exam 4 hr written exam at 21.03.2017 that will count 30% of your final grade.

Exam 4 hr written exam at 01.06.2017 that wil count 70% of your final grade.

COURSE CONTENT

FIRST HALF

Date	Theme	Lecturer
23.01.2017	Introduction	Ole-Johan
25.01.2017	Sampling and quantization	Are
01.02.2017	Geometrical operations	Are
08.02.2017	Graylevel mapping	Are
15.02.2017	Histogram based operations	Are
22.02.2017	Neighbourhood operations I	Ole-Johan
01.03.2017	Neighbourhood operations II	Ole-Johan
08.03.2017	Color images and color spaces	Ole-Johan
15.03.2017	Mid-term repetition	Are, Ole-Johan
21.03.2017	Mid-term exam	

SECOND HALF

Date	Theme	Lecturer
22.03.2017	Fourier Transform I	Are
29.03.2017	Fourier Transform II	Are
05.04.2017	Segmentation	Are
26.04.2017	Compression and coding I	Ole-Johan
03.05.2017	Compression and coding II	Ole-Johan
10.05.2017	Morphology	Ole-Johan
24.05.2017	Repetition	Are, Ole-Johan
01.06.2017	Final exam	

INTRODUCTION

LEARNING OUTCOME

- What does image processing software (Gimp, Photoshop, ImageJ etc) *actually* do to the images.
- What can I do with the images from my digital camera.
- Multimedia background.
    - TIFF, GIF, PNG, PBM, JPEG.
    - Image representation, compression, formats, color spaces.
    - Image improvement
- Solid foundation for further studies in image processing.
    - Image filtering
    - Edge detection
    - Geometrical operations
    - Segmentation
- Programming experience, *learning by doing*.

Image processing is the study of methods producing an *altered* or *compressed* image. Alterations can e.g. be improved image representations or image pre-processing for a certain task. Examples are blurring, contrast enhancement, and JPEG compression.

Image analysis is the study of methods of collecting and analyzing information in an image. Examples being object detection, image classification, and image segmentation. Image processing is often a vital part of an image analysis method.

INF2310 aims to give a thurough introduction to, and a solid understanding of *image processing*.

INF4300 — Digital Image Analysis (fall)

- Feature extraction
- Image segmentation
- Pattern recognition
- Supervised and unsupervised image classification

INF5860 — Machine Learning for Image Analysis (spring)

- Premiere this year (spring 2017)
- Central machine learning algorithms
- Deep neural networks
  - CNN: Convolutional neural networks
  - RNN: Recurrent neural networks
- Learning to use TensorFlow (Google's framework for Deep Learning)

MOTIVATION AND APPLICATIONS

NON-EXHAUSTIVE LIST OF APPLICATIONS

- Medical applications (e.g. ultra sound, MRI, CT)
- Industrial inspection
- Traffic surveillance
- Text recognition, document processing, maps
- Identification with facial recognition, fingerprints or iris detection (biometri)
- Earth observation from satelite images
- Seafloor mapping
- Oil reservoir exploration (seismic, electromagnetic imaging)

- Segment of machine learning that has revolutionalized areas of cognitive computer tasks (such as computer vision, image analysis, natural language processing, game engines, etc.).
- Since 2012 [1] [2], ”every” state-of-the-art method in image analysis is based on artificial neural networks.
- Noticable sub-categories of neural networks are *Recurrent Neural Networks* (RNN) and *Convolutional Neural Networks*.

[1] Krizhevsky, A., Sutskever, I. and Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks NIPS 2012: Neural Information Processing Systems, Lake Tahoe, Nevada

[2] Really some 30-year old idea with a small (compared to now) hype at the end of the 1980s, which somewhat died down. The recent renaissance kicked off with the above cited contribution.

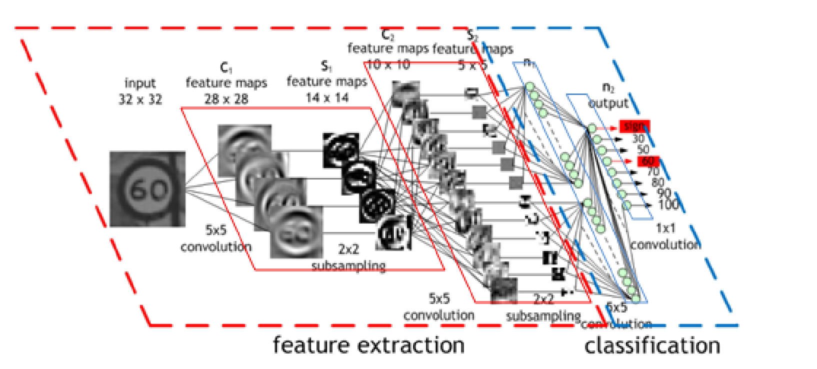

Figure 1: Traffic sign classification (Source: NVIDIA)

IMAGING SYSTEMS

Electromagnetic waves can be described by their frequency, or wavelength

$$f = \frac{c}{\lambda}$$

- f: Frequency
- λ: Wave length
- $c = 299792458 m/s$: Fundamental maximum speed of information in this universe (the speed of light).

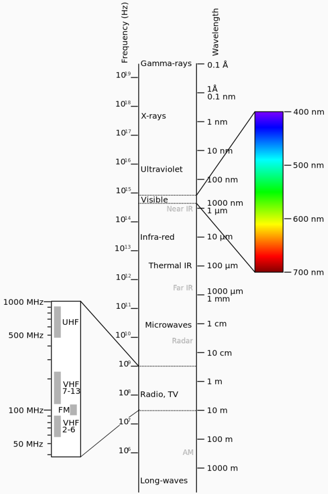

Figure 2: The electromagnetic spectrum (Source: By Victor Blacus - SVG version of File:Electromagnetic-Spectrum.png, CC BY-SA 3.0, `https://commons.wikimedia.org/w/index.php?curid=22428451`)

EXAMPLES

- Photography
- The eye
- Optical satelite: Landsat
- Radar satelite: SAR
- Infrared satelite
- Medical ultra sound
- X-ray (Roentgen) and computed tomography (CT) (multiple X-ray images)
- Magnetic resonance imaging (MRI), (radio waves, magnetic fields, field gradients)
- Sonar and seismic imaging (using sound)
- Microscopy
- Laser distance sensors (range imaging)

OPTICAL SYSTEMS

LA TRAHISON DES IMAGES

Figure 3: *La trahison des images* - René Magritte (1929). English title: *The Treachery of Images*. It is an image of a pipe, with the caption *Ceci n'est pas une pipe*, french for *This is not a pipe*.

POINT SOURCE

- We will study the light ray trajectory from a point source.
- A point will be imaged as a ”blurred” version of itself.
- The reason for this is two unrelated phenomena: *abberation* and *diffraction*.

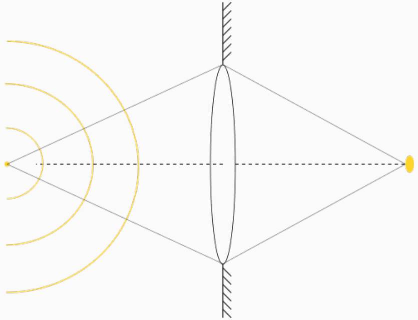

Figure 4: Light wave propagation from point source through circular aperture.

POINT SPREAD FUNCTION

- The point spread function (psf) describes the response from of an imaging system on a point source.
- The resulting image is obtained by convoluting the source with the point spread function.
- Can be used to measure the blurring of a point object, and is a measure of the quality of the imaging system.

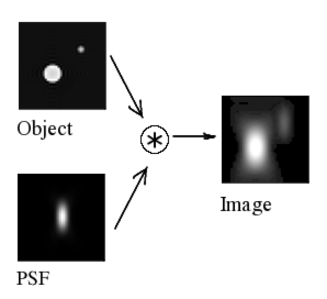

Figure 5: Point spread function. (Source: By Default007 - Own work, Public Domain `https://commons.wikimedia.org/w/index.php?curid=877065`)

OPTICAL ABBERATION

- Distortion in the image formed by an optical system.
- Arise because of limitations in optical components such as lenses and mirrors.
- There are many types of optical abberations, and they can to a large extent be repaired.

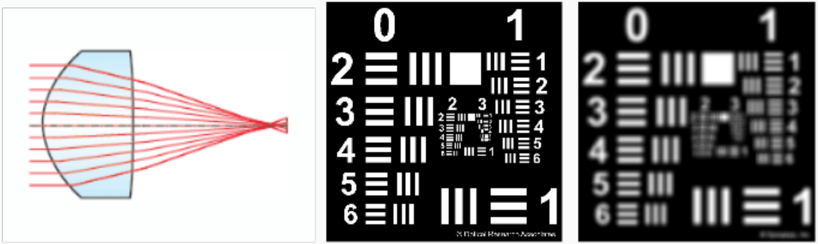

Figure 6: Spherical abberation. Rays from the edge of the lens focus points along a line in stead of at a single point. (Source: `http://www.edmundoptics.com`)

- Caused by wavefronts of propagating waves bending in the neighborhood of obstacles.
- A fundamental limit in imaging.
- Possible to ”avoid” this limit with super-resolution imaging.

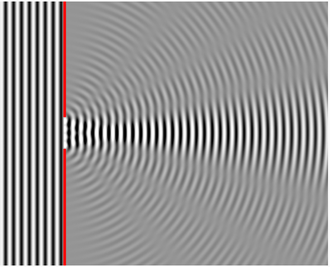

Figure 7: Illustration of diffraction pattern from a slit of width four wavelengths with an incident plane wave. (Source: By Dicklyon at English Wikipedia - Transferred from en.wikipedia to Commons by Shizhao using CommonsHelper., Public Domain, https://commons.wikimedia.org/w/index.php?curid=5699291)

- The abberation in a system can often be handled by adjusting the optics.
- The diffraction is a physical limit due to the wave propagation property of light.
- A system is said to be *diffraction limited* if diffraction is the dominating the point spread function. That is, if the spatial resolution is as good as the instrument's theoretical limit.
- The resulting image from a point source imaged from a diffraction limited optical instrument is termed a *Airy disk*.

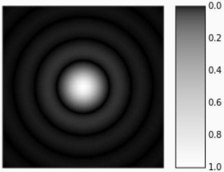

Figure 8: Airy disk

SPATIAL RESOLUTION

- Spatial (angular) resolution describes the ability of an imaging system to distinguish details in an image.
- Because of the blurring, described by the point spread function, there is a limit to the resolving ability of an image.

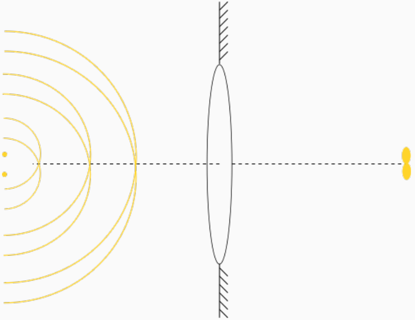

Figure 9: Resulting image from two point sources after being imaged by a circular aperture.

- The Rayleigh criterion describes the smallest distance between two light point sources that we need in order to obseve them as separate objects.
- The Rayleigh criterion: Two objects are just dissolved when the mode of the intensity function from one of the objects overlaps with the first zero of the second object's intensity function.

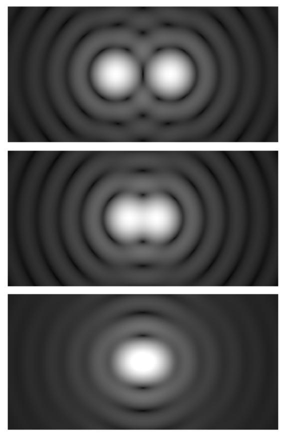

Figure 10: Airy disks of two objects. Well dissolved (top), just dissolved (middle), not dissolved (bottom). (Source: By Spencer Bliven - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=31456019)

FRAUENHOFER DIFFRACTION PATTERN FOR A CIRCULAR APERTURE

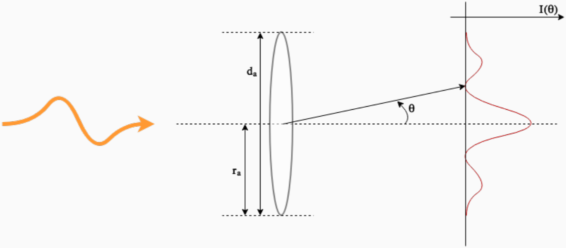

Figure 11: Diffraction pattern from circular aperture.

$$I(\theta) \propto \left[\frac{2J_1(kr_a \sin\theta)}{kr_a \sin\theta}\right]^2 \quad (1)$$

- J_1: A *first order Bessel function of the first kind*
- $k = \frac{2\pi}{\lambda}$: Wave number of the propagating light
- λ: Light wave length.
- r_a: Aperture radius.
- θ Angle to the first zero.

- The intensity function in eq. (1) is zero when the Bessel function is zero.
- The first zero of a Bessel function[1] $J_1(x)$ occurs at $x \approx 3.8317$,
- Thus

$$kr_a \sin\theta \approx 3.8317$$

or

$$\sin\theta \approx 1.22\frac{\lambda}{d_a} \tag{2}$$

which is famously known as the *Rayleigh criterion*.

[1] Bessel functions are solutions to a particular set of differential equations, and the mathematics is somewhat complicated. But feel free to stydy them on your own if you want to understand how the roots are found.

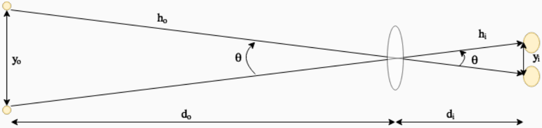

Figure 12: Schematic figure of distances.

For very small θ, $\sin\theta \approx \theta$. We can then approximate eq. (2) to

$$\theta \approx 1.22\frac{\lambda}{d_a}$$

and furthermore, $h_o \approx d_o$, yielding

$$y_o \approx 1.22\frac{\lambda}{d_a}d_o$$

This is the smallest distance we are able to dissolve in an image.

With reference to fig. 13, we use the term *thin lens* when the lens thickness d_l is neglible when compared to the radii of the curvature of the two lens surfaces, r_1 and r_2, respectively.

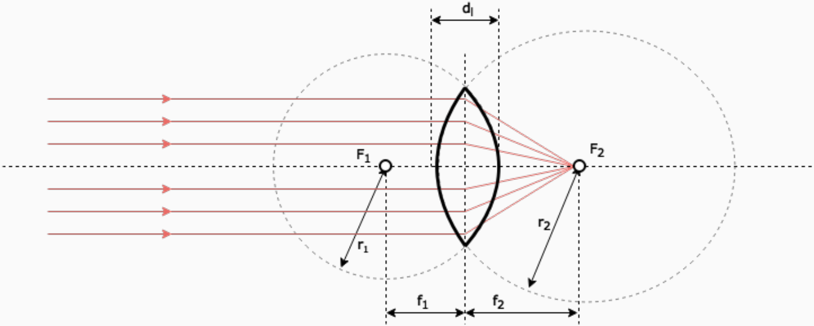

Figure 13: Thin lens

For a lens in air (*refractive index* of about 1), the focal length f_2 is approximated by the *lensmaker equation*[1]

$$\frac{1}{f_2} \approx (n-1)\left[\frac{1}{r_2} - \frac{1}{r_1} + \frac{(n-1)d_l}{nr_1r_2}\right].$$

Here, n is the refractive index of the lens material, and we see that when we invoke the thin lens assumption (that is, $d_l << r_1r_2$), the above approximation is simplified to

$$\frac{1}{f_2} \approx (n-1)\left[\frac{1}{r_2} - \frac{1}{r_1}\right]. \quad (3)$$

[1] This equation is a bit too long to derive here.

PARAXIAL RAY APPROXIMATION

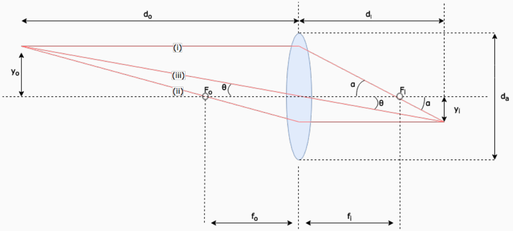

Figure 14: Idealised imaging system

(i) An incident ray which is parallel to the optic axis is refracted through the imaging focal point F_i.

(ii) An incident ray which passes through the object focal point F_o, is refracted parallel to the optic axis.

(iii) An incident ray which passes through the optic center is refracted with the same refraction angle as incident angle (will not change direction).

The rays following the respective patterns are illustrated in fig. 14.
From (*iii*), we see that the angle θ is the same, and therefore

$$\frac{y_o}{d_o} = \frac{y_i}{d_i}. \tag{4}$$

Likewise, from the two equal angles α on ray (*i*), we get that

$$\frac{y_o}{f_i} = \frac{y_i}{d_i - f_i}. \tag{5}$$

Rearranging eq. (4) yields

$$\frac{y_o}{y_i} = \frac{d_o}{d_i}$$

which, when combined with the relation in eq. (5) yelds

$$\frac{d_o}{d_i} = \frac{f_i}{d_i - f_i}. \tag{6}$$

Equation (6) can be rearranged to the familiar form

$$\frac{1}{f_i} = \frac{1}{d_i} + \frac{1}{d_o}. \tag{7}$$

Equation (7) is sometimes referred to as the *thin lens equation*.
Using eq. (4), we get that

$$\frac{1}{f_i} = \frac{1}{d_o}\left(\frac{y_o}{y_i} + 1\right),$$

which we can rearrange to

$$y_i = \frac{y_o f_i}{d_o - f_i}$$

This gives us (for our small, idealised imaging system) the relation between the actual size in the object plane y_o and the observed size in the imaging plane y_i.

Depth of field (DOF) is the distance (in the object plane) that seems to bee in focus. As can be seen in fig. 15, only the green bulb is in focus, and everything in front and in the back of this green bulb is out of focus (blurry).

Figure 15: Depth of field (Source: http://staciviers.blogspot.no/2015/10/depth-of-field.html)

DEPTH OF FIELD AND APERTURE SIZE

In general, the DOF increases with decreasing aperture diameter, as can be seen in fig. 16, but the relationship is not trivial.

Figure 16: Depth of field and aperture size (Source: https://damienfournier.co/dof-and-aperture/)

DEPTH OF FIELD LIMITS

- Under certain assumptions, we can compute the limits of the depth of field, s_N and s_F.
- The lens is symmetrical, and so thin that we can apply the thin-lens assumption discussed above.
- The subject is at a distance s from the lens, and would be in focus at a image distance h.
- c is the *circle of confusion*.

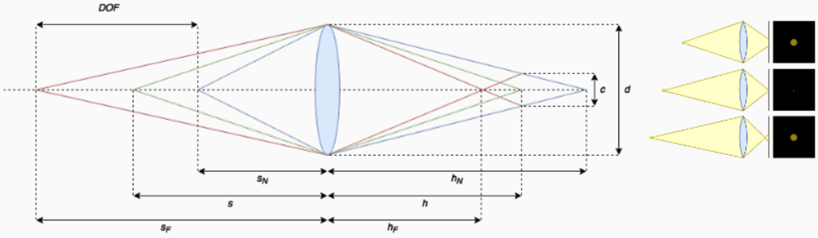

Figure 17: Depth of field model (left). Circle of confusion (right).

From similar triangles in fig. 17 we have that

$$\frac{d}{h_N} = \frac{c}{h_N - h} \iff h_N = \frac{h}{1 - c/d}, \tag{8}$$
$$\frac{d}{h_F} = \frac{c}{h - h_F} \iff h_F = \frac{h}{1 + c/d}. \tag{9}$$

Using the thin-lens approximation in eq. (7), we get the following three relations

$$\frac{1}{f} = \frac{1}{s} + \frac{1}{h} \iff s = \frac{fh}{h - f}, \tag{10}$$
$$\frac{1}{f} = \frac{1}{s_N} + \frac{1}{h_N} \iff s_N = \frac{fh_N}{h_N - f}, \tag{11}$$
$$\frac{1}{f} = \frac{1}{s_F} + \frac{1}{h_F} \iff s_F = \frac{fh_F}{h_F - f}. \tag{12}$$

We can use eq. (8) to eq. (12) to derive relations for the DOF limits s_N and s_F. For s_N, we begin with eq. (11), and substitute in results from eq. (8) and eq. (10). Similarly, for s_F, we begin with eq. (12), and substitute in results from eq. (9) and eq. (10). With this, we arrive at

$$s_N = \frac{sf^2}{f^2 + Nc(s - f)}$$
$$s_F = \frac{sf^2}{f^2 - Nc(s - f)},$$

where we have substituted in the *f-number* $N = f/d$.

THE EYE AND OUR VISUAL SYSTEM

- Providing an ability to sense light, which enables conscious (and unconscious) visual perception.
- The lens is deformable.
- Variable geometric resolution.
- Can distinguish about 10 million colors.
- Images are formed and processed by the brain.

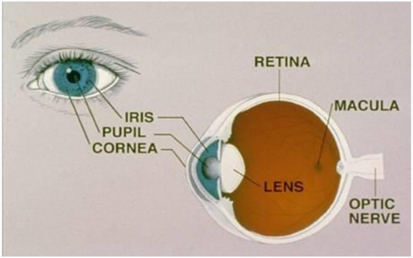

Figure 18: Illustration of the mammalian eye.

- The lens system of the eye focuses the light.
- Focal length, $f \approx 1.5\text{cm}$.
- Often use the unit of measurement *dioptre* $d = 1/f$.
- The eye lens is about $d = 67\frac{1}{\text{m}}$, meaning that the lens focuses parallel rays of light at $1/67\text{m}$.
- The cornea (hornhinne) is about $d = 45\frac{1}{\text{m}}$.
- The eye lens has the ability to change its focal length.
- The ability to change focus quickly (accommodation) decreases with age.

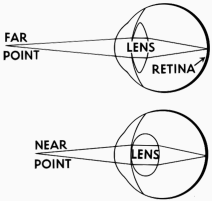

Figure 19: Minimum (top) and maximum (bottom) accommodation. (Source: By derivative work: ZirgueziAccommodation(PSF).png: User:OldakQuill - User:OldakQuill, Public Domain)

Iris

- The colored part of the eye.
- Works as a blender.
- The pattern is used in identity verification.

Pupil

- The black ”opening” in the iris.
- Diameter ≈ 2mm in bright light.
- Diameter ≈ 8mm in weak light.
- Lets light through to the retina (netthinne).
- The light does not escape it.

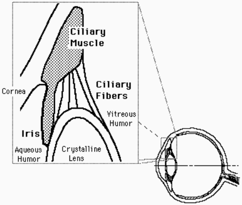

Figure 20: Iris, detailed illustration.

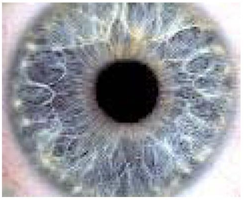

Figure 21: The iris and the pupil.

- The retina is the light sensitive layer at the back of the eye.
- Covers about 65% of the inner surface.
- About 130 million detectors.
- Two types of detectors: rods and cones.
- The detectors faces away from the light.

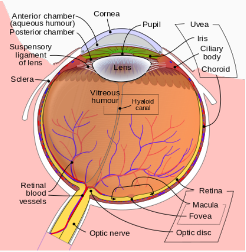

Figure 22: The human eye. (Source: By Rhcastilhos - Schematic_diagram_of_the_human_eye_with_English_
Public Domain, `https://commons.wikimedia.org/w/index.php?curid=1597930`)

RODS

- About 120 million rods.
- Distributed uniformly over the retina.
- Scotopic (low light) vision, luminance level 10^{-6} to 10^{-3}cd/m^2.
- Graylevel vision
- Slow: takes about 30 min to change to scotopic vision.
- Mostly sensitive light with wavelengths of about 498 nm (green/blue). Insensitive to wavelengths above 640nm (red).

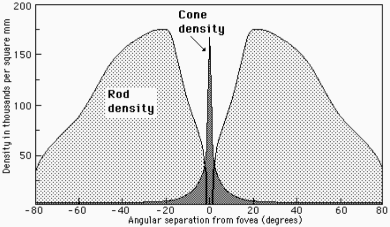

Figure 23: Distribution of rods anb cones on the retina.

CONES

- About 7 million cones.
- Concentrated location at the fovea.
- Photopic (bright light) vision, luminance level 10 to $10^8 cd/m^2$.
- Fast: takes about 5 min to change to photopic vision.
- Three types of cones to sense light in three color bands. Allows color perception.

Questions?