S

Miroslav Korbelář

Masaryk university

Neurogeometry seminar
Brno 2012

Notification

The aim of this presentation is to show that the visual perception is a very complex and complicated process in contrast to how it is usually treated (*'What should be so complicated about it? I just open my eyes and see, don't I?'*).

This complexity is demonstrated on optical illusions, but due to an enormous amount of different phenomena only a small part of them is mentioned. The main source is the book by Francis Crick: *Astonishing hypothesis - A scientific search for soul*.

In this book the author tries, by studying visual perception, to support his hypothesis that "You" are the behavior of a vast assembly of nerve cells and their associated molecules.

Introduction

The visual perception does not only mean a kind of an 'automatical image representation' (within the brain) but a complex creating process, which involves not only realizing what we see:

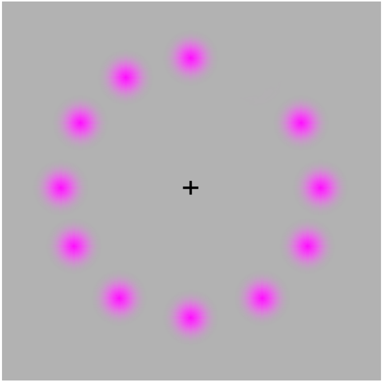
- for instance: a moving cat is recognized as one object (i.e. differently moving parts of it belong in fact together)

but also an ability to create something that does not exist:

- the optical illusions.

Such illusions might be a good guide for testing what does exactly the visual perception mean for the brain. They also show that a naive idea of what seeing is - i.e. transferring and simple combining images from both eyes within the brain in a way: camera $\rightarrow$ computer (or screen) - is too simplified or wrong.

An interesting example



- If you follow the rotating pink dot, you will only see one color - pink.
- If you watch + in the center, the moving dot turns to green.
- If you concentrate on +, after a short time all the pink dots will slowly disappear and you will only see a green dot rotating.

Causes

Optical illusions are caused partially by the eyes themselves:

- the blind spot (a feature of vertebrates, e.g. octopuses do not have it)
- projecting the 3D world onto the curved 2D retina
- different reactions of the stamens and cones at different wave lengths of light
- using one eye only instead of the binocular vision

and partially by processing (so far not understood) the signals, coming from the eyes, by the brain.

Remark: Evidence for non-trivial processing is the fact that even if the light in the eye goes through a non-homogeneous medium (capillaries, vitreous humor) the resulting image is very clear and without distortion. The brain hence has to be able to eliminate disturbing effects (including eye movements) to create a "right" representation of the outside world. And sometimes something is added ...

Basic characteristics of sight

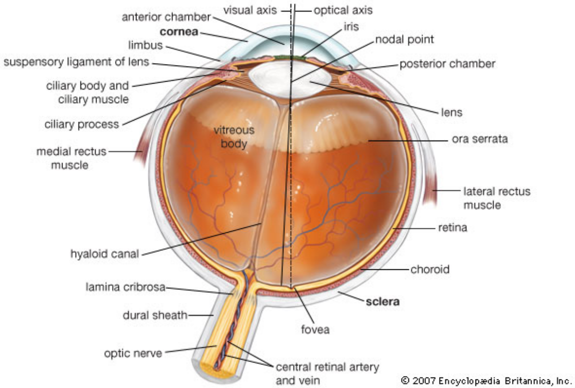
Sight does not only serve for mapping our surroundings but mainly for detecting what is going on around us (necessary for survival) $\Rightarrow$ The brain does more than mediate the image - it tries and needs to make an immediate interpretation of it. Such a feature is always presented no matter what we look at.

Despite this incredible evolution ability it turns out that

1. The optical system deceives you easily.
2. The visual information provided by eyes needs not to be uniquely interpreted.
3. Seeing is a constructive process.

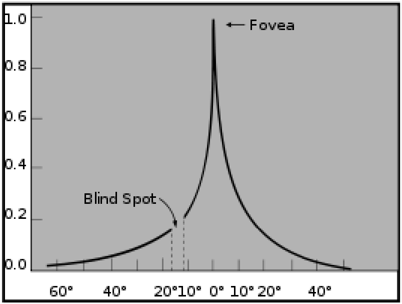
Important: What we see needs not to be real, but it is something the brain believes to be real.

- We don't see everything clearly, only that which is in our visual axis (i.e. in the yellow spot (fovea) on the retina). The opposite impression is caused by movement of the eyes.



- Minimal angle of view is in the yellow spot ($\sim$ 1 arc minute). The distance of images on the retina is then $\sim 5\mu m$ (between two stimulated cells is one not stimulated).

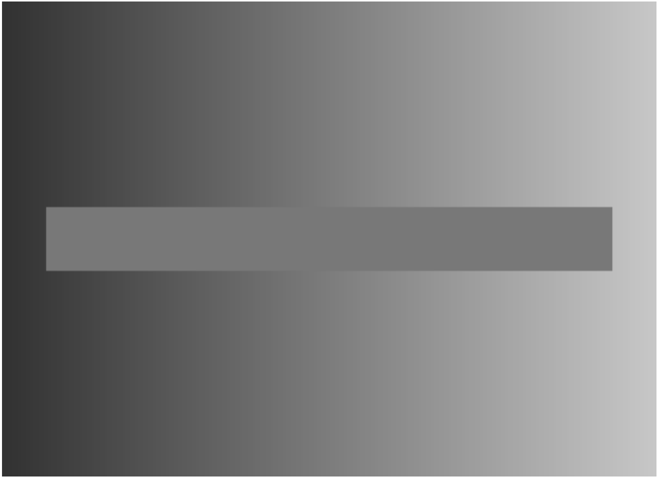
1. The optical system deceives you easily: Illusion of seeing everything clearly.



The diagram shows the relative acuity (%) of the left human eye (horizontal section) in degrees from the fovea.

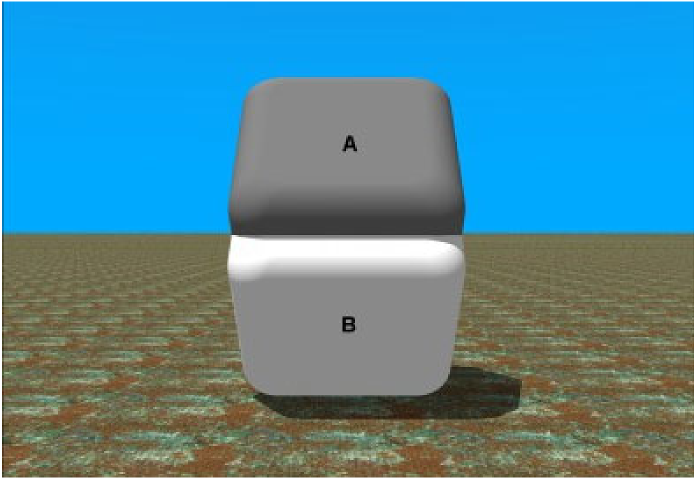
- Visual acuity decreases rapidly by moving away from the visual axis. Tiny letters can be read in the visual axis, but not when you try to read them focusing on a finger a bit further away.
- At the visual periphery we recognize only motion with certainty but not shape. A bit closer we then see shape before we determine the color.

1. The optical system deceives you easily: Different shades illusion



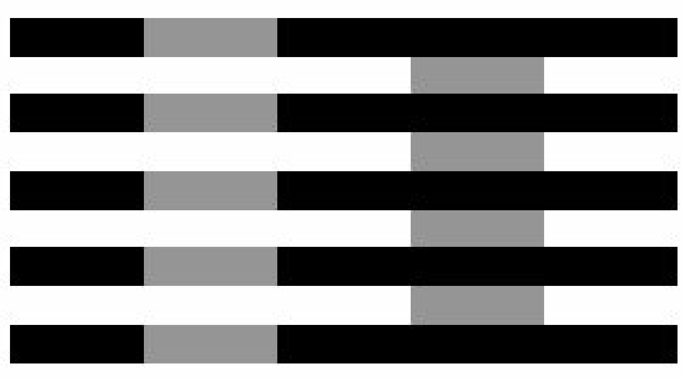
Does it seem that the middle strip changes brightness? In fact it is homogeneous - cover the other ones!

1. The optical system deceives you easily: Different shades illusion



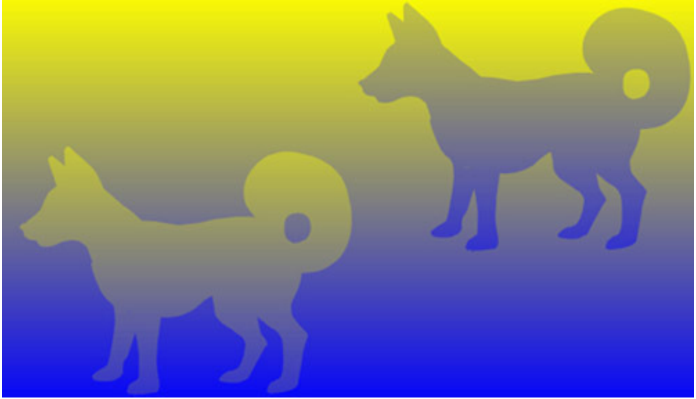
Although it doesn't look like it, the buttons in this picture are the exact same shade. Cover up the middle part!

1. The optical system deceives you easily: Different shades illusion



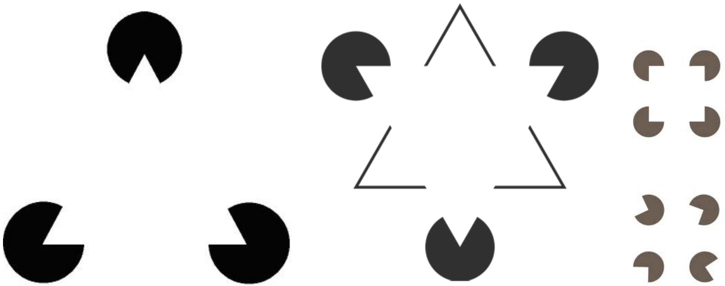
Every gray rectangle in this image is the same shade. Use some editing software or print and cut them!

1. The optical system deceives you easily: Different colors illusion



A blue dog and a yellow dog? Both dogs are the same color. Proof: same as before.

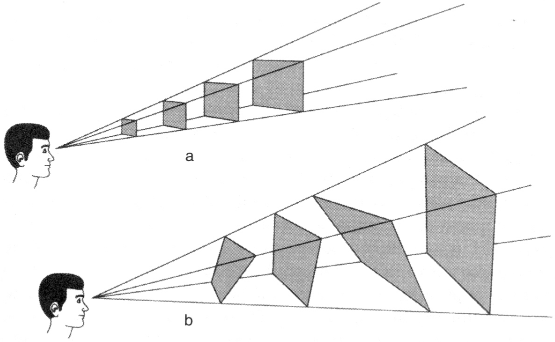
1. The optical system deceives you easily: Illusory contours (Kanisza triangle)



Do you see white triangle(s) (square, resp.)?

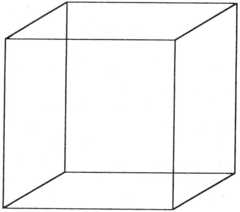
- Illusory contours (or subjective contours) are perceived without a luminance or color change across the contour. (If you cover a part of the picture - the black circles - you won't see the contours.)
- Moreover, the inner object seems to be brighter then the rest. Changing position of objects the illusion disappears.

- To have a unique interpretation of what we see, we need some constraining conditions that choose out of more possibilities the only one (or the most likely one). Sometimes such information is missing.
- When there are more plausible interpretation the brain usually can not decide which of them to choose and constantly switches between the possibilities, but never makes a mixture of them.



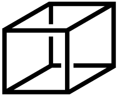
- Looking with one eye only without moving the head (or looking at a photo) we can not determine the distance (**a**) or the real shape (**b**) of objects with the same projection into the eye. (Imagine that there is a white background behind, for instance.).
- Even in this situation we are still able to feel the 3D depth.

3D anamorphic street paintings: illusionary 2-dimensional images that appear 3-dimensional when viewed from a fixed point through a camera lens.

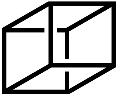


Watch the cube for a while. Does the interpretation change?

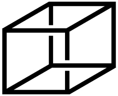
- The Necker cube is sometimes used to test computer models of the human visual system to see whether they can arrive at consistent interpretations of the image the same way humans do.



Most common.



Less common.



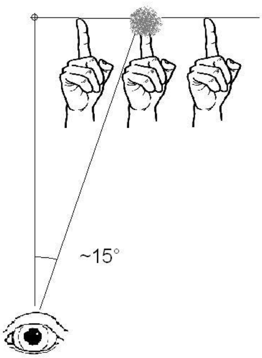
Inconsistent.

Remark: A similar inconsistent object is the Penrose triangle (impossible in ordinary Euclidean space, but existing in certain Euclidean 3-manifolds).

Penrose triangle sculpture in East Perth, Western Australia

3. Seeing is a constructive process: Filling the blind spot.

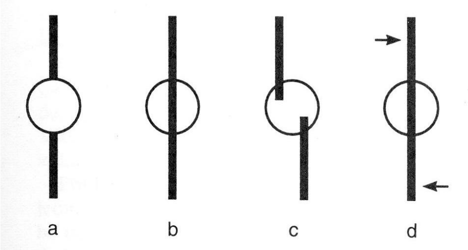
- Looking with right eye only, move the finger horizontally to the right starting from the visual axis. At one point you won't see the fingertip.



- Looking at a monochromatic surface (a wall) we don't see any hole in it (that should correspond to the blind spot) ⇒ Brain fills the place in by the approximation it thinks is the "best". In fact, if there will be any object in that spot we will not see it.

A similar (but slower) process of 'approximating' was observed (Ramachandran, Gregory) in people who have either damaged a spot of the retina or the corresponding part of the visual cortex.

- Picture **a** shows what was presented and **b** shows what was seen in the end. The circle corresponds to the missing part of the visual field.



- More surprising was showing the picture **c** of two parallel lines which were seen as moving to one another ending as one single line - picture **d**.

Hierarchy of the visual perception

- The representation of an image in the brain is not as simple as a pixel matrix only (e.g. as in the TV), otherwise we would not be able to recognize a particular face for instance (what TV doesn't), but a symbolic representation of higher degree. (Currently there are programs able to recognize a general human face.)
- On the one hand the brain decomposes the picture into parts (e.g. there are mountains in the background, closer there are trees and houses and people and cars are moving in the front). Similar representation is also used for computer game simulations, for instance.
- On the other hand the recognized parts of objects are grouped into bigger units (e.g. legs + hands + head + ... = a human etc.)
- The way how we recognize objects (which even need not be uniquely demarcated for us) is not a process given to us from the outside, but has to be learned.

Remark: The recognition can be complicated by a partial hiding of the objects or by a confusing background. The brain has to do all this work in real time.

Decomposing and grouping

The picture is divided into four parts, but we don't see four different parts of a face (that eventually do not belong to each other), but one face in the back and a window in the front.

Non-uniquely demarcated objects

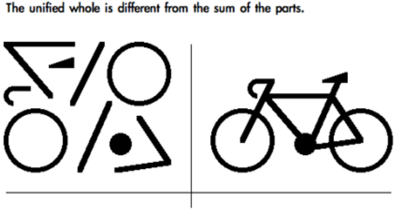
OMG!!! A TWO-HEADED ZEBRA!!!...... Or not?

(A real photo, no Photoshop!)

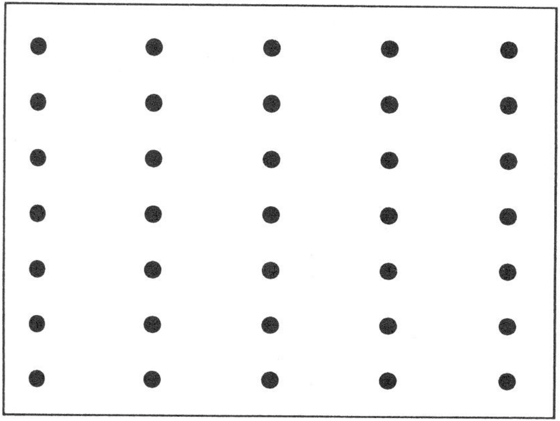
Principles of grouping

- The brain groups parts of objects together based on their interactions.
- This was the ground of the *Gestalt psychology* (originated 1912 in Germany). *Gestalt* (= form, appearance) here means an organized entirety, where every part influences the rest and hence *the entirety is other than a simple sum of its parts* (Kurt Koffka).
- Interactions between parts are classified according to Gestalt laws into:

(1) proximity
(2) similarity
(3) good continuation
(4) closure
(5) common fate
(6) symmetry
(7) Prägnanz (pithiness)
(8) figure and ground
(9) past experience
(10) uniform connectedness (Law of Unity)
(11) common region
(12) parallelism



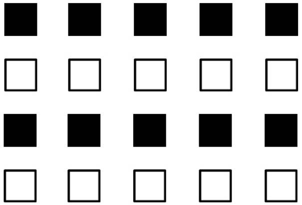
(1) proximity

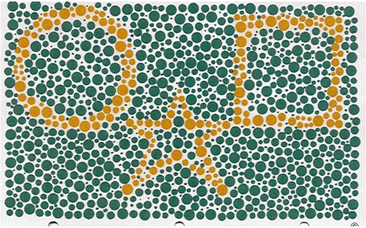


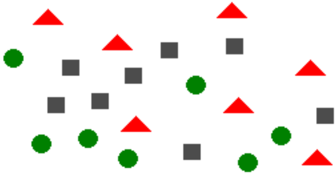
Do you see rather vertical or horizontal lines?

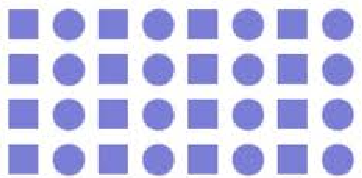
We group rather objects that are closer than further (similar) things. This proximity is in fact given by closeness on the retina (e.g. star constellations).

(2) similarity





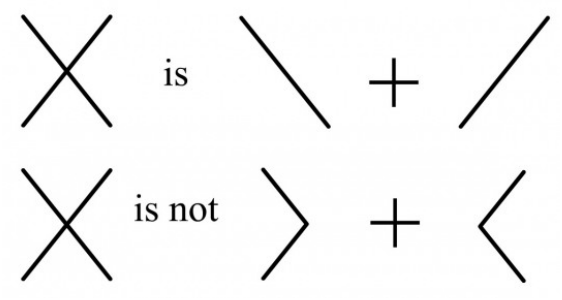




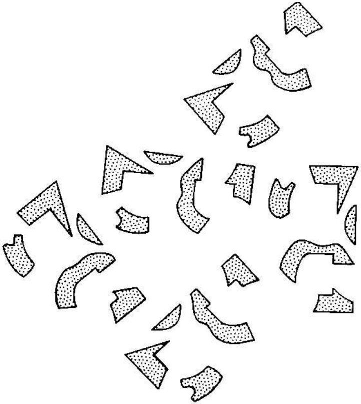
How would you describe what you see on each of the pictures above?

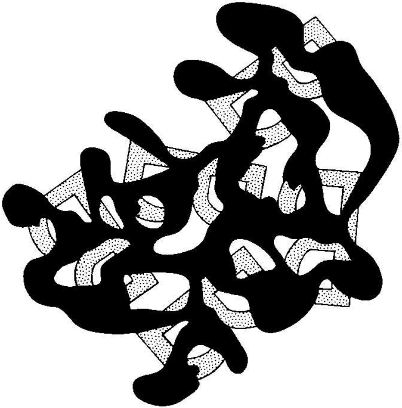
- We tend to group objects with an obvious visual property in common such as shape, color, shading or other qualities.

(3) good continuation



- When there is an intersection between two or more objects, we tend to perceive each object as a single uninterrupted object.
- We prefer the lines or curves that follow an established direction over those defined by sharp and abrupt changes in direction.





Can you recognize something?

(Many independent objects...)

What about now?

(...divide into 5 equal group of objects.)

(4) closure

Fig.1

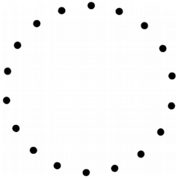


Fig. 2

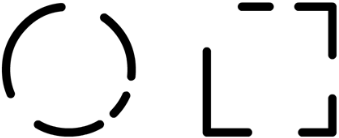
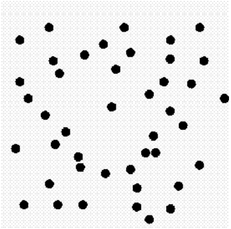
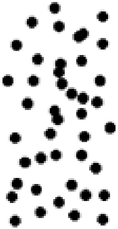


Fig. 3

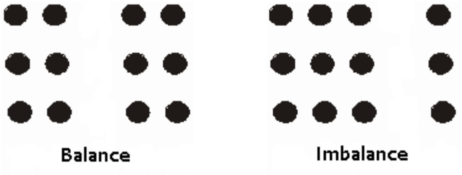
- We tend to ignore gaps and complete contour lines.
- We perceive one object (Fig. 1), even if consisting only of distant dots (Fig. 2). In Fig. 3 we perceive two groups of lines.

(5) common fate

- We group together objects which have the same trend of motion.
- When some of the dots move (the 3rd picture) you will recognize a familiar figure in motion even when there is no apparent contour (the evolutionary need to distinguish a camouflaged predator from its background). When motion ends the figure ceases to be visible.

(6) symmetry



- It is more pleasing when objects can be divided into an even number of symmetrical parts. These are perceived collectively, in spite of distance (e.g. we see 3 pairs of symmetrical brackets rather than 6 individual brackets).
- Impression: A visual object will appear as incomplete if it is not balanced or symmetrical.

(7) Prägnanz

Fig. A

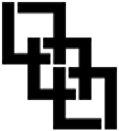


Fig. B

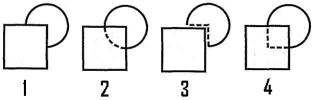
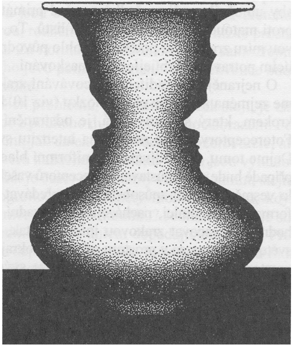
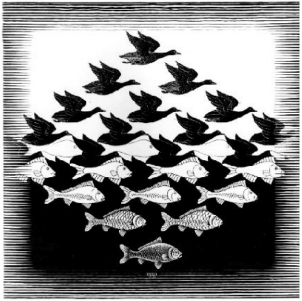


Fig. C

- We attempt to organize the visual information into the simplest form possible (in the spirit of Occam's Razor) that is symmetrical, simple and with the fewest number of shapes.
- Fig. A: The picture is perceived as a series of circles rather than as of many more complicated shapes.
- Fig. B: We see 3 overlying squares, although the picture in fact consists of 6 independent capital **L**'s.
- Fig. C: An ambiguous pattern (1) is recognized as the fewest simple shapes possible (circle and rectangle overlapping in 2) and not complex shapes (3,4).

Sometimes recognizing what is the figure and what is the ground can be difficult...





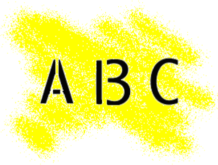
M.C. Escher: *Sky and Water I* 1938 woodcut

...sometimes ambiguous...

...or even inconsistent...

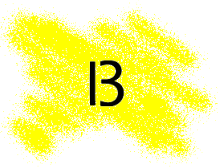
(1 vase or 2 face profiles? Saxophonist or woman's face? What is background in the middle part?)

...but we never see two interpretations at once.



Seeing first this..

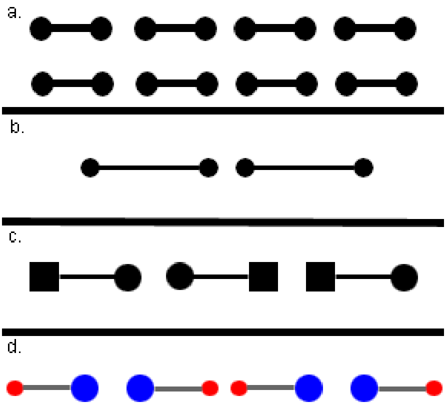
...or this...



...influences perceiving of this symbol.

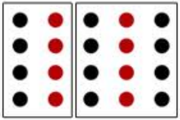
- Past experience and context have an effect of how we interpret and group elements (e.g. you may perceive either **ONE** letter (**B**) or **TWO** digits (**13**) on the last picture).

(10) uniform connectedness (Law of Unity)



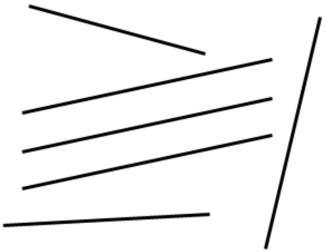
- We group those objects that appear to be connected with a line (a). This principle is stronger than proximity (b) or similarity (c,d).

(11) common region



- We group objects located within the same closed region.

(12) parallelism



- Elements that are parallel to each other appear more related than elements not parallel to each other.

Gestalt grouping principles: Conclusion

Gestalt laws of perception should not be regarded as rigid laws but as helpful heuristics. The way the brain groups objects depends on many hints.

(Example: compact objects (proximity) + recognizable boundary (closure) + the same movement direction (common fate) + the same color (similarity) = red ball)

In the result brain prefers "rational" (i.e. the most simple and stable) interpretations prior to "fluctuating" (i.e. unstable under small changes).

Applications: character- and face- recognizing, CAPTCHA, user interface design, advertisement

Criticism & Limitations

- The gestaltists believed that perception came from innate processes, while the current view is that it is a combination of learnt and innate processes.
- People can view the same figure in different ways.
- The gestalt view is descriptive rather than explanatory (explanations tend to be post-hoc, so it is difficult to make predictions).
- It uses vague language (e.g. it's not clear what a good or simple shape is).
- The approach is a bottom-up approach, while there is much evidence that top-down processes, based on prior knowledge, affects perception.

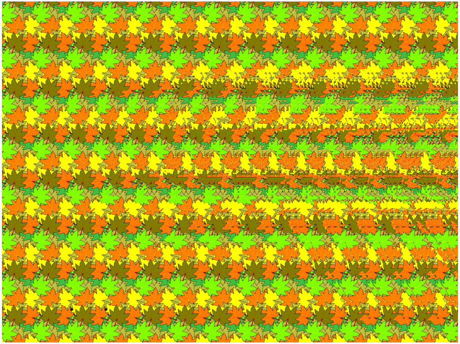
3D perception

One of the most difficult problem for the brain is to determine the 3D depth from a perceived 2D retinal image. For this purpose we use various cues including:

- stereopsis (i.e. eyes placed frontally, e.g. herbivores usually don't have this) Note that images from both eyes have to be combined in a right way, which is not always the case for everybody.
- motion of eyes and head (or of the objects, resp.)
- shades of objects
- perspective principles
- occlusion (covering objects with other objects)
- relative size of objects
- texture gradient
- aerial perspective

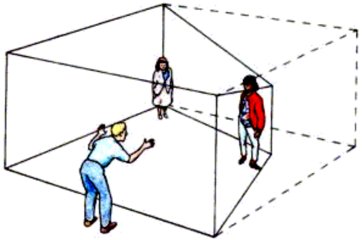
Having (+) or missing (−) these cues cause on one hand having additional impression of 3D depth for 2D plain images or, on the other hand, losing information about the 3D space.

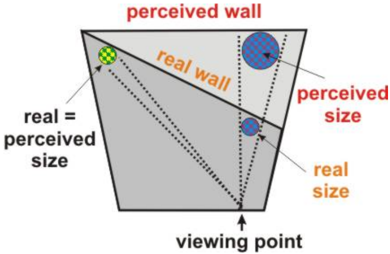
NOTE: 3D perception of *OBJECTS* is not sufficient to tell which of them are closer and which further. We need to perceive in 3D the whole *SPACE*!



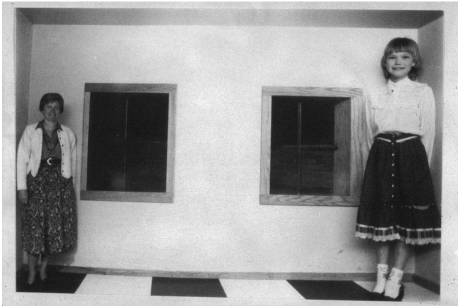
An ant (supposedly...)

- Stereograms are 3D images hidden within another picture. To see the hidden picture most people need only to unfocus their eyes and look through the stereogram, others need different techniques.

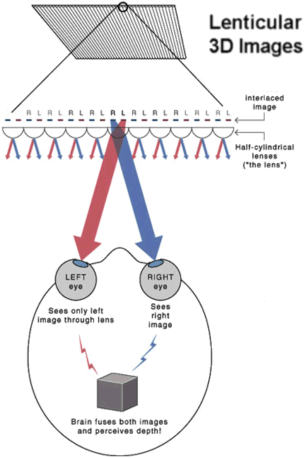




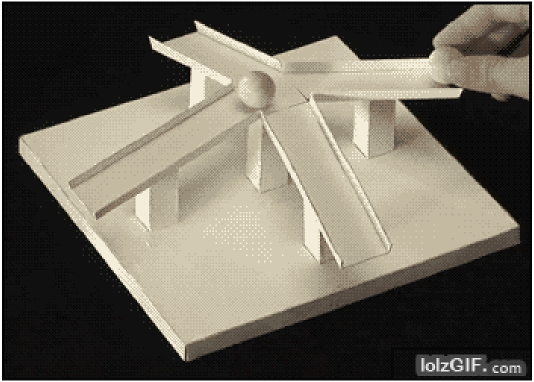
- a trapezoidal room appearing to be cubic-shaped when watched through a peephole with only one eye (constructed in 1934 by Adelbert Ames)
- Ames' original design contained a groove with a ball in it, that appeared to roll uphill, against gravity (possible explanation of "magnetic" or "gravity" hills in the landscape - see uphill illusion).

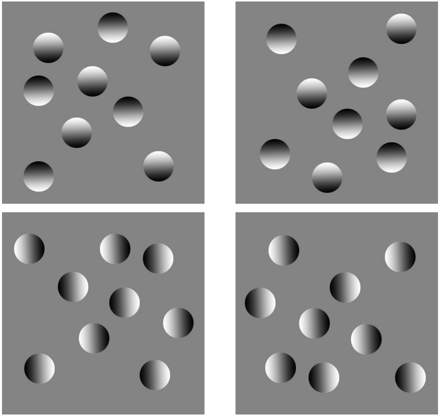


- The person in one corner appears to be a giant, while the person in the other corner seems to be a dwarf. It is so convincing that a person walking from left to the right and back appears to grow or shrink.
- Used in TV and movies for special effects (e.g. Lord Of The Rings)
- Honi phenomenon: some married persons (often women than men) perceive here less size distortion of the spouse than of a stranger.



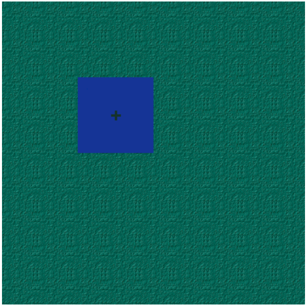
- Lenticular images use lenticular lenses to produce printed images with an illusion of depth, or the ability to change or move as the image is viewed from different angles.


lolzGIF.com

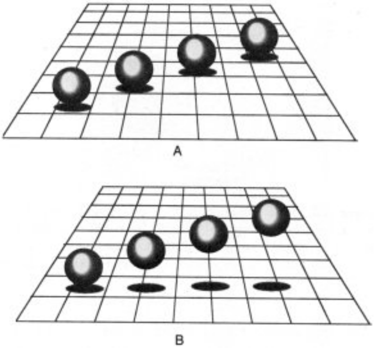


Which discs are convex and which concave?

- Pictures in each row are rotated by 180 degrees with respect to each other. This means that each disc changes how it is perceived, because of consistency (a possible explanation is due to Bayesian analysis).
- The first row pictures are often perceived as there is 'light from above'.

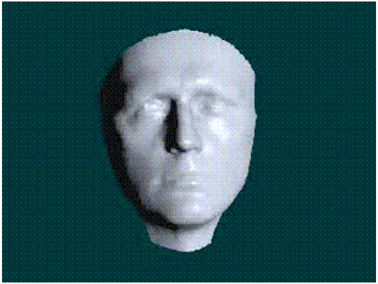


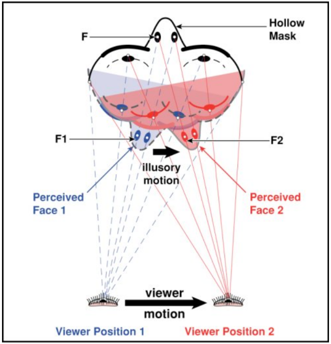
Is the square moving?



Have the balls changed their positions?

- Positions of shadows change the perception substantially.
- Despite the strong perception that the square (in the 1st picture) is going up and down, the only thing actually moving is the shadow.



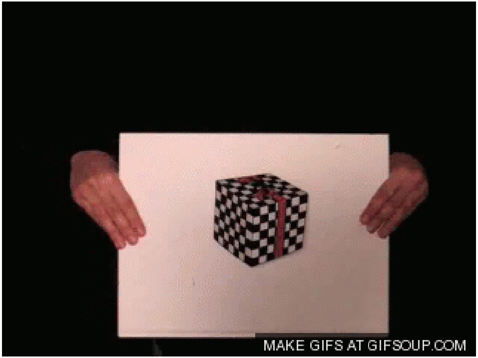


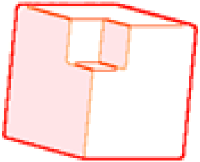
Does the rotation switch the direction?

- This psychological bias of seeing faces as convex is so strong it suppresses even the depth cues, such as shading and shadows.
- Schizophrenia sufferers are not fooled by this optical illusion.
- Lighting a hollow face from below to reverse the shading cues can reinforce the illusion.

Does the middle ballerina turn to the right or to the left?

- Silhouette is more often seen rotating clockwise (as seen from above) than counter-clockwise (as seen from below). According to a research with varying camera elevations this is rather caused by a viewing-from-above bias than by a clockwise bias.





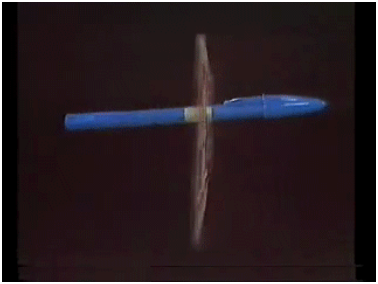
Do you see rather a cube or an (inner) corner?

- We expect often that a bounded object should be 'convex' or bulging towards to us.

(3D perception) (+) Perspective

Are the rails, corresponding to each other on both pictures, parallel?

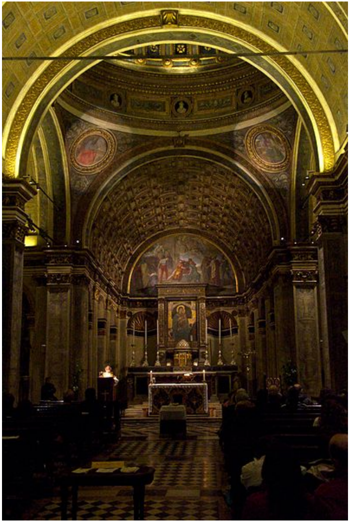
(These two pictures are exactly the same and put one next to each other without any rotation.)

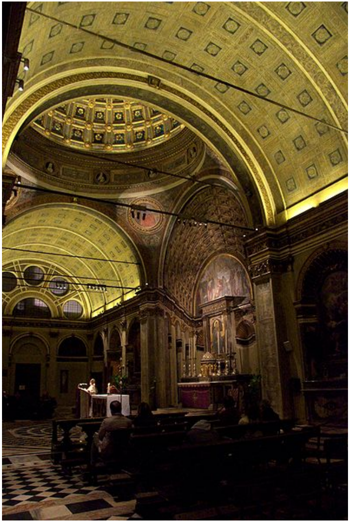


- It is not a rectangular window but a trapezoid.
- When observed the window appears to rotate through less than 180 degrees (this varies with the dimensions of the trapezoid).
- The rotation seems to stops momentarily and reverses its direction.

- The Potemkin Stairs in Odessa extend for 142 meters, but give the illusion of greater depth since the stairs are wider at the bottom than at the top.
- Forced perspective: a technique to make an object appear farther away, closer, larger or smaller than it actually is (used in photography, filmmaking and architecture).

(3D perception) (–) Perspective: Forced (deformed) perspective

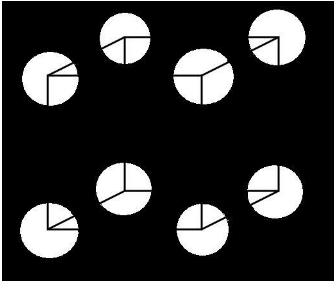


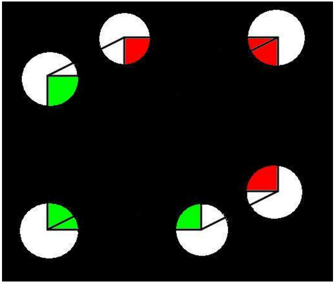


Illusion of a large space at the apse of Santa Maria presso San Satiro in Milan, Italy

(real space is just around one meter deep)

- Occlusion (also referred to as interposition) means overlapping of objects. The overlapped object is considered further away.
- Partial covering of objects may cause them to seem closer (first picture) or bigger (second picture).

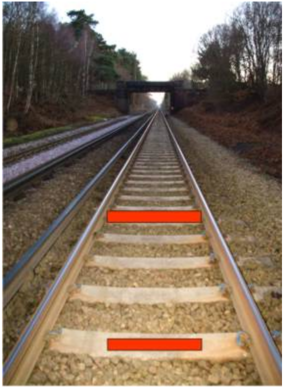




From which angle we see (from above) a transparent (possibly colored) cube?

Can you see the hidden animals?

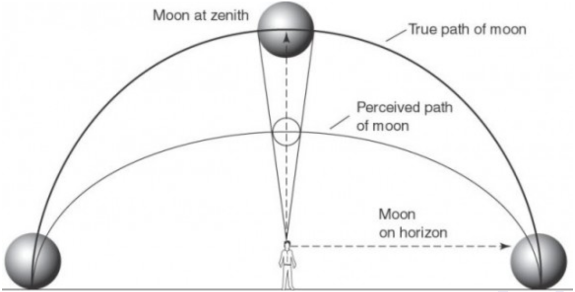
What kind of car do you see?

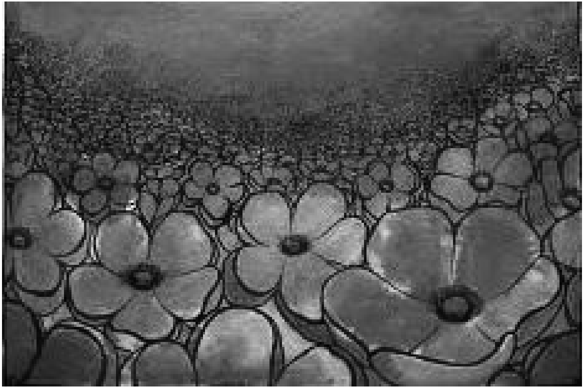


Are the lines (cars) of different size?

- If one of two objects with the same (absolute) size subtends a larger visual angle on the retina than the other, then it appears to be closer. Vice versa: if we think that one of two objects with the same (retinal) size is further, then it appears to be bigger.


Actual
Multiple Exposure
Moon Illusion
The moon seems larger in angular size when it is near the horizon than when it is high in the sky

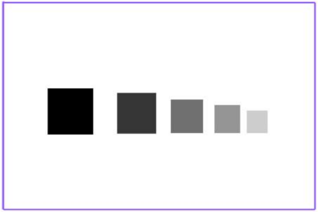

Moon at zenith
True path of moon
Perceived path of moon
Moon on horizon

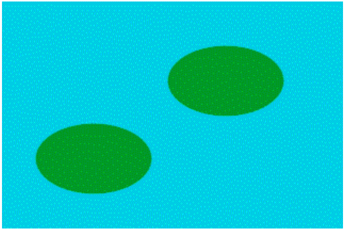


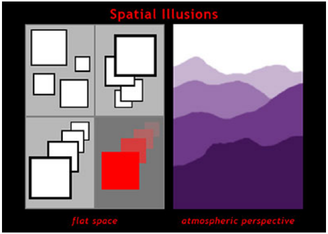
- Texture gradient: fine details on nearby objects can be seen clearly, whereas such details are not visible on faraway objects. This cue provides illusion of depth even on simple 2D drawings.

(3D perception) (−) Texture gradient: Tilt-shift photography

- Tilt-shift photography: use of camera movements - rotation of the lens plane relative to the image plane (tilt) and movement of the lens parallel to the image plane (shift). Selective focus via tilt is often used to simulate a miniature scene from real scenery.







- Due to light scattering by the atmosphere, far away objects have lower luminance contrast and lower color saturation (shifted toward the blue end of the spectrum). Objects differing only in their contrast with a background appear to be at different depths.

Raw and detailed perception

- To recognize (and understand) what we see we need to perceive both raw contours and fine details together.
- If we could only see bigger spatial changes, our world would seem to be blurry. On the other hand, seeing only details we would not be able to realize mutual interconnections.
- These scalings are called *spatial frequencies* - *high* for the details and *low* for the raw contours. Different parts of the brain response to different spatial frequencies.
- Similarly, in computer graphics, images are functions f(x,y) describing how colours or grey values (intensities, or brightness) vary in spatial coordinates (x,y). The Fourier theory provides a decomposition of these functions into particular frequencies. Using a *low-pass filter* (that passes low-frequency signals and reduces the amplitude of signals with frequencies higher than the cut-off frequency) we may obtain picture corresponding to lower frequencies. An analogous result we obtain by using *high-pass filter*.

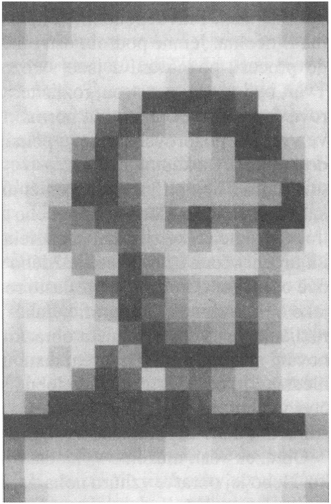
High and low frequencies

Original image

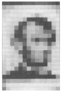
High-pass filtering

Low-pass filtering

Raw and detailed perception



Can you recognize what is in the picture?



What about now?

(the same picture but smaller)

Raw and detailed perception

Can you recognize a familiar picture?

(Try to squint your eyes...)

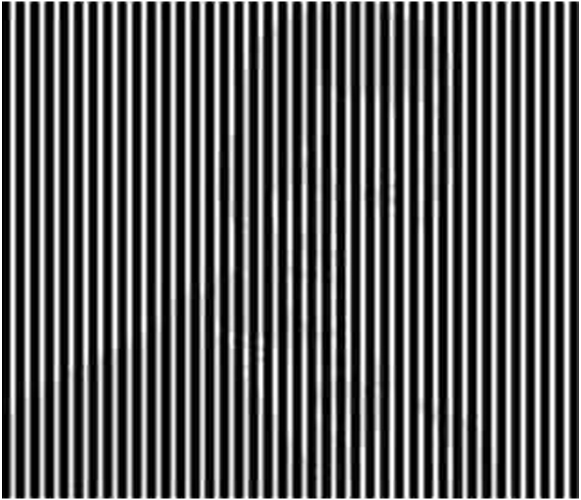
Can you find the lion?

What do you see?

Raw and detailed perception: Mrs. Angry and Mrs. Calm

Who is who?

- At close range it seems that "Mrs. Angry" is on the left and "Mrs. Calm" on the right. From a bigger distance they switch their positions.



Can you see something?

(Cross your eyes, blur your stare or look from a bigger distance...)

References:

- FRANCIS CRICK: Astonishing hypothesis - A scientific search for soul
- JOSEF PELIKÁN: Optikcé kalmy (slides)
- WIKIPEDIA and other internet sources

Further reading:

- CATHERINE Q. HOWE, DALE PURVES: Perceiving Geometry - Geometrical Illusions Explained by Natural Scene Statistics
- GAIL A. CARPENTER, STEPHEN GROSSBERG (EDS.): Neural Networks for Vision and Image Processing

Thank you for your attention!