Friedrich-Alexander-Universität
Department Artificial Intelligence
in Biomedical Engineering | AIBE

Machine Learning
Data Analytics

Chapter 3:

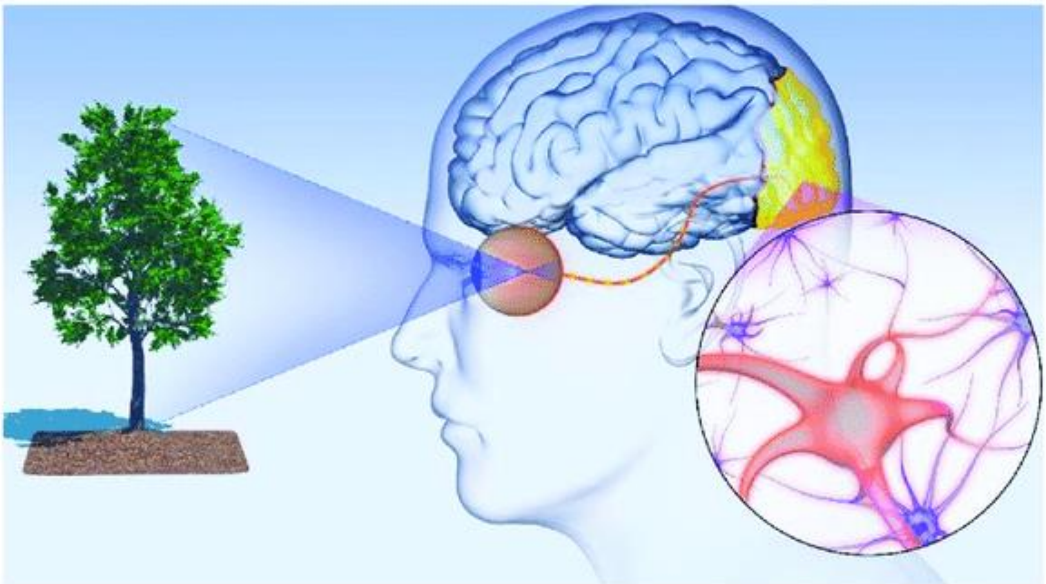
Visual Perception, Optical Illusions, and Gestalt Laws

Overview

1. Visual Perception

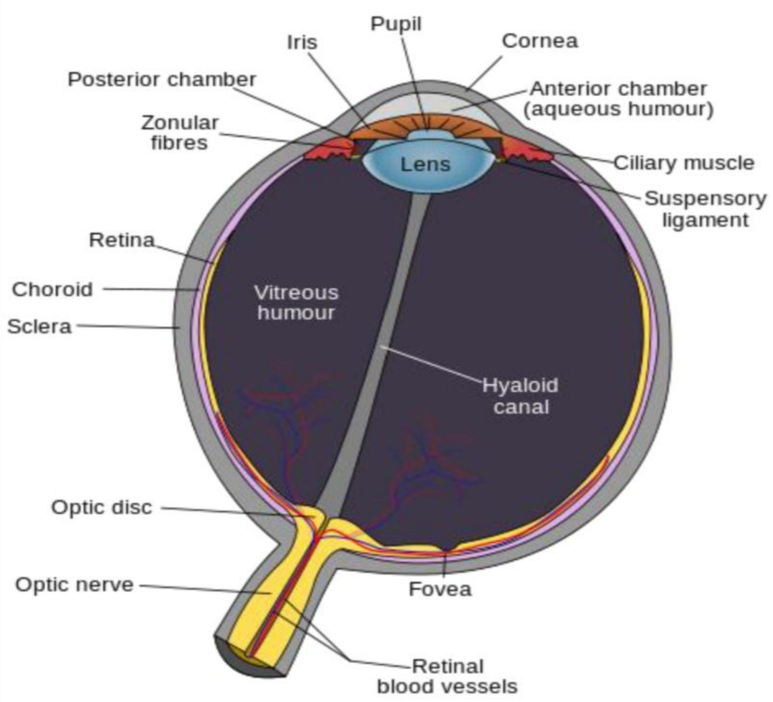
Visual perception is one of the most important sources of information.
Approximately 60-80% of all information is perceived visually.
We can define three terms that will clearly distinguish the processes in visual perception:

Reception describes the transformation of the stimulus (light) into electrical energy.
Cognition describes the *"Understanding"* in the brain.
Perception describes the sensors (receptors) and signal processing happening in the eyes and in brain.



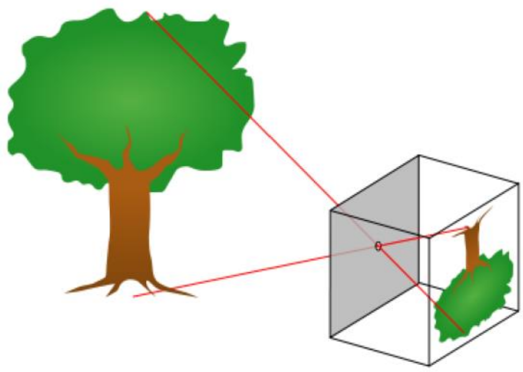
Deng, Wei & Zhang, Xiujuan & Jia, Ruofei & Huang, Liming & Jie, Jiansheng. (2019). Organic molecular crystal-based photosynaptic devices for an artificial visual-perception system. NPG Asia Materials. 11. 77. 10.1038/s41427-019-0182-2.

Sensory Organ – The Human Eye



The human eye has some very basic attributes:

- Very high dynamic range
- Bad color vision in dark conditions
- Best contrast perception in red/green
- Limited temporal resolution (reaction speed) – The human is said to be blind when moving the eyes.
- Good resolution and color in central area (macula)
- Maximum resolution and color only in the very center (fovea)
- Retina contains **rods** for low light vision and **cones** for color vision (they transform light into electrical energy) → Receptor for light stimuli
- Ganglion cells inside the retina are already part of the brain and detect patterns and movements
- Pinhole camera (everything we see is upside-down projected on the retina)



Interpreting the signal

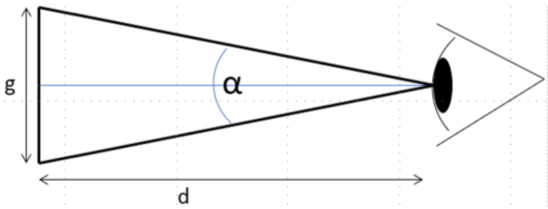
Here we will cover the basic and first signal interpretations:

- Size and depth
- Brightness
- Color

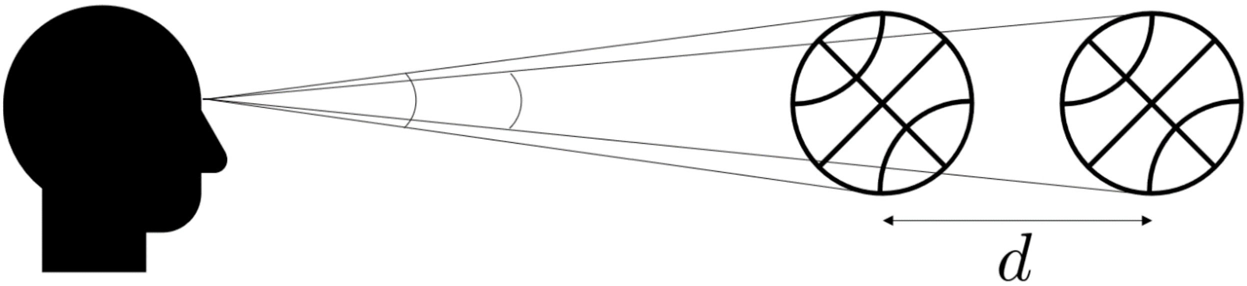
Size and depth:

Visual acuity (VA) is the ability to perceive details (smallest resolvable object size *g* at given distance *d*). This ability is limited and can change over time. The better, the more precise is our interpretation of size.

$$VA = \frac{1}{arctan(\frac{g}{d}) \cdot 60}$$



The visual angle indicates how much of our field of view an object occupies (relates to size and distance from eye)



Visual angle of an object (ball) maps onto the retina. The same object spans a different visual angle based on the distance to the eye.

Familiar objects are perceived as constant size (despite changes in visual angle when far away). This one is closely related to the image above. The two balls are interpreted as having the same size.

Our depth perception mainly relies on *depth cues*.
Thereby we can distinguish two types of *cues:* **monocular and binocular depth cues**

- Monocular depth cues (depth cues we can perceive with one eye)
  Examples for monocular depth cues are:

  - **Accomodation** is tension of the muscles for the lens of the eyes -> change focal length, helps to correctly map the image onto the retina
  - **Monocular Movement Parallax** -> by moving your head you can perceive depth
  - **Retinal image size** -> when object size is known, smaller objects are perceived farther away (see above)
  - **Linear perspective** -> railroad tracks that meet in infinity
  - **Texture gradient** -> closer means more detailed (standing at a tree and looking up -> rough bark of the tree loses details) -> relates to visual acuity
  - **Overlapping** -> closer objects block objects that are further away
  - **Aerial perspective** bluish fog or hazy
  - **Shadows** give a hint when there is only one light source

- Binocular depth cues (depth cues we can only perceive with both eyes)
  Examples for binocular depth cues are:

  - **Convergence** when the eyes are moved inward to focus on a close object
  - **Binocular parallax** are differences in the perspective onto a scene or object caused by the distance between the eyes (different viewing locations)

Brightness:

We have a very subjective reaction to levels of light. However, our reaction is still affected by the luminance of an object. Our eyes "measure" the luminance by just noticeable differences.

Interestingly our **visual acuity increases with luminance** as does flicker (fast changes in luminance).

Rods have a **lower density at the fovea but a higher density temporal and nasal to the fovea**. Thus, they contribute more to the peripheral vision. They cannot detect color.

Color:

Our color perception is made up of:

- **Hue**
- **Intensity**
- **Saturation**

Cones are sensitive to color wavelengths, whereby the acuity to blue color is lowest.

The spectrum of visible light for humans is quite small:

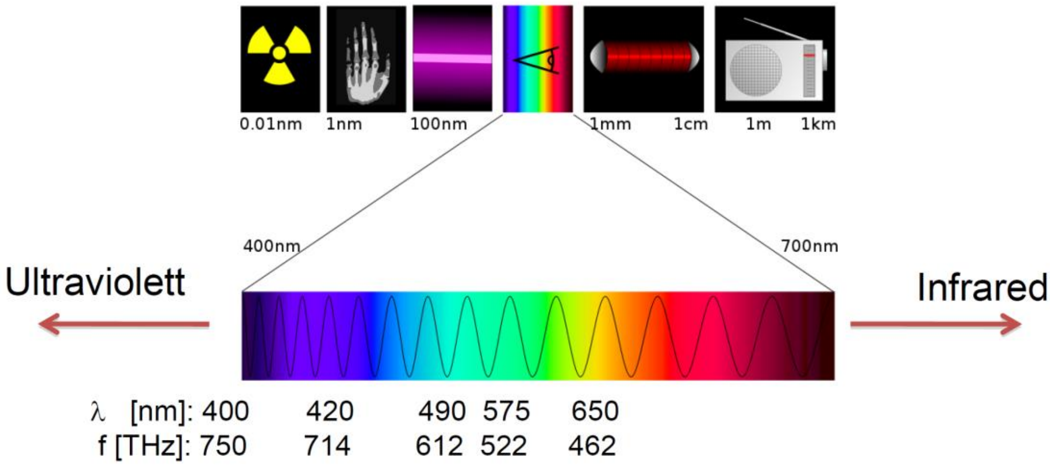
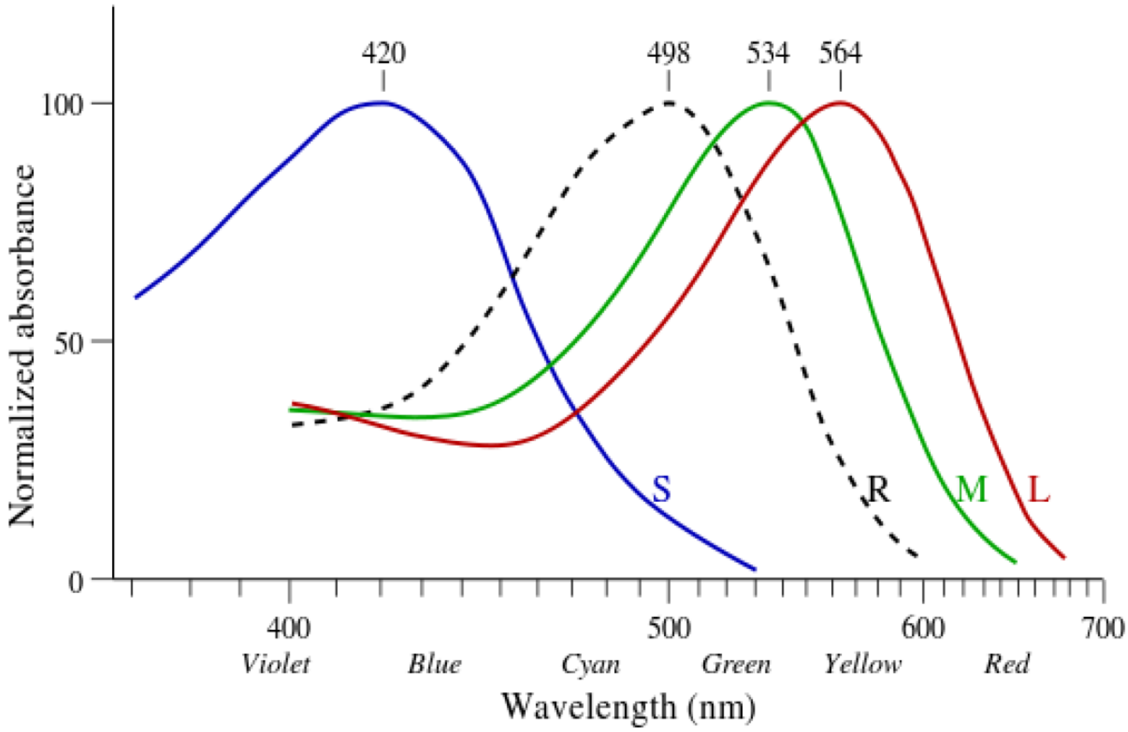


Image by Tatoute and Phrood (CC BY-SA 3.0)
https://commons.wikimedia.org/wiki/File:Spectre.svg

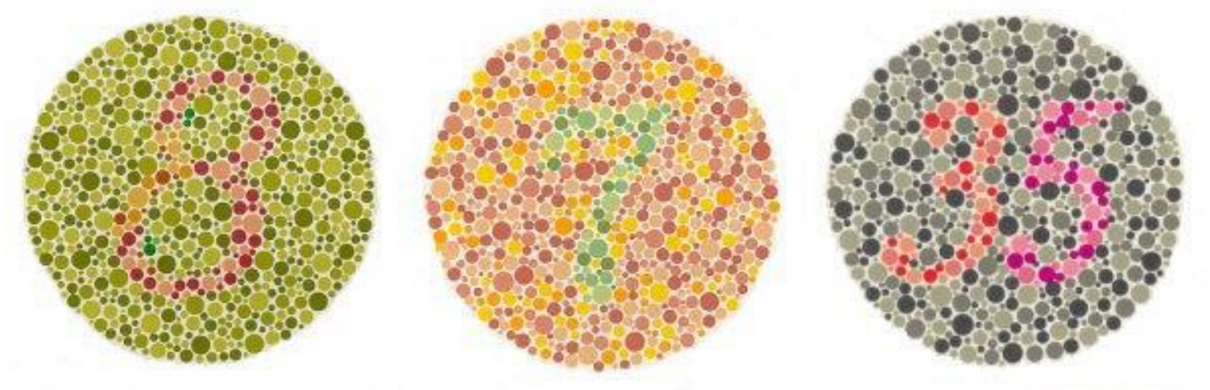
We only can perceive wavelengths between 400 and 700 nm. The Tristimulus theory (**trichromaticity**) describes that humans have three different cones that are differently sensitive to wavelengths:

- Red (Long)
- Green (Medium)
- Blue (Short)
- Rods: Dashed line → cannot detect color, sensitive to all wavelengths



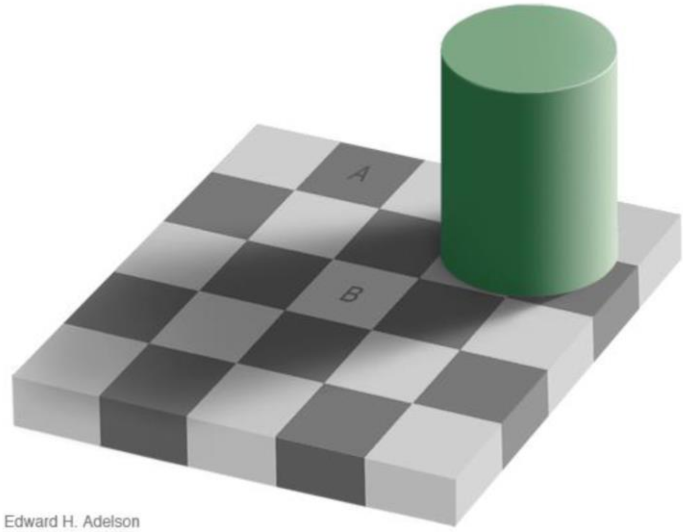
https://de.wikipedia.org/wiki/Datei:Cone-response-de.svg
Bowmaker J.K. and Dartnall H.J.A., "Visual pigments of rods and cones in a human retina." J. Physiol. 298: pp501-511 (1980).

Some people cannot detect colors in general. They are color blind or cannot detect or distinguish between specific colors. Color blindness is a hereditary disease and is more prevalent in male (**8-10% males and just 1% females are color blind**). The ype of color blindness can be detected with the following images:



Left:	Under normal color viewing condition: 8	Red/Green blind: 3 or nothing
Middle:	Under normal color viewing condition: 7	Color blind: nothing
Right:	Under normal color viewing condition: 35	Red blind: 5; Green blind: 3

Our perception can be changed by the environment. A very good example is the *Adelson's checkerboard illusion.* This illusion makes use of a phenomenon in which the perceived brightness and color seems to play tricks on us. The brightness and color of tile *A* and *B* seem to be different. Are they different?



Video Link: https://www.youtube.com/watch?v=GALLMJxLvgA

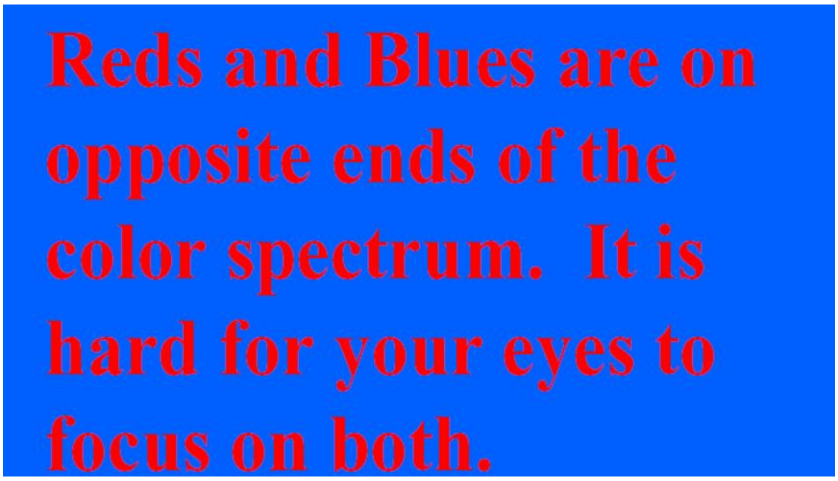
Thus, using color keys can be difficult and misleading!!

Another phenomenon is the afterglow of the opposite color. Following the link below, you will find an animation. You must fixate on the star in the middle. When you do that, a green dot will appear on the spots where the purple dot is missing.

Video Link: https://appliedperception.wordpress.com/2015/09/22/hintons-lilac-chaser/

Based on this, we can also define rule of thumbs for good designs especially for- and background colors:

1. Do not use opposite colors because...

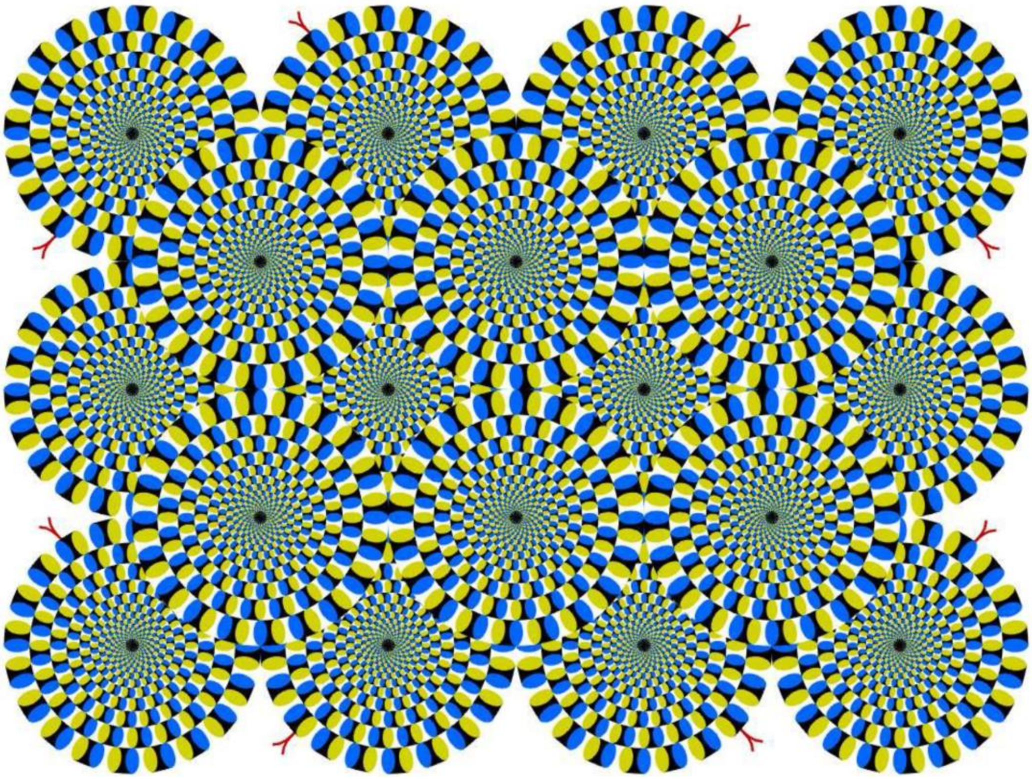


2. Rule of thumb: lightness difference > 0.2 results in good contrast.

2. Optical Illusions and Gestalt Laws

Our visual system compensates for movements and changes in luminance. The Context is used to resolve ambiguity. However, optical illusions sometimes occur due to overcompensation of our visual system.

We – for example – have a different perception in focus region and in peripheral view leading to motion artifacts. When you look at the following image in full screen you get the impression that the circles are moving although they're not. This is caused by the difference in the focus and peripheral view.



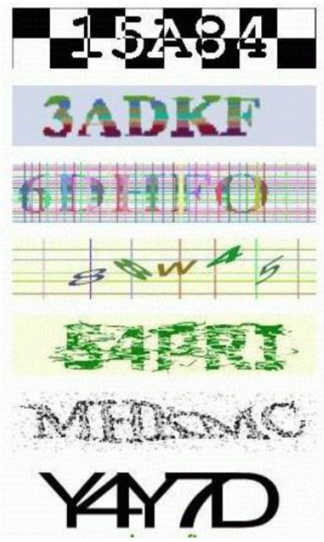
An illusion is also the Escher Waterfall. This picture shows a reality that cannot be real. The waterfall is driving a wheel and after that the water flows back "up" into the tower to again fall and drive the wheel. This is impossible.

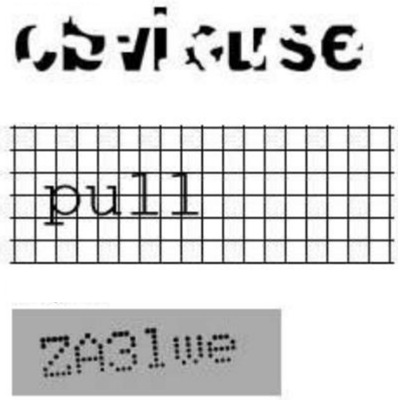
Our visual system tries to make sense out of the visual information we get. This behavior can also be used in HCI to create a special kind of interaction that can be interesting for many different applications.
Good examples of this are the Gestalt Laws and surely all of us already got in touch with those…

Gestalt Laws

Why can we decode Captchas?
What are the implications for user interface design?

Look at the following captchas.





It is simple for us to read the words or digits. However, for a computer it is not that easy. Here we can make use of our behavior to make sense out of given visual information. All these Captchas follow Gestalt Laws.

Have a closer look at the warning signs before the staircases:

The left one reads “Keep Red” and “Off Line” because we tend to group closer things together. The right one reads “Keep off Red Lines”. Just by placing the words differently the perception changes.

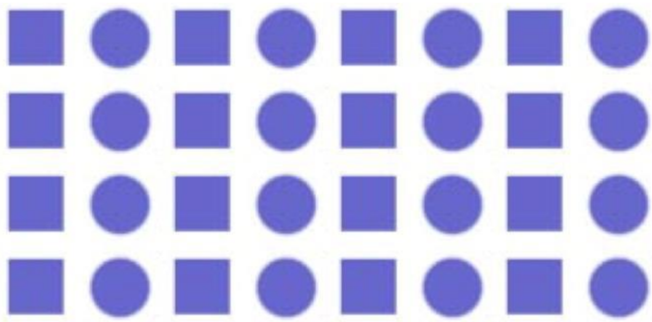
There are many different Gestalt Laws, and we will stick to seven Laws namely:

- Law of Similarity
- Law of Proximity
- Law of Continuity
- Law of Closure
- Law of Pragnanz
- Law of common fate
- Law of Symmetry

There are more Gestalt Laws like *Figure and Ground* or *Smallness Area*. However, we will only cover the seven Laws above.

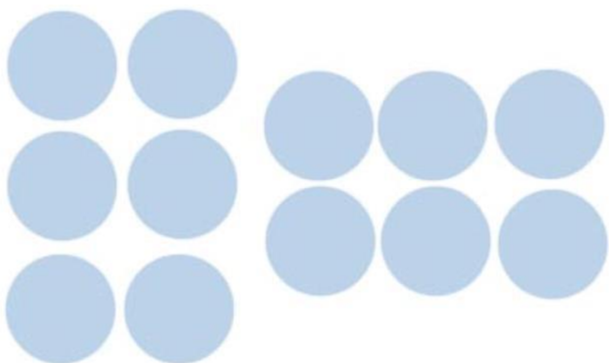
Law of Similarity

- Items that are similar tend to be grouped together
- In the image, most people see vertical columns of circles and squares



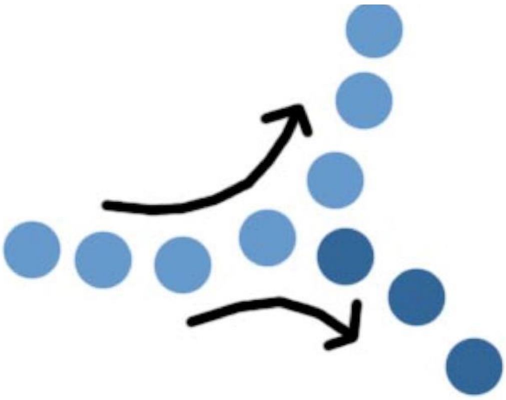
Law of Proximity

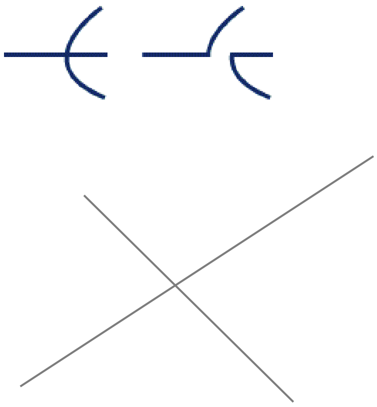
- Objects near each other tend to be grouped together
- The circles on the left appear to be grouped in vertical columns, while those on the right appear to be grouped in horizontal rows



Law of Continuity

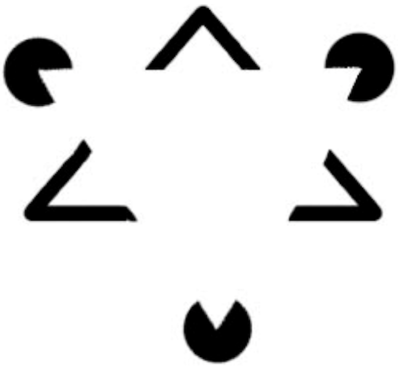
- Lines are seen as following the smoothest path
- In the left image, the top branch is seen as continuing the first segment of the line. This allows us to see things as flowing smoothly without breaking lines up into multiple parts.

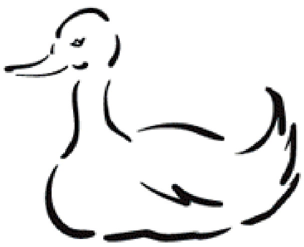




Law of Closure

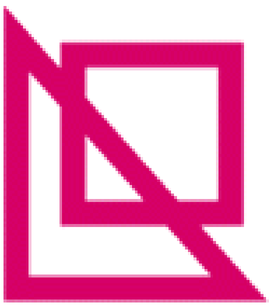
- Objects grouped together are seen as a whole
- We tend to ignore gaps and complete contour lines. In the left image, there are no triangles or circles, but our minds fill in the missing information to create familiar shapes and images.





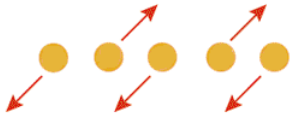
Law of Pragnanz (Law of Simplicity/ Law of good shape)

- Reality is organized or reduced to the simplest form possible
- E.g., we see the left image as a series of circles rather than a much more complicated shape



Law of common fate

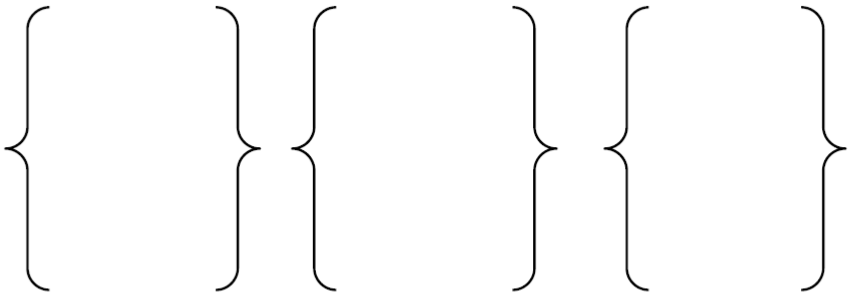
- Elements with the same moving directions are perceived as a collective or unit



Video Link: https://www.youtube.com/watch?v=LZMaTtPHBMk

Law of symmetry

- Symmetrical images are perceived collectively, despite their distance to each other



Other interesting Illusions

There is one interesting phenomenon in perception called the *Change Blindness*. It describes that even large changes in a scene are not noticed. Reasons for this are short distractions caused by *Mud splashes, brief flicker* or *cover boxes.*

Some further explanation can be found here:
http://nivea.psycho.univ-paris5.fr/ECS/ECS-CB.html
https://www.youtube.com/watch?v=5O71iuTPTTI

Examples for change blindness:
http://nivea.psycho.univ-paris5.fr/ECS/kayakflick.gif
https://www.youtube.com/watch?v=FWVDi4aKC-M

Summary

Why do we need to know about all the illusions and Gestalt Laws?
We can better guide the **Distribution of attention and perception.**

The Distribution depends on:

- Culture (cultural background)
- Custom / habit
- Perception
- Processing
- Experience

If we know how to organize elements in a scene or we know which visual phenomenon could result in an illusion that breaks attention, we can account for this and create better User Interfaces or Products in general that can be used easily.