COMP7503 Multimedia Technologies

Media Sensor and Reproduction - Part1

Dr. Bill LUO

THE UNIVERSITY OF HONG KONG

DEPARTMENT OF

COMPUTER SCIENCE

Introduction to Sensors

- Information about the environment is always obtained though sensors
- A sensor is a device that measures a physical quantity and converts it into a signal that a human or a machine can use
- The most important sensors for multimedia computing
    - Sound sensor
    - Light sensor

Types ofSensors

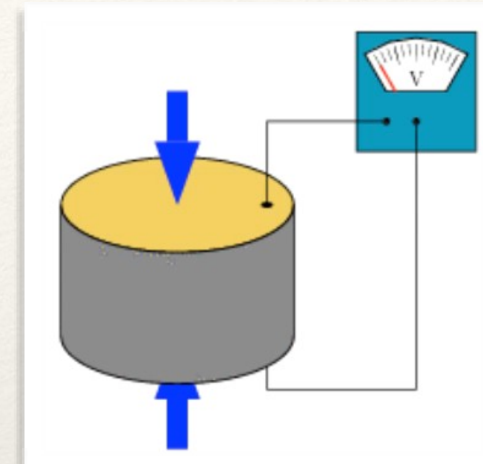
- Sensors that imitate human sensors
- Sensors that do not imitate human sensors
  - Measure aspects of life
- Sensors that measure physical facts about the environment

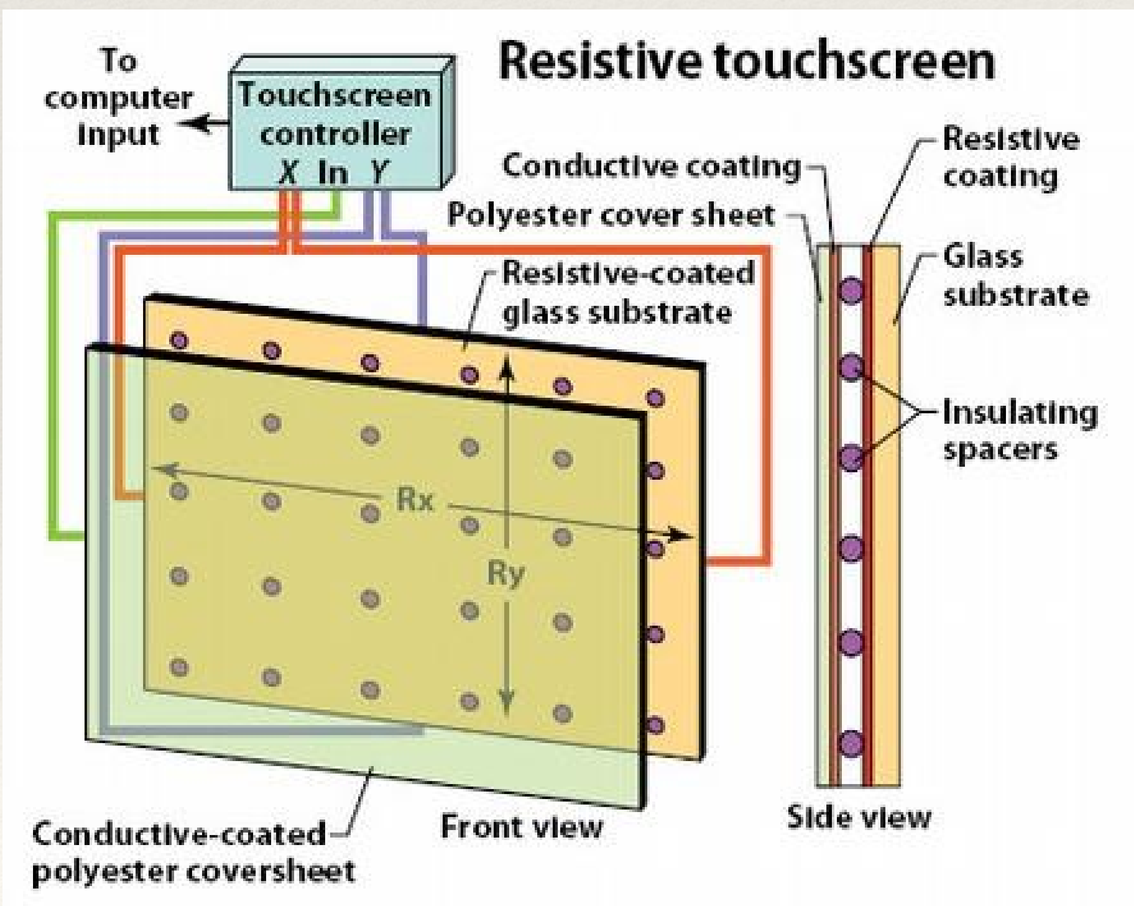
Human Sensors

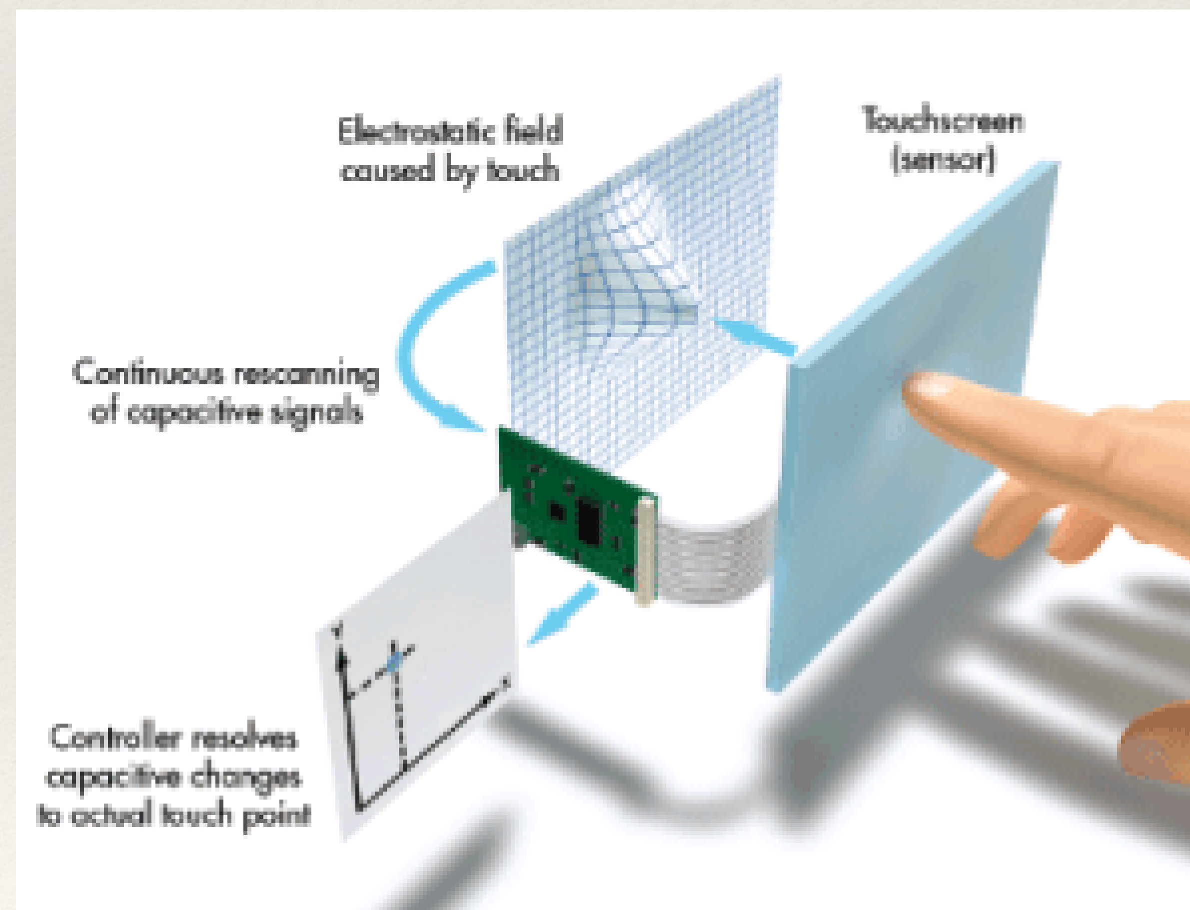
- Correspond to 5 senses
  - Touch
  - Sight
  - Hearing
  - Smell
  - Taste

Touch Sensor

- Sense organ for touch
  - Skin - The largest organ on human body
- Sensors that imitate human skin
  - *Tactile sensors* - Translate pressure into signals
  - Pressure Sensor, Resistive Sensor and Capacitive Sensor







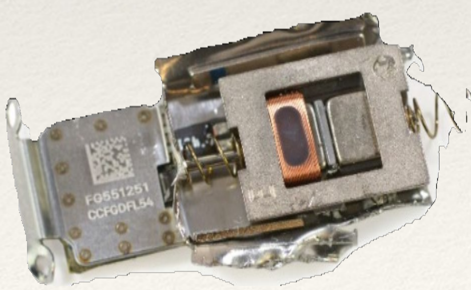
Source: http://mymechatronics.pbworks.com/w/page/63688051/Resistive%20Touch%20Screen%20Sensor

Touch Sensor

3D Touch Technology

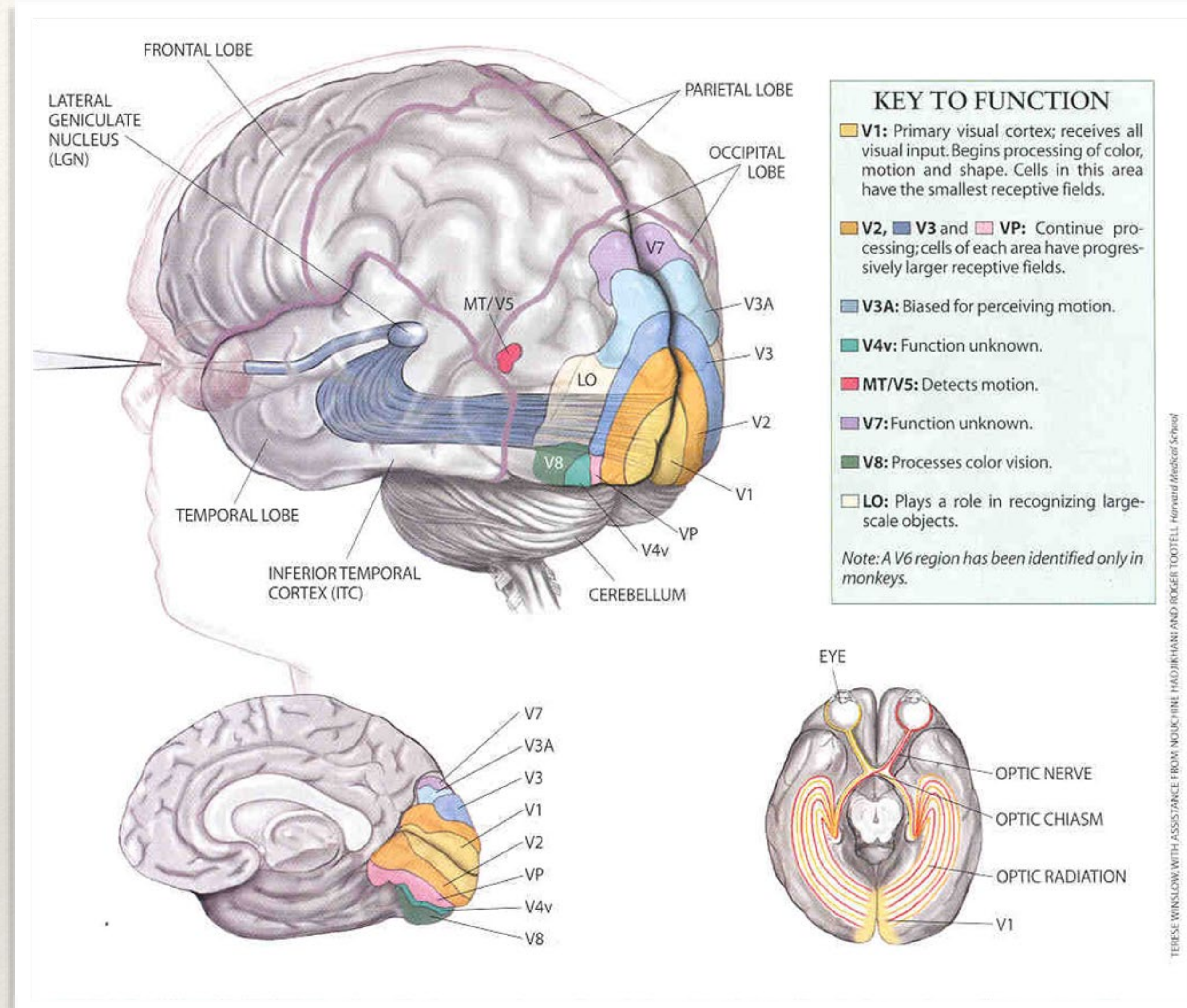
Touch Sensor

- Feedback is provided by ***actuators***
    - Vibration module or Taptic Engine



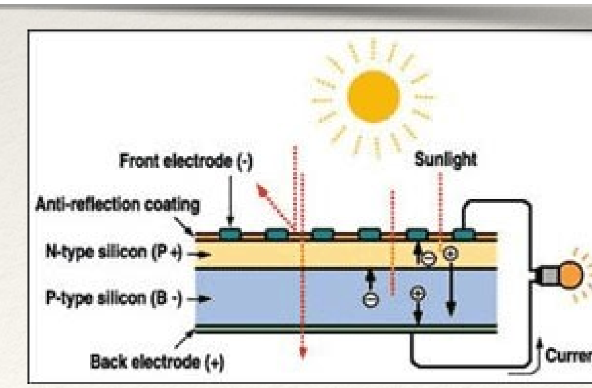
Sight Sensor

- Sense organ for sight
  - Eyes
  - A big portion of human brain is devoted to the sight sensation
- Sight sensors are collectively called cameras
- Sensing elements
  - Based on electrical photovoltaic principles
    - Charge-Coupled Device (CCD)
    - CMOS Sensor
  - The more sensor elements, the higher would be the resolution



References:

1. https://covdblog.wordpress.com/2010/07/06/the-hole-in-the-hand-a-look-into-the-neural-software-of-the-visual-brain/
2. http://www.engineering.com/SustainableEngineering/RenewableEnergyEngineering/SolarEnergyEngineering/Photovoltaics/tabid/3890/Default.aspx



Spot the differences

Source: http://www.spotthedifference.com

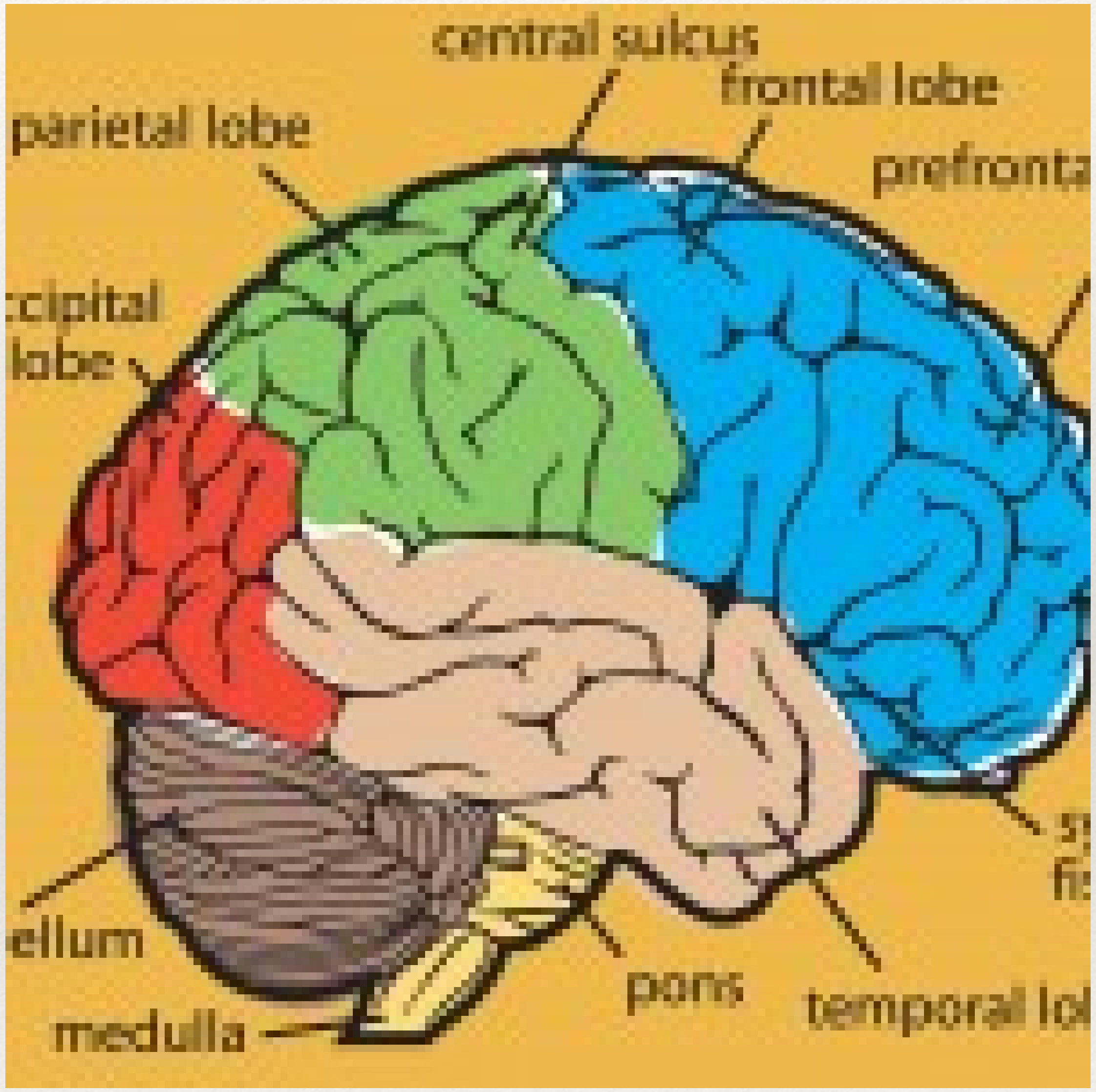
Spot the differences

Source: http://www.spotthedifference.com

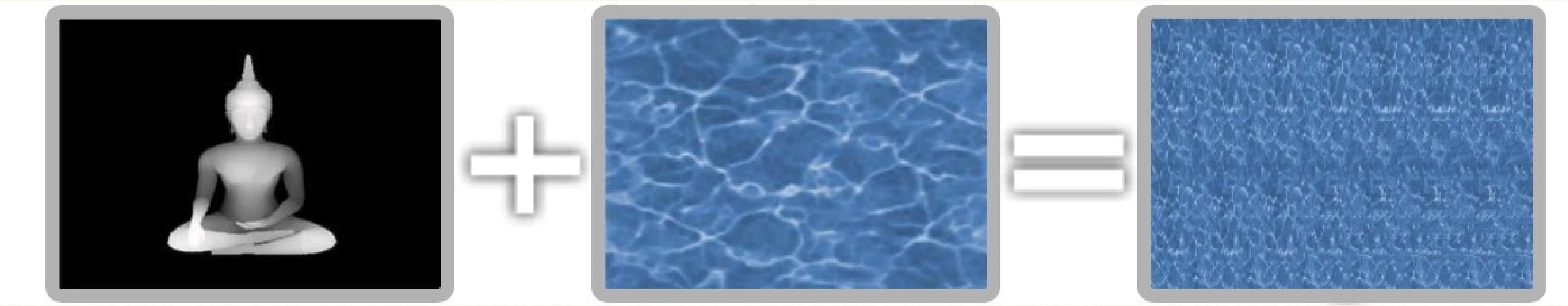
Spot the differences

Which brain areas you used just now

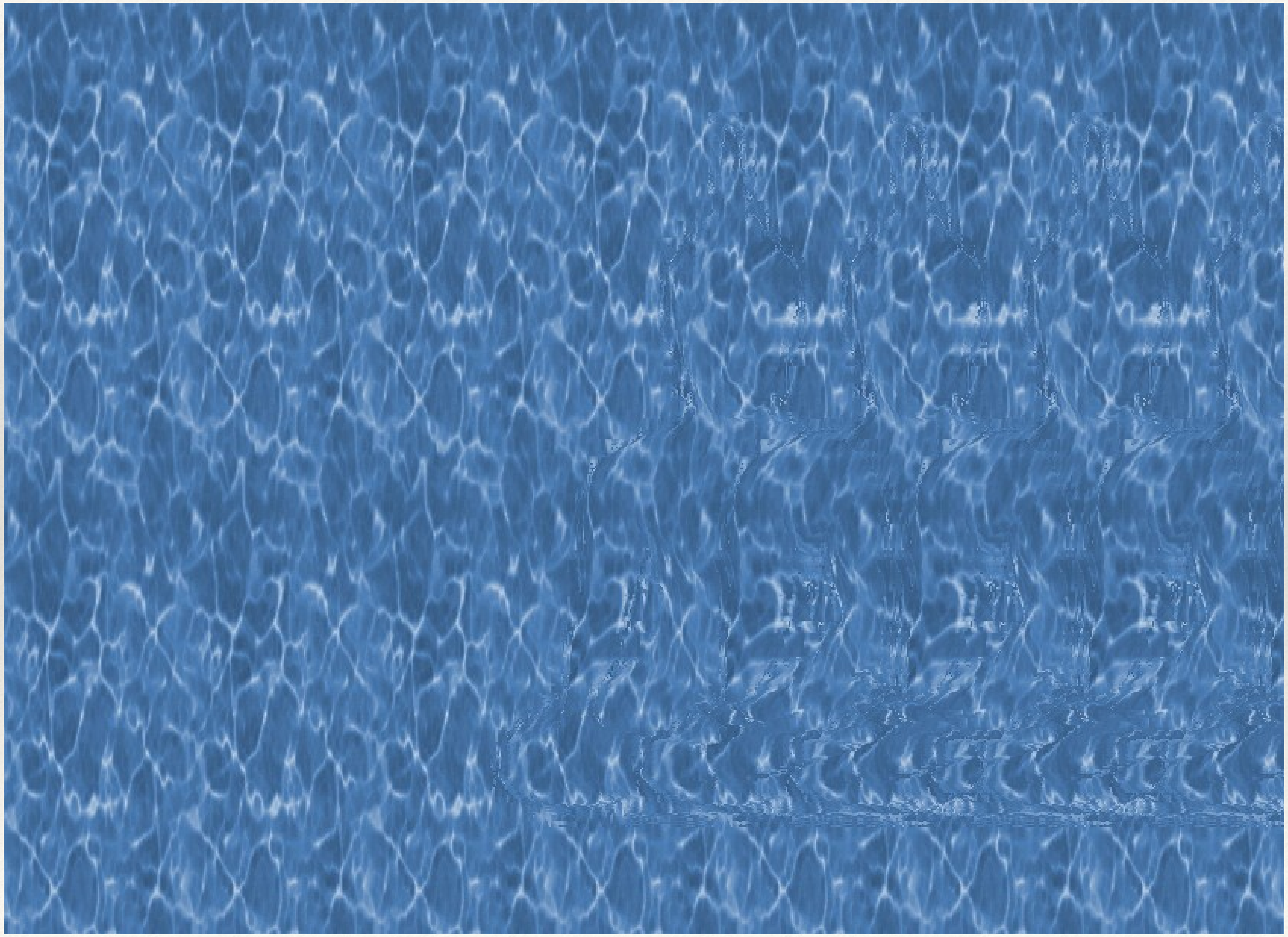
- *Red region*: You have to identify the objects that you see: this involves your occipital lobes.
- *Green region*: You have to analyzed the spatial relationships between the objects that you see: this involves your occipital and parietal lobes.
- *Blue region*: You have to remember what you see in one picture and compare it to what you see in the other picture, that is you have to use your short-term memory: this involves your frontal and parietal lobes.
- *Blue region*: You have to mark down the locations where you see a difference: this involves mostly your frontal lobes (for the movement)



https://sharpbrains.com/blog/2008/08/12/brain-teasers-spot-the-difference/



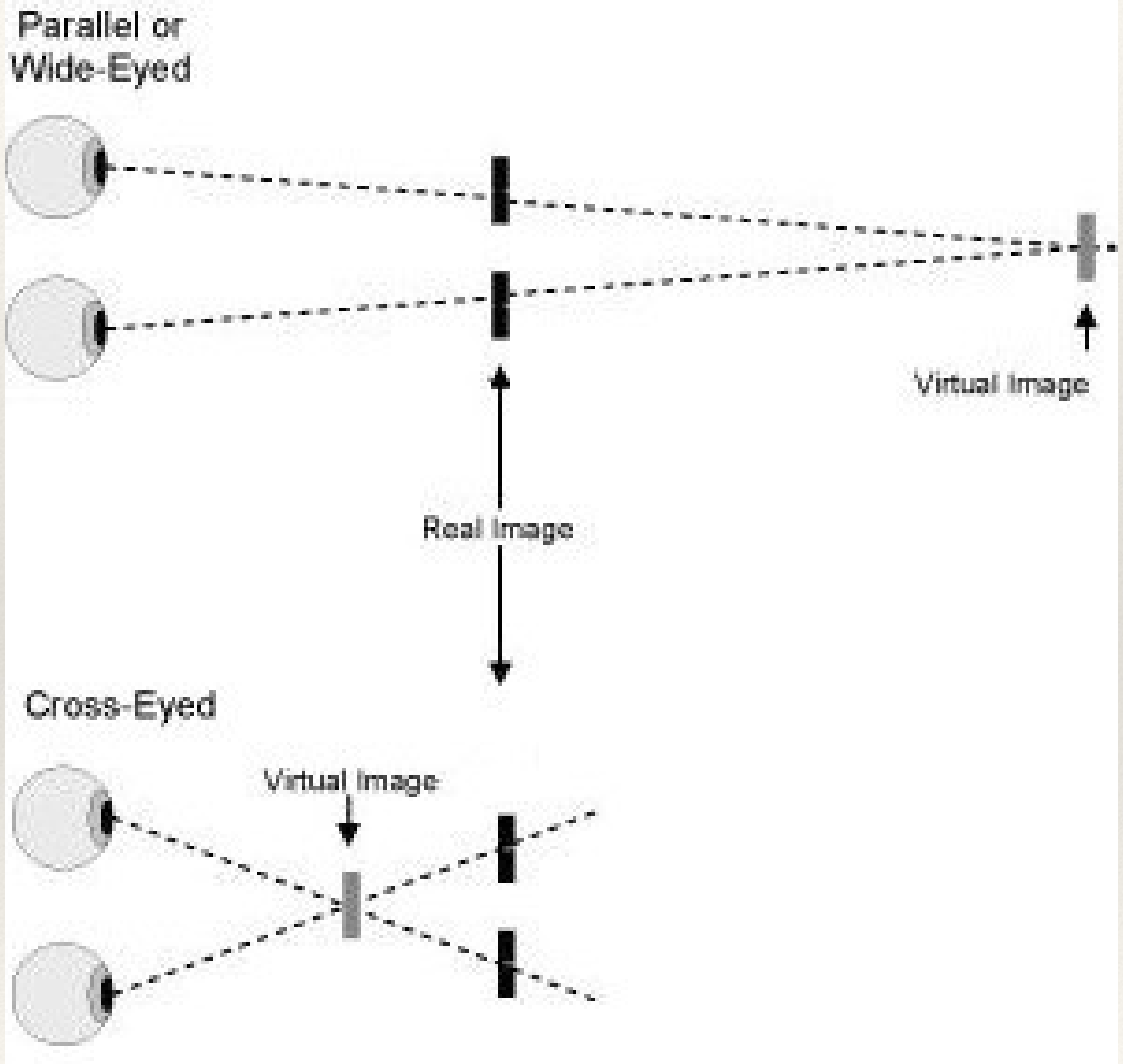
- Binocular Vision

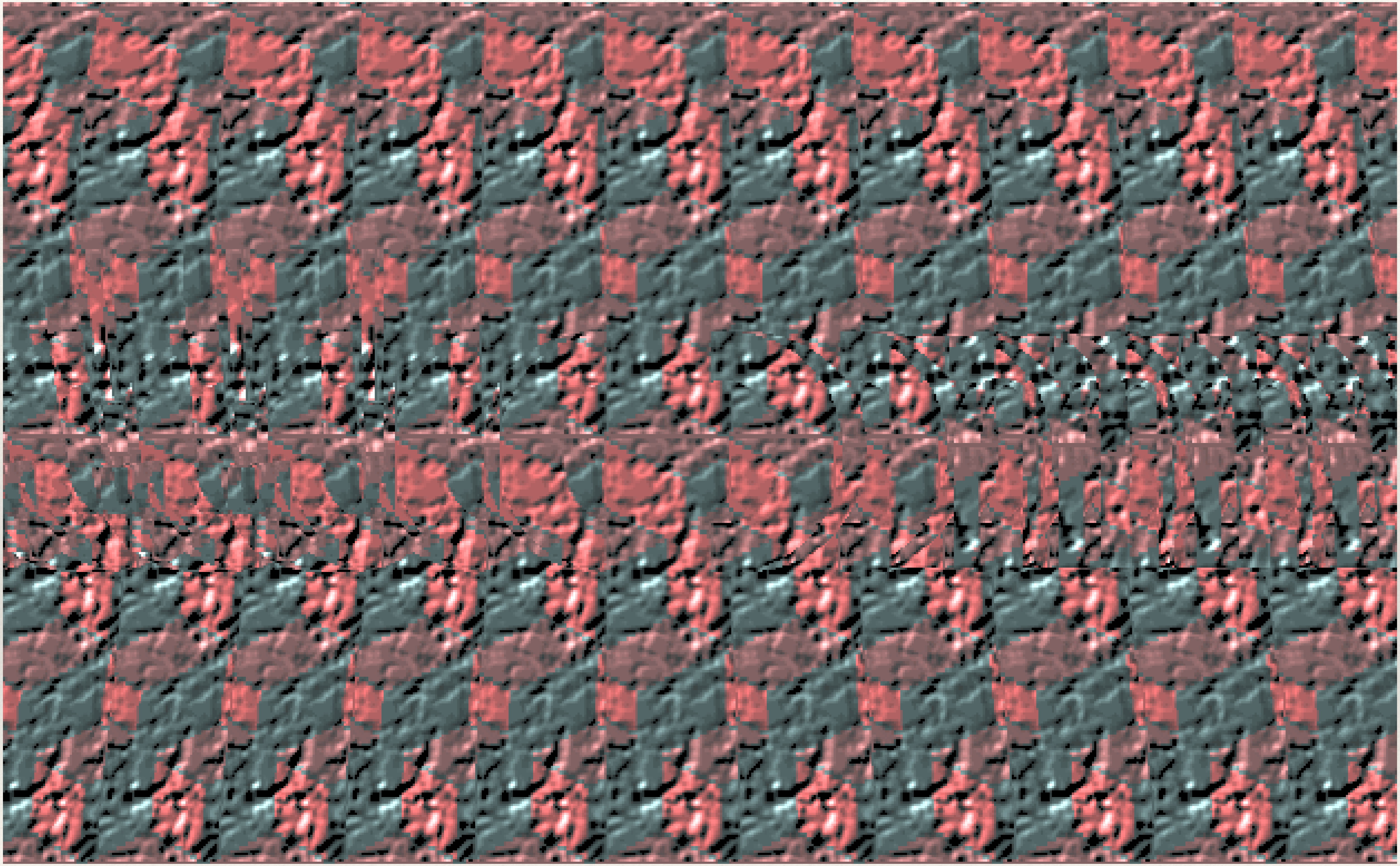


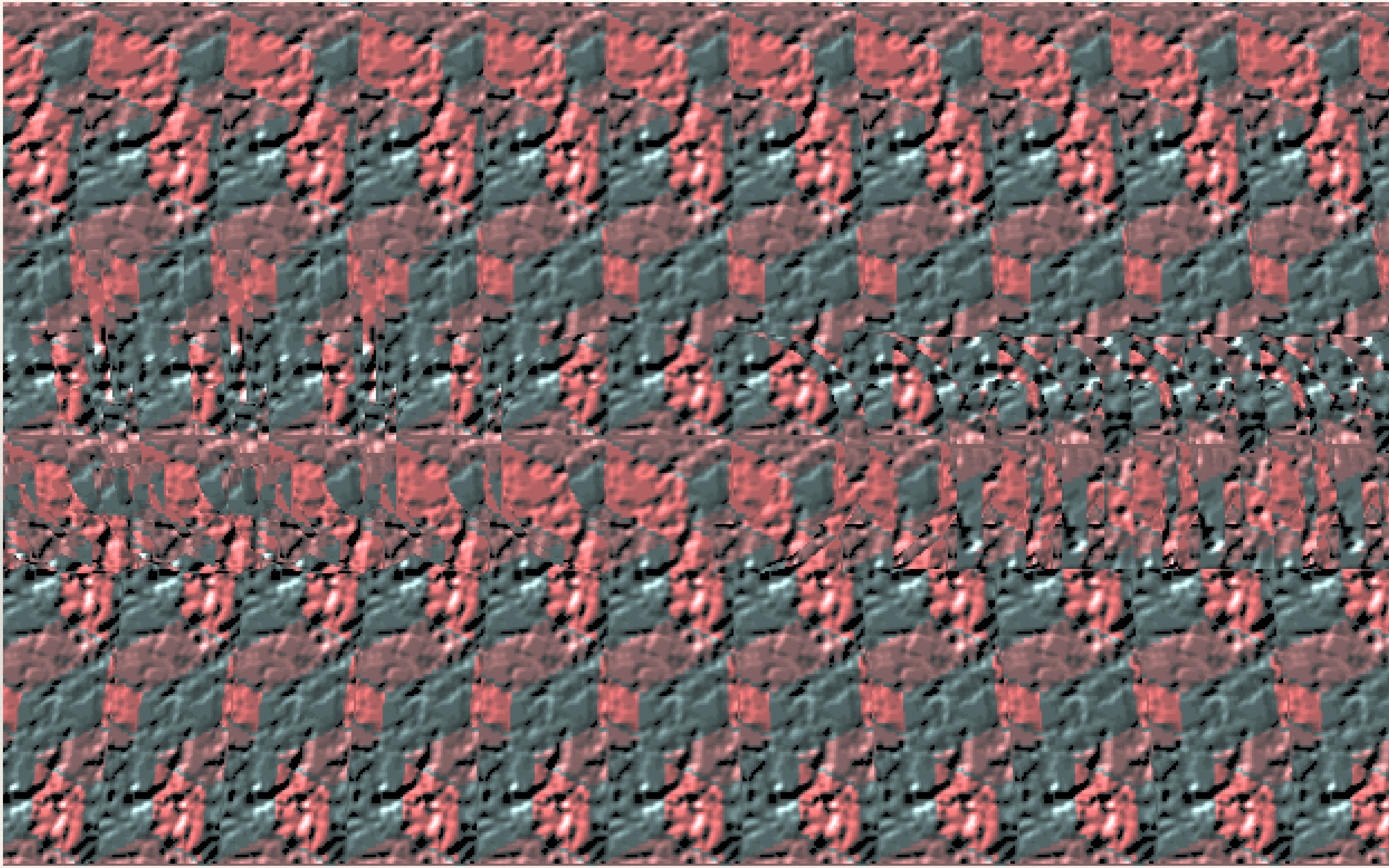
 Ref: https://www.easystereogrambuilder.com

How to see a Stereogram

- Bring the stereogram image really close to your eyes (until you touch it with your nose). At this distance your eyes cannot focus on the image and they look somewhere behind the image. Now, slowly push the image away from you, while trying to keep the eyes off focus. At some point you will see the hidden image.
- Another method is to take an object and put it behind the image (about half of meter behind it). Now, focus on the object behind the image while keeping the eyes looking at the image.



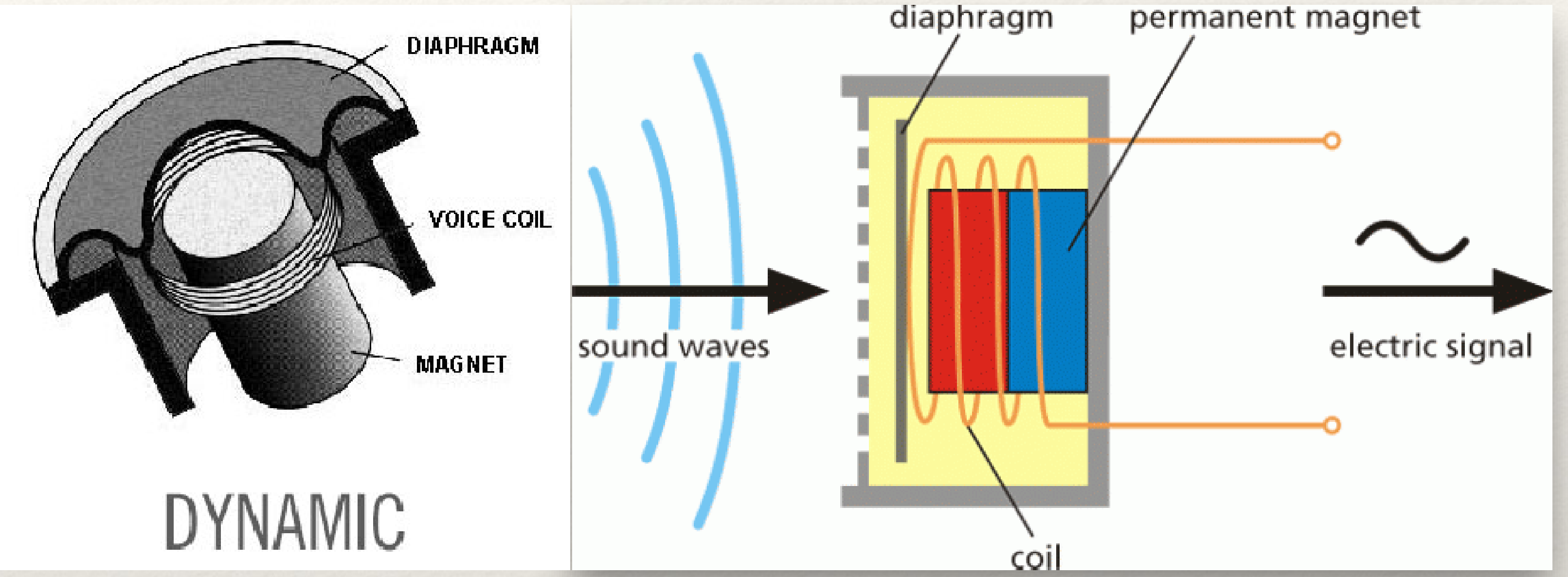


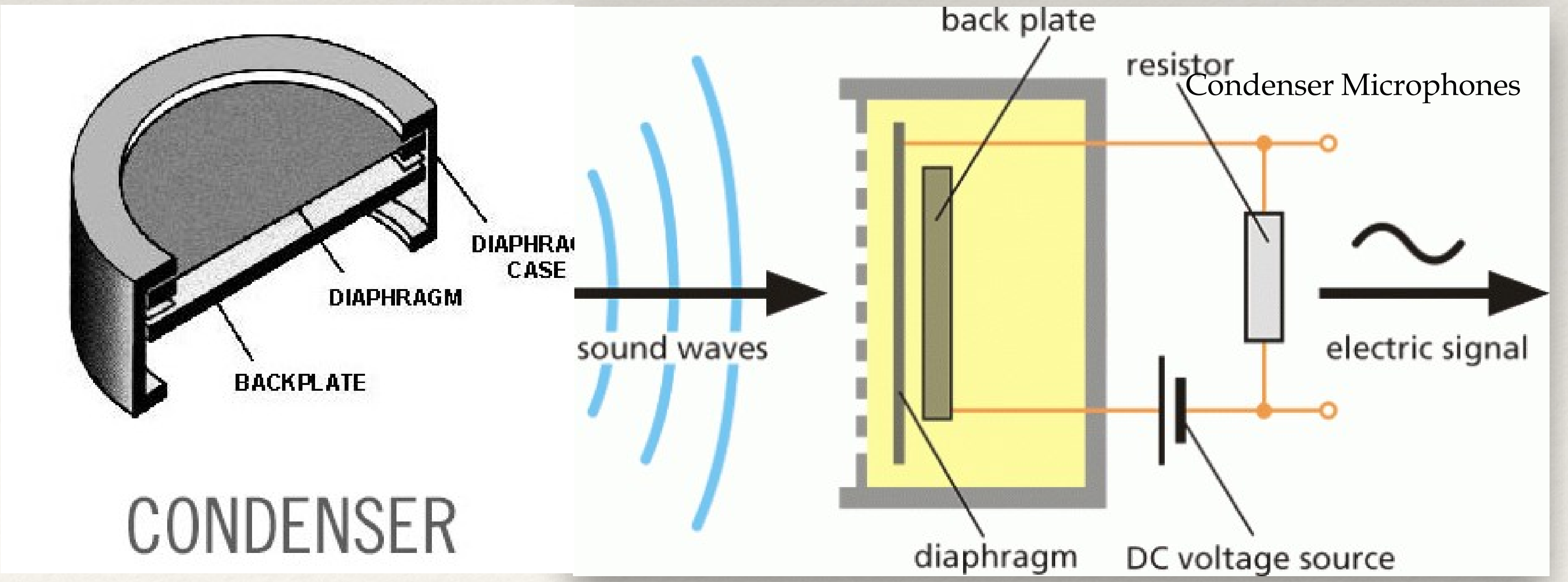


Ref: http://www.tanos.co.uk/portfolio/graphics/randomdot/

Hearing Sensor

- Sense organ for audio
  - Ears
- Audio sensors are usually called microphones
  - Converting change in air pressure into electrical signals





Condenser Microphones

Reference:
https://backstagebitsgoa.wordpress.com/2014/03/29/speakers-microphones/
http://shureblog.co.uk/microphone-basics-transducers-polar-patterns-frequency-response/

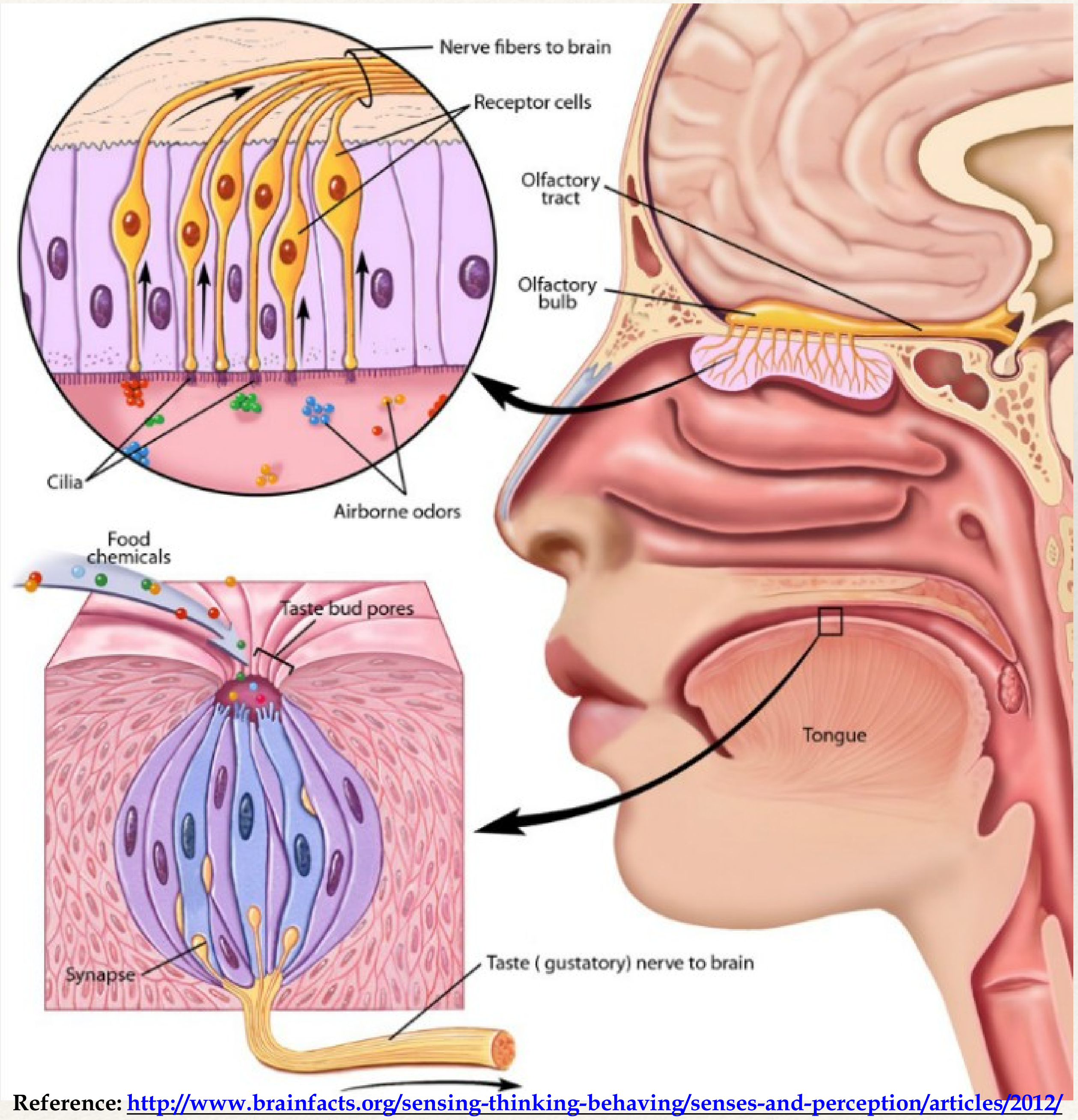
Hearing Sensor

2021 Best Selling Dynamic

2021 Best Selling Condenser

Smell & Taste Sensors

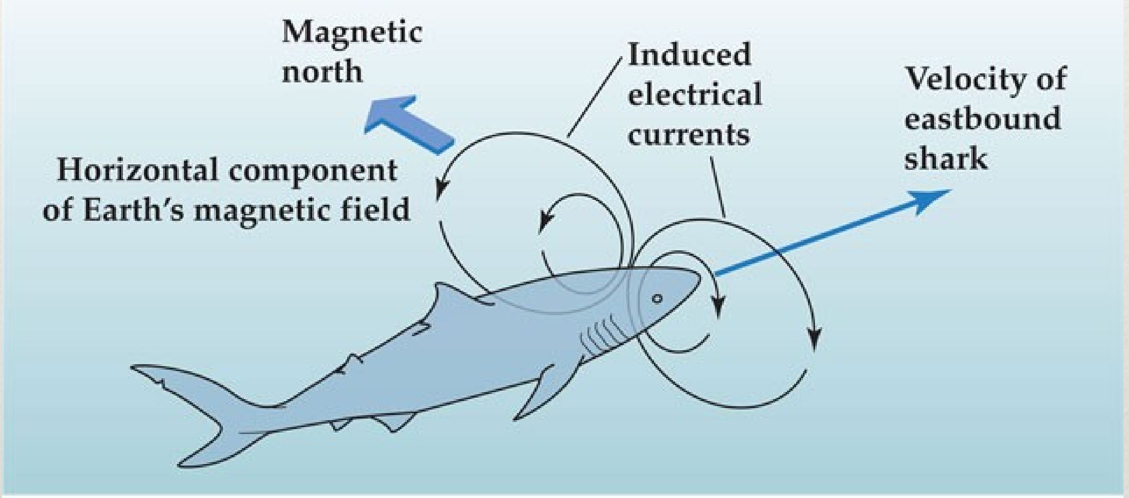
- Sense organs for smell and taste
  - Nose & Tongue
- Chemistry and physics of smell and taste are not well understood yet
  - No artificial noses and tongues have been properly produced yet
- For this course, we shall focus primarily the sense of sight and hearing



Reference: http://www.brainfacts.org/sensing-thinking-behaving/senses-and-perception/articles/2012/taste-and-smell/~/media/Brainfacts/Article%20Multimedia/Sensing%20Thinking%20and%20Behaving/Senses%20and%20Perception/Taste/TasteSmellLarge.ashx

Sensors in other Species

- Other animals have sensors for the following
  - Temperature
  - Gravity
  - Magnetic fields
  - Humidity
  - Infrared



Source: http://sites.sinauer.com/animalphys3e/boxex/AnPhys3e-BoxEx-18-01-B-0.jpg

Sensors about Life

- Sensors used to measure life vitals or health statistics
- Weight sensor
  - To measure body weight
  - Together with height, gender, and age, we can also estimate the BMI values
- Temperature sensor
  - To measure body temperature
- Noise sensor
  - To detect loud environment
- Heart rate sensor
  - To measure heart rate
- ECG sensor
  - To detect Sinus Rhythm
- Movement sensor
  - To measure motion/acceleration magnitudes
    - Together with heart rate, estimation of burned calories is possible
- These sensors are becoming more abundantly available, so so they are likely to be very important for future multimedia computing

Sensors about Life

一線通 平安鐘

長者安居協會

Care-on-Call Service for senior citizen

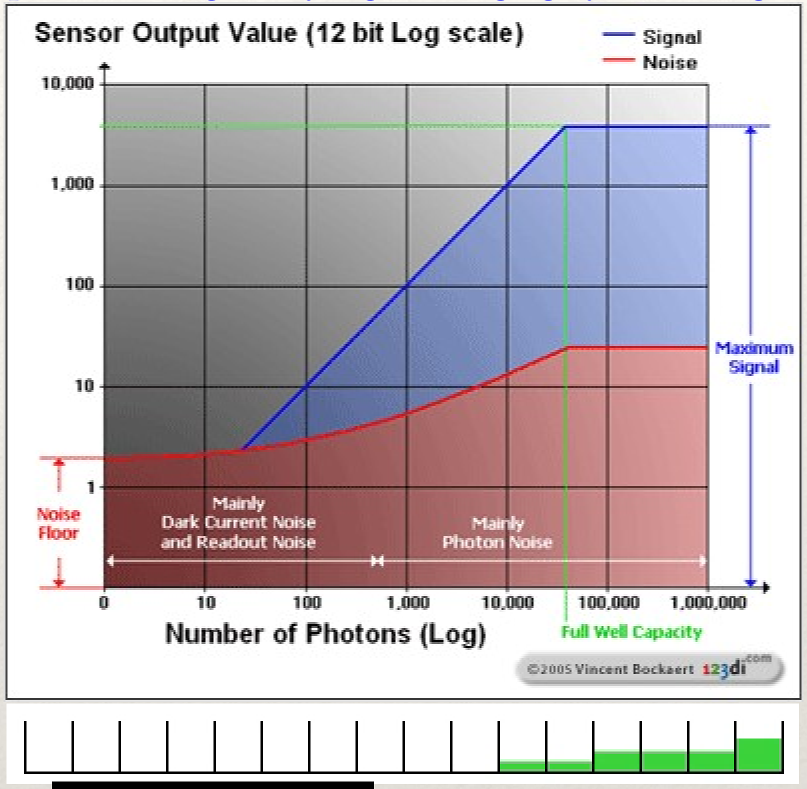
Environment Sensors

- Accelerometers
  - To sense device orientation
  - Can provide data for pedometer
- Gyroscopes
  - To sense device rotation motion
- GPS receiver
  - For location sensing
- Radiation sensor
  - For sensing radiation
- Altimeter sensor
  - For sensing altitude
- Sometimes, environment sensors is more reliable than human perception
  - A person may suffer from spatial disorientation

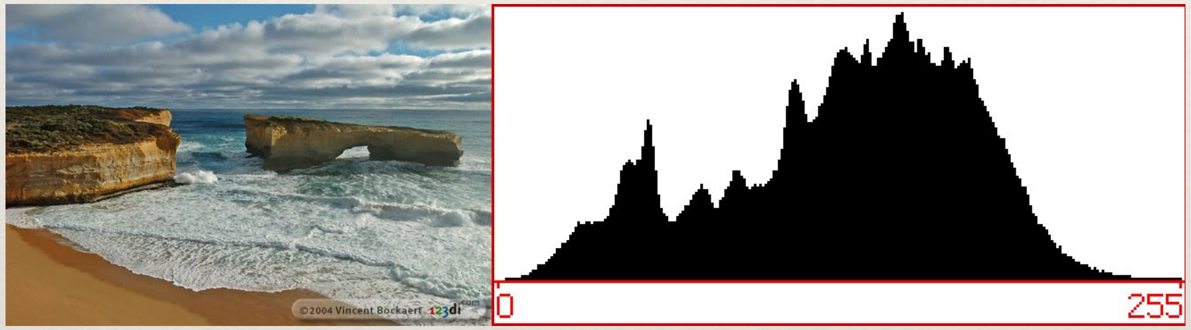
Reference: http://www.aboutflight.com/handbook-of-aeronautical-knowledge/ch-7-flight-instruments/gyroscopic-flight-instruments

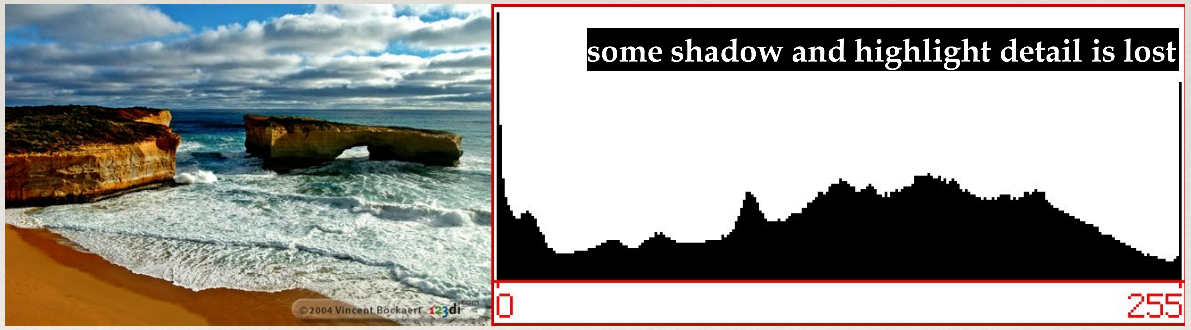
Sensors Limitations

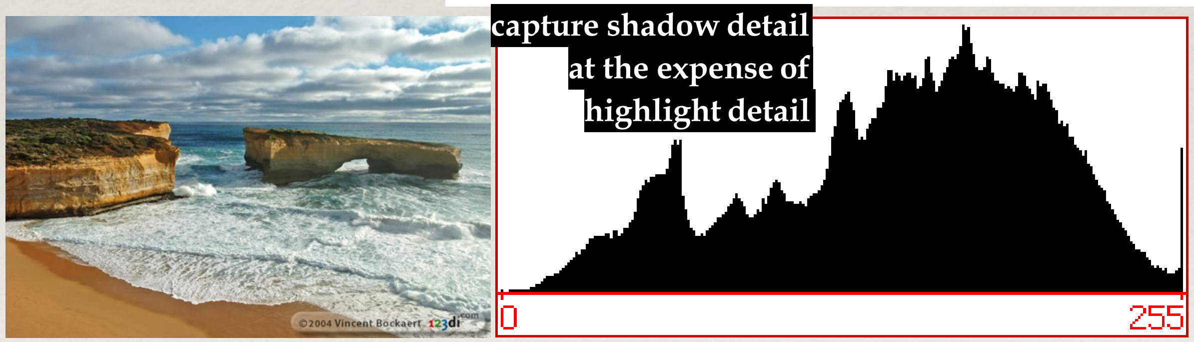
Reference: http://www.dpreview.com/glossary/digital-imaging/dynamic-range

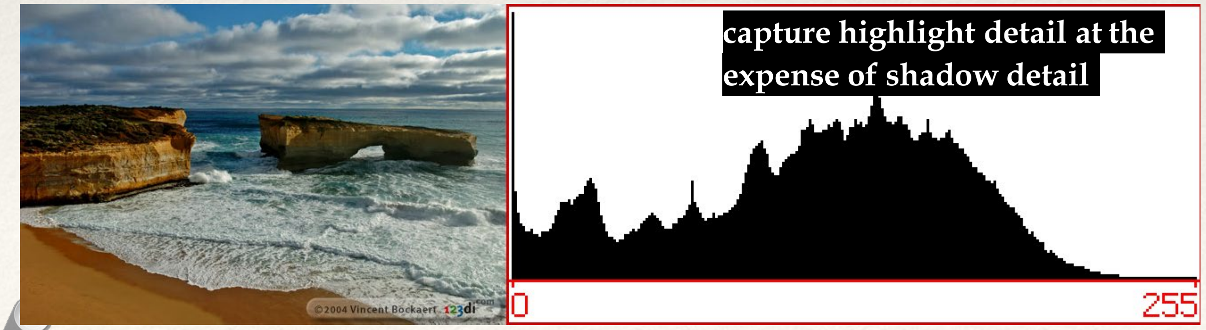


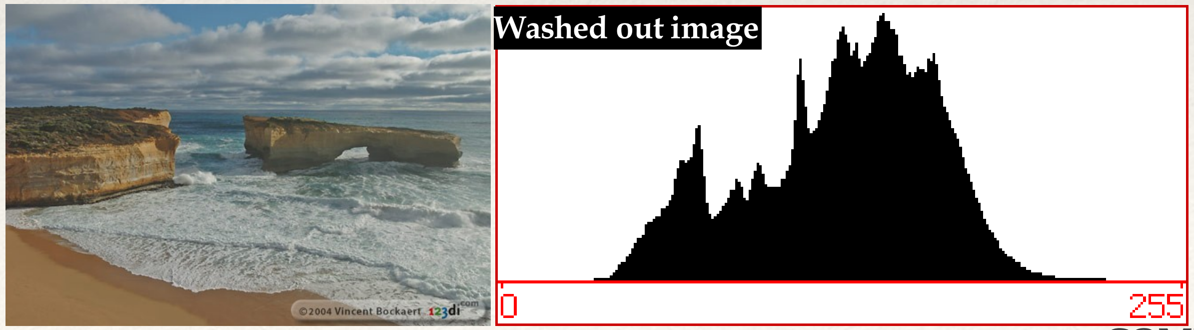
- ***Limited dynamic range***
  - Possible intensities of the input signal lie within a certain interval





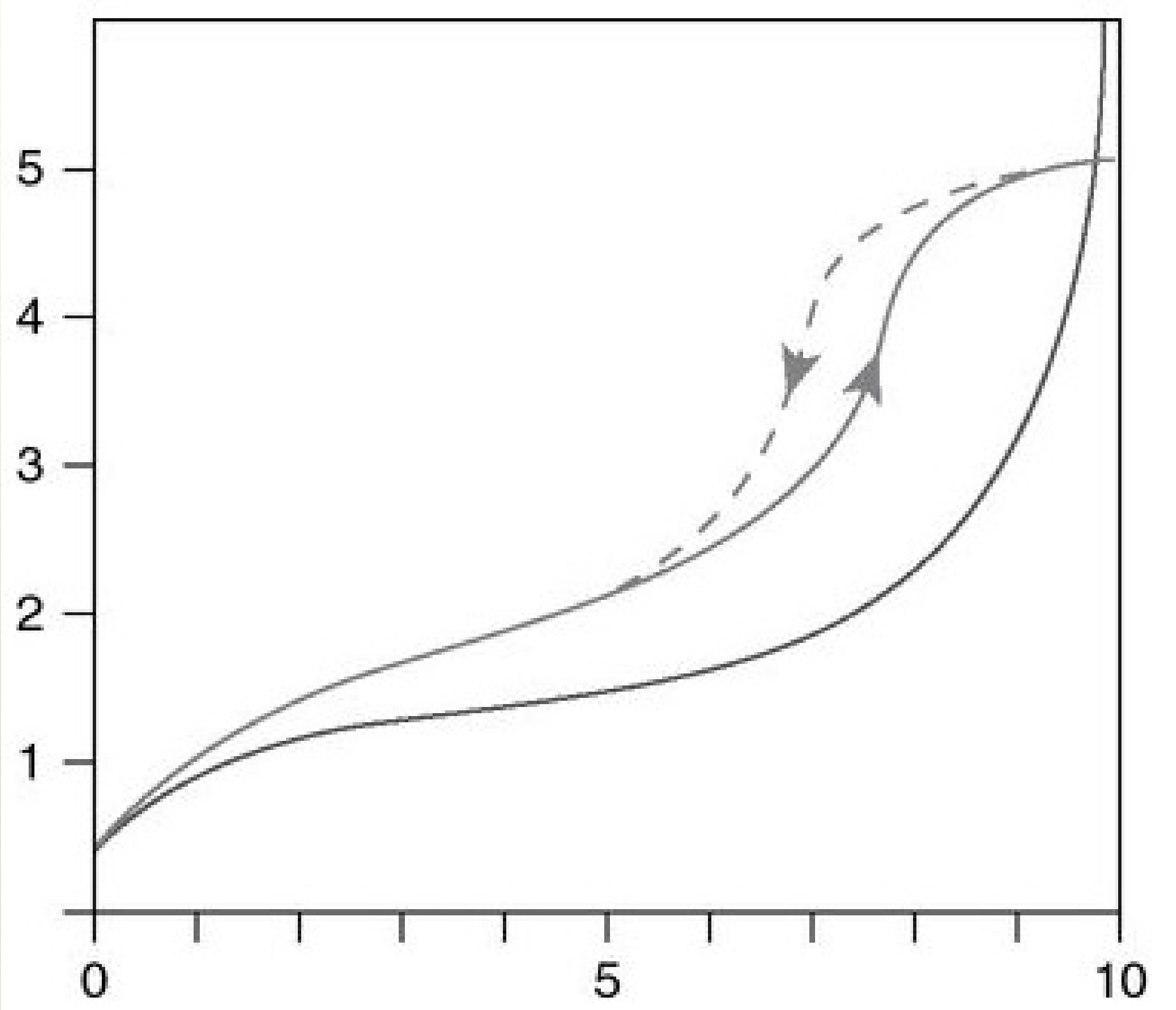






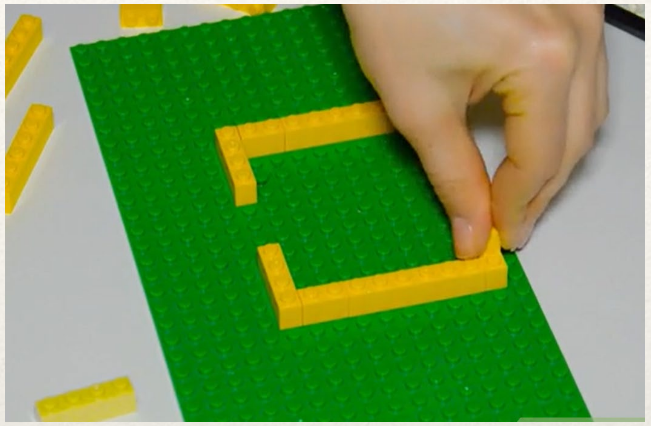
Sensors Limitations (Cont'd)

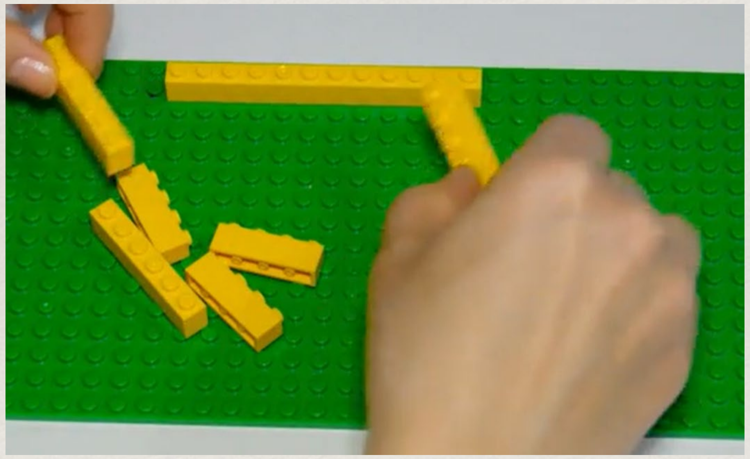
- ***Offset or bias***
  - The sensor output is not zero when the input is zero
- ***Nonlinearity***
  - A linear increase in input signal does not result in a linear increase in the sensor output
  - Some sensor can be tuned to behave linearly inside an ***operational range***
- ***Hysteresis***
  - Sensor Response Deviation over time

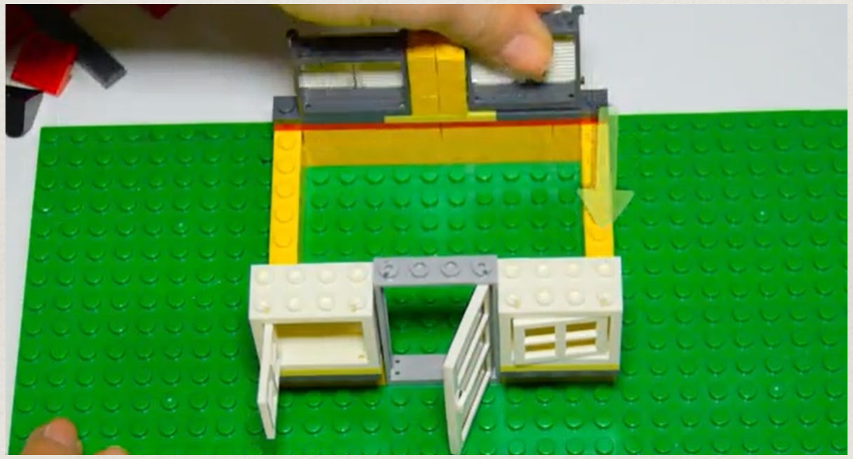


A general example for the hysteresis concept. The bottom curve is the sensor input, and the upper curve is the output. As the arrows show, the curve behaves differently when the measured entity decreases compared to when it increases.

How to build a LEGO House

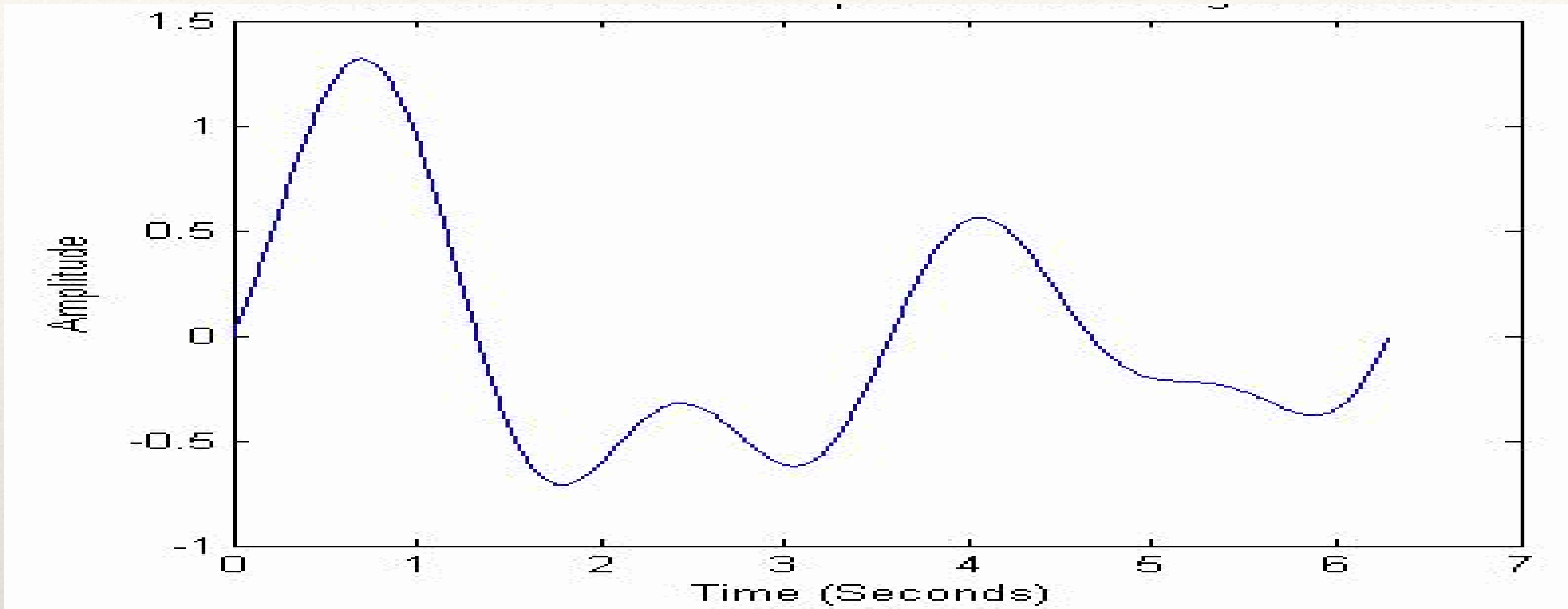




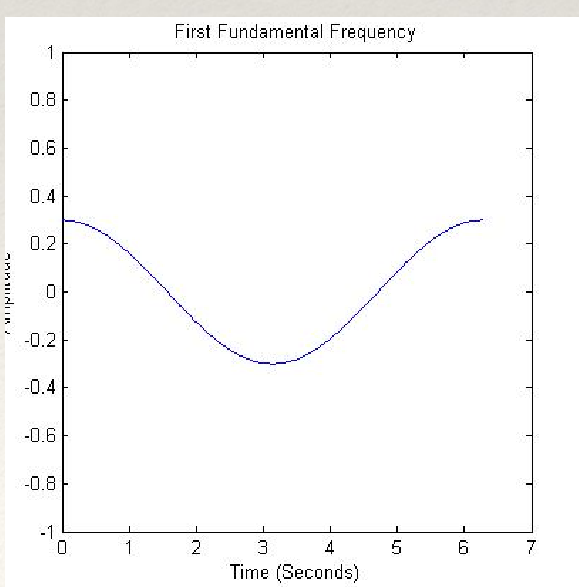


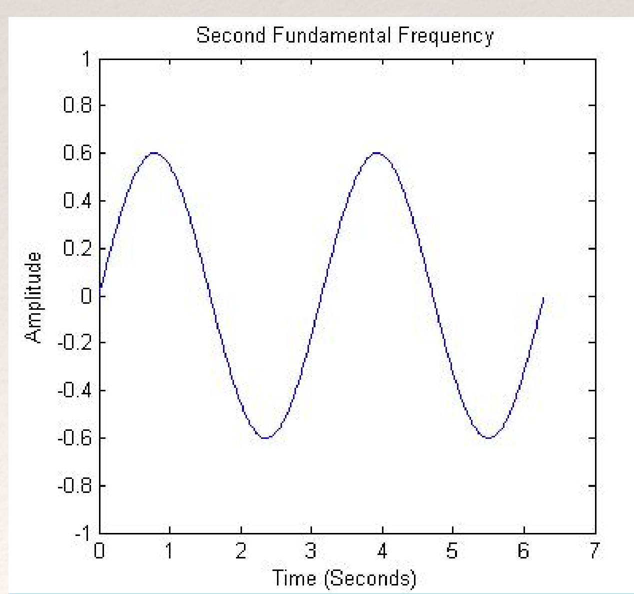
https://www.wikihow.com/Build-a-LEGO-House

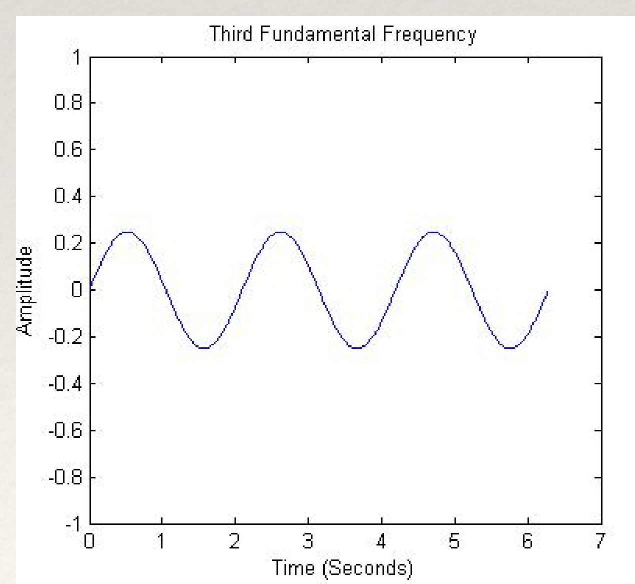
Sensor Output Analyze

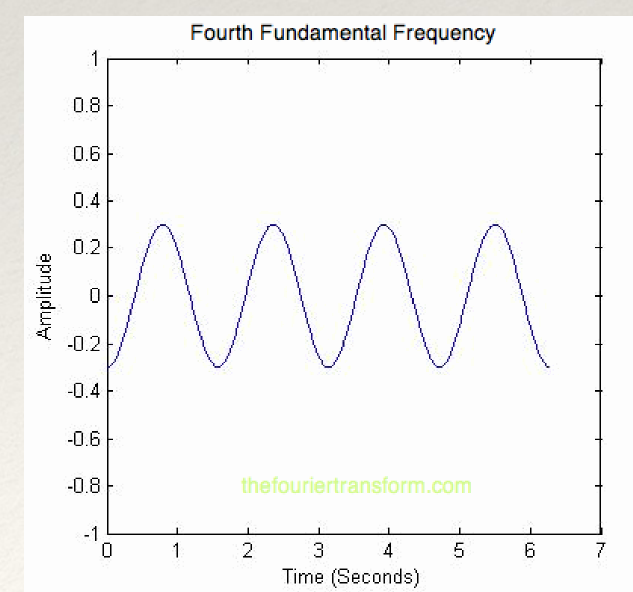


t

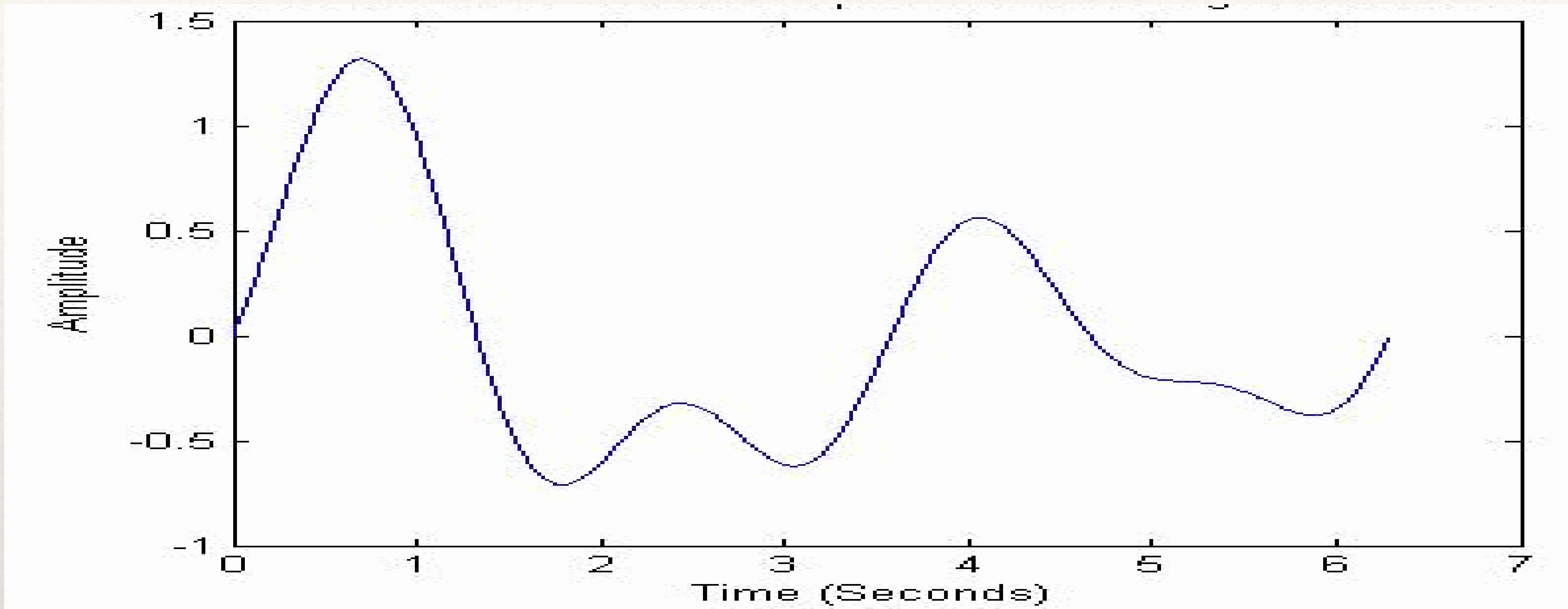




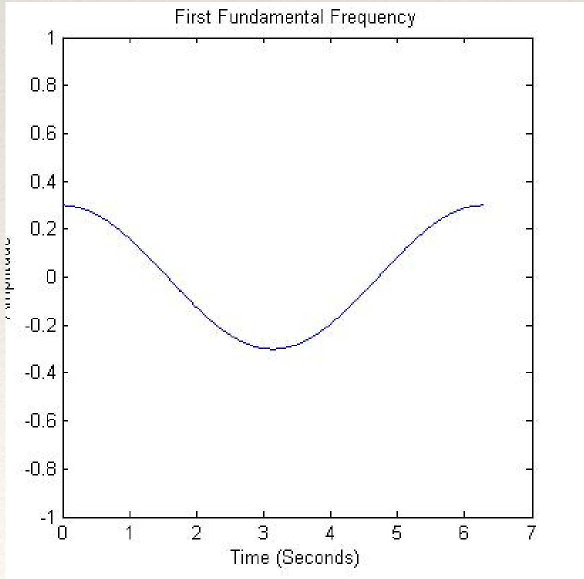


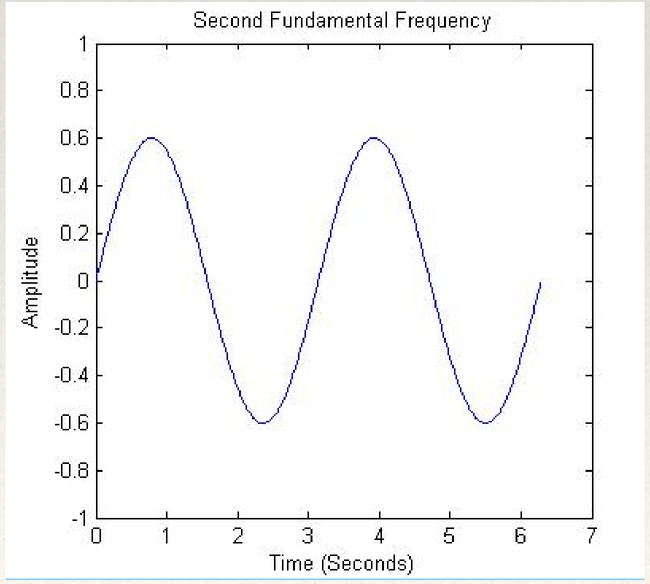


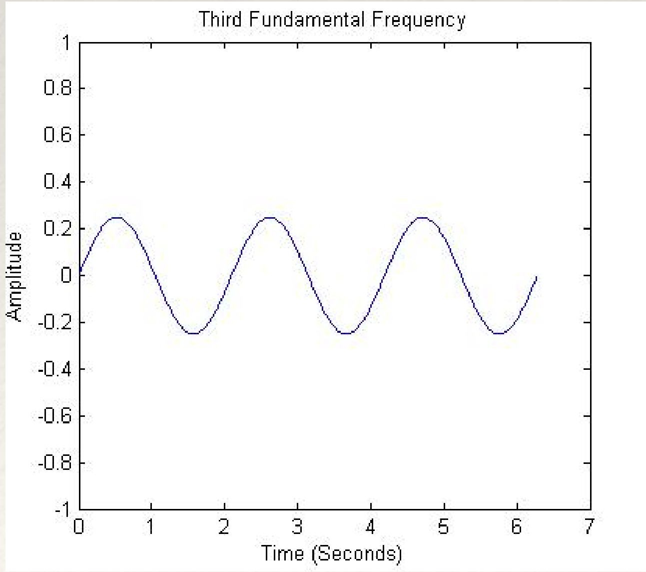
Sensor Output Analyze

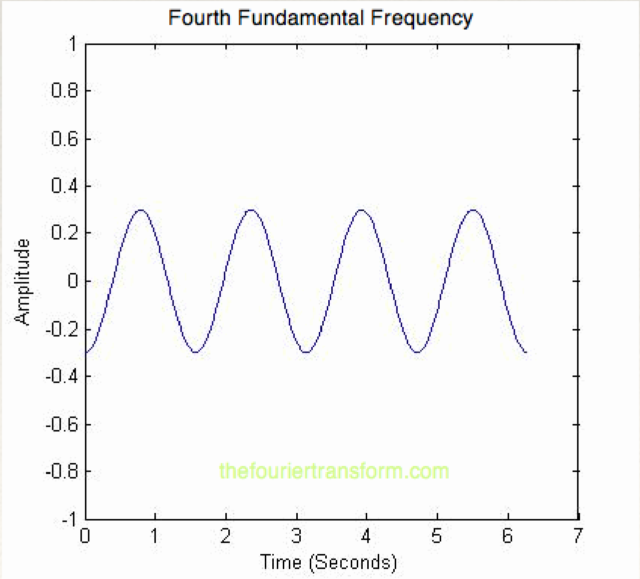


t



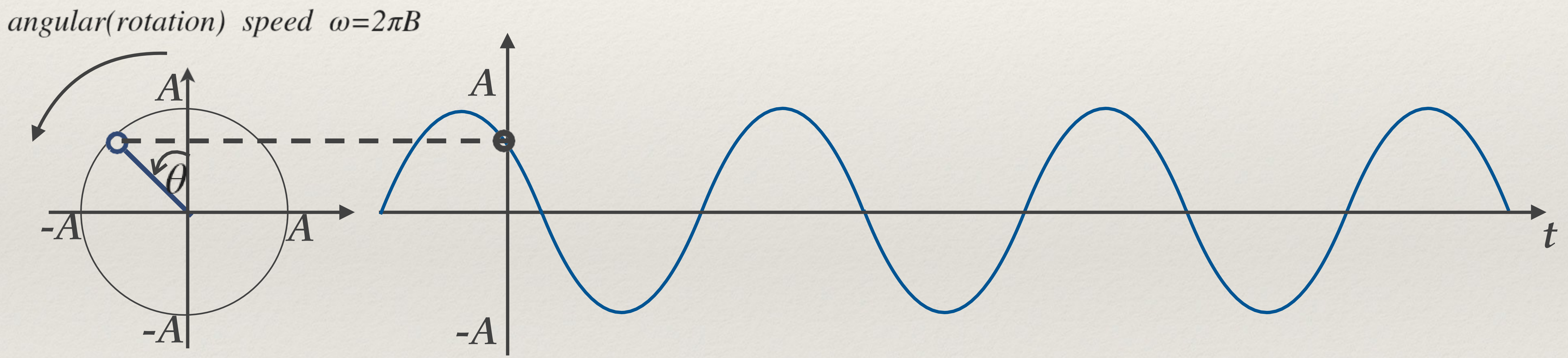






Sinusoid Signal

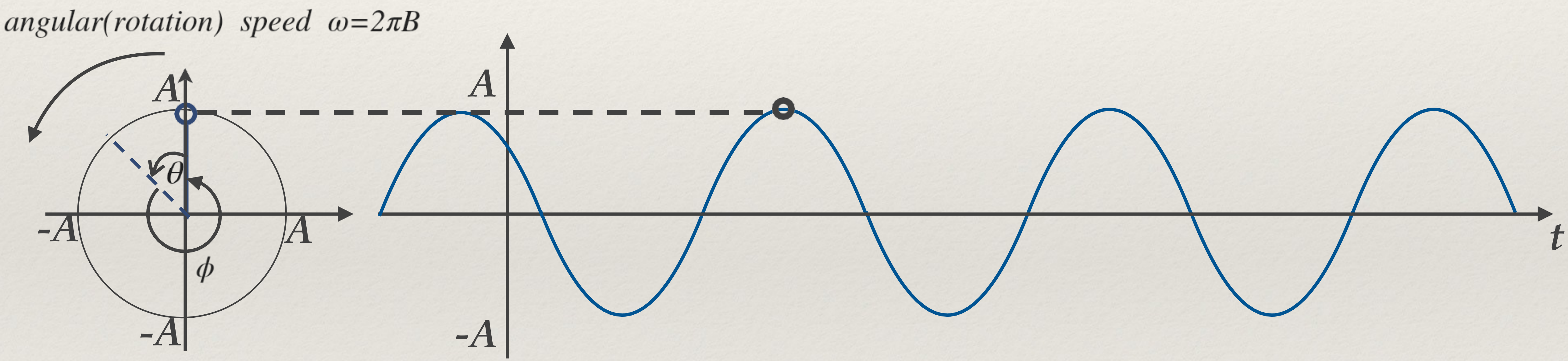
- A sinusoidal signal can be regarded as the projection on a plane of the path of a point moving around a circle at uniform speed



$$x(t) = A\cos(2\pi Bt + \theta)$$

Sinusoid Signal (Cont'd)

- A sinusoidal signal can be regarded as the projection on a plane of the path of a point moving around a circle at uniform speed



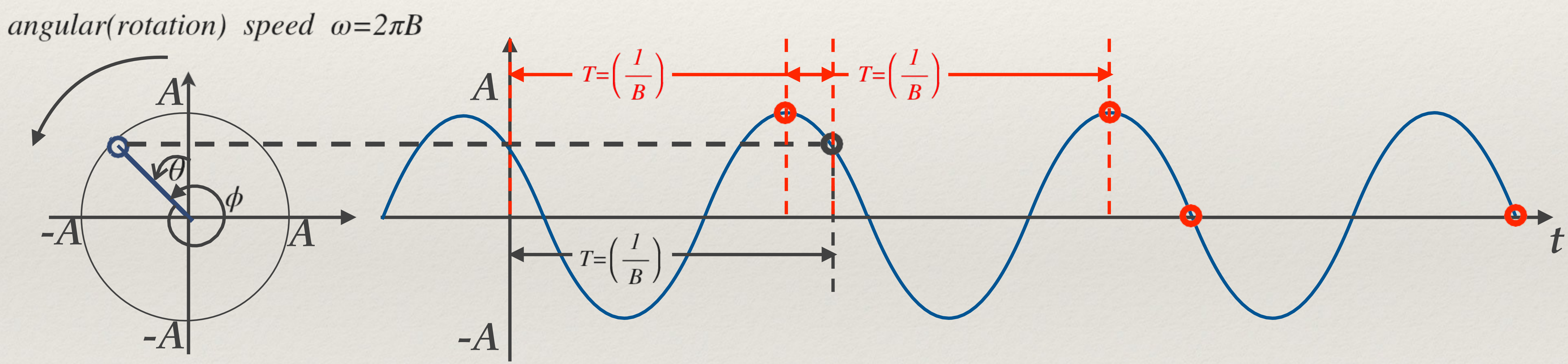
$$x(t+\Delta t) = A\cos(2\pi B(t+\Delta t)+\theta) = A\cos(2\pi Bt + (2\pi B\Delta t+\theta))$$

$$x(t+\Delta t) = A\cos(2\pi Bt + (\phi+\theta)),\quad \text{where}\quad \phi = 2\pi B\Delta t$$

$$\text{when}\quad \Delta t = \frac{2\pi-\theta}{2\pi B},\quad \text{then}\quad \phi+\theta = 2\pi$$

Sinusoid Signal (Cont'd)

- A sinusoidal signal can be regarded as the projection on a plane of the path of a point moving around a circle at uniform speed



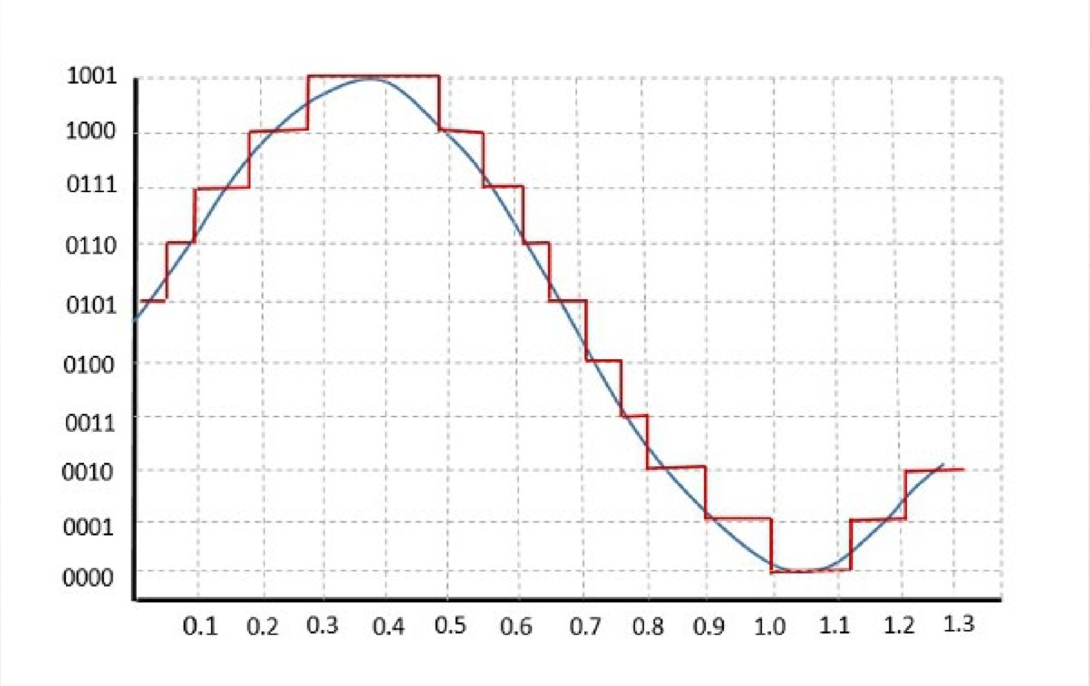
$$x(t+\Delta t) = A\cos(2\pi B(t+\Delta t)+\theta) = A\cos(2\pi Bt+(2\pi B\Delta t+\theta))$$

$$x(t+\Delta t) = A\cos(2\pi Bt+(\phi+\theta)), \quad where \quad \phi = 2\pi B\Delta t$$

$$when \quad \Delta t = \frac{2\pi+\theta-\theta}{2\pi B} = \frac{1}{B} = period, \quad then \quad \phi = 2\pi$$

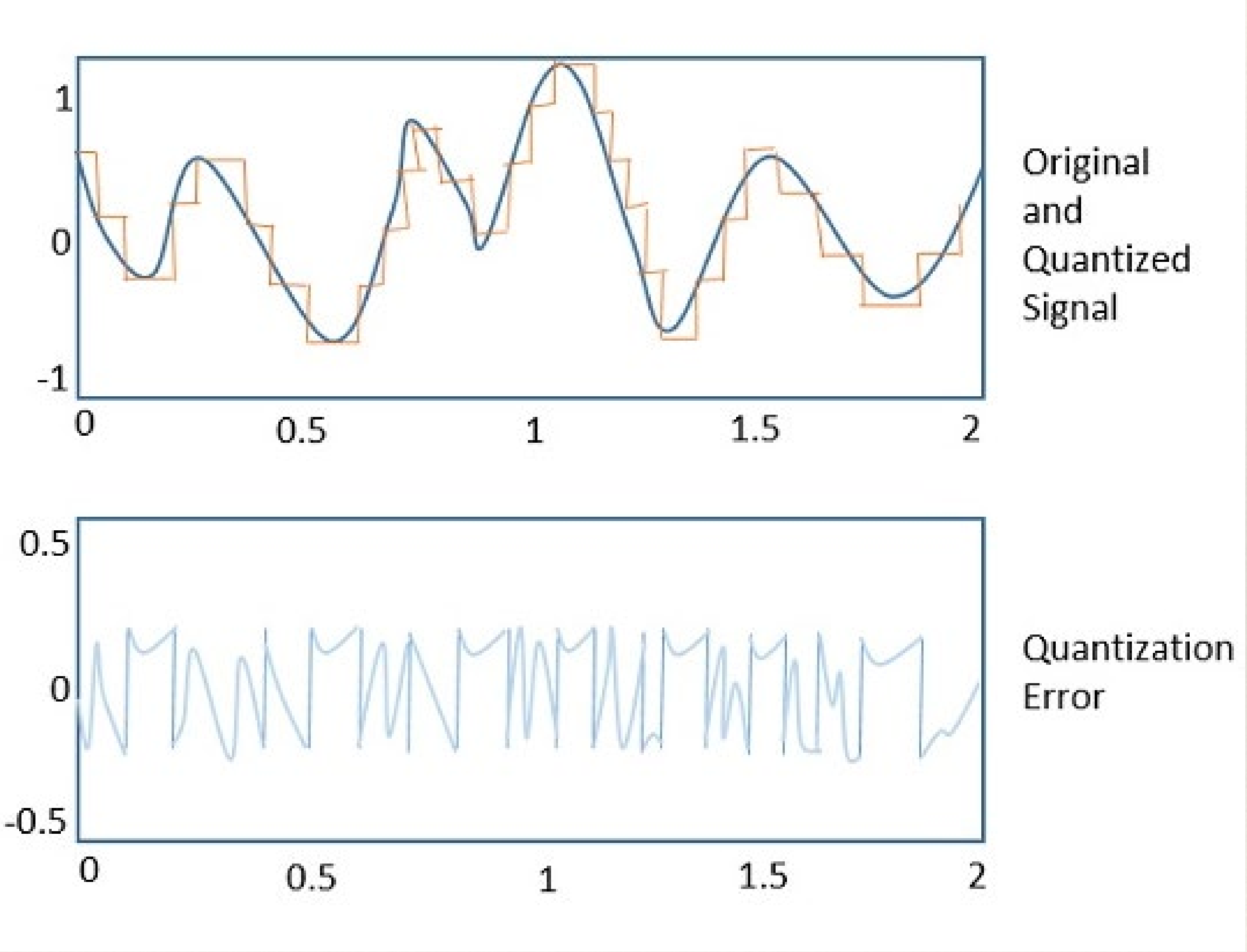
Digitization

- Sensor outputs continuous electrical signal (voltage/current), and these signal needs to be digitised for computer to operate
- Digitization is a process for converting continuous signal into a discrete set of its samples
  - It happens in both the time and signal values
  - In a *sampling* step, the analog signal is captured at regular time intervals (the frequency of taking sample is called sampling rate)
  - Each reading is called a sample, and each sample is rounded to a fixed set of integer numbers
    - This is called ***quantization*** process
- Digitized sample values are much less susceptible to distortions than analog signals



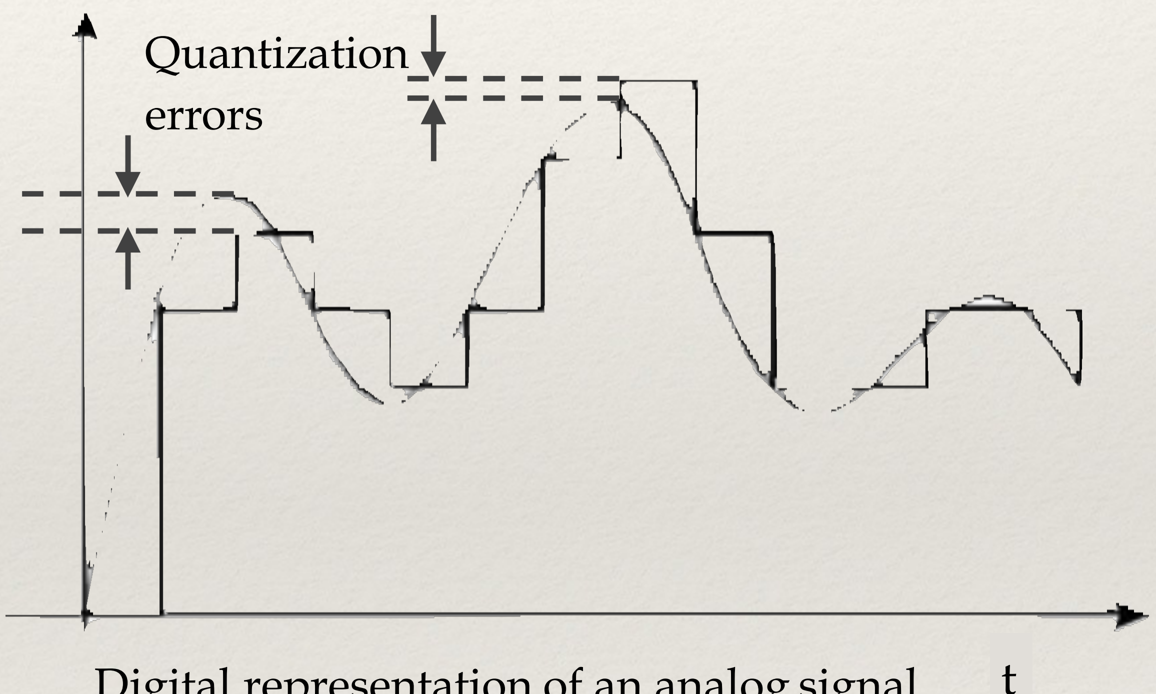
Digital representation of an analog signal.
Both amplitude and time axis are discretised.

Digitization (Cont'd)



Digitization (Cont'd)

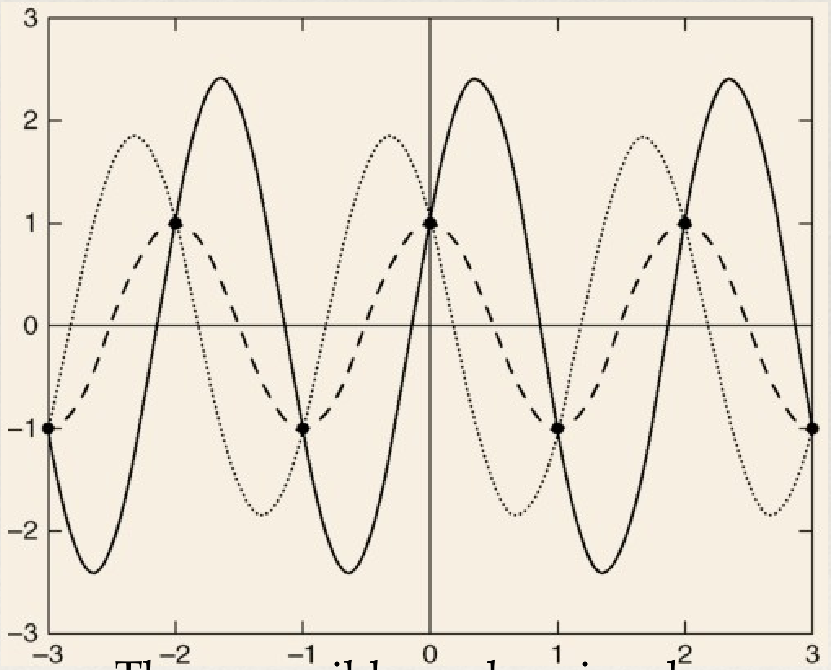
- The error the quantisation introduces is the ***quantization noise***
  - It affects how accurately the amplitude can be represented
  - Few bits ==> Coarse representation
  - More bits ==> Finer representation
- For audio and video, eight, sixteen, twenty-four and thirty-two bits per sample are used
- The error the sampling rate introduces is the ***discretization error***
  - It determines the maximum frequency that can be represented in the signal
  - The maximum frequency is called ***Nyqvist Frequency***, which is half the sampling frequency (rate) *fs*
- If a function $x(t)$ contains no frequencies higher than B hertz, $x(t)$ is completely determined by its ordinates at a series of points spaced $1/(2B)$ seconds apart



Digital representation of an analog signal.
Both amplitude and time axis are discretised.

Nyquist Frequency

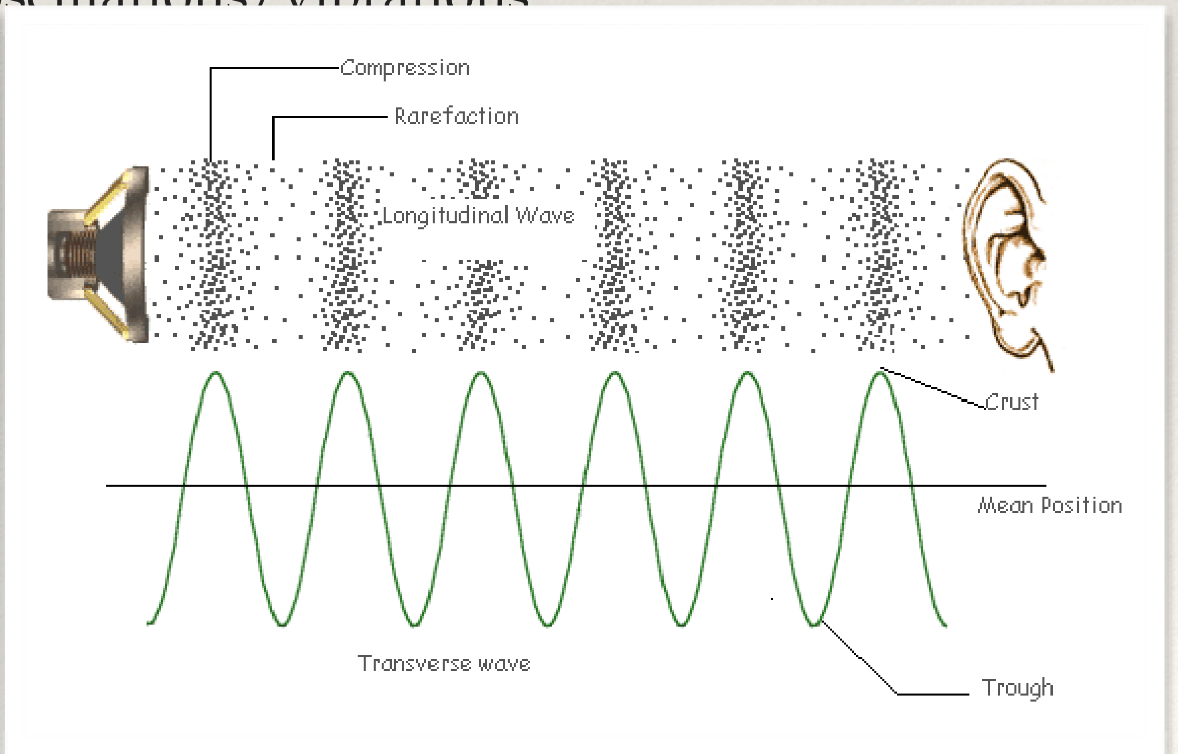
- Necessity of *fs* > 2*B*
  - Consider the sinusoid
    - $x(t) = A\cos(2\pi Bt + \theta) \equiv \cos(2\pi Bt)\ \cos(\theta) - \sin(2\pi Bt)\ \sin(\theta)$
  - With *fs* = 2*B* (i.e. *T* = 1/(2*B*)), the samples are given by
    - $x(nT) = \cos(\pi n)\ \cos(\theta) - \cancelto{0}{\sin(\pi n)}\ \sin(\theta) = \cos(\pi n)\ \cos(\theta)$
  - which render samples that are indistinguishable from the samples of
    - $y(t) = \cos(2\pi Bt)\ \cos(\theta)$



Three possible analog signals for the same sampling points

Physical Properties of Sound

- Sound
  - "A travelling (longitudinal) wave which is an oscillation of pressure transmitted through a solid, liquid, or gas, composed of frequencies within the range of hearing and of a level sufficiently strong to be heard, or the sensation stimulated in organs of hearing by such vibrations."
- Transmission
  - Propagate in physical medium by mechanical oscillations / vibrations
  - Cannot travel through vacuum
- Speed of sound is
  - Mach 1 or 343 ms^{-1} in dry air at 20°C
  - 1,482 ms^{-1} in water at 20°C
- Human hearing frequency range
  - 12Hz ~ 22kHz



Reference: http://www.mediacollege.com/audio/01/sound-waves.html

Sound Pressure Level(SPL)

- Sound Pressure Level (SPL), or L_p , is defined as
  - $$L_p = 10\log_{10}\left(\frac{p^2}{p^2_{ref}}\right) = 20\log_{10}\left(\frac{p}{p_{ref}}\right)\ dB$$
  - Measurement on logarithmic scale using decibels
  - where p is the RMS sound pressure and p_{ref} is a reference sound pressure
  - Typical reference sound pressure for silence is 20 µPa in air and 1 µPa in water
    - p_{ref} is often considered as the threshold of human hearing (roughly the sound of a mosquito flying 3 m away)
- The sound pressure perceived by human ear is non-linear
- Human ear, does not react equally to all frequencies
  - The same sound pressure at a different frequency will be perceived as a different volume level

A-weighting scheme

- Sound-pressure measurements are often frequency weighted to match perception. A-weighting is applied to instrument-measured sound levels in an effort to account for the relative loudness perceived by the human ear, as the ear is less sensitive to low audio frequencies. It is employed by arithmetically adding a table of values, listed by octave or third-octave bands, to the measured sound pressure levels in dB. Sound-pressure levels weighted by A-weighting scheme are usually labeled as **dBA** or **dB(A)**

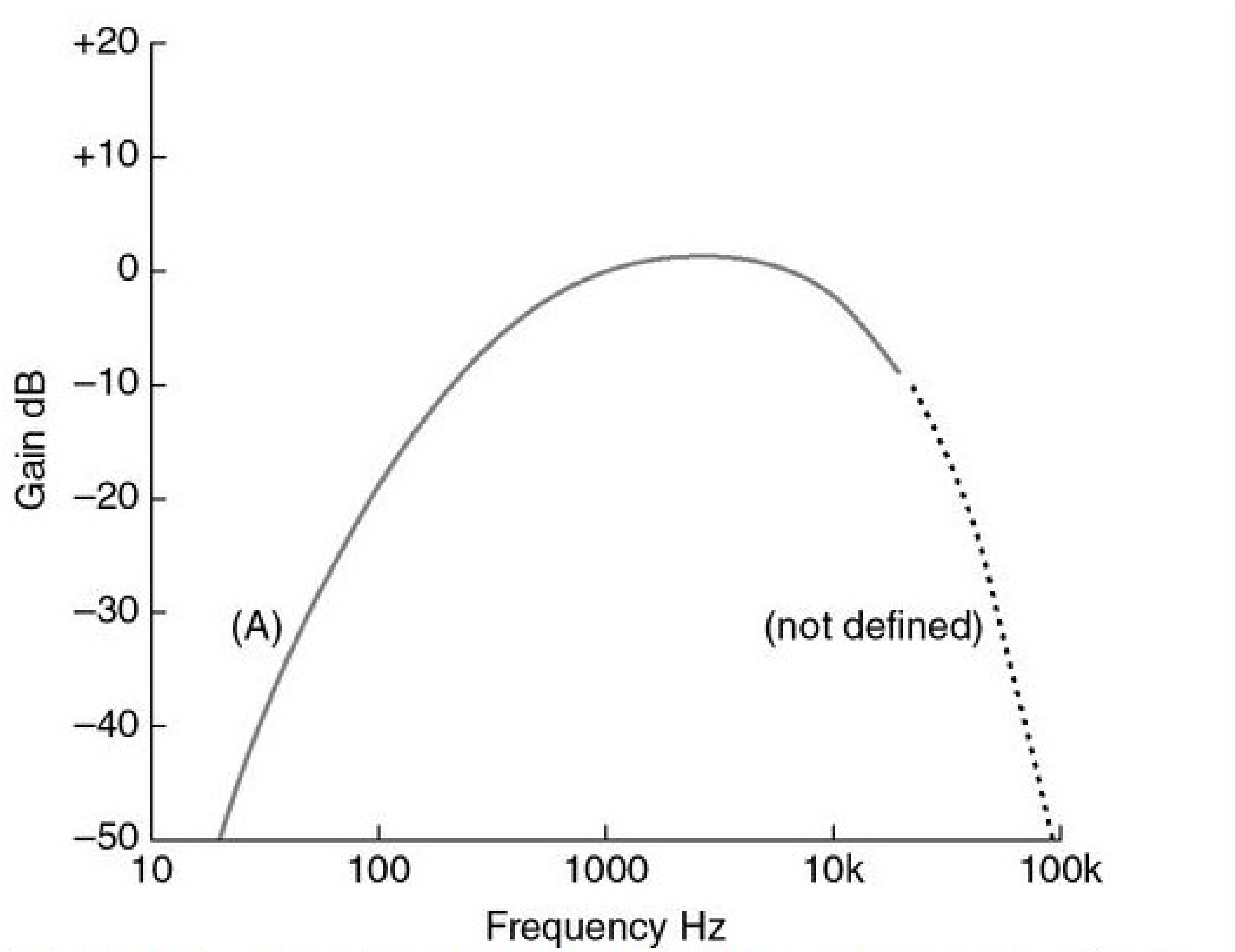
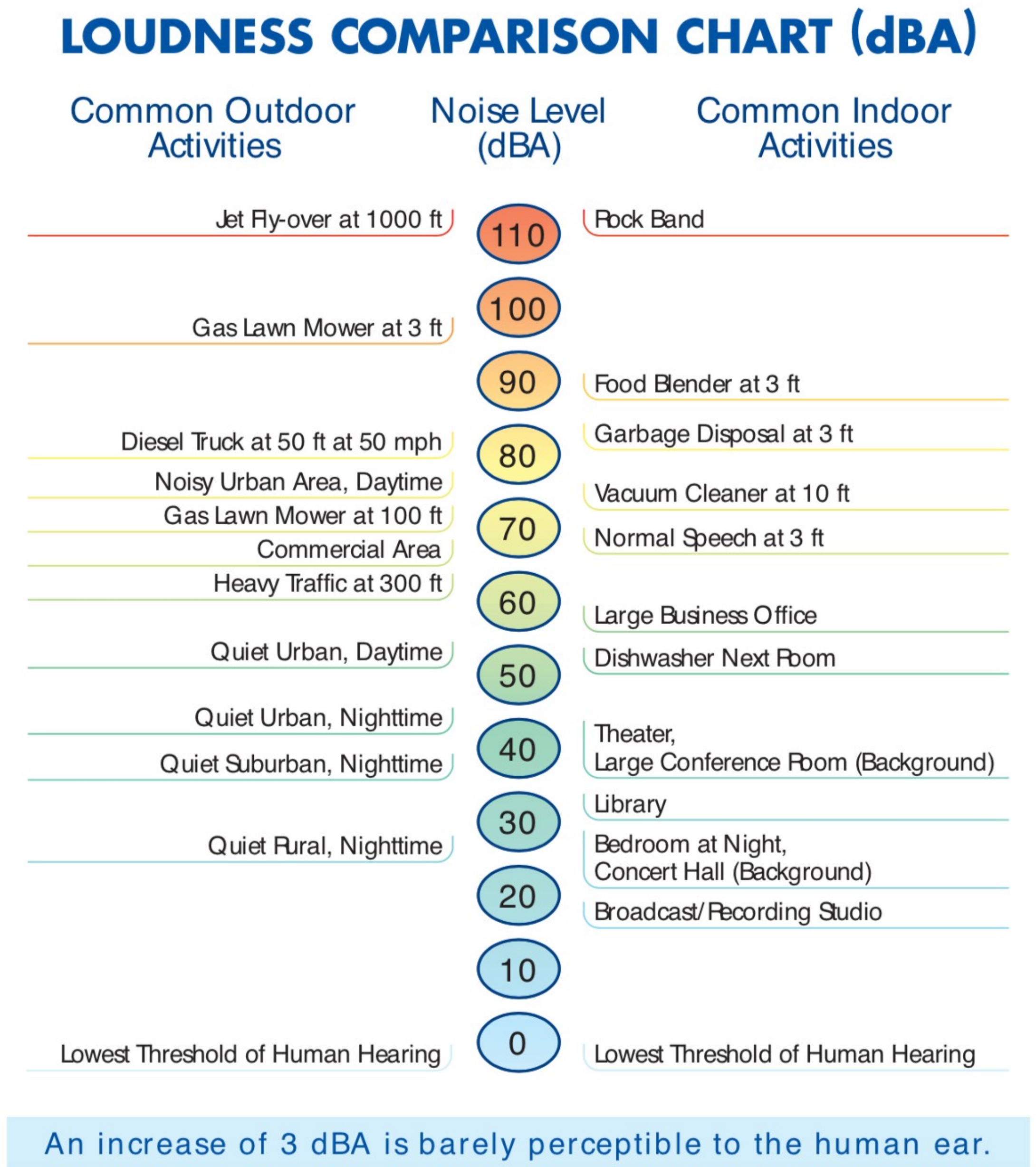


Figure 5.1. Sound pressure A-weighting scheme according to IEC 61672:2003.



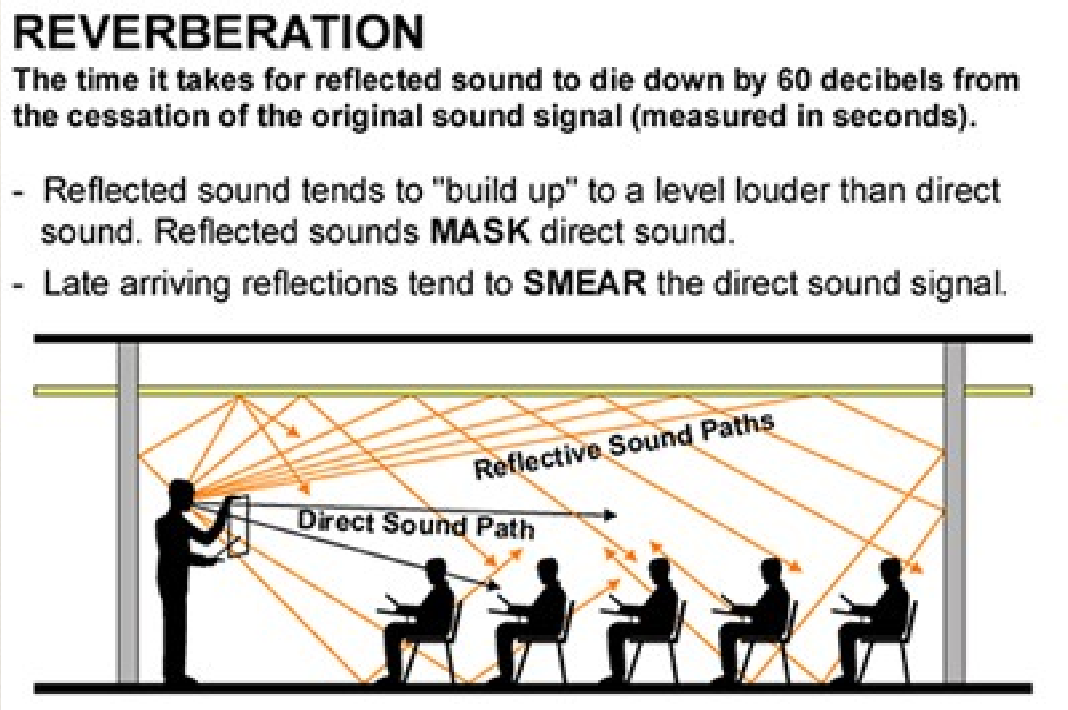
Ref: https://johnsanzone.com/2017/10/31/loudness-comparison-chart-decibels-dba/

Undesirable Sound Effects

- Echo, Reverberation, and Interference make automatic processing and recognition of sounds difficult
- Echo
  - Echo is a reflection of sound, perceived by the listener some time after the original
    - e.g. Echoes produced by the bottom of a water well, or by mountain enclosures
    - Echoes are present in every environment because most materials easily reflect sounds
  - Human ear cannot distinguish between the echo and the original sound if the delay is less than 1/10 secs
    - But a machine-based system trying to recognise sound might still have an issue with these echoes

Reverberation

- Reverberation is created when a sound is produced in an enclosed space
    - Numerous echoes build up and slowly decay as the environment absorbs the sound
- Multimedia content analysis techniques often suffer from not accounting for reverberation, even when it is inaudible



Reference: http://www.acousticalsurfaces.com/acoustic_IOI/reverberation.htm

Interference

- Interference is the superposition of two or more waves that results in a new wave pattern
- Destructive interference can be useful for certain scenario, can you think of any?

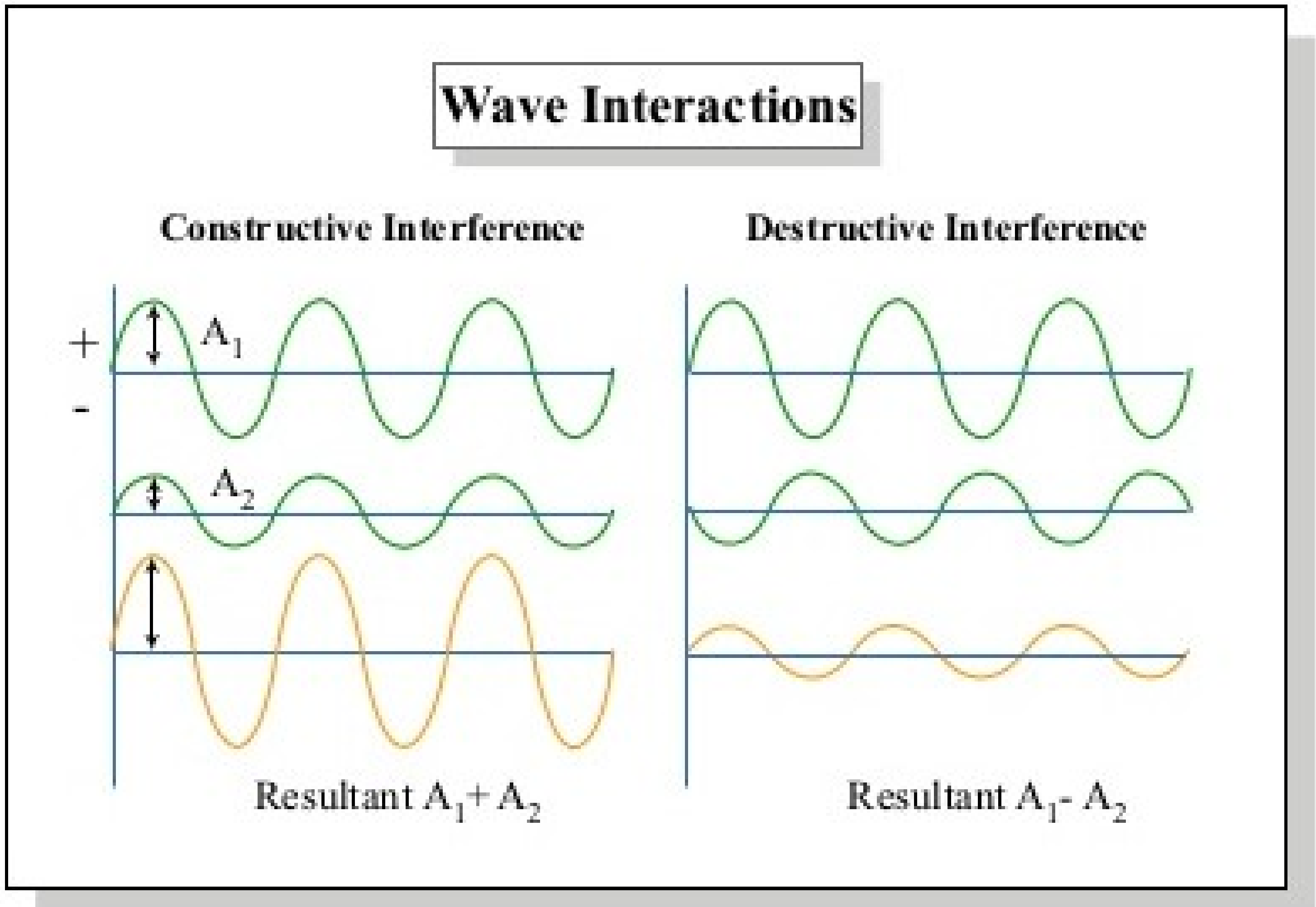
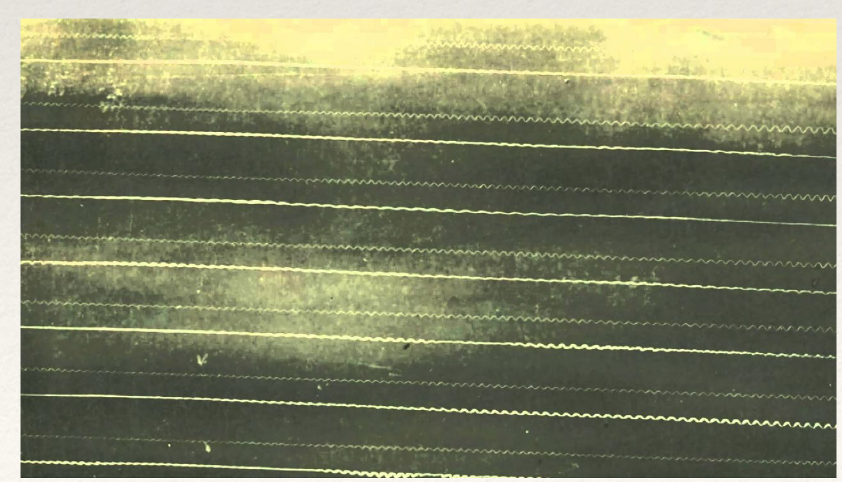


Figure by MIT OCW.

Primitive form of Recording and Reproduction of Sound

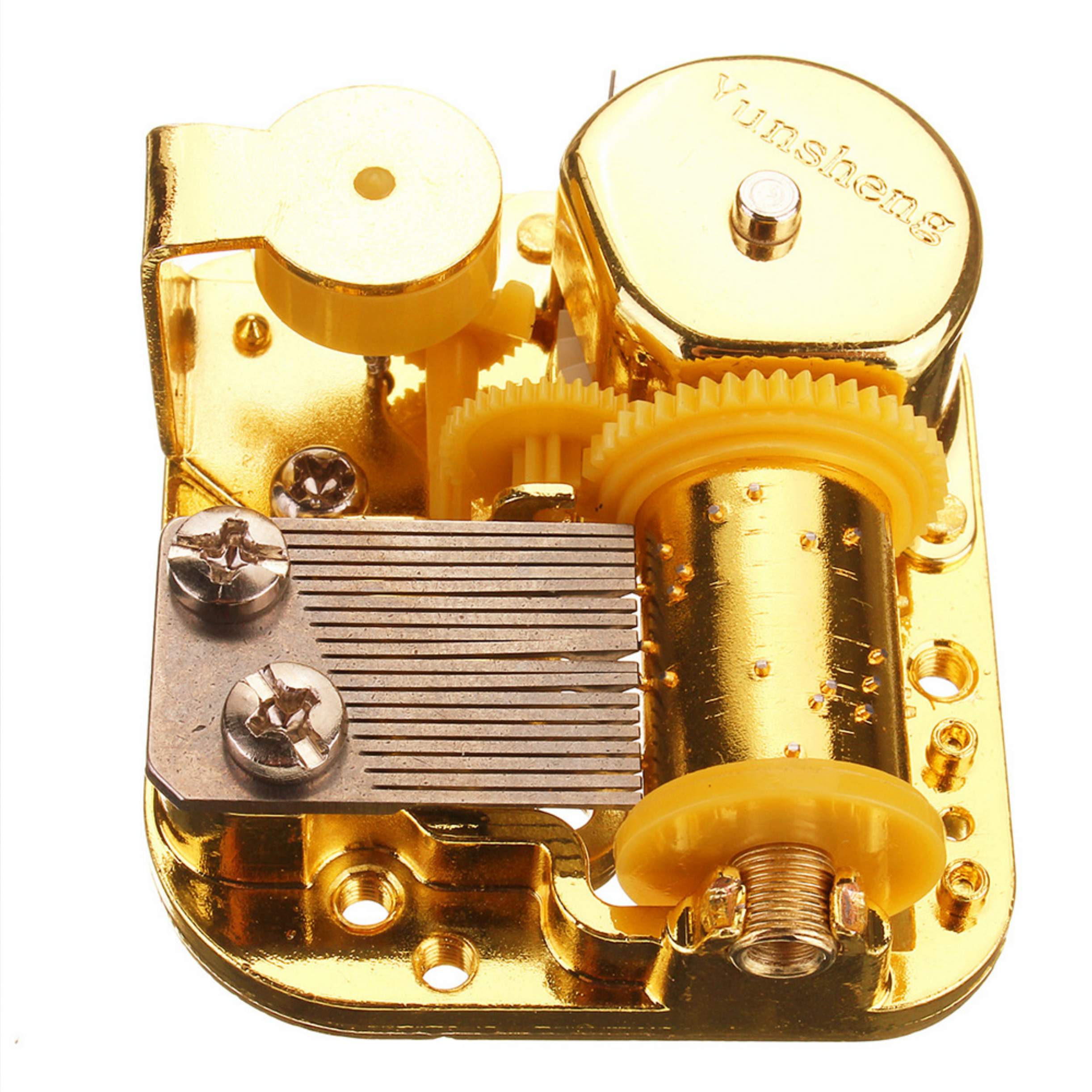
- Sound pressure variation are captured and mechanically engraved on the outside surface of a strip of tinfoil wrapped around a rotating cylinder
- To playback, a needle ran along the cylinder, to convert mechanical engravings into sound saves that would be mechanically amplified

Reference: http://www.audiohistory.com/files/documents/Firsts_NYT.htm



Primitive form of Recording and Reproduction of Sound (Cont'd)

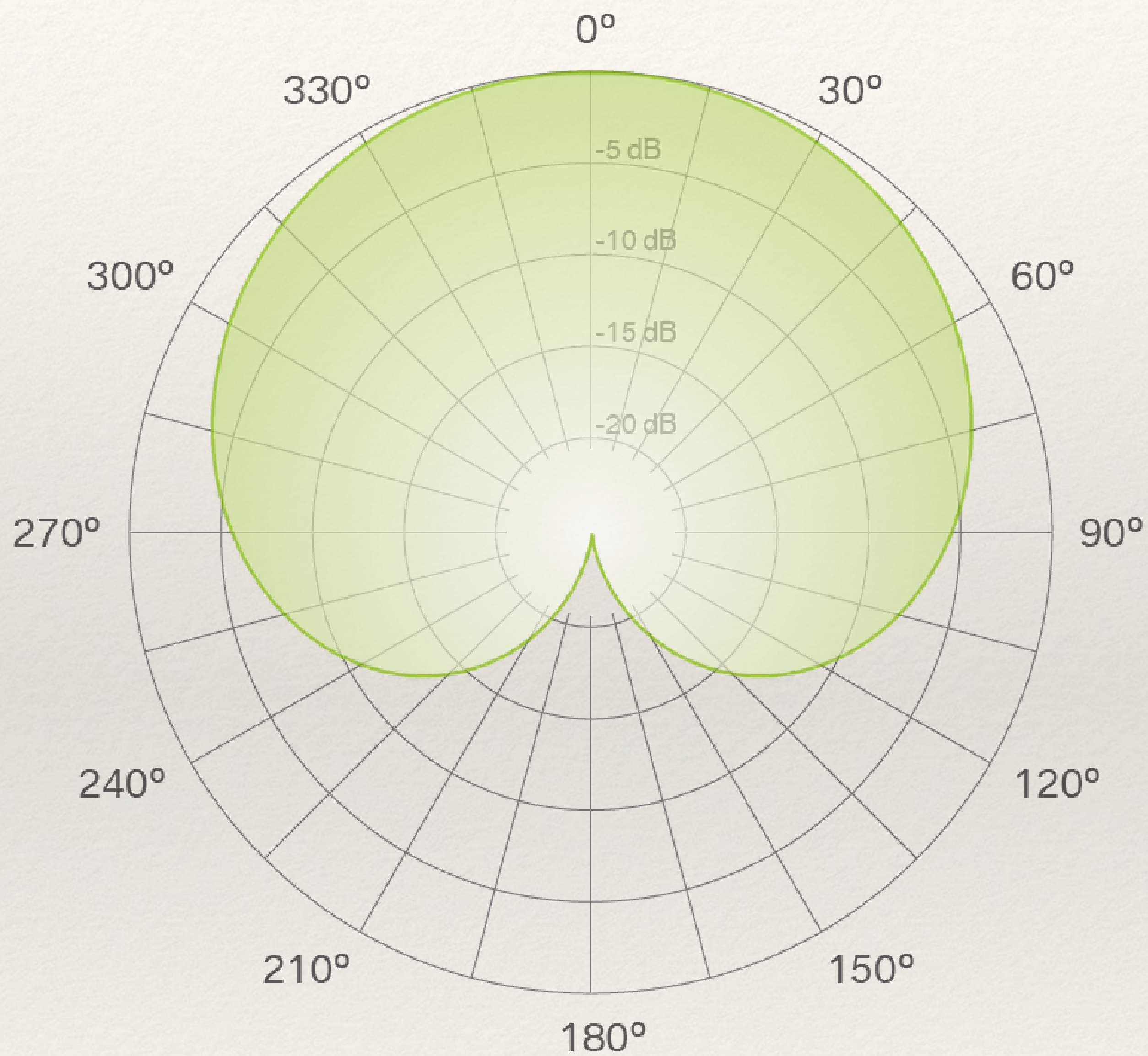
Eighteen-note music box



Microphones

- Microphone is an acoustic sensor that converts sound into an electrical signal
  - A membrane is exposed to sound pressure, which create electrical properties variations (e.g. resistance, voltage, capacitance, etc.)
- Directionality
  - This is an important characteristic that indicates the sensitivity to pressure waves arriving at different angles
  - It is usually represented by a polar pattern, visualising the locations that produce the same signal level output in the microphone if a constant sound-pressure level is generated
    - In real world, polar patterns are a function of frequency

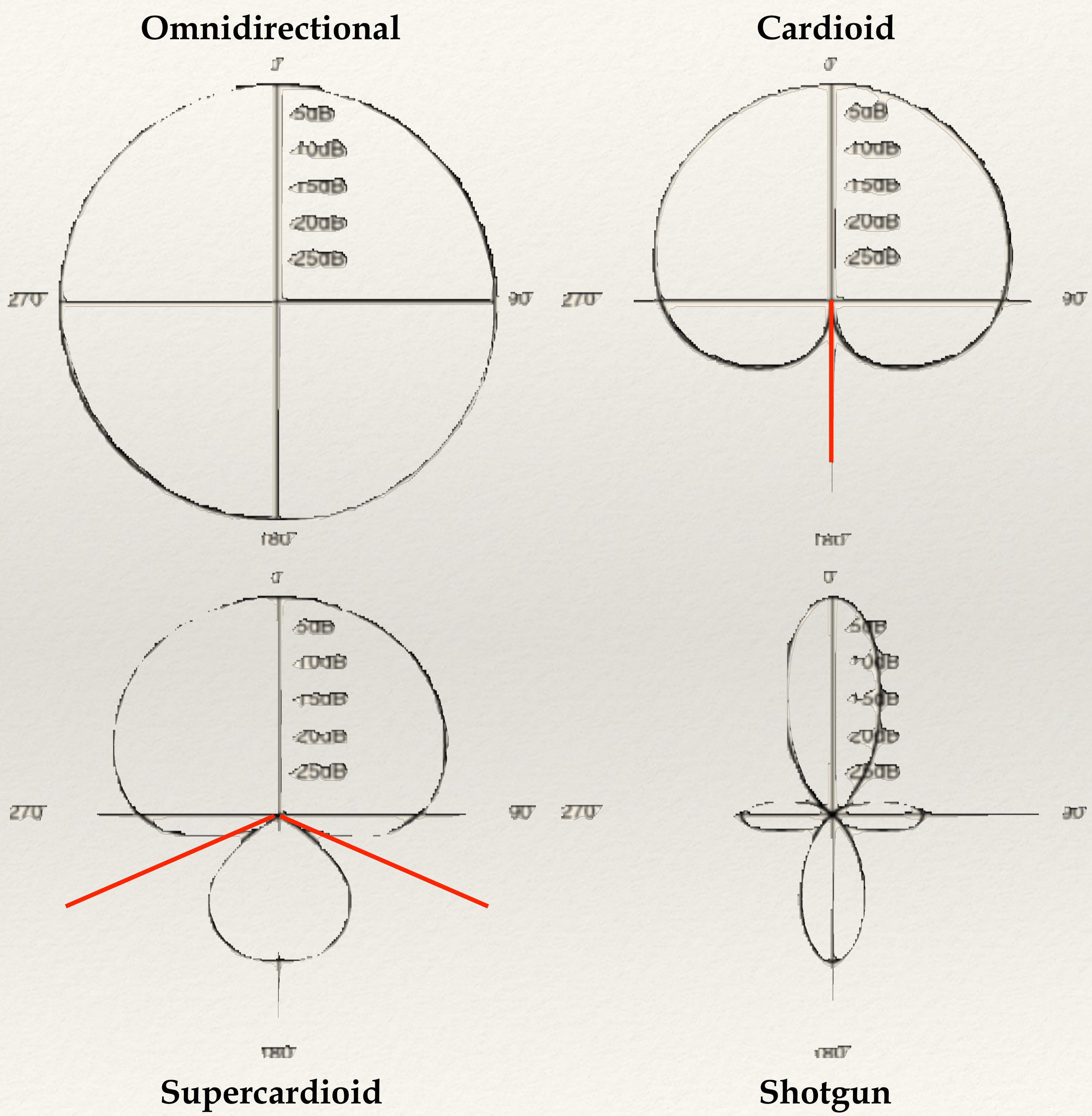
Polar Patterns



Cardioid ref: https://www.lewitt-audio.com/blog/polar-patterns

Polar Patterns

- Omnidirectional
    - Omni mics are equally sensitive to sound arriving from all angles
    - Good for studio environments without noise
    - Bad in rejecting background noise
- Cardioid
    - Most commonly used directional polar pattern
    - Maximum rejection at 180 degrees off-axis
- Supercardioid
    - Narrower front pick angles than cardioid
    - Maximum rejection at 120 degrees
- Shotgun
    - Highly Directional Microphones



Mic Frequency Response

- The electrical voltage output from microphone should be ideally directly proportional to the sound pressure
- In reality, sound recording has a linear area for certain sound pressure levels and frequency ranges only

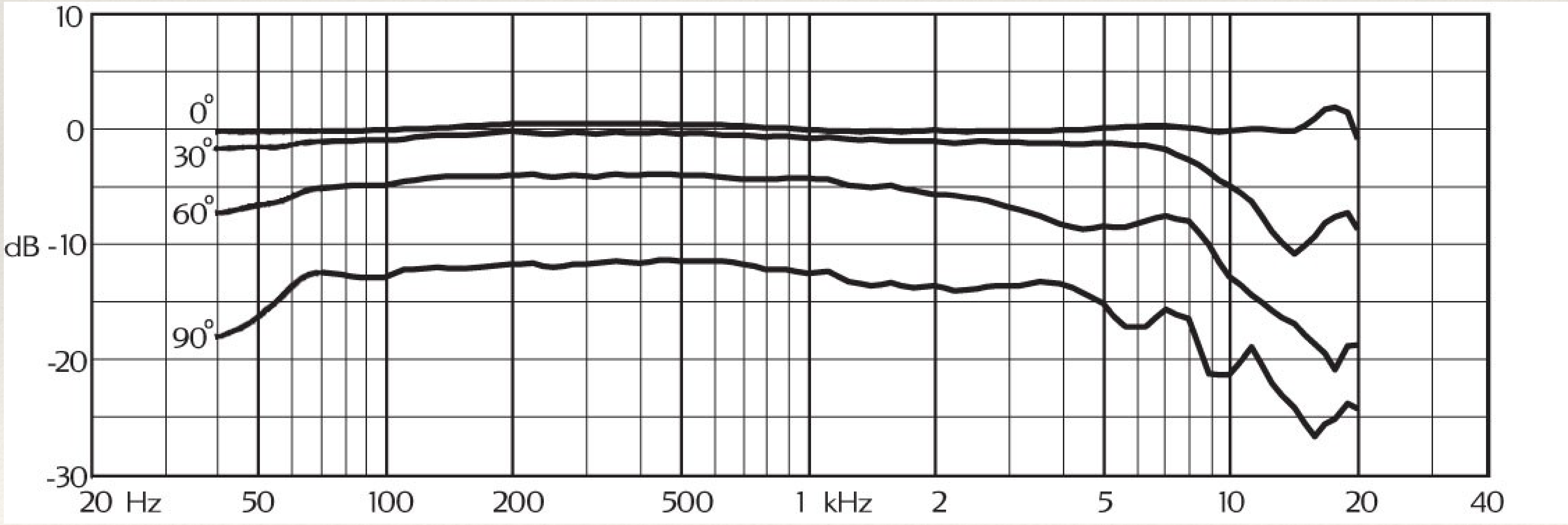


Figure adopted from http://www.dpamicrophones.com/en/products.aspx?c=Item&category=233&item=24402#specifications

Microphone Output

- Analog Signal Output
  - Microphone's analog output is feed to an amplifier before it is digitised
  - If the amplifier gain is not set properly
    - Signal can be overdriven
      - Signal with high amplitude can be clipped out, leading to distortion
    - On the other hand
      - Signal with small amplitude can zero out after A/D conversion

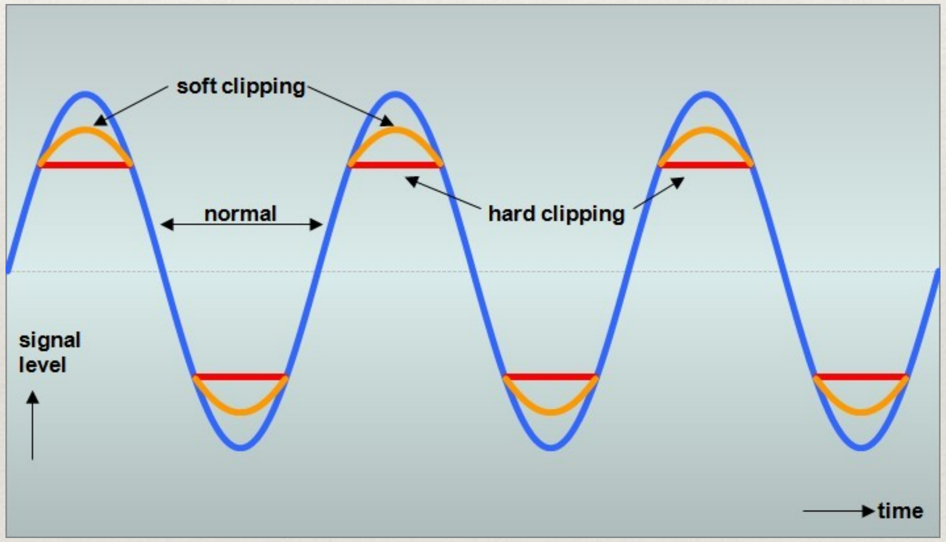
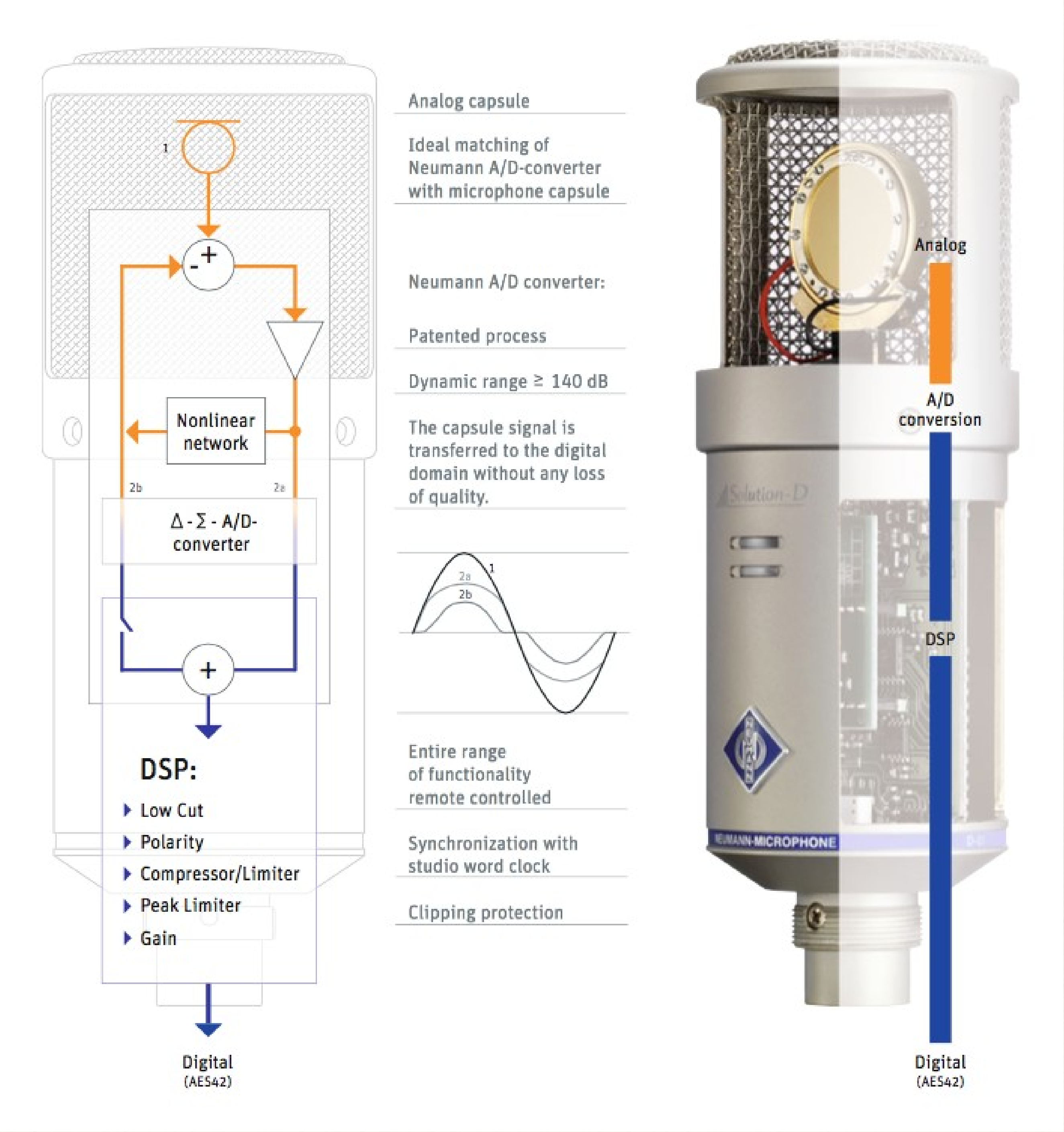


Image adopted from http://guitargear.org/2009/11/05/overdrive-vs-distortion/

Microphone Output (Cont'd)

- AES 42 Digital Output
  - No interference problem
  - Configurable Settings
    - Polar-pattern
    - Low-cut filter
      - 40Hz, 80Hz, 120Hz
    - Gain
      - Unity to +63dB
    - Mute
- For more details, please refer to
  - http://www.taperssection.com/reference/pdf/AES%20-%20Digital%20Microphones%20-%20AES42%20and%20all%20that.pdf



Reference: https://www.neumann.com/img/Linkgraphics/Solution-D_E.pdf

Human Audible Frequency Range

- Human Auditory System
  - Hearing range is from 20 Hz to 22 kHz
- Professional audio-recording equipment usually employs sampling frequencies of at least 44 kHz
  - Though 48 kHz and 96 kHz might also be used, even if the human ear cannot perceive them
  - Digital filtering and machine learning, might use these higher frequencies