Universität Stuttgart

Pattern Recognition
Chapter 2: Human Visual System

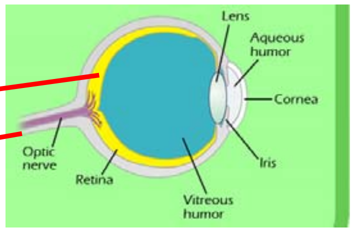
Prof. Dr.-Ing. Uwe Sörgel
soergel@ifp.uni-stuttgart.de

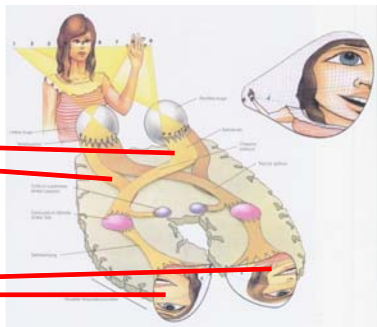
Apl. Prof. Dr.-Ing. Norbert Haala
haala@ifp.uni-stuttgart.de

ifp

Human Visual System

- Important parts:
  - Eye
    - Retina
    - Optic nerve
  - Optic track
  - Visual cortex





Contents

Signal processing

- Retina
  - Receptor cells are stimulated by light
  - Early signal processing steps, e.g.:
    - Contrast enhancement at edges
- Brain (including visual cortex)
  - Visual perception, e.g.:
    - Recognition of edges and their slope
    - Object recognition
    - Determine depth
    - Multi-level processing
- The processing is done in parallel
- How humans perceive the environment is subject of research and still not fully understood yet.

Related Sciences

Neurophysiology:
„How does human vision work?"
(e.g. mode of operation of cones/rods)

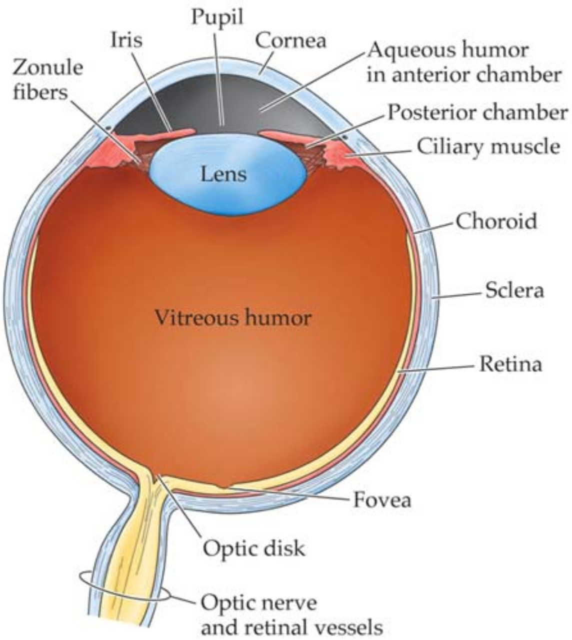
认知
Cognitive psychology:
„How does human visual perception work?"
(e.g. Gestalttheorie)

Contents

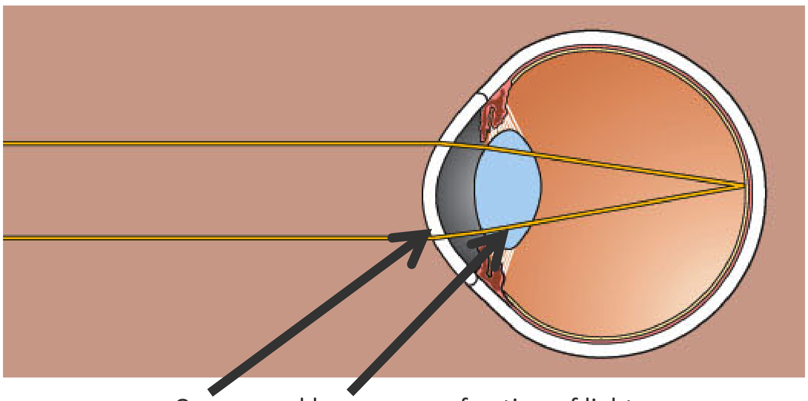
- Image acquisition
  - Eye
  - Retina
  - Photoreceptor cells
- Adaption of eye to brightness
- Signal processing in retina
- Further processing
- "Gestalttheorie"
- Optical illusions

Human eye

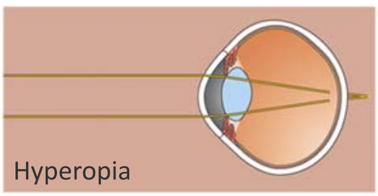
- Pupil → aperture
- Ciliary muscle causes contraction of lens to control accommodation (focus)
- Retina
  - *Rods* work at dim light and are sensitive to brightness
  - *Cones* support perception of color, three types for red, green and blue

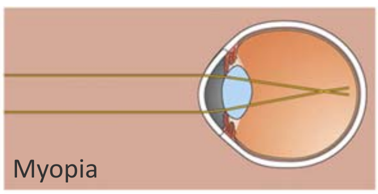


Accommodation of Eye

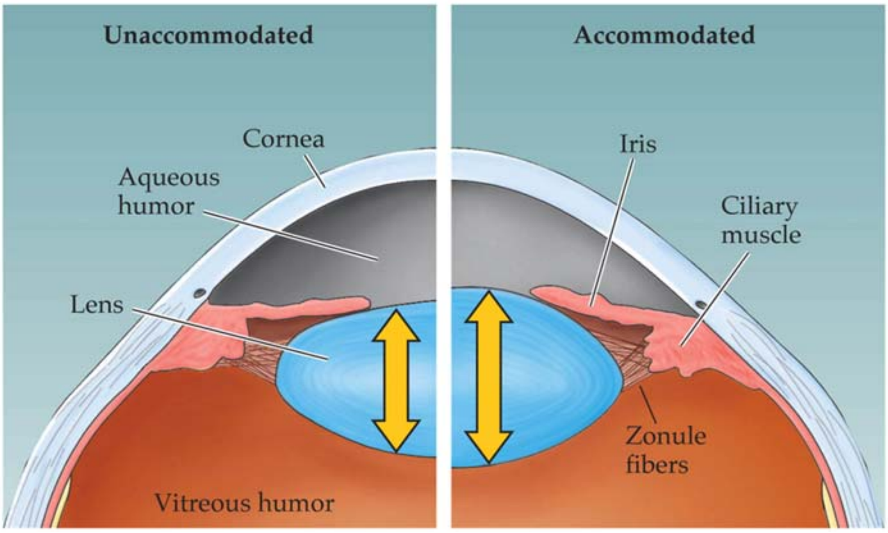


Cornea and lens cause refraction of light.
Change of curvature of lens by ciliary muscle enables focus on certain depth.





Accommodation of Eye

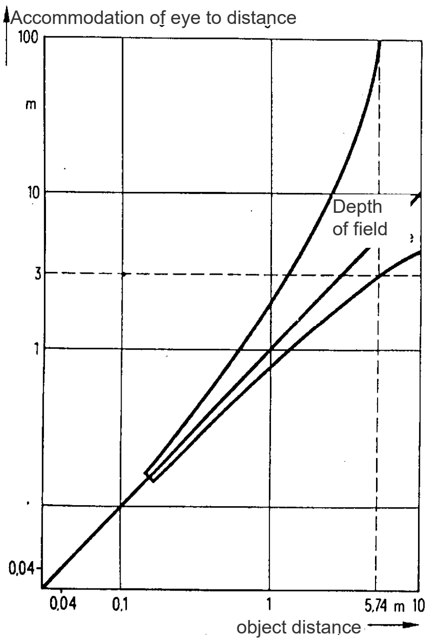


- Continuous adaption of refractive power leads to variable depth of field.
- Enables sharp image on retina for objects located in different distance to observer

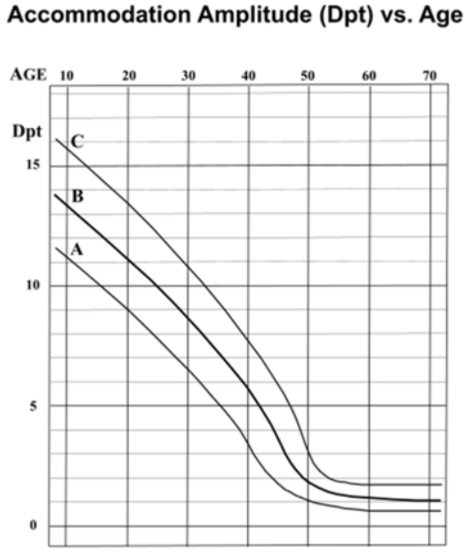
Accommodation of Eye

- Depth of field after accommodation (for mid level pupil diameter).
- Accommodation at hyperfocal distance of 5,74 m produce acceptably sharp images in a depth range between 3 m to ∞

→ Fix-focus camera



Accommodation ability fades with age

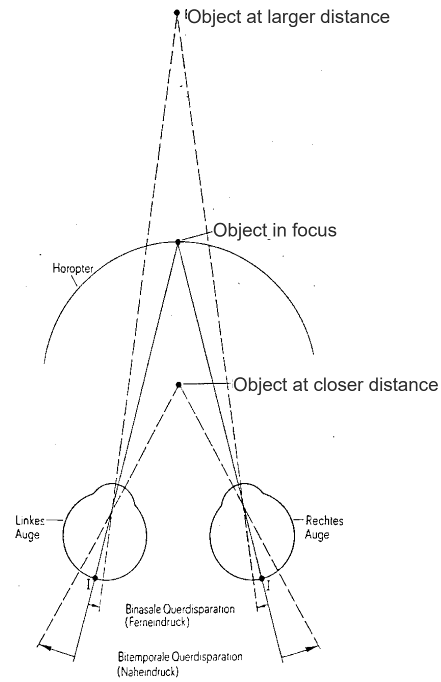


Diopter:=
1 / focal length

Mean (B) and approximate lower (A) and upper (C) standard deviations are show

Principle of stereoscopic vision (I)

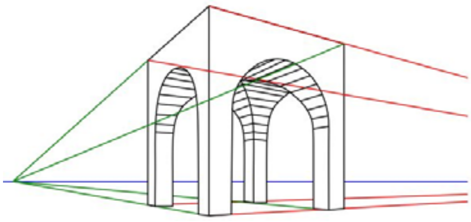
- Convergence: Rotation of eyes so that the object appears at the center of the retina in both eyes.
- Accommodation onto object of interest.
- Rays of objects either closer or further away are mapped laterally shifted on retina compared to fixated object *I*.
- The direction of lateral disparity is for objects further away in both eyes towards nose, whereas the image of closer objects is shifted towards temples. This disparity is evaluated in the brain for depth perception.



Principle of stereoscopic vision (II)

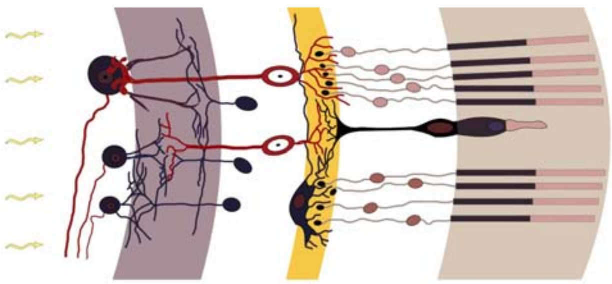
- Besides convergence and accommodation for depth perception also other concepts are used:

  - Occlusions
  - Perspective
  - Knowledge about object size and structure
  - Comparison with objects nearby of known size



Structure of Retina

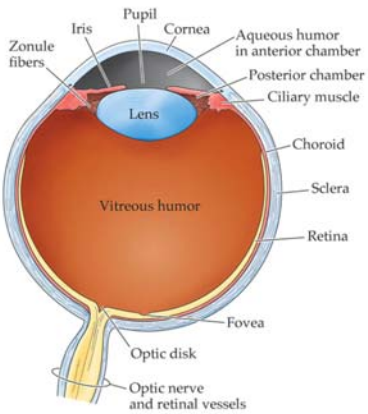
- The cell layer of the retina is about 300 μm thick.
- About 130 million receptor cells (rods and cones) are connected with ca. 1 million fibers of optic nerve.
- There are vertical and horizontal connecting structures.
- Drawing below: Rods, cones and nerve layers in the retina. The front (anterior) of the eye is on the left. Light (from the left) passes through several transparent nerve layers to reach the rods and cones (far right). A chemical change in the rods and cones send a signal back to the nerves. The signal goes first to the bipolar and horizontal cells (yellow layer), then to the amacrine cells and ganglion cells (purple layer), then to the optic nerve fibres.

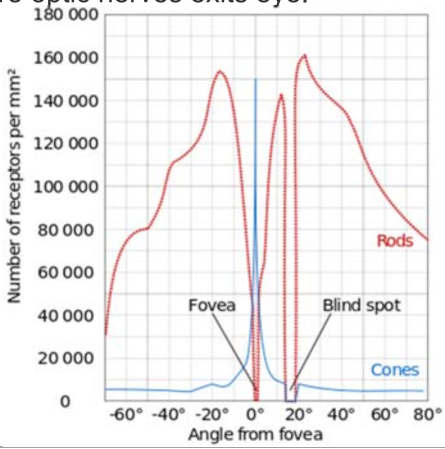


Wikipedia

Distribution of rods and cones

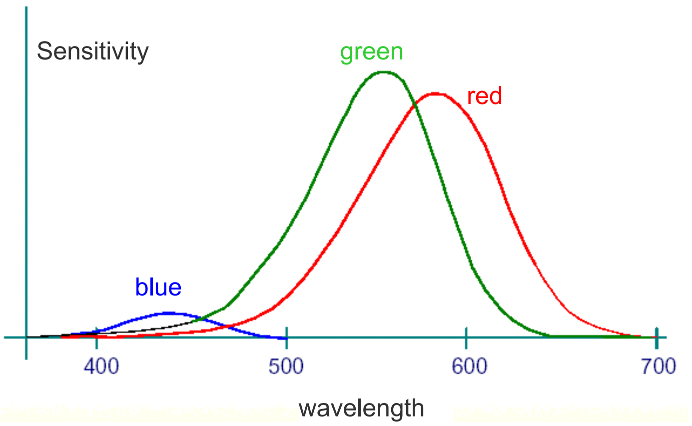
- Highest density of cones in fovea, the region of highest sensitivity and finest spatial resolution of human eye.
- Decay of color sensitive cone density with rising angle to fovea, black-and-white sensitive rods take over perception in periphery.
- The blind spot coincides with location where optic nerves exits eye.





Spectral sensitivity of cones

- Light in the wavelength range 400-700 nm is perceived as different color.
- Over a wide range of the visible spectrum humans can distinguish colors apart as small as 1-2 nm.
- Maximum of sun radiance in green, Rayleigh-Scattering causes bluish appearance of clear sky.



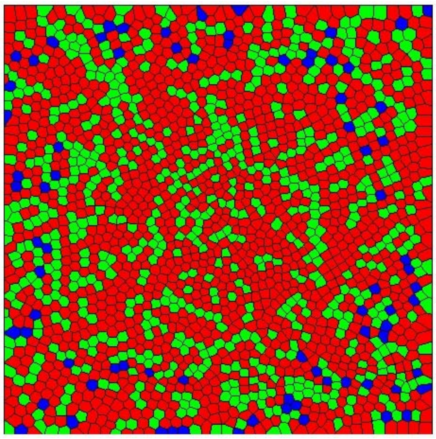
Distribution of rods in macula of retina

- In the center of macula (the fovea) are only cones for red and green light.
- Focus, as a function of wavelength, is optimized for green light.
- Less blue rods (about 10 %)

➔ Less sharp image for blue

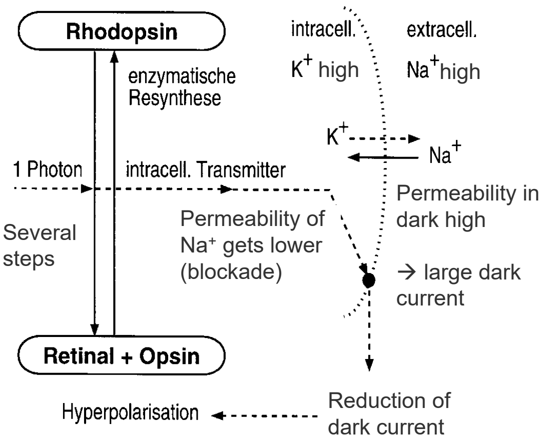
Simulation of rod mosaic in Fovea

Source http://www.allpsych.uni-giessen.de/karl/teach/farbe.html



Function of photoreceptor cells (primary transduction)

- Phototransduction is a process by which *light is converted into electrical signals* in the photoreceptor cells by chemical reaction of light-sensitive receptor protein:
  - Rhodopsin (visual purple) for rods
  - Three different pigments for color perception (cones)

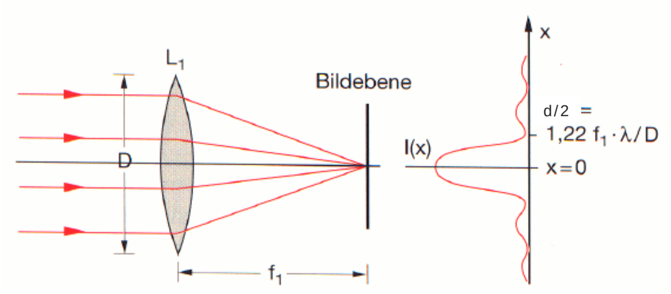


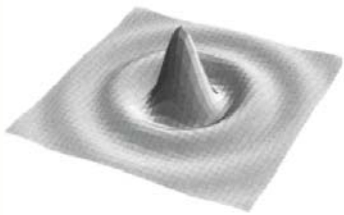
1. Absorption of a photon causes rhodopsin to fall apart into retinal und opsin.
2. The lower concentration of rhodopsin causes lower permeability of membrane for Na^+.
3. This results in a change of the potential (voltage) of the membrane (this is our signal).
4. In case of constant photon absorption rate, decomposition and re-synthesis of rhodopsin are in equilibrium.

Resolution limit of human eye (I)

- Diffraction limits spatial (angular) resolution
- A point-like object is mapped on retina as intensity pattern (Airy disk): *sin(x)/x*
- The diameter d of the main lobe is:

$$d \approx 2{,}44 \frac{\lambda}{D} \cdot f$$



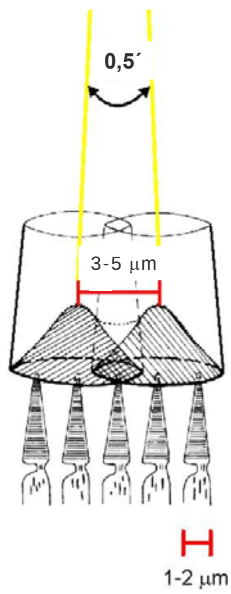


Eye:

Wavelength	λ = 500 nm	
Aperture	D = 0.8 cm	$d \approx 6\lambda \approx 3\mu m \approx 0{,}5'$
focal length	f = 2 cm	

Resolution limit of human eye (II)

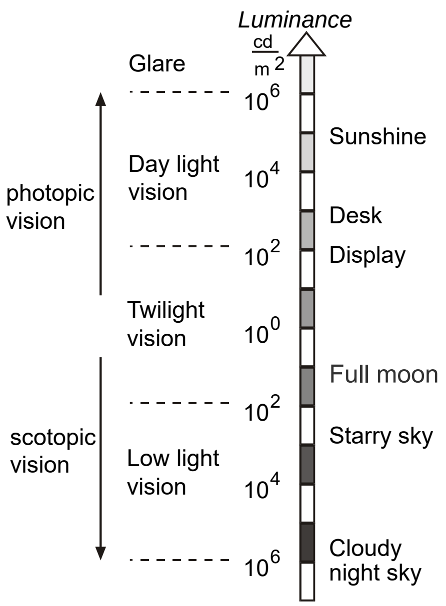
- The smallest cones are about 1 µm wide
- Highest density of cones in fovea (~160000/mm^2, hawk ~1 Mio/mm^2)
- Improvement of resolution by about one order of magnitude by:
  - Steering of head and eyes such that the image of object of interest falls on fovea in both eyes.
  - Fusion of both images (stereoscopy)
  - Small and quick motions of eyes often across edges
  - „Sub pixel accuracy“ by interpolation



Contents

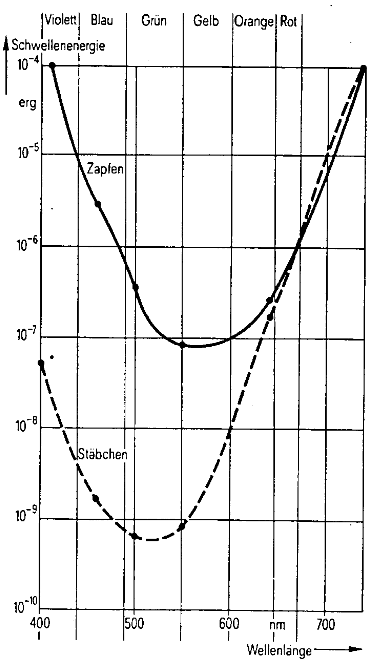
Luminance range of human eye

- The human eye is capable to work over an enormous range of luminance spanning 12 orders of magnitude.
- This large span is possible by use of two different populations of receptor cells in retina:
  - Rods are used for vision at low light conditions (scotopic vision).
  - The cones enable vision at day light and provide color perception (photopic vision).



Sensitivity limit

- Limits of intensity sensitivity of human eye and monochromatic light.
- The lower curve is for scotopic vision (rods, low light), the upper for photopic vision (cones).
- The absolute limit of rods in case of full adaption is at 507 nm.



Adaption of eye to brightness

- Pupil diameter
  - By decrease of diameter the amount of incoming radiance can be reduced by a factor of 16
- Change of concentrations of light-sensitive receptor proteins
- Two different receptor populations
- Variation of the size of **receptive fields**

Retinal adaption (receptor protein)

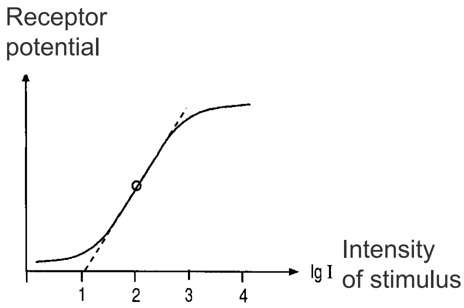
Change of concentrations of receptor proteins:

- Broad daylight
  - Preferred reaction – decomposition of Rhodopsin
    → Concentration of Rhodopsin decreases
  - Due to the lower concentration, the probability of further decomposition drops.

  → **Decrease of sensitivity**

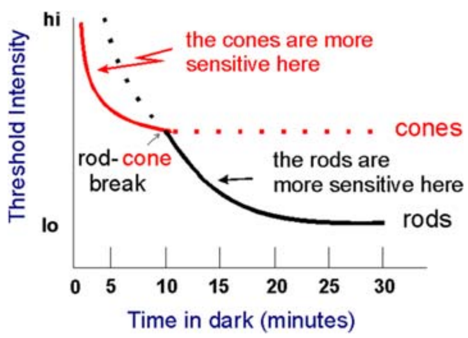
- Low light
  - Preferred reaction – re-synthesis.
    → Concentration of Rhodopsin rises.
  - Due to the higher concentration, the probability of further decomposition increases.

  → **Rising sensitivity**



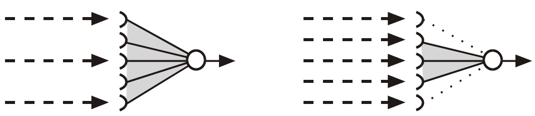
The position of the equilibrium between decomposition and synthesis of rhodopsin depends on luminance on retina.

Retinal adaption (Receptor type, receptive field size)



- Switch from cone to rod system

- Another mechanism for adaption is change of the size of receptive fields in the retina by lateral inhibition.
  The spatial summation is larger at low light and smaller in daylight condition.



dark bright

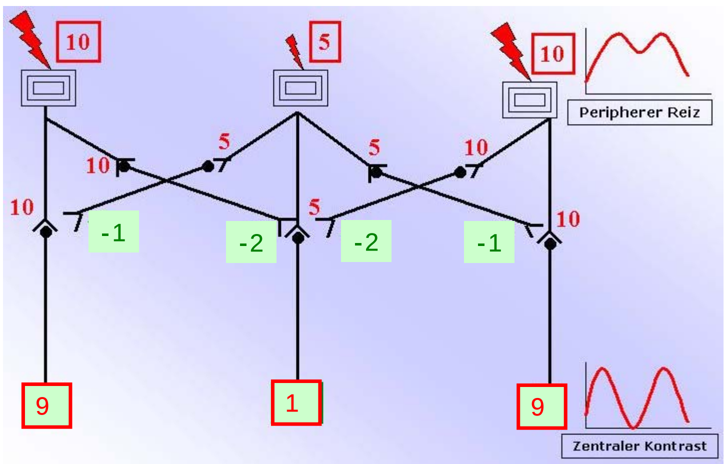
Contents

- Image acquisition
  - Eye
  - Retina
  - Photoreceptor cells
- Adaption of eye on brightness
- Signal processing in retina
- Further processing
- “Gestalttheorie”
- Optical illusions

Local signal processing in retina

Lateral inhibition at edges:

- Neighbored cells are connected by so-called horizontal cells.
- Large stimulus of a certain cell leads to inhibition of neighbor cells
- Aim: contrast enhancement to highlight object contours

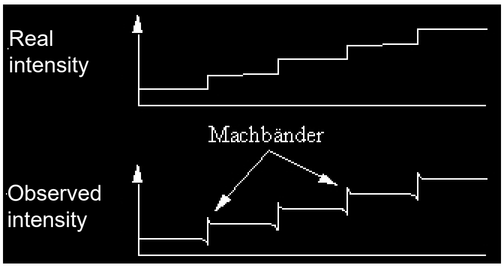


Principle of lateral Inhibition:

The stimulus is weighted (here: 0.2) inverted, and then added to signal of neighbor cell.

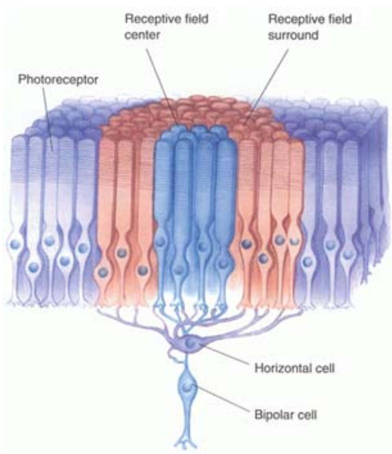
Lateral inhibition: Mach Bands

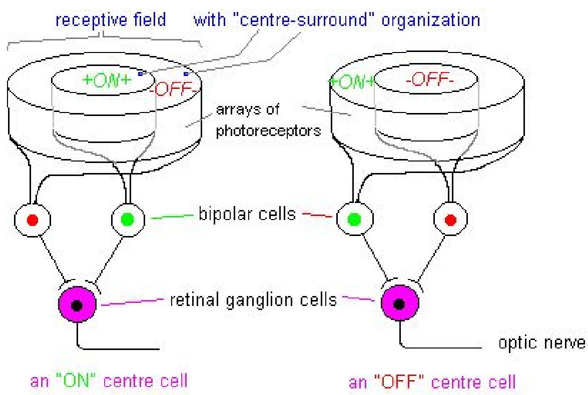
A sequence of bands of different but homogenous grey value



Seemingly wedge-shaped brightness gradient caused by lateral inhibition

2D signal processing in the Retina (I)

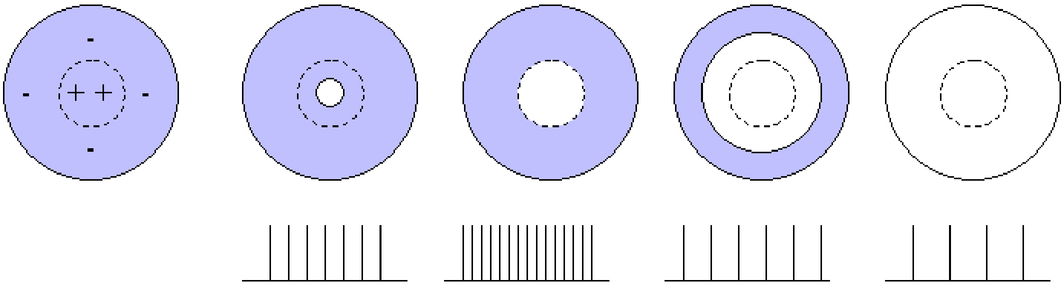




The receptor cells in the retina are pooled to so-called *receptive fields* (RF).

- Size of RFs rise with distance from fovea.
- We distinguish stimuli of the RF's center from stimuli of the RF's surround.
- 2 Types:
  - On-cell: In case of exposure to light, the center "fires" and the surround inhibits.
  - Off-cell: In case of exposure to light, the center inhibits and the surround fires.

2D signal processing in the Retina (II)

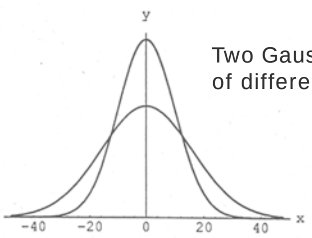


Answer of an on-cell (exposure to light excites center and inhibits surround) to different stimuli (bright tone)

- Simplified: as soon as a stimulus exceeds a threshold (action potential), the cell “fires” (emits a single impulse in the sketch).
- After a certain minimum time required to recover, the cell may shoot again if there is still a large enough stimulus.
- **The strength of the stimulus is transformed into impulse frequency.**
- Excitation and inhibition lead to rising or decreasing frequency, respectively.

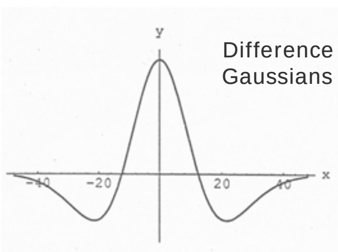
Hierarchical processing

- Such processes are repeated in different levels of processing.
- In image processing similar hierarchical methods have been developed.
- For example, *Difference of Gaussian* Operator (DoG) for edge detection.



Two Gaussians of different σ

Original image



Difference of Gaussians

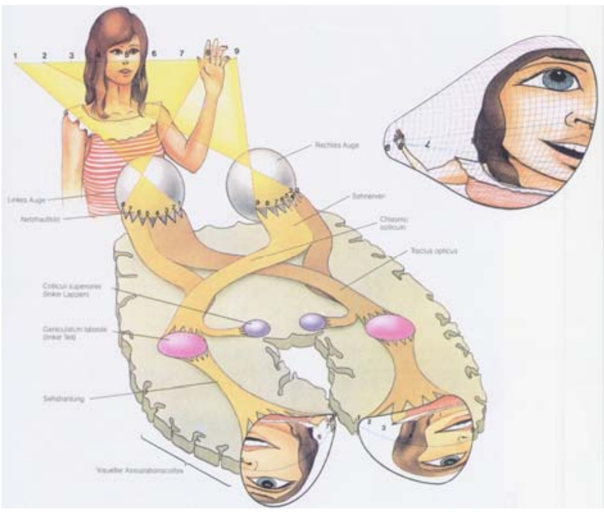
Result of filtering

Contents

- Image acquisition
  - Eye
  - Retina
  - Photoreceptor cells
- Adaption of eye on brightness
- Signal processing in retina
- Further processing
- “Gestalttheorie”
- Optical illusions

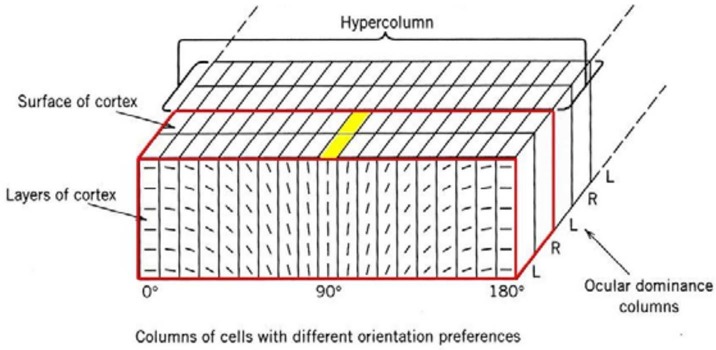
Schema of optic track

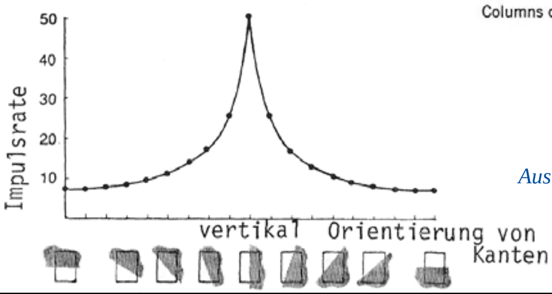
- The optic tracks are crossing:
- The left half of the visual field is processed in the right half of brain and vice versa → prerequisite for stereoscopy
- Focus of processing on center parts



Recognition of edge orientation (I)

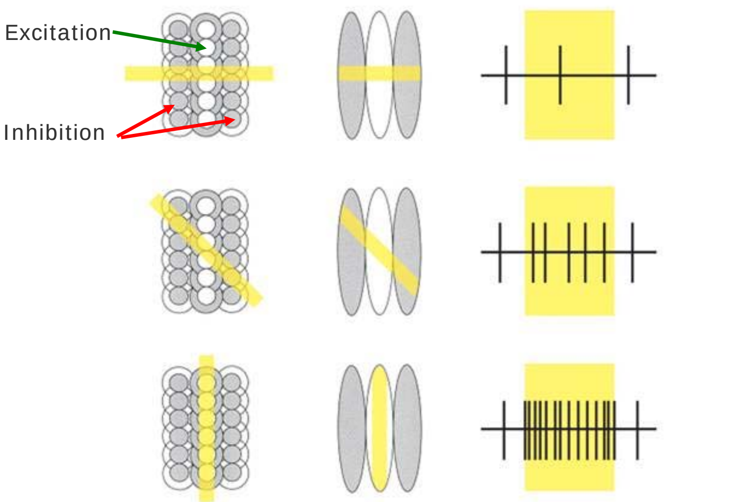
- In visual cortex RF are pooled into so-called hypercolumns such to be sensitive to orientation of edges.





Aus: (Korn, 1981: BV durch das vis. System)

Recognition of edge orientation (II)



Pooling of simple RF to columns

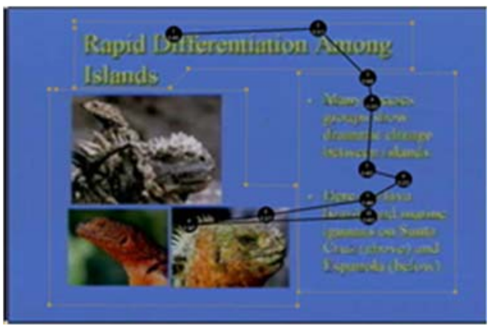
Signal frequency

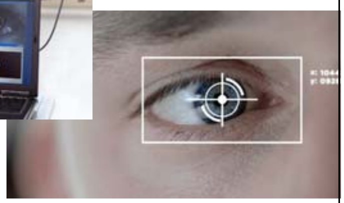
From: http://www.allpsych.uni-giessen.de/karl/teach/aka.htm

Human visual perception is rotation variant

Sequential perception of environment

- In a distance of one meter the size of the area of sharpest imaging (macula) cis about as large as a stamp.
- We permanently scan our environment by very quick motion of both eyes, so-called saccades; in this manner the “image” is collected step by step.
- During saccade the optical flow is so enormous that we are practically blind, however, we not aware of it.





Recording of eye motion (Eyetracker)

Recognition of motion of objects

Recognition of motion of objects is of great importance (motion ≡ potential hazard):

- Complex process involving several parts of visual system
- Specialized neurons detect certain direction of motion
  - Linear motion
  - Rotation or spiral motion
- Analysis of changes of observed angles (object gets e.g. larger)

Contents

- Image acquisition
  - Eye
  - Retina
  - Photoreceptor cells
- Adaption of eye on brightness
- Signal processing in retina
- Further processing
- “Gestalttheorie”
- Optical illusions

Gestalttheorie (or Gestaltpsychology, part of Cognitive psychology)

Principles of grouping
- Similarity
- Good continuation
- Proximity

• "Prägnanz"
- Simplicity
- Symmetry
- Wholeness

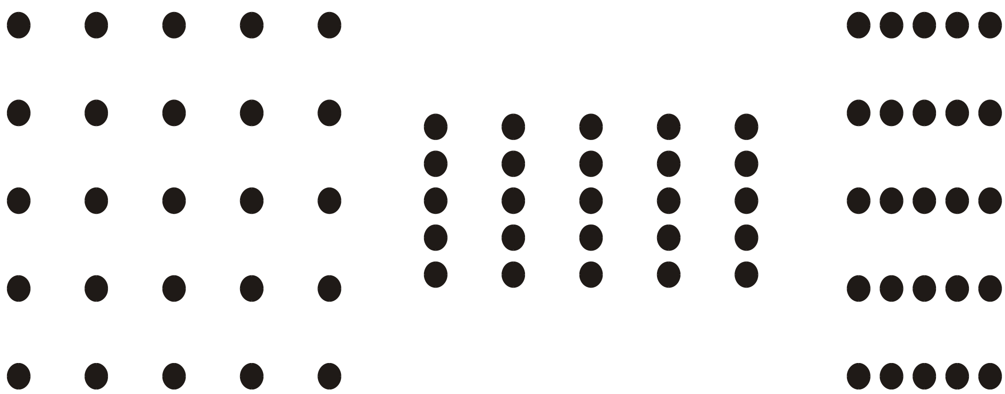
• Figure-ground separation
• Size constancy
• Whole and parts

• Optical illusions
- Brightness
- Geometry
- Contour
- Motion

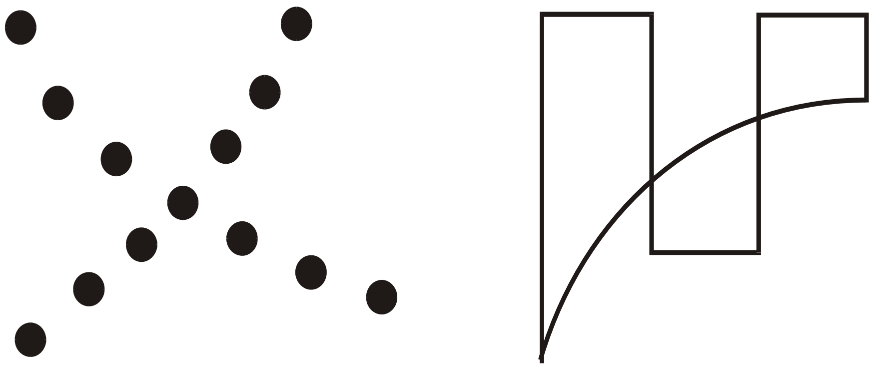
• Max Wertheimer (1880-1943)
• Wolfgang Köhler (1887-1964)
• Kurt Koffka (1886-1941)

Nice website:
http://www.michaelbach.de/ot/

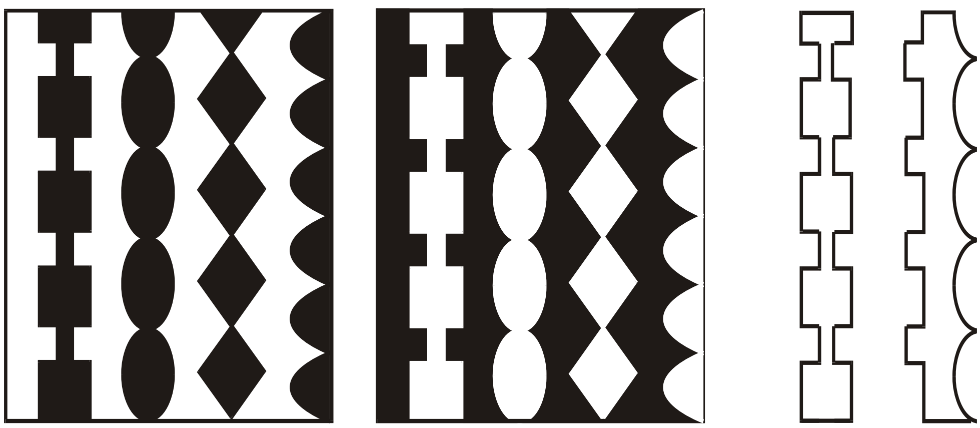
Grouping: Proximity



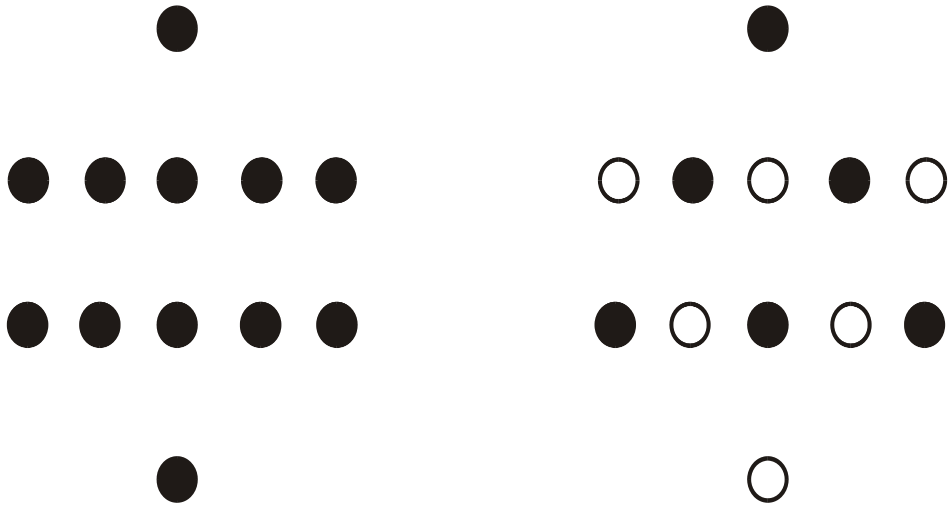
Grouping: Good continuation



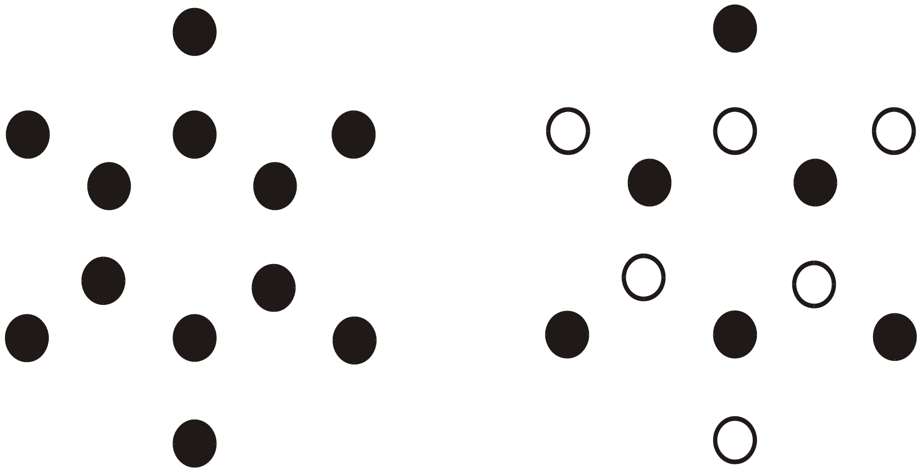
Grouping: Symmetry



Grouping: Proximity and similarity (I)



Grouping: Proximity and similarity (II)



Grouping: Symmetry versus good continuation

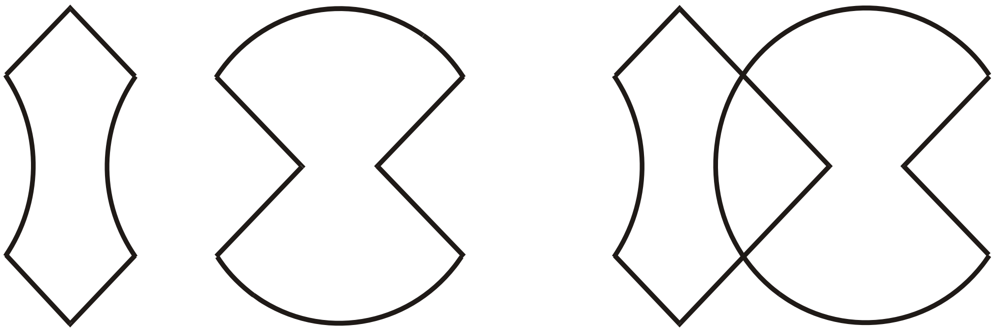


Figure-ground separation (I)

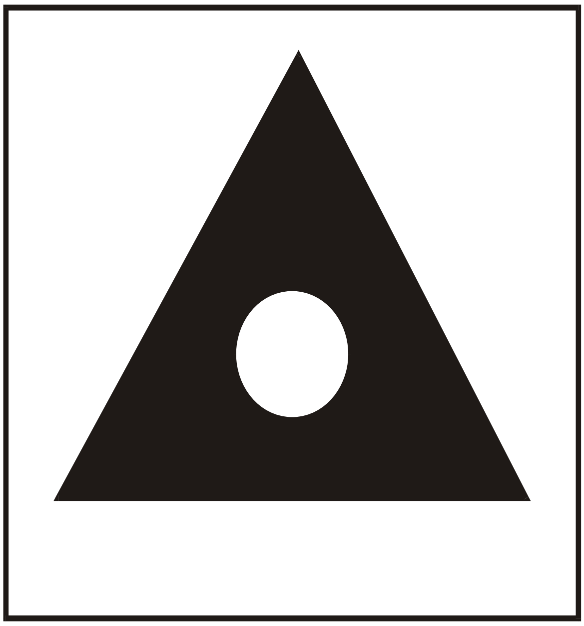
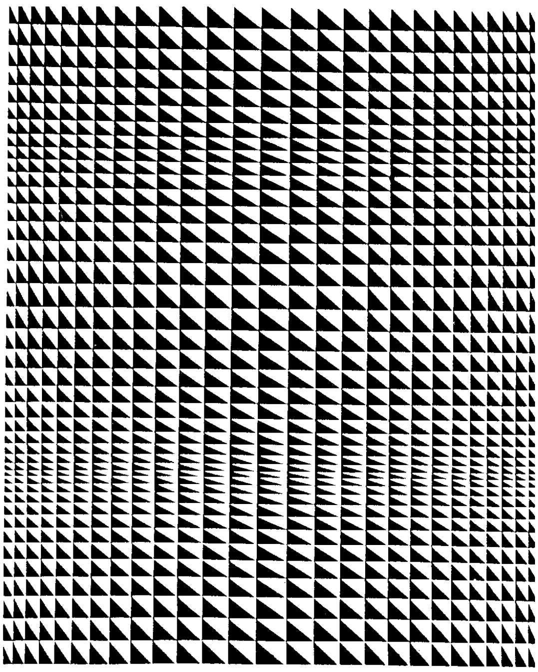


Figure-ground separation: Face profiles or vase

Figure-ground separation: Young or old lady

Size constancy

Whole and parts: „straight“ curves (painting by Riley)

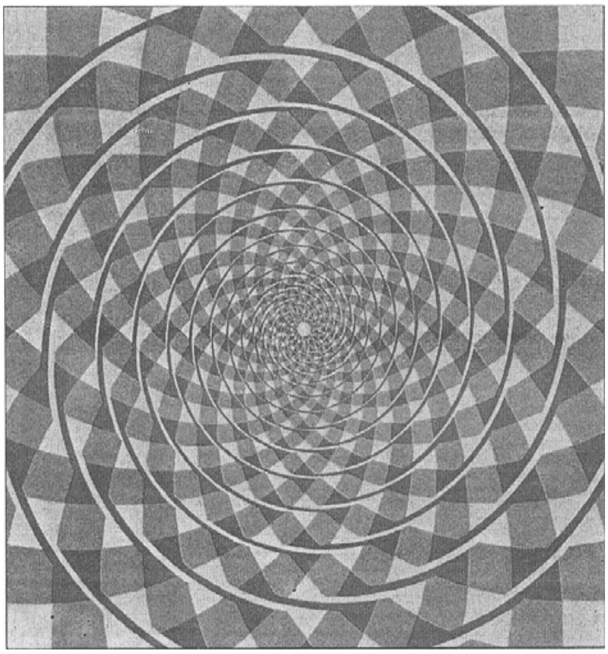


The whole influences perception of parts:

• The wave-alike appearance of the image forces triangle edges to look rounded.

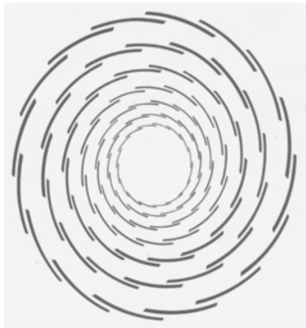
Kurt Koffka: "The whole is *other* than the sum of the parts"

Whole and parts: Fraser's Spiral

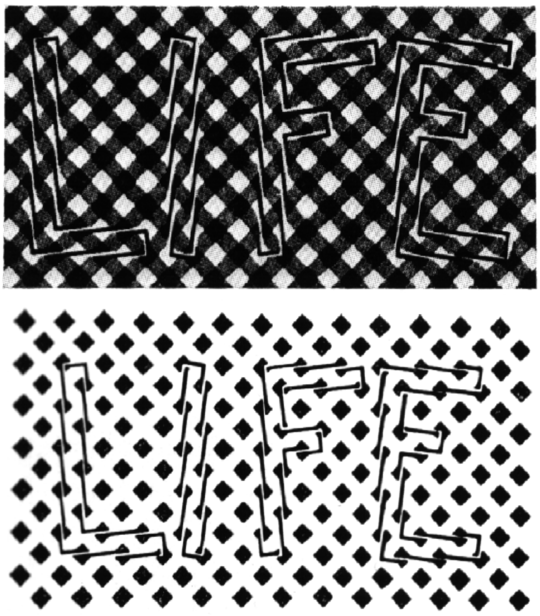


The parts influence perception of the whole:

Brightness gradient on circles causes illusion of a spiral.



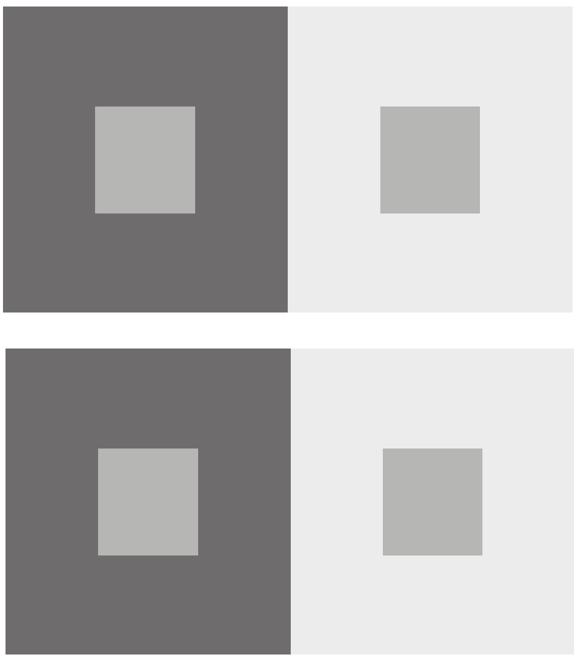
Whole and parts: Life



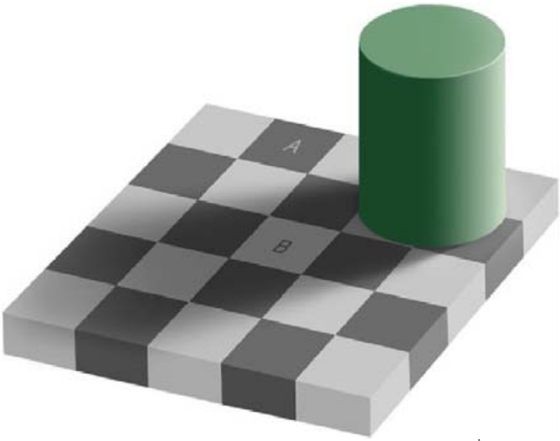
Contents

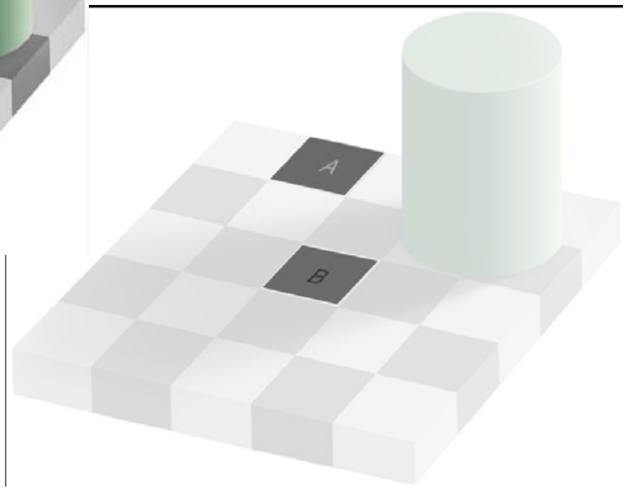
- Image acquisition
  - Eye
  - Retina
  - Photoreceptor cells
- Adaption of eye on brightness
- Signal processing in retina
- Further processing
- “Gestalttheorie”
- Optical illusions

Brightness illusion (I)

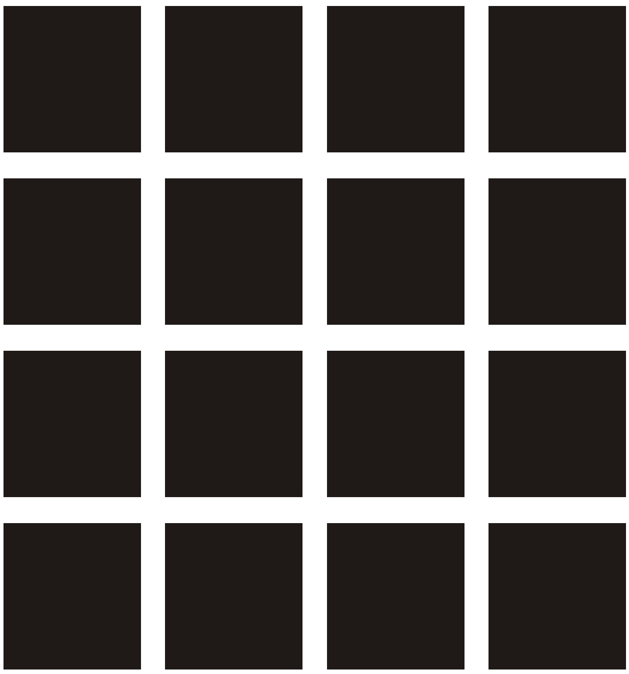


Brightness illusion (II)

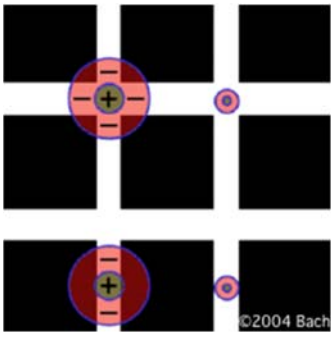




Brightness illusion: Hermann‘s Grid



Hermann's Grid: Cause? Standard answer:



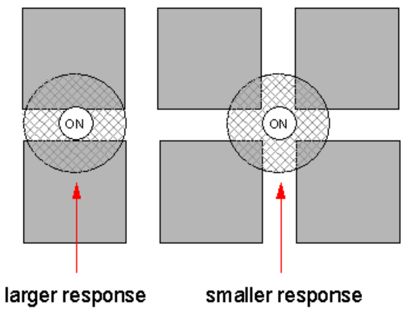
1. Why do we see the dark patches?
Look at the left part of the left diagram and assume an on-center retinal ganglion cell. Its receptive field is indicated by the reddish disk. When the ganglion cell is, by chance, looking at the grating so that its centre ('+') is positioned at a crossing (left-top), there are 4 bright patches in the inhibitory surround. A ganglion cell looking at a street (left-bottom) however only gets 2 inhibitory patches, so it will have a higher spike rate then the one at the crossings.

2. Why don't we see the patches when we look right at them?
Because then we direct the fovea at the crossings, and in the fovea the receptive fields are much smaller (see the small reddish disks on the right of the left figure). With such small receptive fields it obviously does not matter whether they are at the crossings or not.

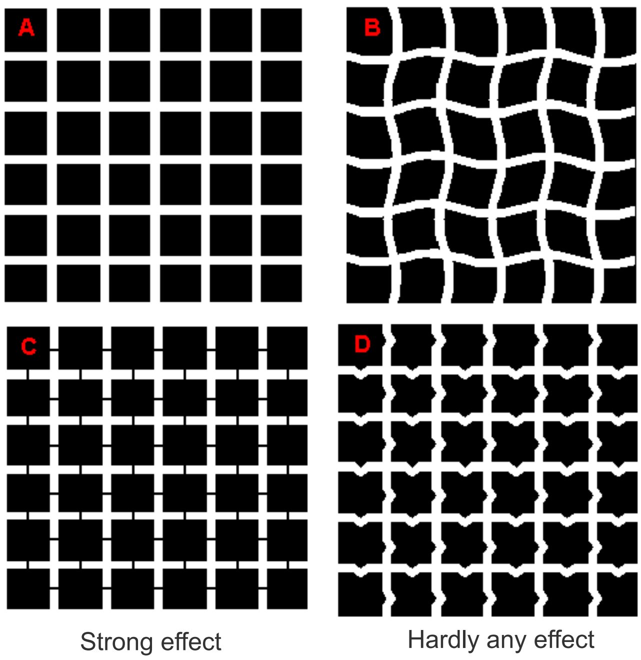
http://web.mit.edu/bcs/schillerlab/research/A-Vision/A15-7.htm

Hermann's Grid: Cause?

Standard answer:



Does not tell whole story:



http://web.mit.edu/bcs/schillerlab/research/A-Vision/A15-7.htm

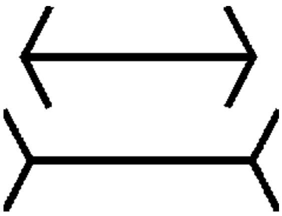
Geometric illusions (I)

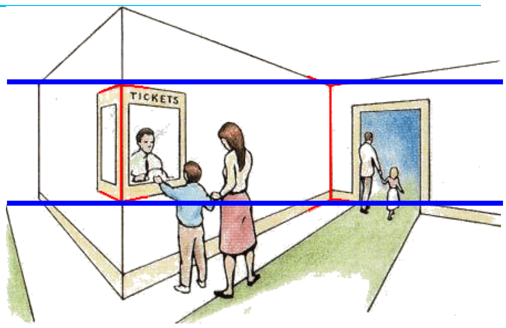
Vertical-horizontal illusion

Ponzo

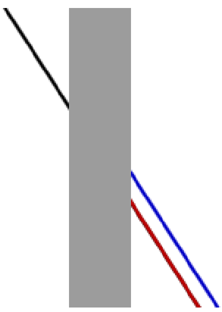
Geometric illusions (II)

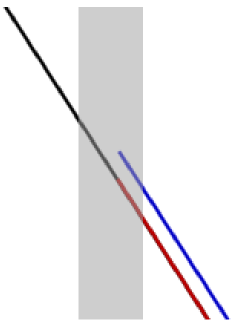
Müller-Lyer



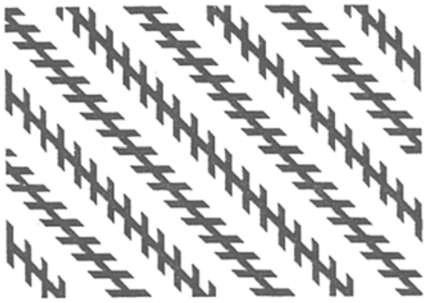


Poggendorff

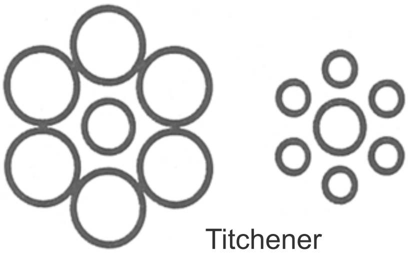




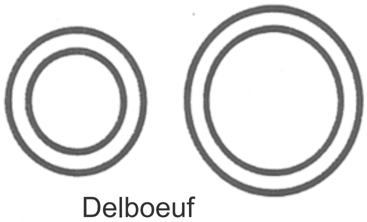
Geometric illusions (III)



Zöllner

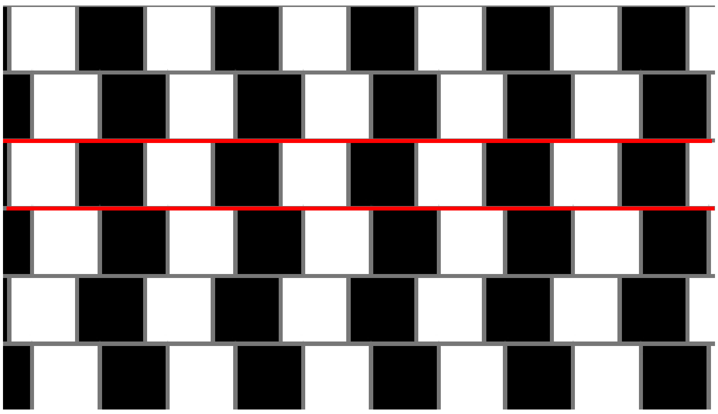


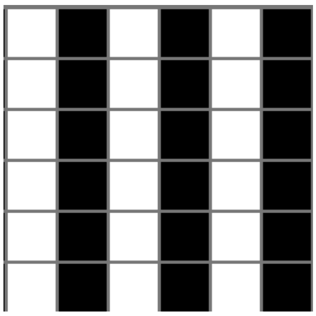
Titchener

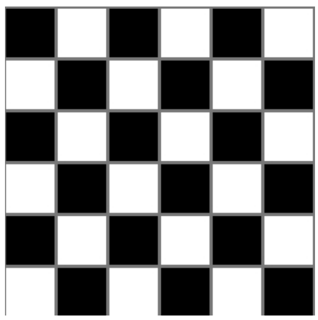


Delboeuf

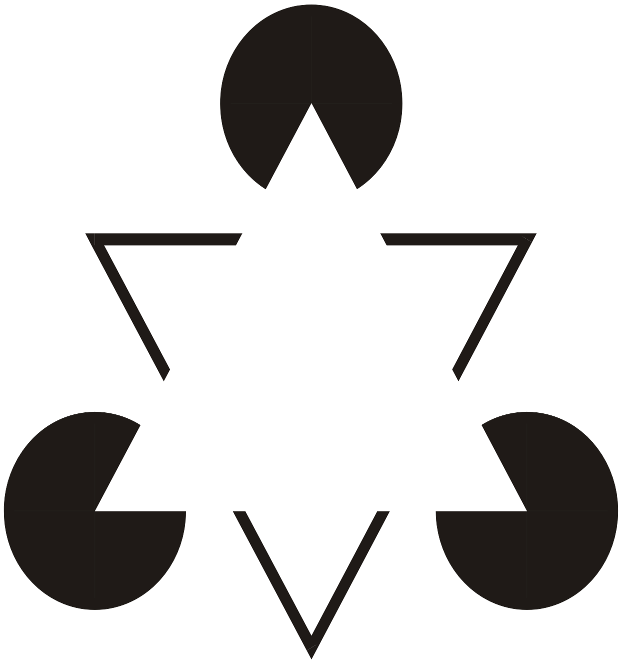
Café-Wall illusion



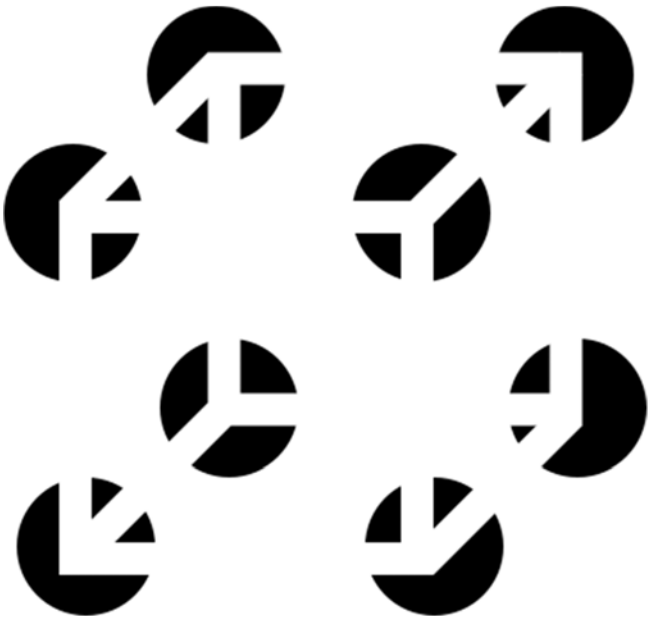




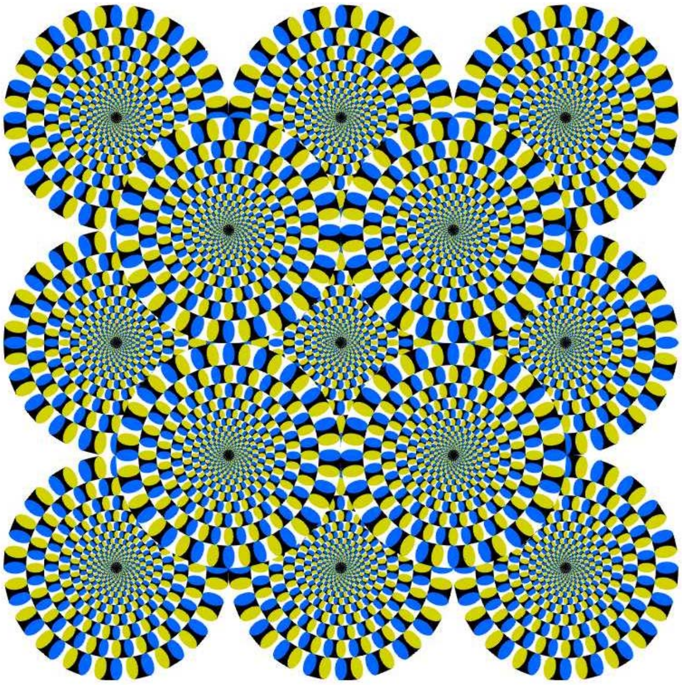
Illusory contours: Kanizsa



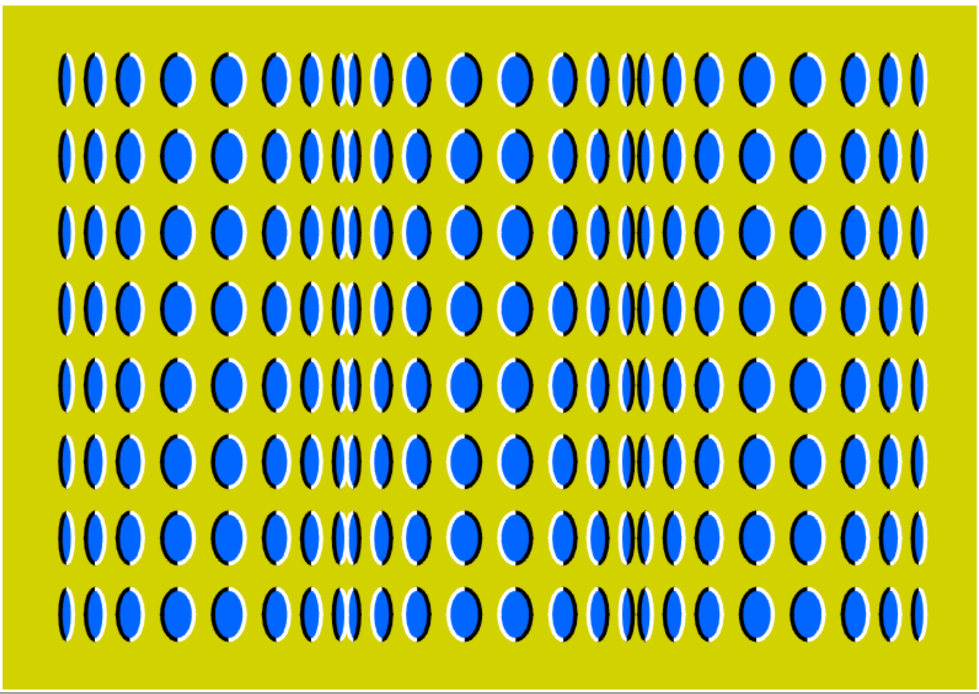
Illusory contours: „Hallucinating“ a cube



Motion illusion (I)



Motion illusion (II)



Challenge of automated object recognition

What is figure (object), what is background?