Digital Data Representation

Topics for the next few weeks

Representing Numbers

Representing Text

Representing Audio

Representing Images

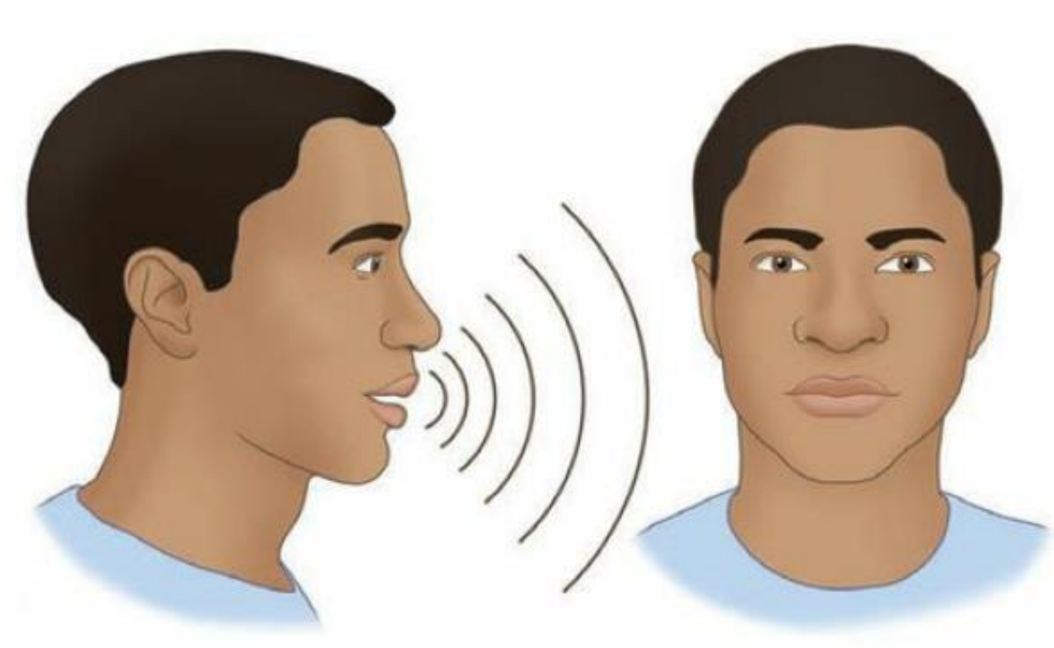
Representing Video

Digital Media

Representing audio, images, and video
in a computer system

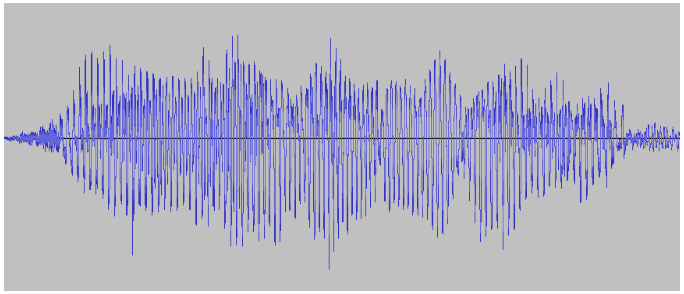
Audio

Can you hear me now?



We perceive sound when a series of air compressions vibrate a membrane in our ear, which sends signals to our brain.

Audio as a Sound Wave



- Microphone is a transducer, converting audio (sound waves) to electrical signals. sometimes immediately digitizing it

A **transducer** is an electronic device that converts energy from one form to another. Common examples include microphones, loudspeakers, thermometers, position and pressure sensors, and antenna

Audio as a Sound Wave

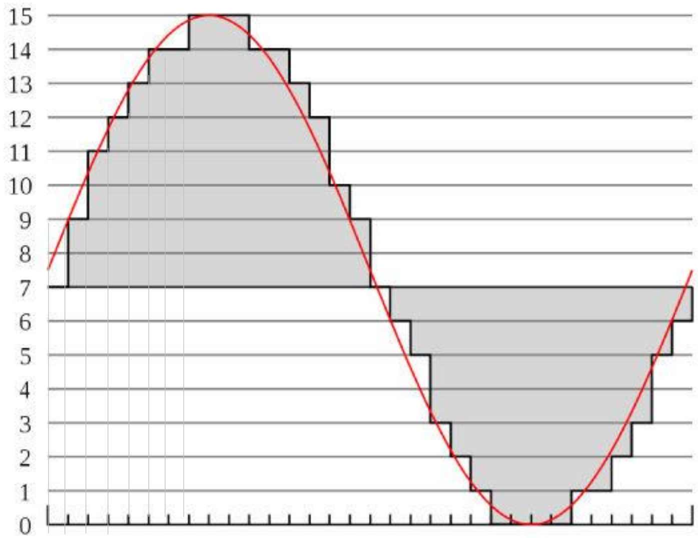
Sound waves (analog) → Electrical signals (digital)

- Microphone is a transducer, converting audio (sound waves) to electrical signals

A **transducer** is an electronic device that converts energy from one form to another. Common examples include microphones, loudspeakers, thermometers, position and pressure sensors, and antenna

Ways of Creating Digital Audio

Sampling

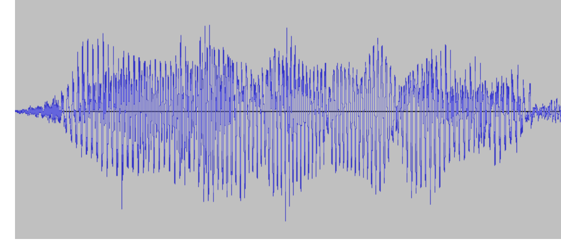


Each digitized sample of audio is assigned a value that corresponds to the amplitude of the analog wave.

Synthesis

Mathematic model
create artificial sound

Anatomy of Audio:
Sine Wave as function of time



$$f(t) = a\,\sin\big(2\,\pi\,f\,(\,t - \theta\,)\big)$$

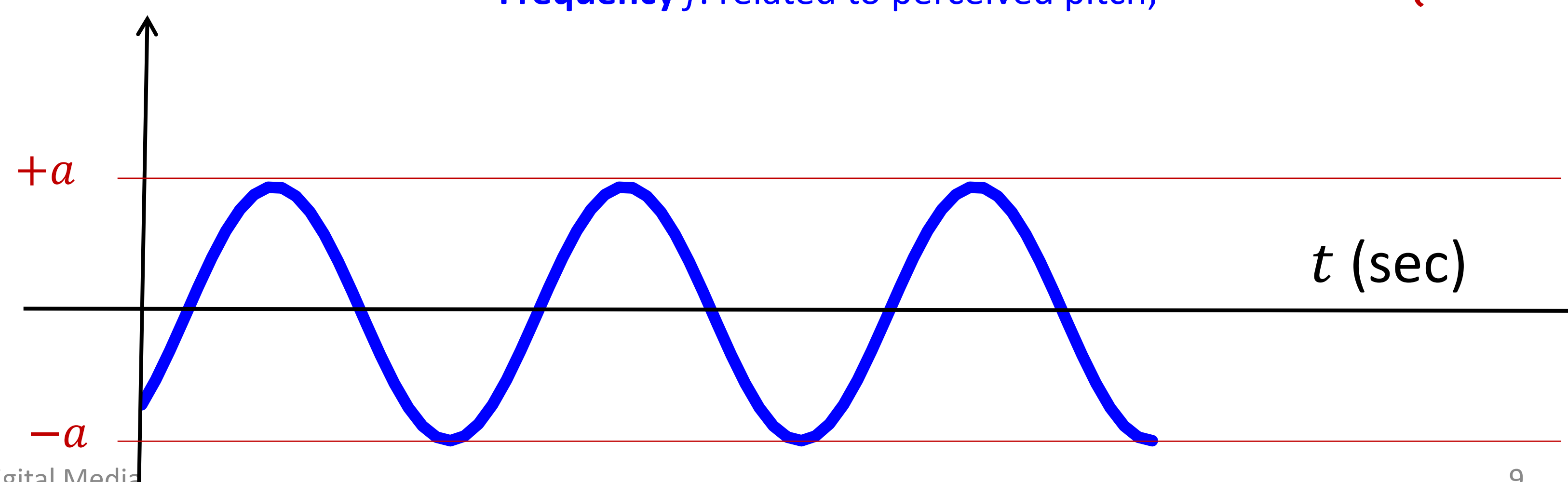
amplitude

frequency

time

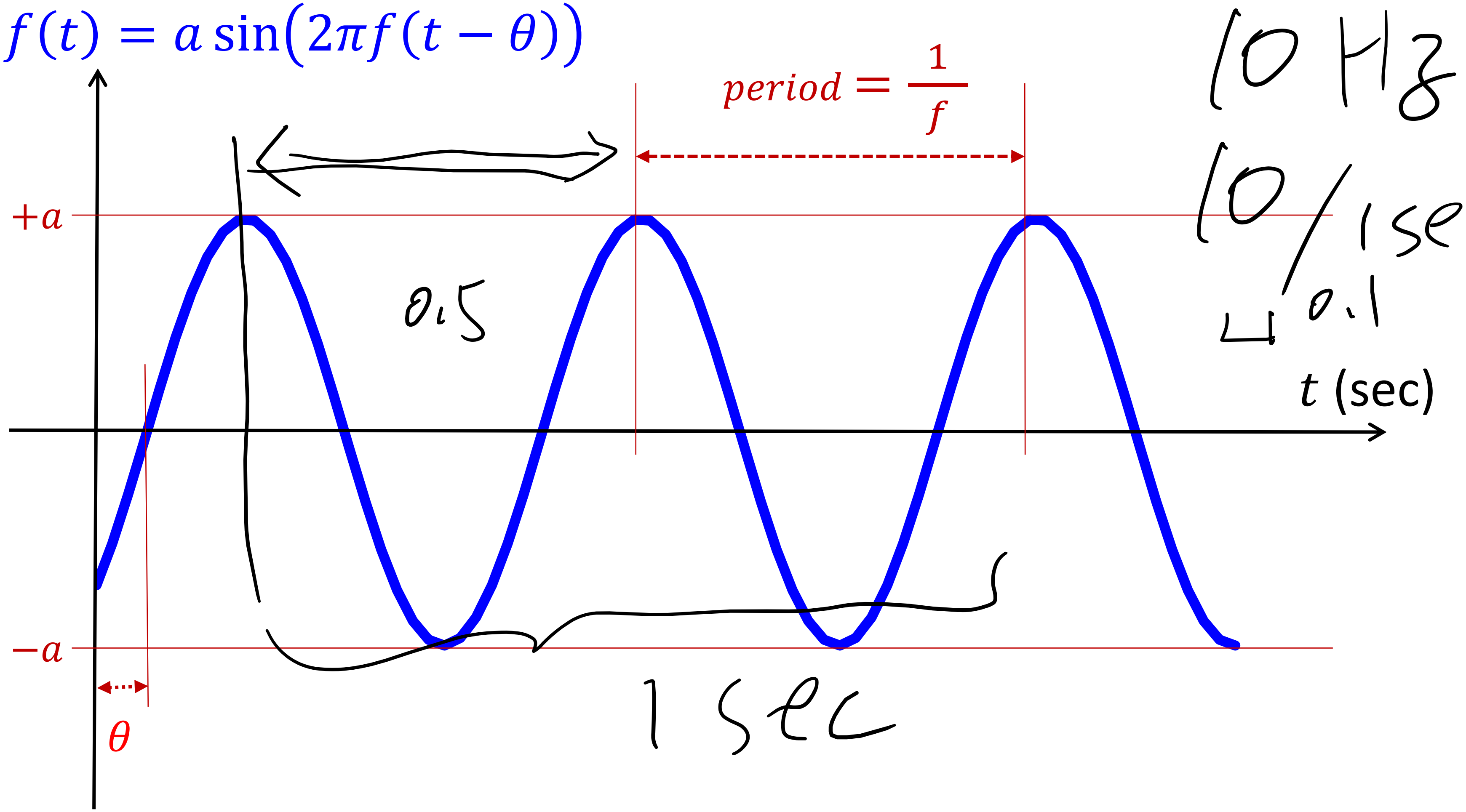
phase: time shift

Amplitude a: related to intensity (loudness);
Frequency f: related to perceived pitch;



Anatomy of Audio:
Sine Wave as function of time

$$f(t) = a\sin\big(2\pi f(t-\theta)\big)$$



Amplitude *a*: related to intensity (loudness);
Frequency *f*: related to perceived pitch;

Hearing **Intensity** Limits

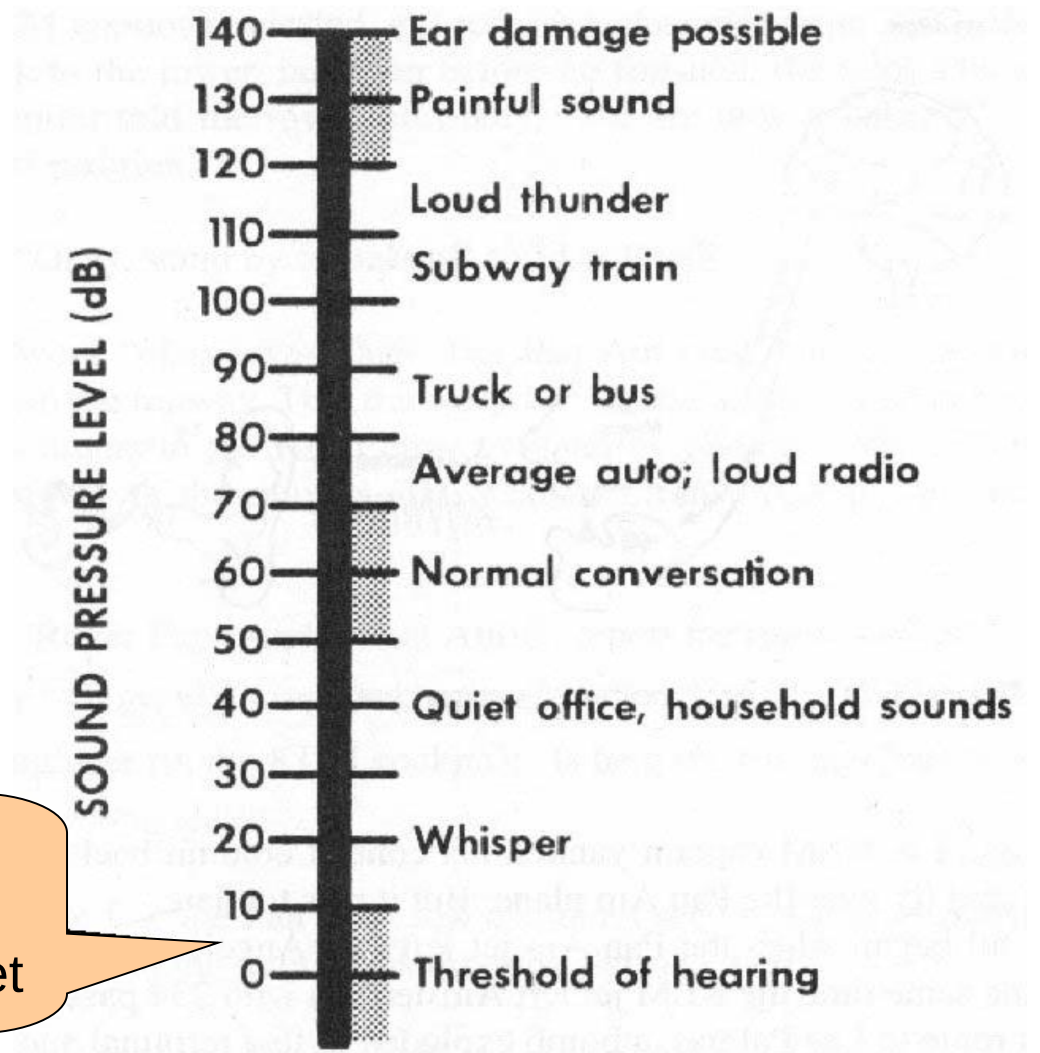
Threshold of hearing:
RMS sound pressure of **20 μ*Pa***
(0.98 pW/m^2 at 1 atm, and 25°C)

dB-SPL (sound pressure level) is a *relative logarithmic* unit relative to the **20 μ*Pa*** value above

$$\text{dB-SPL} = 20 \log_{10} (\text{sound} / \text{threshold})$$

Without using dB, the numbers vary by a factor of 100000 or more!

Tick of a watch under quiet conditions at 20 feet



More info: https://en.wikipedia.org/wiki/Decibel

Sound Pressure *(Pa)*

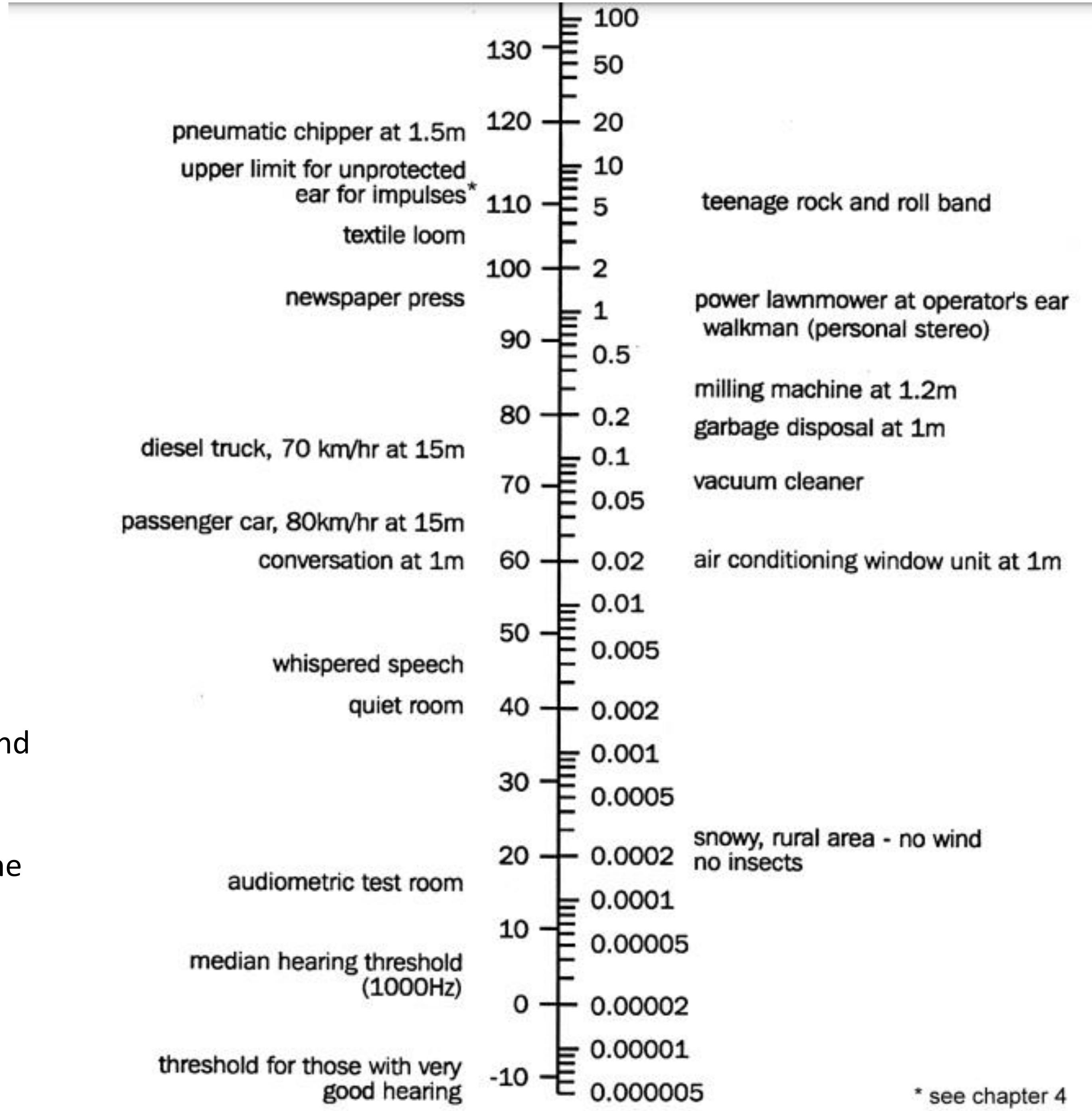
Sound Pressure is the sound force (*N*) acting on the surface area (m^2) perpendicular to the direction of the sound

Wide range $2*10^{-5}$ ~ 60

Chang a lot for us to hear the difference,

Due to this range it is convenient to express sound pressure with a logarithmic [decibel scale]() related typically to the lowest human hearable sound - $2\ 10^{-5}$ *Pa* or *0 dB*.

Small number represent large change 0 10 100 1000



Can you hear me now?

http://www.noiseaddicts.com/2009/03/can-you-hear-this-hearing-test/

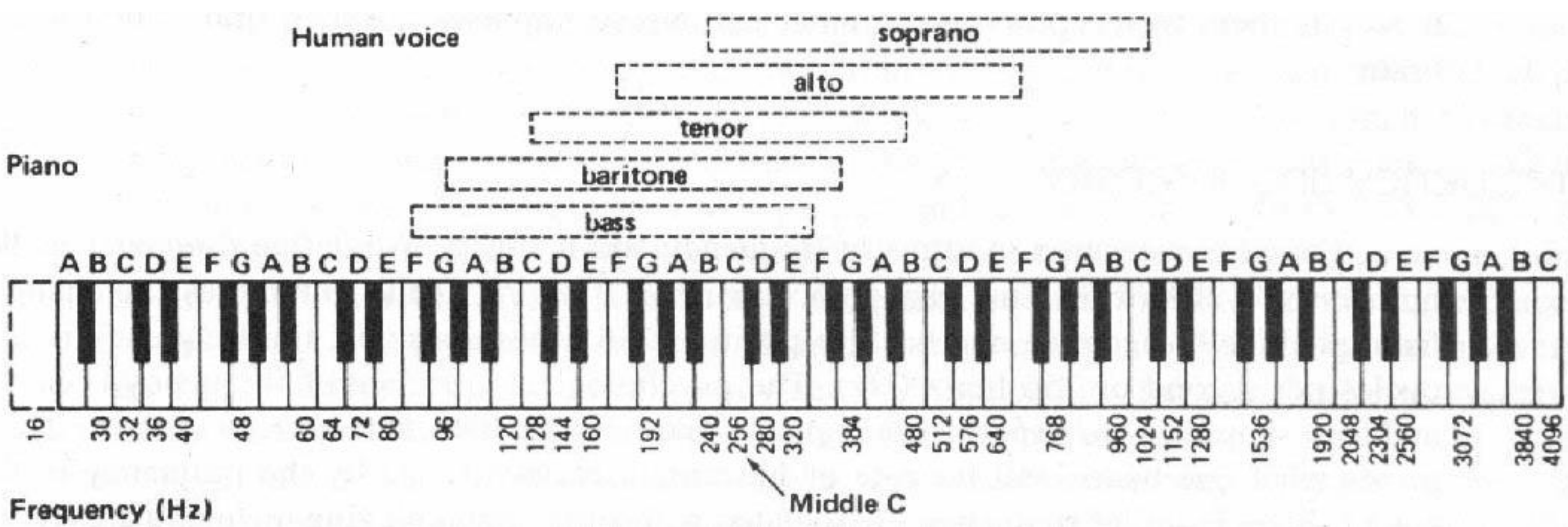
Only hear certain range "16 Hz – 20 kHz" , if beyond the range, don't need to store!

Hearing **Frequency** Limits

“16 Hz – 20 kHz” 20yr 15kHz 50: 12kHz 90: 5kHz

Upper boundary is variable (!) and steadily decreases (!) with age

High frequencies usually change the timbre (coloration) of sound



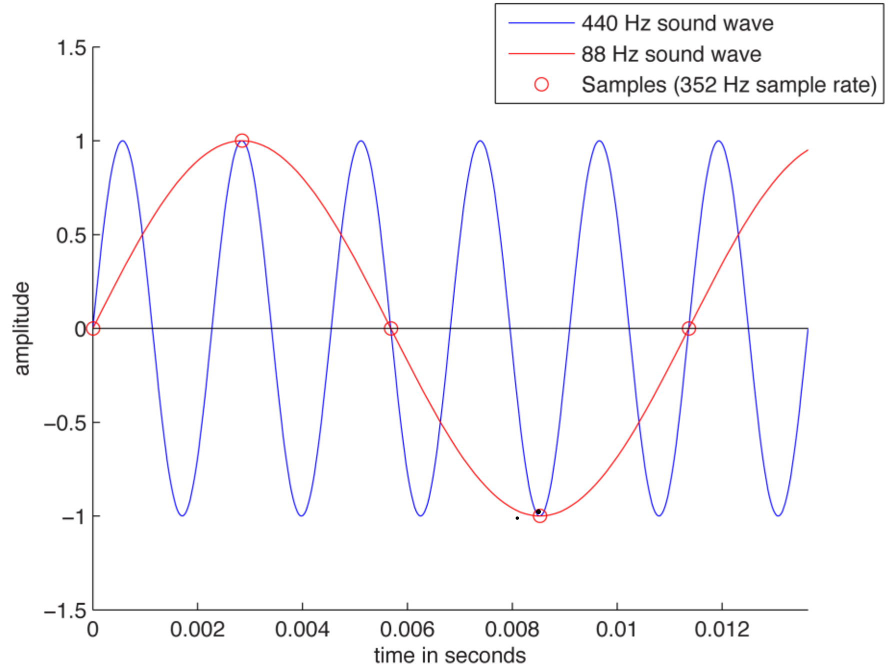
Sampling Rate and Bit Rate

- Digitizing sound involves taking **samples** (measurements) at a fixed rate (**sampling rate**) and recording them
- What rate should be used?

Too fast → large file

Too slow → inaccurate

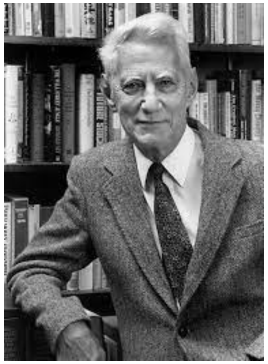
Low sampling rate reproduces a sound (red) that does not match the original frequency (blue)

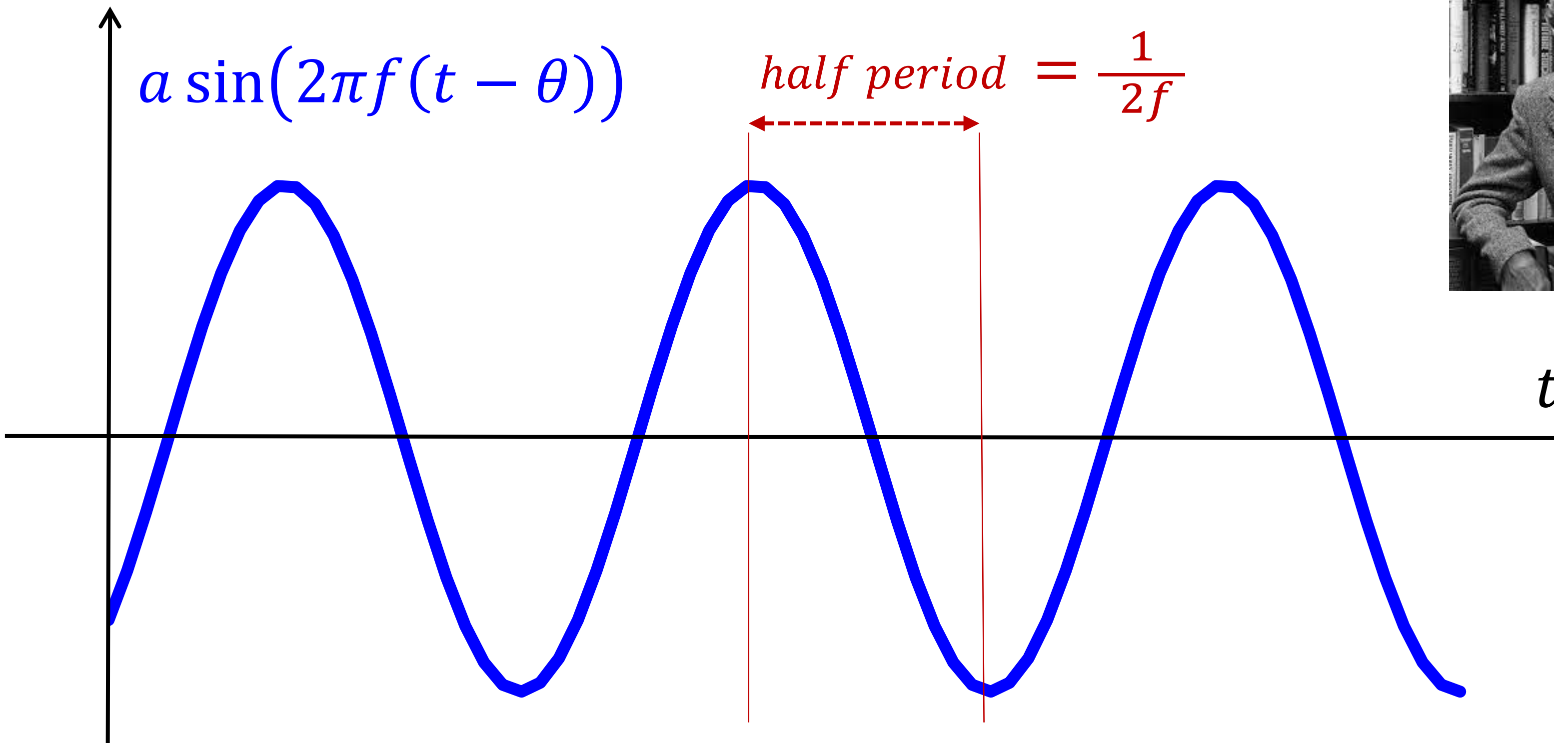


Courtesy of digitalsoundandmusic.com

Q: how often (frequent) should we sample?

Nyquist-Shannon: Sample interval $\leq$ half period
to be able to catch peaks and valleys: ↓
i.e., sampling-rate $f_s \geq 2f$ $\qquad p_s \leq \frac{1}{2}p$





Q: given f=10kHz, what is sampling rate? What is sampling period/interval?

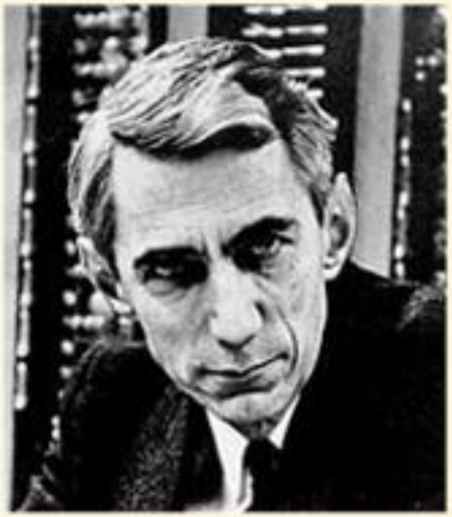
< 1/20000 20k samples 20kHz freq

Therefore, the following question arises: Which is the minimum necessary sampling frequency for a given type of signal, that will not distort the underlying information and/or allow its accurate reconstruction? The answer is given by the **Nyquist-Shannon sampling theorem**, that may be simply stated as follows:

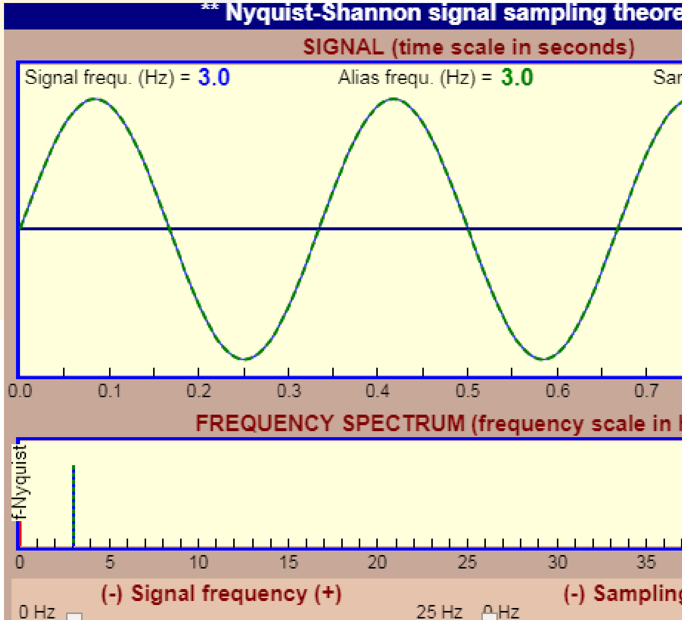
The minimum sampling frequency of a signal that it will not distort its underlying information, should be double the frequency of its highest frequency component.

If f_S is the **sampling frequency**, then the **critical frequency** (or **Nyquist limit**) f_N is defined as equal to $f_S/2$.

Harry Nyquist
(1889-1976)

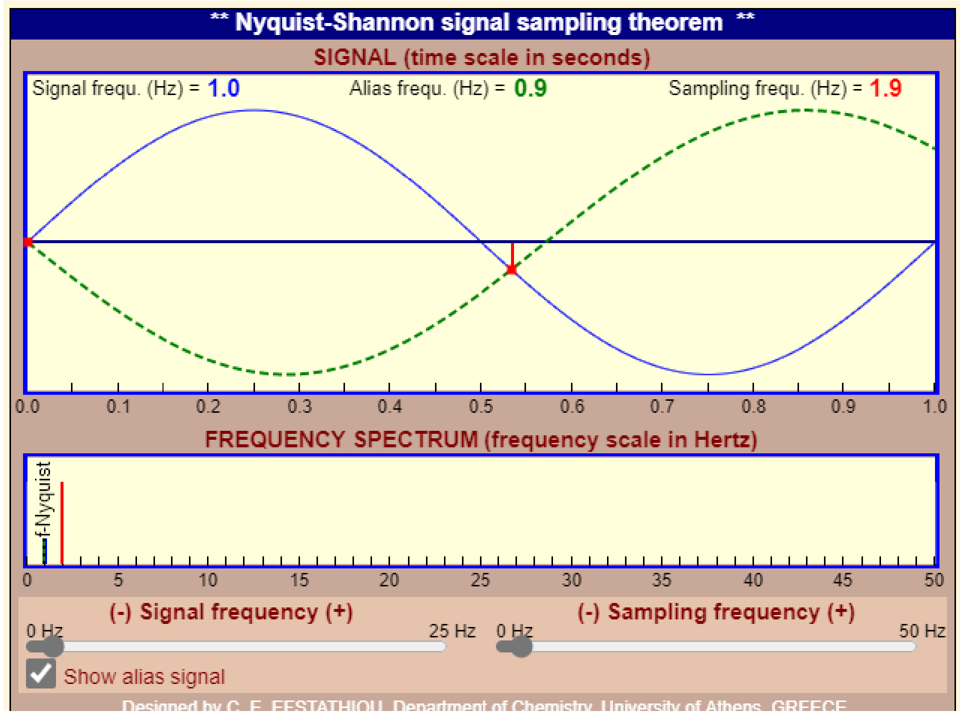


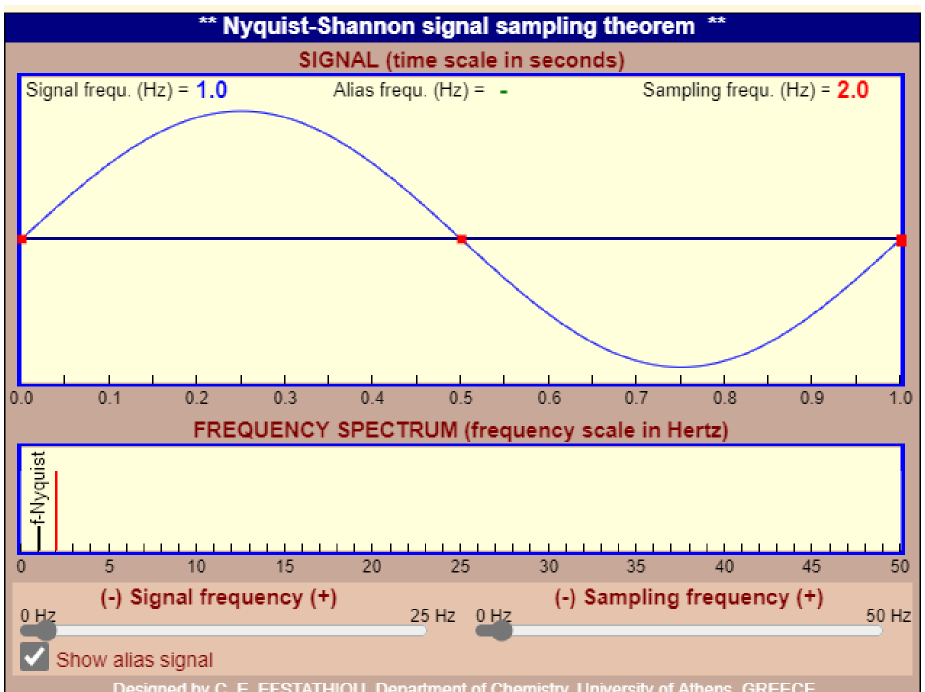
Claude Shannon
(1916-2001)

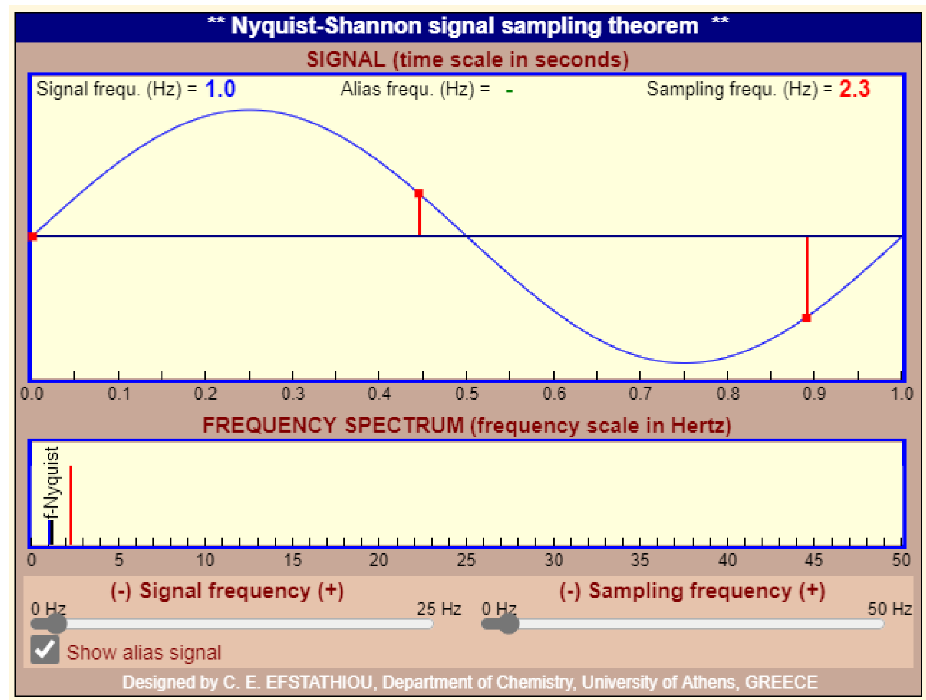


Q: given f=10kHz, what is sampling rate? What is sampling period/interval?

< 1/20000 20k samples 20kHz freq


** Nyquist-Shannon signal sampling theorem **
SIGNAL (time scale in seconds)
Signal frequ. (Hz) = 1.0
Alias frequ. (Hz) = 0.9
Sampling frequ. (Hz) = 1.9
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
FREQUENCY SPECTRUM (frequency scale in Hertz)
f-Nyquist
0 5 10 15 20 25 30 35 40 45 50
(-) Signal frequency (+)
0 Hz
25 Hz
(-) Sampling frequency (+)
0 Hz
50 Hz
Show alias signal


** Nyquist-Shannon signal sampling theorem **
SIGNAL (time scale in seconds)
Signal frequ. (Hz) = 1.0
Alias frequ. (Hz) = -
Sampling frequ. (Hz) = 2.0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
FREQUENCY SPECTRUM (frequency scale in Hertz)
f-Nyquist
0 5 10 15 20 25 30 35 40 45 50
(-) Signal frequency (+)
0 Hz
25 Hz
(-) Sampling frequency (+)
0 Hz
50 Hz
Show alias signal
Designed by C. E. EFSTATHIOU, Department of Chemistry, University of Athens, GREECE


** Nyquist-Shannon signal sampling theorem **
SIGNAL (time scale in seconds)
Signal frequ. (Hz) = 1.0
Alias frequ. (Hz) = -
Sampling frequ. (Hz) = 2.3
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
FREQUENCY SPECTRUM (frequency scale in Hertz)
f-Nyquist
0 5 10 15 20 25 30 35 40 45 50
(-) Signal frequency (+)
0 Hz
25 Hz
(-) Sampling frequency (+)
0 Hz
50 Hz
Show alias signal
Designed by C. E. EFSTATHIOU, Department of Chemistry, University of Athens, GREECE

Sampling Rate and Bit Rate (2)

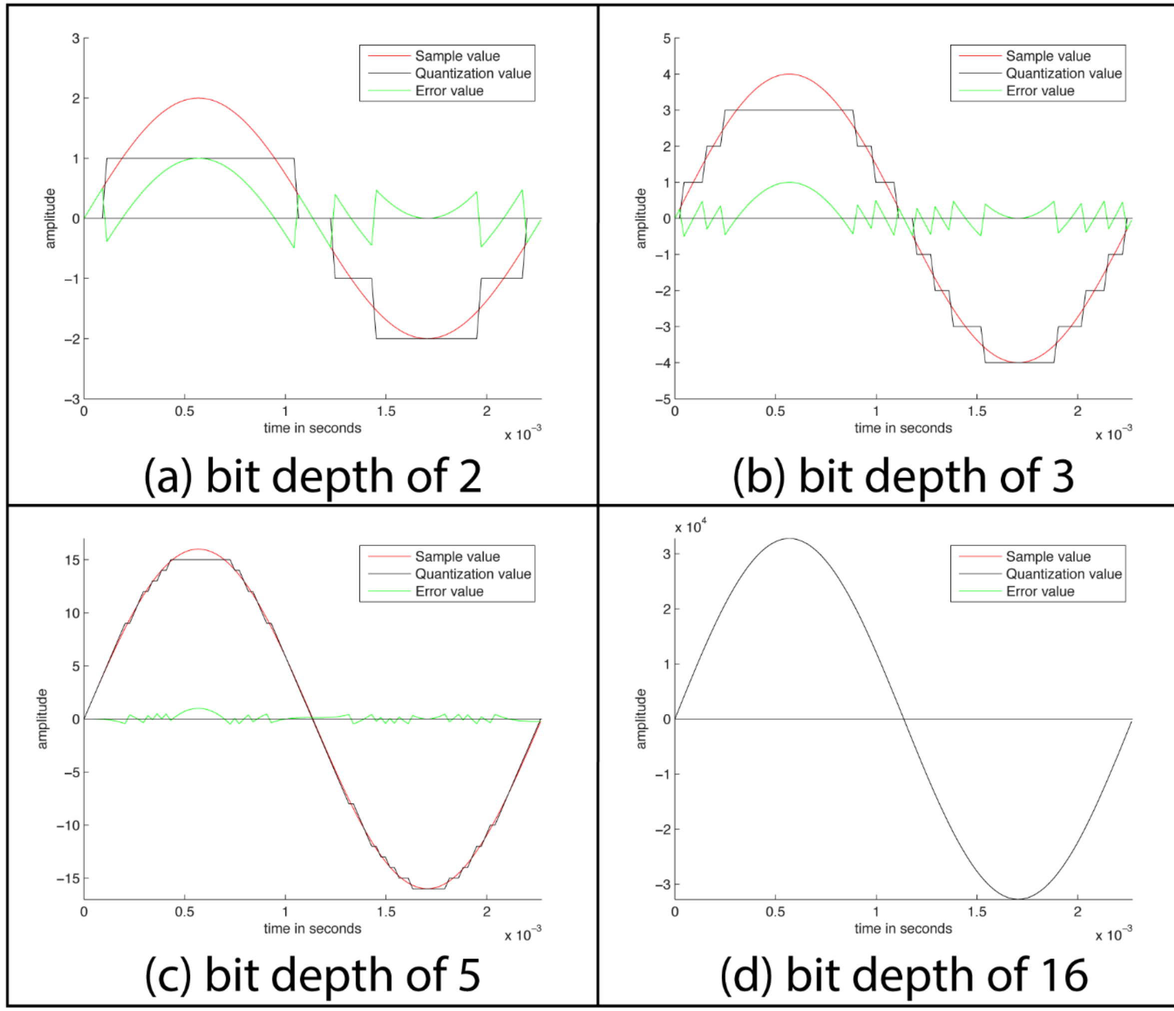
- Each sample is recorded using a fixed number of bits, **bit depth**
- Higher bit depth means a more accurate sample

Quiet sounds will be better preserved!

Red: original

Black: sampling

Green: difference
AKA *quantization noise*



Courtesy of digitalsoundandmusic.com

2bits: -2 ~ 1
3bits: -4~3
5bits: -16~15

Sampling Rate and Bit Rate (2) one more example

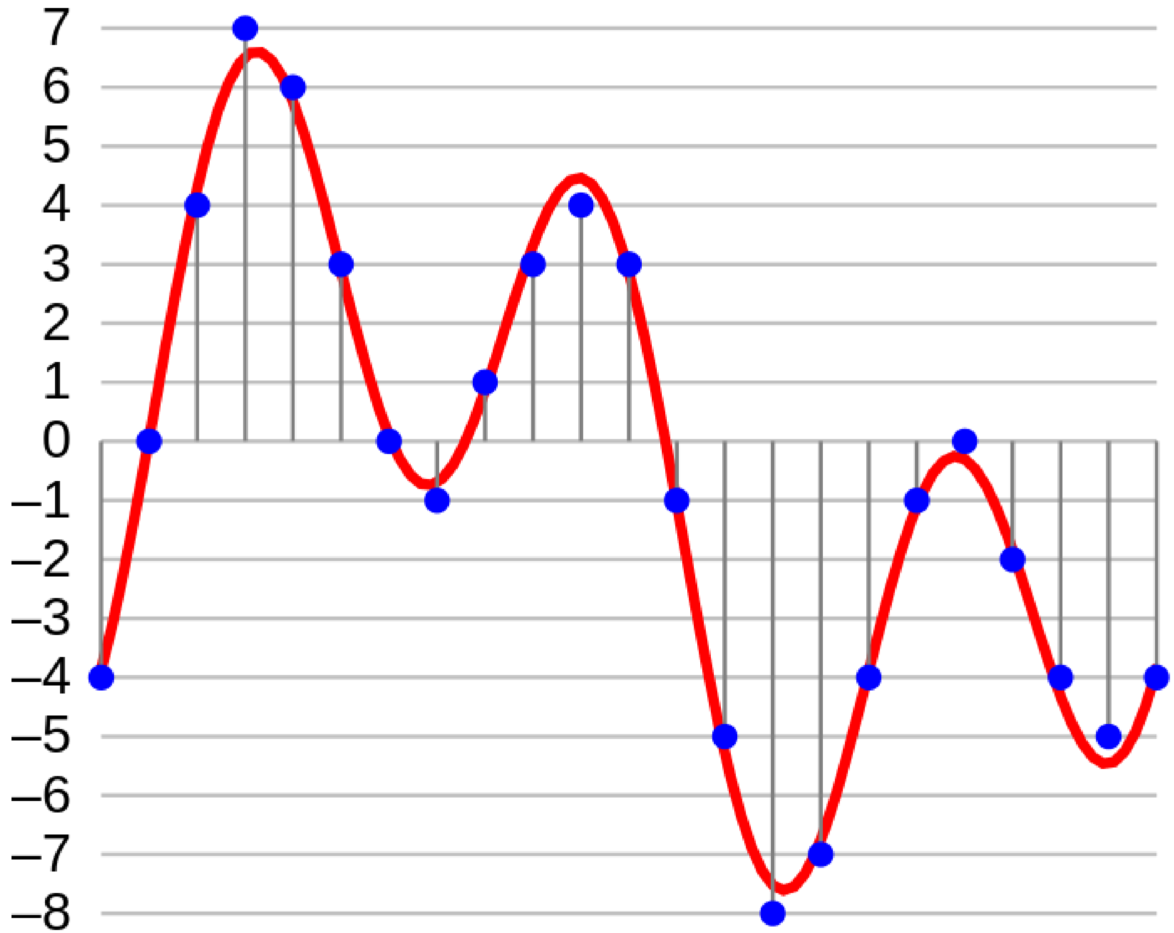
- Each sample is recorded using a fixed number of bits, **bit depth**
- Higher bit depth means a more accurate sample

Quiet sounds will be better preserved!

Red: original

blue: sampling

4bits: -8 ~ 7



An analogue signal (in red) encoded to 4-bit digital samples (in blue); the bit depth is four, so each sample's amplitude is one of 16 possible values.

Bit Depth in the Practice

<8 bits

- Not used for sound*; Can be used to record physical processes like blood pressure, heartbeat, motion due to walking or running

8 bits

- Common in telephony
- Quantization noise can be perceptible *sometimes*

16 bits

- Most high-quality sound (CDs, DAT, MP3, movie audio tracks)

24+ bits

- Even higher quality/dynamic range (DVD-Audio, DTS)
- Often used before or during sound processing/editing (*mastering*)

*– exception: SACD

Sampling Rate and Bit Rate (3)

- Combining bit depth and (x) sampling rate gives a **bit rate**
- Higher bit rate = better quality = larger file size

- Range of human hearing: 16 Hz – 20 kHz

  Nyquist-Shannon theorem recommends sampling rate $\geq 2 \times$ max frequency for accurate reproduction

- CD-quality audio (industry standard):

  16 bits bit depth (per channel, stereo has 2 channels)

  44.1 kHz sampling rate

  1.4 Mbits/second bit rate 2x16x44.1k=1411200 bits/sec

  e.g. phone: 8bit 1ch 8000hz 8 * 8000 = 64000 bits/sec

- Uncompressed audio typically saved in a .wav format

1411200 bits/sec 4 minutes song? A 2 hour concert?

1.41M*60*4/8~40M 1.41M*60*120/8 ~ 1G

Compression is important!

TV Standard and Digital Audio

- Digital audio was developed during the 70s
- CD was released in 1982

- Problem: storage of data
  - CD data capacity (74 min audio format mode): ~780 MB
  - Hard disks at that time: ~10 MB (!)
- The only practical way to store massive amount data was video tape (lack of random access was not a concern)
  - It was easier to tweak digital audio parameters than to modify an existing TV standard

Audio Compression Techniques

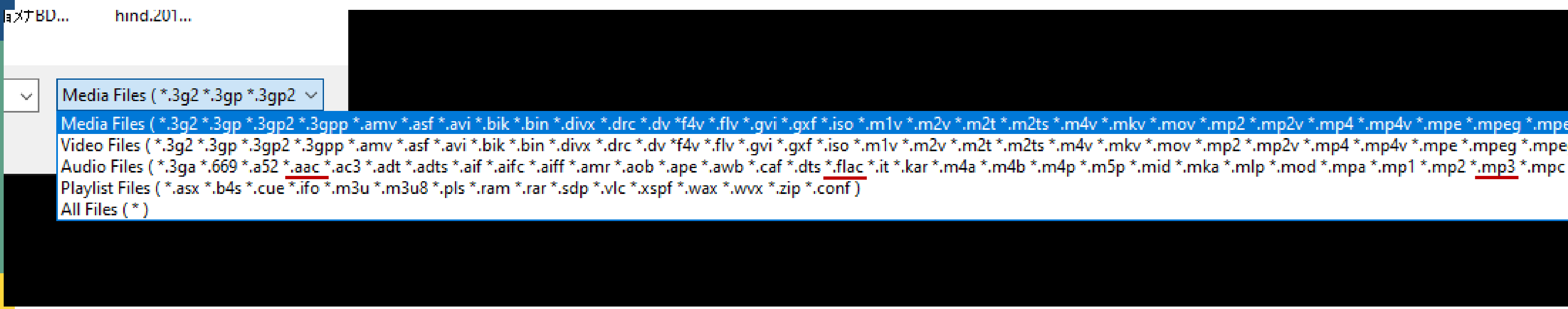
- **Lossless** audio compression
  - Similar to text compression techniques
- **Lossy** audio compression
  1. Remove imperceptible sounds (using psychoacoustics models)
     - Out of range
     - In range but hidden by louder
  2. Reduce bit rate (less accurate reproduction of the original)

- **Codecs**
  - Compression and decompression algorithms for audio (and video)
  - Portmanteau (word combination) of “**co**der” and “**dec**oder” or (as in this case) “**co**mpressor” and “**dec**ompressor”

Example codecs

- **FLAC** (Free Lossless Audio Codec)
  - **Lossless**, using a combination of run-length and Huffman encoding
  - Compression ratio around 62% [reference]
  - Typically used for archiving high quality audio
- **MP3** (Moving Pictures Experts Group, Audio Layer 3)
  - **Lossy** **(+ lossless)**, with a compression ratio around 13% [reference]
  - Uses psychoacoustics (discard imperceptible info), Huffman encoding, and lower bit rates
  - Popular on most mobile devices
- **AAC** (Advanced Audio Coding)
  - **Lossy**, **(+lossless)** with a compression ratio around 14% [reference]
  - Uses psychoacoustics, Huffman encoding, and lower bit rates
  - Better perceived quality than MP3 [reference]
  - Popular on Apple devices and most smartphones

10M file → 6.2M, 1.3M, 1.4M


hind.201...
Media Files (*.3g2 *.3gp *.3gp2
Media Files (*.3g2 *.3gp *.3gp2 *.3gpp *.amv *.asf *.avi *.bik *.bin *.divx *.drc *.dv *f4v *.flv *.gvi *.gxf *.iso *.m1v *.m2v *.m2t *.m2ts *.m4v *.mkv *.mov *.mp2 *.mp2v *.mp4 *.mp4v *.mpe *.mpeg *.mpe
Video Files (*.3g2 *.3gp *.3gp2 *.3gpp *.amv *.asf *.avi *.bik *.bin *.divx *.drc *.dv *f4v *.flv *.gvi *.gxf *.iso *.m1v *.m2v *.m2t *.m2ts *.m4v *.mkv *.mov *.mp2 *.mp2v *.mp4 *.mp4v *.mpe *.mpeg *.mpe
Audio Files (*.3ga *.669 *.a52 *.aac *.ac3 *.adt *.adts *.aif *.aifc *.aiff *.amr *.aob *.ape *.awb *.caf *.dts *.flac *.it *.kar *.m4a *.m4b *.m4p *.m5p *.mid *.mka *.mlp *.mod *.mpa *.mp1 *.mp2 *.mp3 *.mpc
Playlist Files (*.asx *.b4s *.cue *.ifo *.m3u *.m3u8 *.pls *.ram *.rar *.sdp *.vlc *.xspf *.wax *.wvx *.zip *.conf)
All Files (*)

summary

- f vs p   f: frequency, rate   p: period, interval
- (1) Sampling: sample rate f_s Shannon $f_s \geq 2f$ $p_s \leq p/2$
- (2) Truncating: bit depth
- Bitrate: = f_s × bit depth

- Codecs: Compression and decompression algorithms
  - Lossless: FLAC
  - Lossy: MP3, AAC

  - **Lossy** audio compression
    1. Remove imperceptible sounds (using psychoacoustics models)
       - Out of range
       - In range but hidden by louder
    2. Reduce bit rate (less accurate reproduction of the original)

Digital Data Representation

Topics for the next few weeks

Representing Numbers

Representing Text

Representing Audio

Representing Images

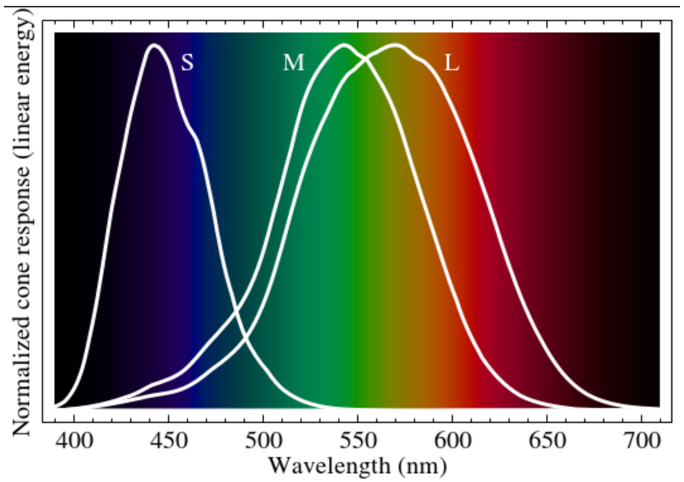
Representing Video

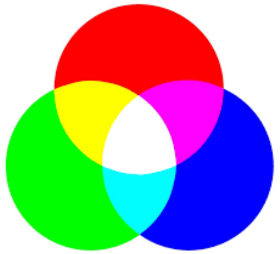
Graphics (Images)

A picture is worth a thousand words

Representing Colour (1)

- Human retina contains light-sensitive cells of 4 types:
  - 1 for low-light “night” vision: rods, ignored here
  - 3 for “regular” vision
    short/medium/long
    Most birds: 4
    Dogs, cows…: 2



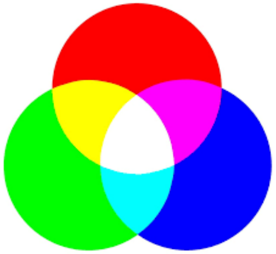


- When all three are stimulated by light: neutral white (or grey) colour
- **Idea**: pick colours that can stimulate individual cones separately!
  - **R**ed, **G**reen, **B**lue
- Most of the *important* perceptible colors can be formed by combining these RGB components

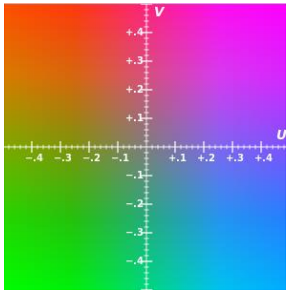
Representing Colour (2)

- RGB colour model
  - Three integers (usually in the range 0..255*) describe the amount of Red, Green, and Blue required to reproduce the colour

255,255,255	White	0,0,0	Black
255,0,0	Red	0,255,0	Green
0,0,255	Blue	255,255,0	Yellow
255,165,0	Orange	128,0,128	Purple



- Other colour models
  - **C**yan, **M**agenta, **Y**ellow, and blac**K** (CMYK) – e.g., colour printers
    - Start with white [paper], then *subtract* individual components
  - **H**ue, **S**aturation, **L**uminance (HSL); Luminance, Blue, Red (YUV)
    - Similar to colour wheel principle
    - Map better to how human brain interprets colours



* – the mapping of numbers to light intensities is NOT linear!

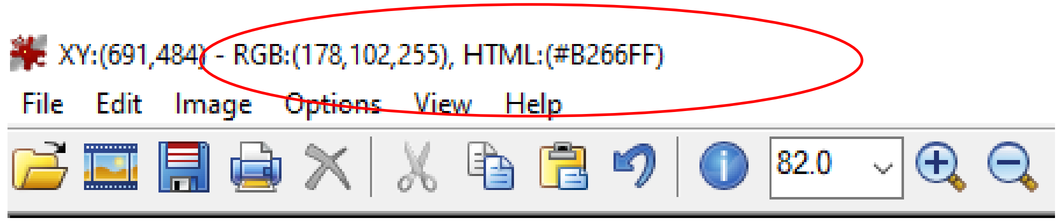
Representing Colour with Hexadecimal

- Each R, G, and B value is in the range 0..255 (1 byte*)
  - Looks familiar? 256x256x256 = 2^{24} = 16Million colors more than enough
- Convert the decimal value to hexadecimal
  - 256 values = ? bits/bytes = ? hexadecimal digits
  - 6 hexadecimal digits used to represent the entire colour: RRGGBB
- Example **this colour** is RGB = (178,102,255)

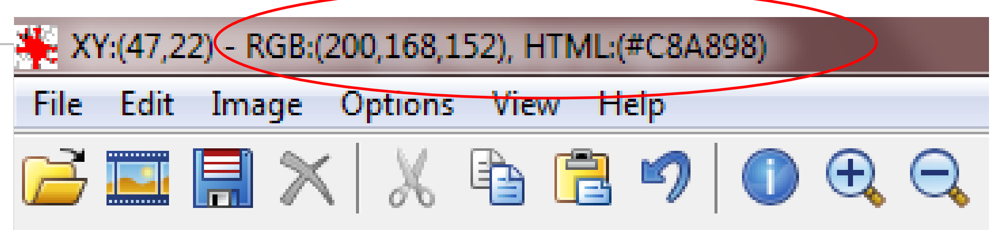
– Mor
ange)
All
accurate and realistic representation

Representing Colour with Hexadecimal

- Each R, G, and B value is in the range 0..255 (1 byte*)
  - Looks familiar? 256x256x256 = 2^{24} = 16Million colors more than enough
- Convert the decimal value to hexadecimal
  - 255 values = 1 byte = ? hexadecimal digits
  - 6 hexadecimal digits used to represent the entire colour: RRGGBB
- Example **this colour** is RGB = (178,102,255)
  - RR = 178_{10} = $B2_{16}$
  - GG = 102_{10} = 66_{16}
  - BB = 255_{10} = FF_{16}
  - So using hexadecimal, it is written: #B266FF (# indicates hexadecimal)



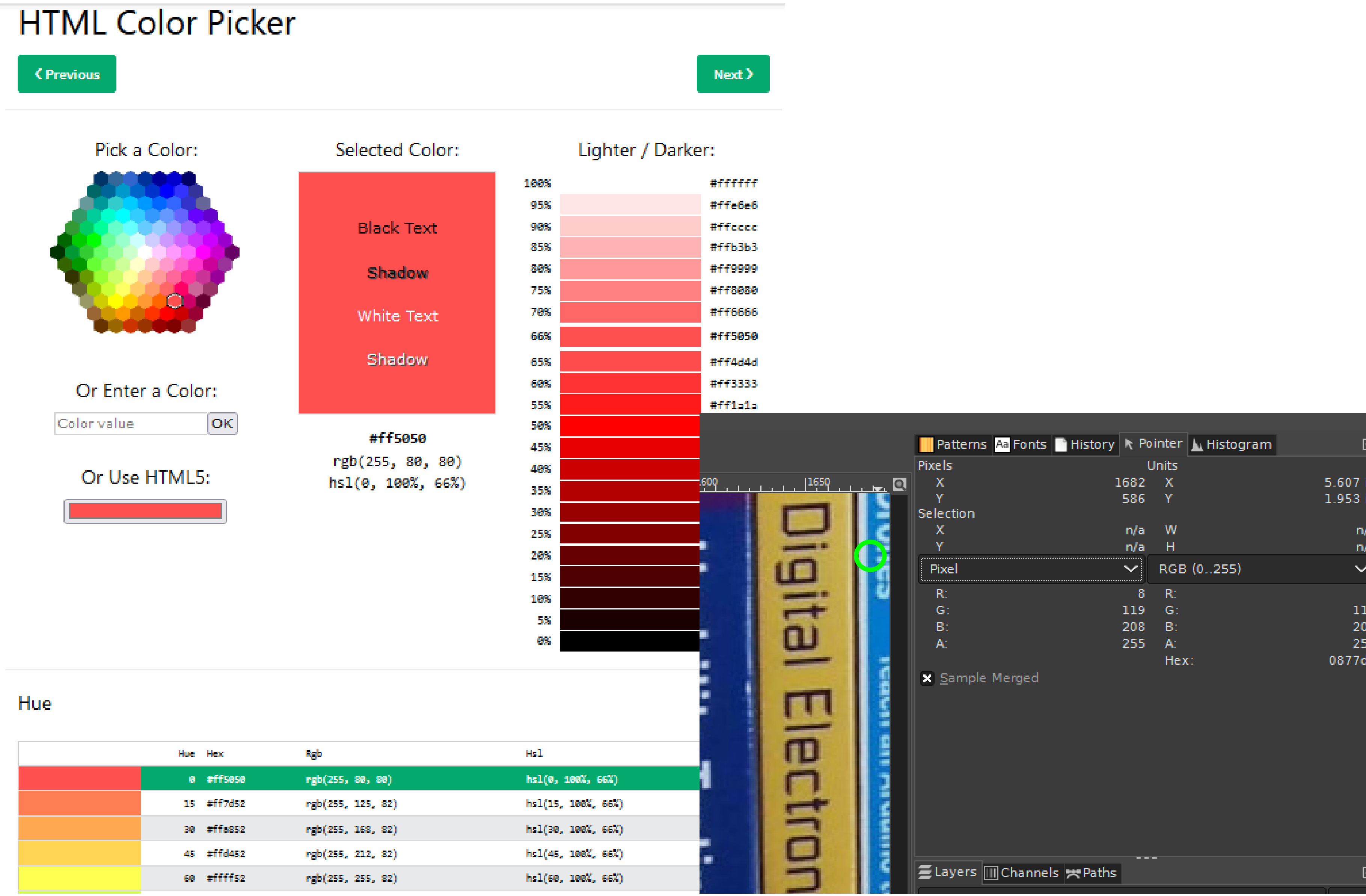
- [Colour code chart]()



– More bits for HDR (high dynamic range) and *wide colour gamut* images

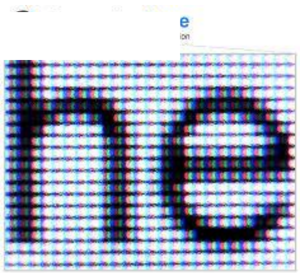
Allows for more perceptually accurate and realistic representation

Colour Values



Digitized Images and Graphics

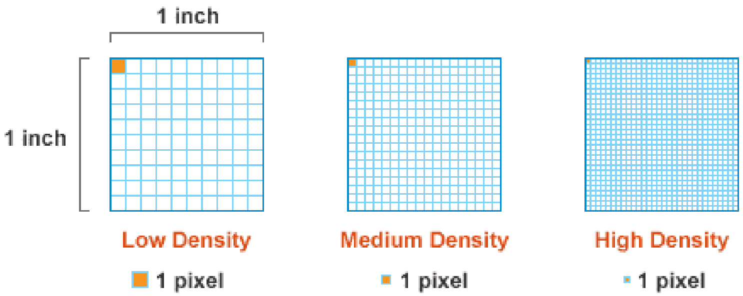
- Pixel (**pi**cture **el**ement)
  - Digitizing: act of representing as a collection of individual dots -- Dots of colour in an image (or on a display)
  - Each pixel is a single color



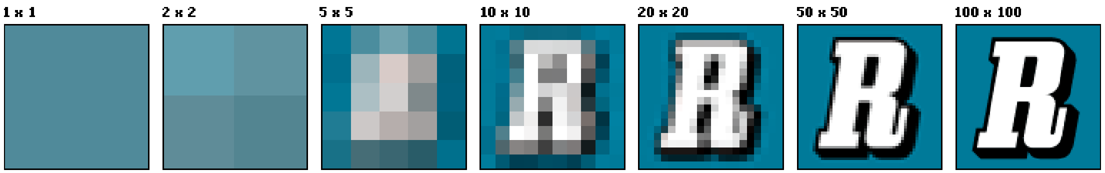
- Resolution
  - Number of pixels in an image (or on a display)
  - Size of an image, measured in pixels   e.g., 125x256

125x256 x 3 bytes

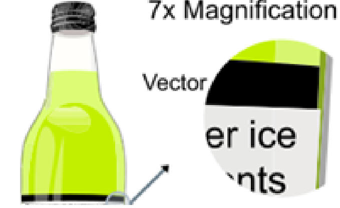
  - Sometimes refers to the pixel density, the number of pixels per unit of distance (e.g., pixels per inch or ppi)



- Raster-graphics
  - Bitmap images
- Vector graphics

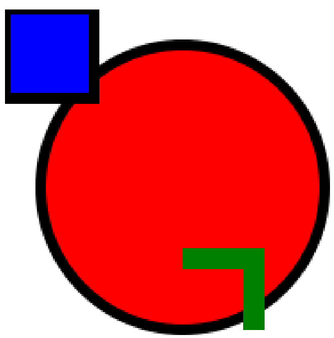


Vector Graphics



- An image that is defined using mathematical equations representing geometric shapes: lines, curves, and polygons
- Can be enlarged without loss of detail, or change in file size
- Images created using **drawing applications** or text editors

- **SVG** (Scalable Vector Graphics)
  - Popular vector graphics format
  - Files are text-based and can be compressed accordingly (must loseless,why?)

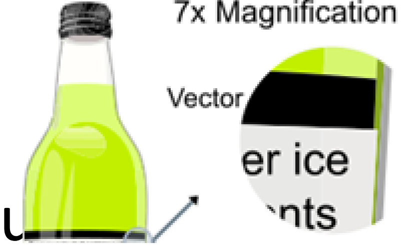


```
<svg height="100" width="100">
  <circle cx="50" cy="50" r="40" stroke="rgb(0,0,0)" stroke-width="3" fill="red" />
  <rect width="25" height="25" style="fill:rgb(0,0,255);stroke-width:3;stroke:black" />
  <polyline points="50,70 70,70 70,90" style="fill:none;stroke:green;stroke-width:6" />
</svg>
```

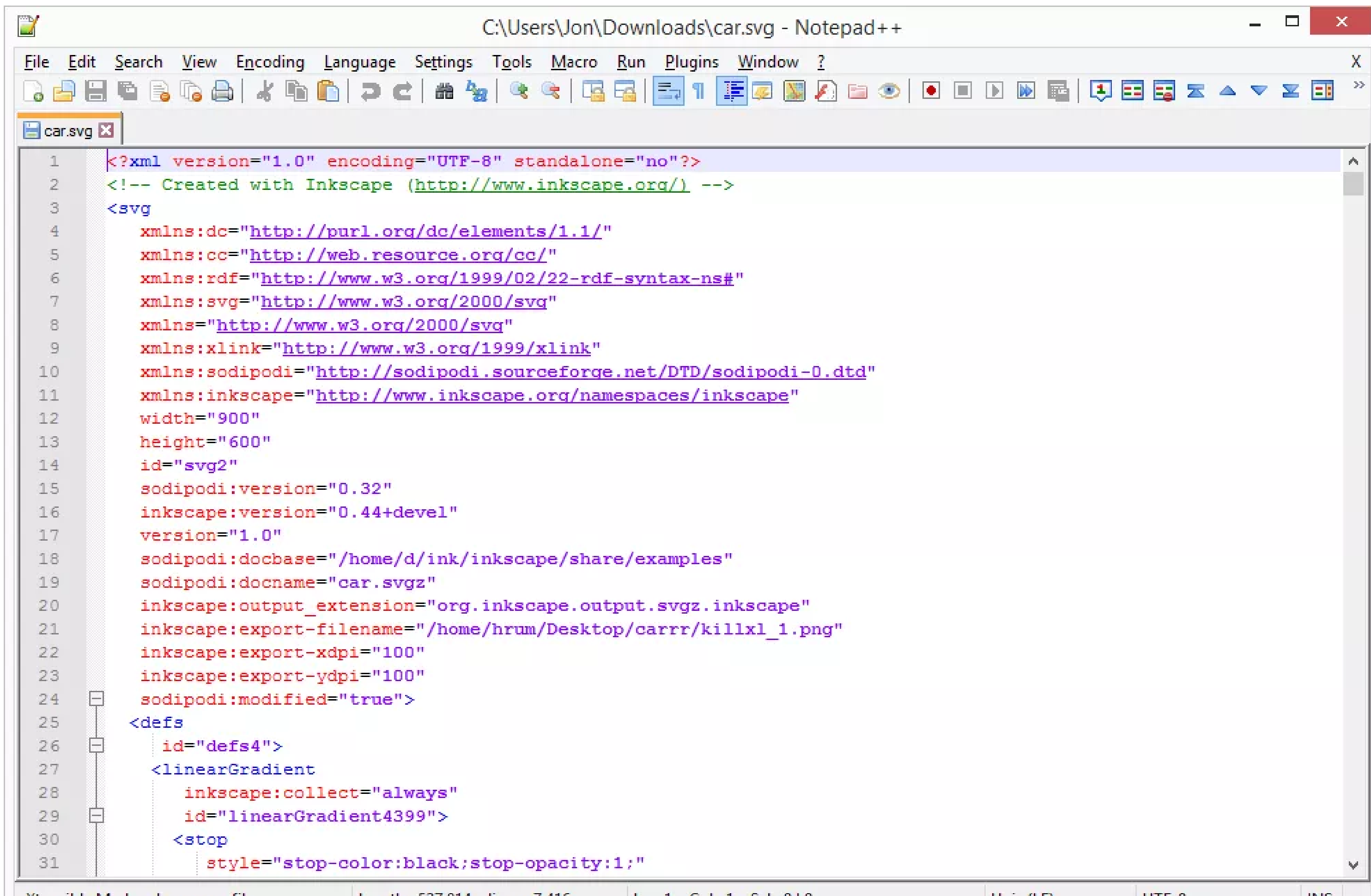
Size in storage? #texts * 7/8/16 bits

Hard to represent real objects

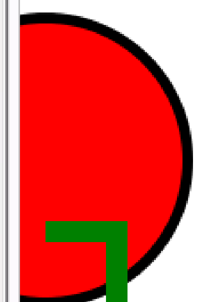
Vector Graphics



- An image that is defined using mathematical equ... representing geometric shapes: lines, curves, and polygons
- Images created using <u>**drawing applications** or text editors</u>




```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!-- Created with Inkscape (http://www.inkscape.org/) -->
<svg
   xmlns:dc="http://purl.org/dc/elements/1.1/"
   xmlns:cc="http://web.resource.org/cc/"
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:svg="http://www.w3.org/2000/svg"
   xmlns="http://www.w3.org/2000/svg"
   xmlns:xlink="http://www.w3.org/1999/xlink"
   xmlns:sodipodi="http://sodipodi.sourceforge.net/DTD/sodipodi-0.dtd"
   xmlns:inkscape="http://www.inkscape.org/namespaces/inkscape"
   width="900"
   height="600"
   id="svg2"
   sodipodi:version="0.32"
   inkscape:version="0.44+devel"
   version="1.0"
   sodipodi:docbase="/home/d/ink/inkscape/share/examples"
   sodipodi:docname="car.svgz"
   inkscape:output_extension="org.inkscape.output.svgz.inkscape"
   inkscape:export-filename="/home/hrum/Desktop/carrr/killxl_1.png"
   inkscape:export-xdpi="100"
   inkscape:export-ydpi="100"
   sodipodi:modified="true">
  <defs
    id="defs4">
   <linearGradient
      inkscape:collect="always"
      id="linearGradient4399">
     <stop
        style="stop-color:black;stop-opacity:1;"
```



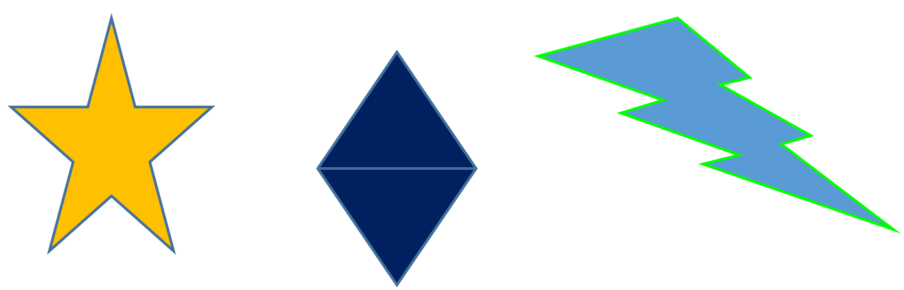
d" />
black" />
idth:6" />

esent

Other Vector Graphics Examples

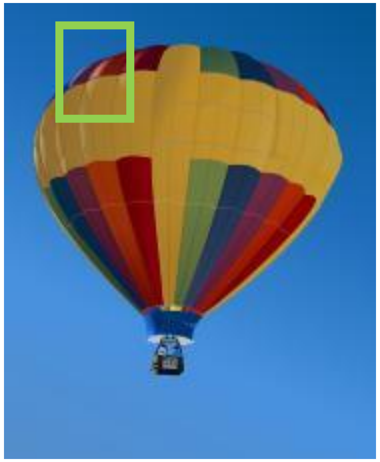
- Computer typefaces (fonts)

- Geometric shapes in various apps



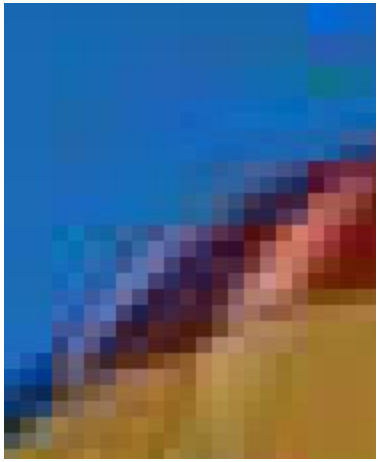
Raster Graphics (bitmap images)

- An image that is comprised of a matrix of pixels
- Stored on a pixel-by-pixel basis
- Images created/edited using **painting applications** (e.g., gimp)
- Very good for photographs, but suffers from **pixelation**
  - When enlarged, the rectangular shape of pixels becomes apparent

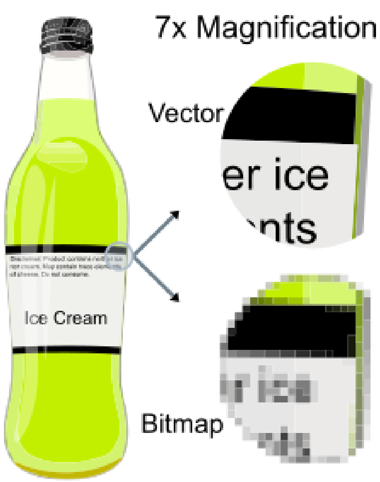


Original

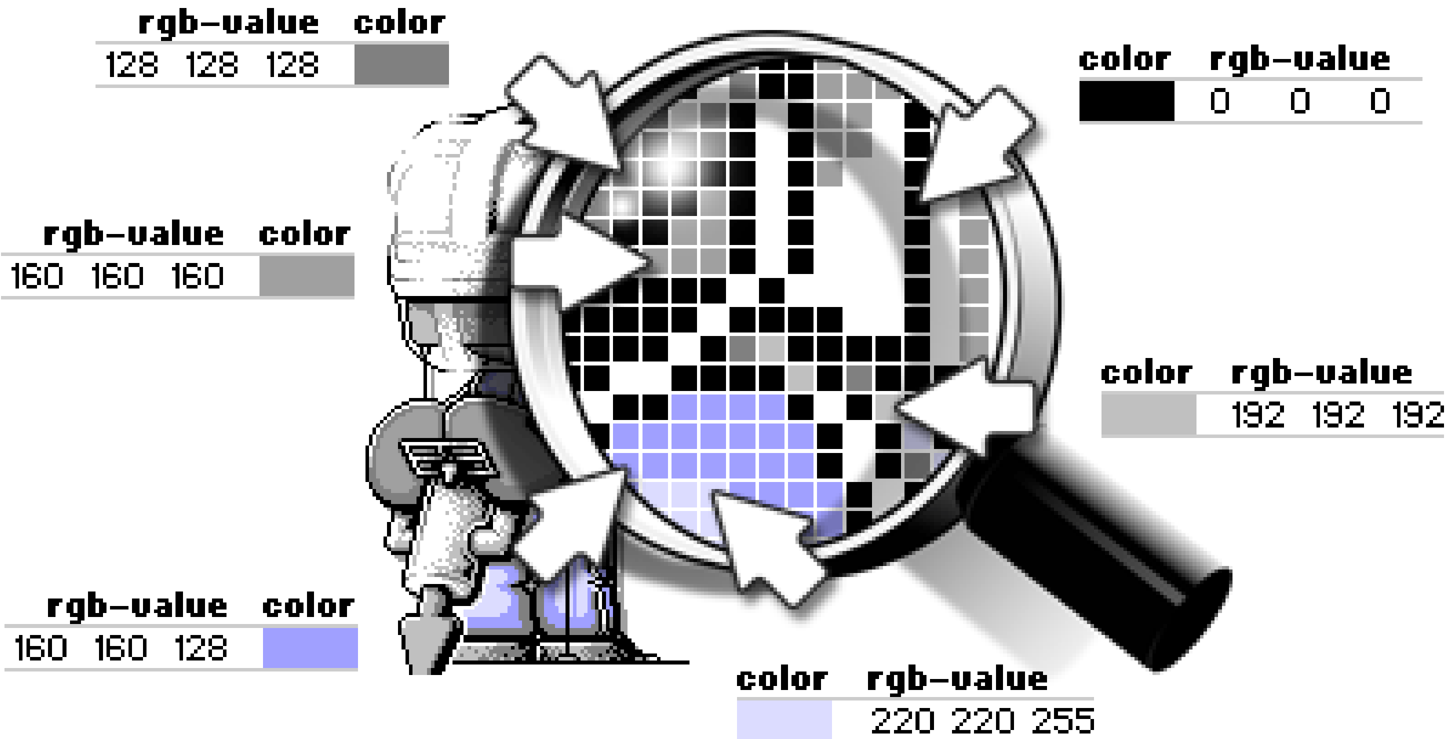
Enlarged Vector



Enlarged Raster



Images courtesy of intmath.com


rgb–value color
128 128 128
rgb–value color
160 160 160
rgb–value color
160 160 128
color rgb–value
0 0 0
color rgb–value
192 192 192
color rgb–value
220 220 255

Rasterised (bitmap) image

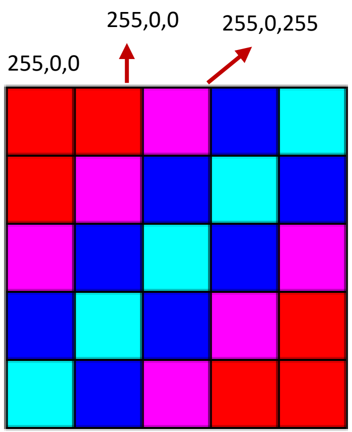


Image (5 x 5 pixels)

R grid (array)

255	255	255	0	0
255	255	0	0	0
255	0	0	0	255
0	0	0	255	255
0	0	255	255	255

G grid (array)

0	0	0	0	255
0	0	0	255	0
0	0	255	0	0
0	255	0	0	0
255	0	0	0	0

B grid (array)

0	0	255	255	255
0	255	255	255	255
255	255	255	255	255
255	255	255	255	0
255	255	255	0	0

Contents of (uncompressed) image file

Size: 25 x 24 = 600 bits 25 x 3 = 75 bytes

256

125

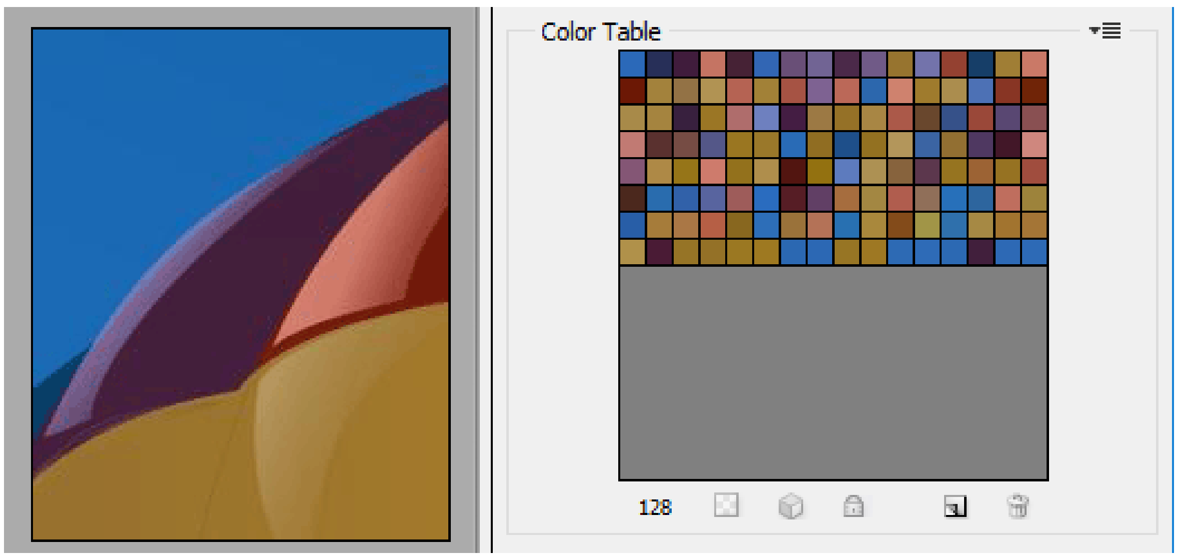
125 x 256 x 3 bytes = 96000 byte

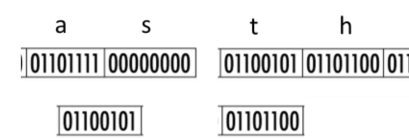
even only few colors exist

Indexed Colour

- Popular technique for compressing (Raster/bitmap) images
- Instead of representing all colours, save only those used
- Similar to compressing text with **keyword encoding**, but with colours instead

Word	Symbol
as	^
the	~
and	+
that	$
must	&
well	%
those	#



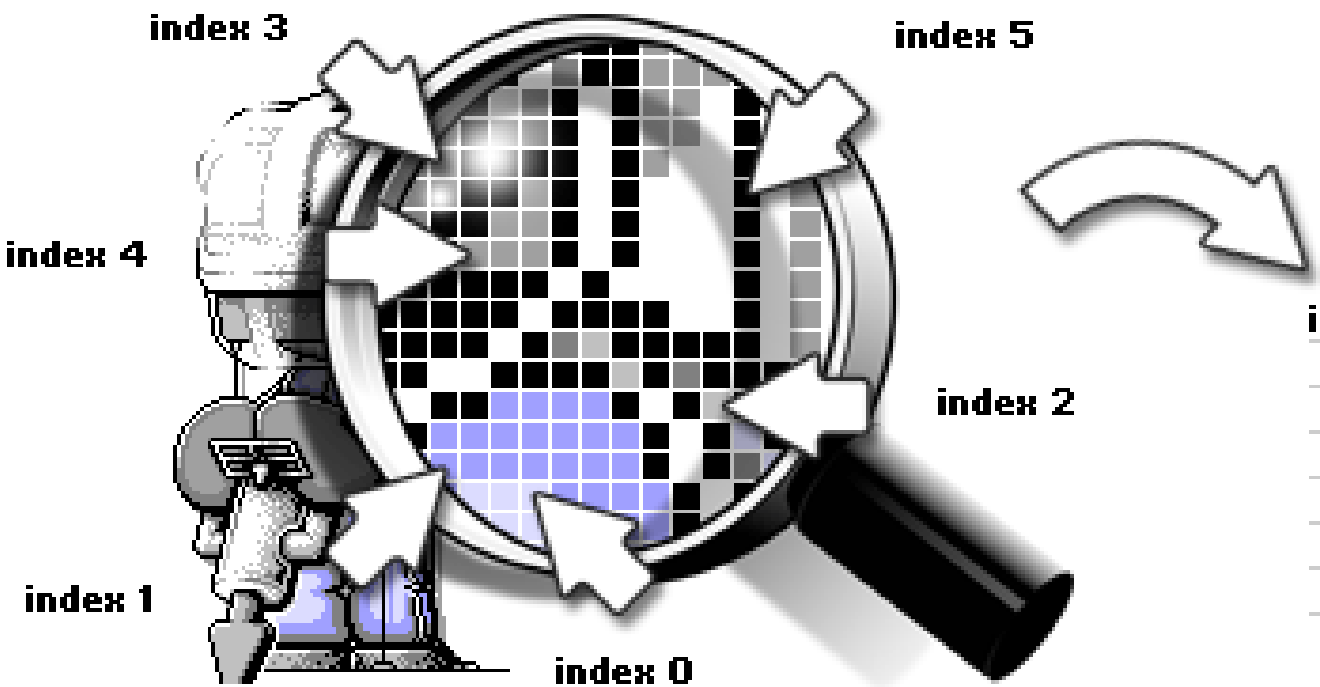


The human body is composed of many
independent systems, such ^ ~ circula
~ respiratory system, + ~ reproductive
only & each system work independen
interact + cooperate ^ %. Overall healt
function of ~ %- being of separate syst
how # separate systems work in conce

An image with its table (pallet) of 128 indexed colours

Word	Symbol
as	^
the	~
and	+
that	$
must	&
well	%
those	#

The human body is composed of many independent systems, such ^ ~ circulatory syst ~ respiratory system, + ~ reproductive system. only & each system work independently, they interact + cooperate ^ %. Overall health is a function of ~ %- being of separate systems, ^ % how # separate systems work in concert.



index	rgb-value			color
0	220	220	255	
1	160	160	128	
2	192	192	192	
3	128	128	128	
4	160	160	160	
5	0	0	0	

Indexed image example

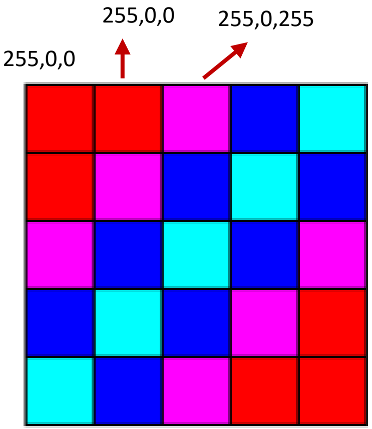


Image (5 x 5 pixels)

R grid (array)

255	255	255	0	0
255	255	0	0	0
255	0	0	0	255
0	0	0	255	255
0	0	255	255	255

G grid (array)

0	0	0	0	255
0	0	0	255	0
0	0	255	0	0
0	255	0	0	0
255	0	0	0	0

B grid (array)

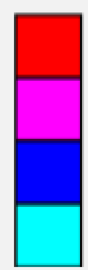
0	0	255	255	255
0	255	255	255	255
255	255	255	255	255
255	255	255	255	0
255	255	255	0	0

Contents of uncompressed image file

Pixel grid

0	0	1	2	3
0	1	2	3	2
1	2	3	2	1
2	3	2	1	0
3	2	1	0	0

Colour table (pallet)



	R	G	B
0	255	0	0
1	255	0	255
2	0	0	255
3	0	255	255

Contents of indexed (compressed) image file

Indexed image example

For your info

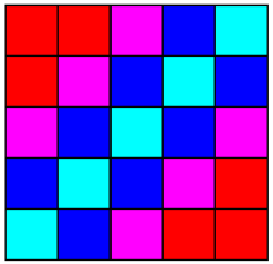


Image (5 x 5 pixels)

R grid (array)

255	255	255	0	0
255	255	0	0	0
255	0	0	0	255
0	0	0	255	255
0	0	255	255	255

G grid (array)

0	0	0	0	255
0	0	0	255	0
0	0	255	0	0
0	255	0	0	0
255	0	0	0	0

B grid (array)

0	0	255	255	255
0	255	255	255	255
255	255	255	255	255
255	255	255	255	0
255	255	255	0	0

Contents of uncompressed image file

Pixel grid

0	0	1	2	3
0	1	2	3	2
1	2	3	2	1
2	3	2	1	0
3	2	1	0	0

Colour table (pallet)

	R	G	B
0	255	0	0
1	255	0	255
2	0	0	255
3	0	255	255

Contents of indexed (compressed) image file

Assume total 256 colors,
8 bits
3bitsR 8
3bitsG 8 8x8x4=256
2bitsB 4

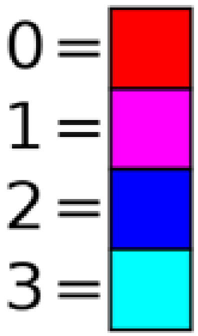
Number of bits	
Uncompressed image	Compressed (indexed) image
25 pixels x 24 bits per pixel = **600** bits	(25 pixels x 8 bits per pixel) + (4 colours x 24 bits per colour) = **296** bits

- Actual compression ratios depend on the image resolution and colour content

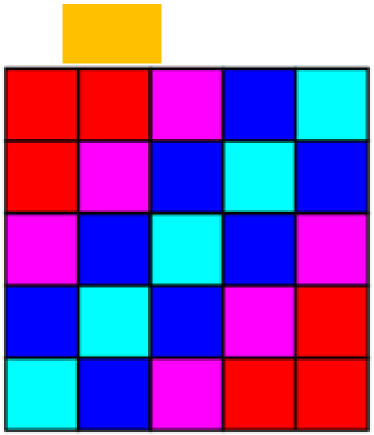
If 4 colors, 25*2+4*24 =146 00 01 10 11

If 16 colors, 25*4+4*24 =196

Indexed Colour



0	0	1	2	3
0	1	2	3	2
1	2	3	2	1
2	3	2	1	0
3	2	1	0	0



4-color

2 bits

16-color older PC

4 bits

256-color old PC

8 bits

3bits R 8
3bits G 8 $2^{(3+3+2)}$ = 8x8x4=256
2bits B 4

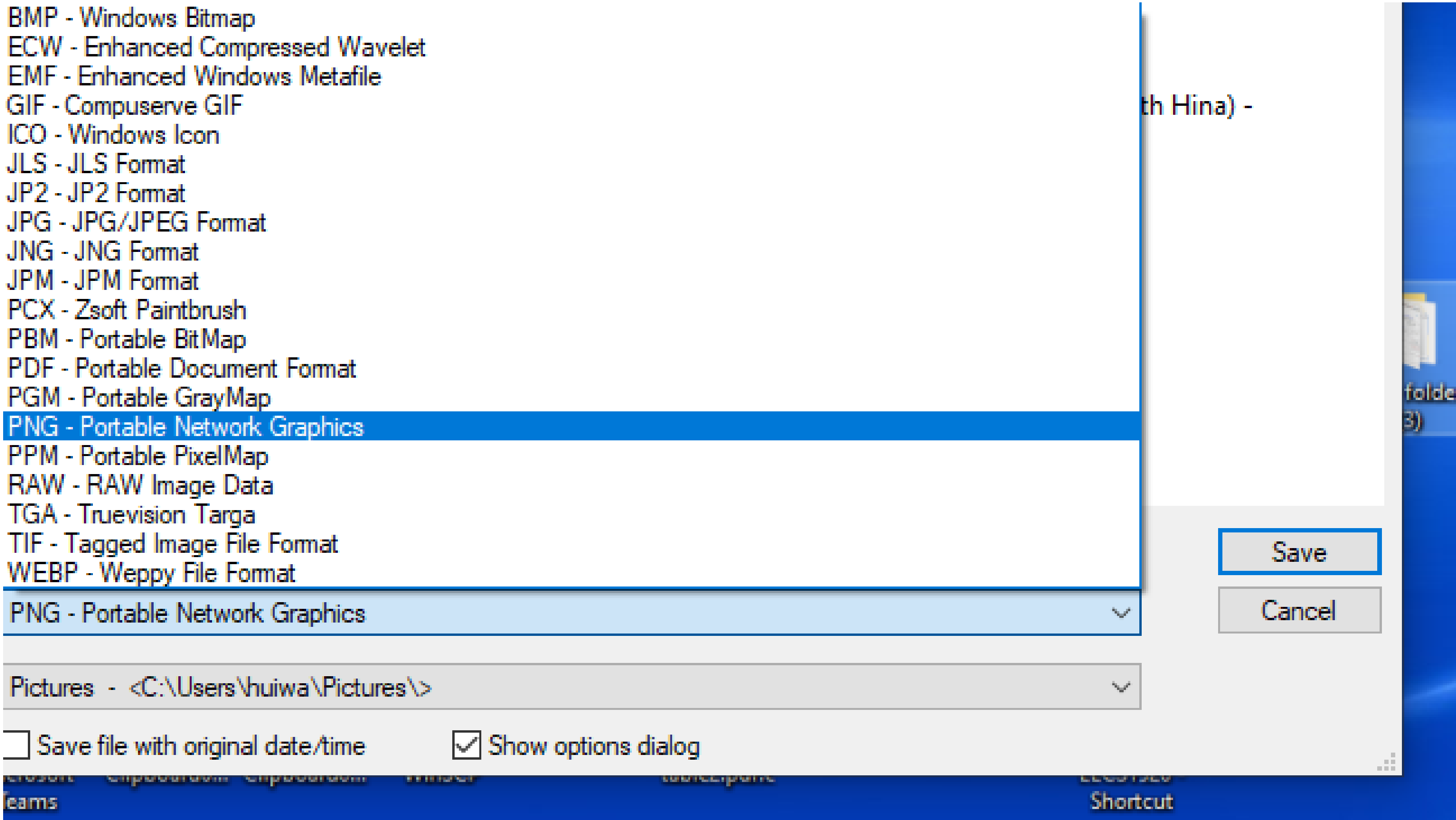
Truecolor New PC

24 bits

8bits R 256
8bits G 256 256x256x256=16M
8bits B 256

For your info

Example Image Formats (Raster)

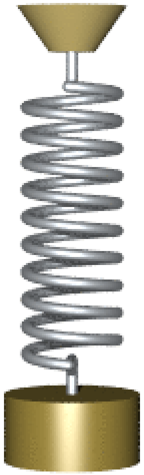


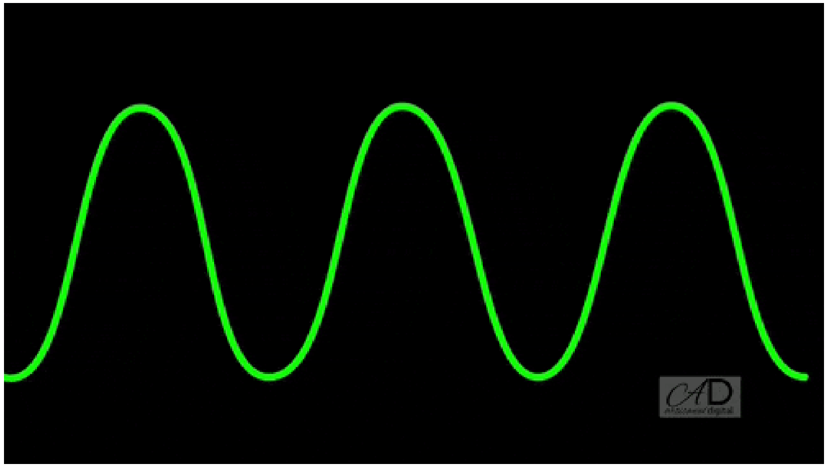
Example Image Formats (Raster)

GIF (Graphics Interchange Format, pronounced "Jiff")

- **Lossless**, indexed colours (maximum of 256 colors), coding related to Huffman's
- Compression ratios of 10% [reference]
- Allows transparency and animation

256-color





Example Image Formats (Raster)

GIF (Graphics Interchange Format, pron. “Jiff” or “ghif”)

- **Lossless**, indexed colours (maximum of 256), coding related to Huffman’s
- Compression ratios of 10% [reference]
- Allows transparency and animation

PNG (Portable Network Graphics)

- **Lossless** compression using indexed colours, run-length encoding
- Compression ratios of 7% [reference]
- Allows transparency
- PNG supports 8-bit color (256 color) like GIF, but also supports 24-bit true color RGB (256x256x256), like JPG does

$2^{(3+3+2)}$ 8*8*6

PNG and GIF best used for line drawings, logos, or diagrams

Photos: GIF – low quality (256 col’s) PNG – too large (lossless);

Consider lossy JPEG

Example Image Formats (2)

JPEG (Joint Photographic Experts Group)

- Ideal format for photos, **bad for images with text or sharp lines (lossy)**
- Compression ratios of 1-10%, depending on quality [reference]
- **Lossy**; compression technique:

  Divide the entire image into blocks of 8 x 8 pixels

  For each block:

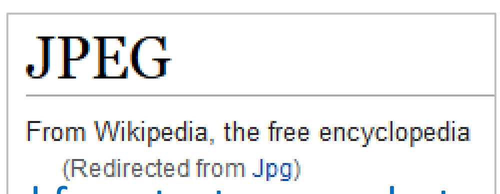
  Store the average intensity

  For each pixel in the block:

  Store the difference between each pixel's intensity and the average intensity of the block

  Differences smaller than a threshold are ignored (this is the lossy part of the compression)

  The threshold depends on quality setting chosen by the user

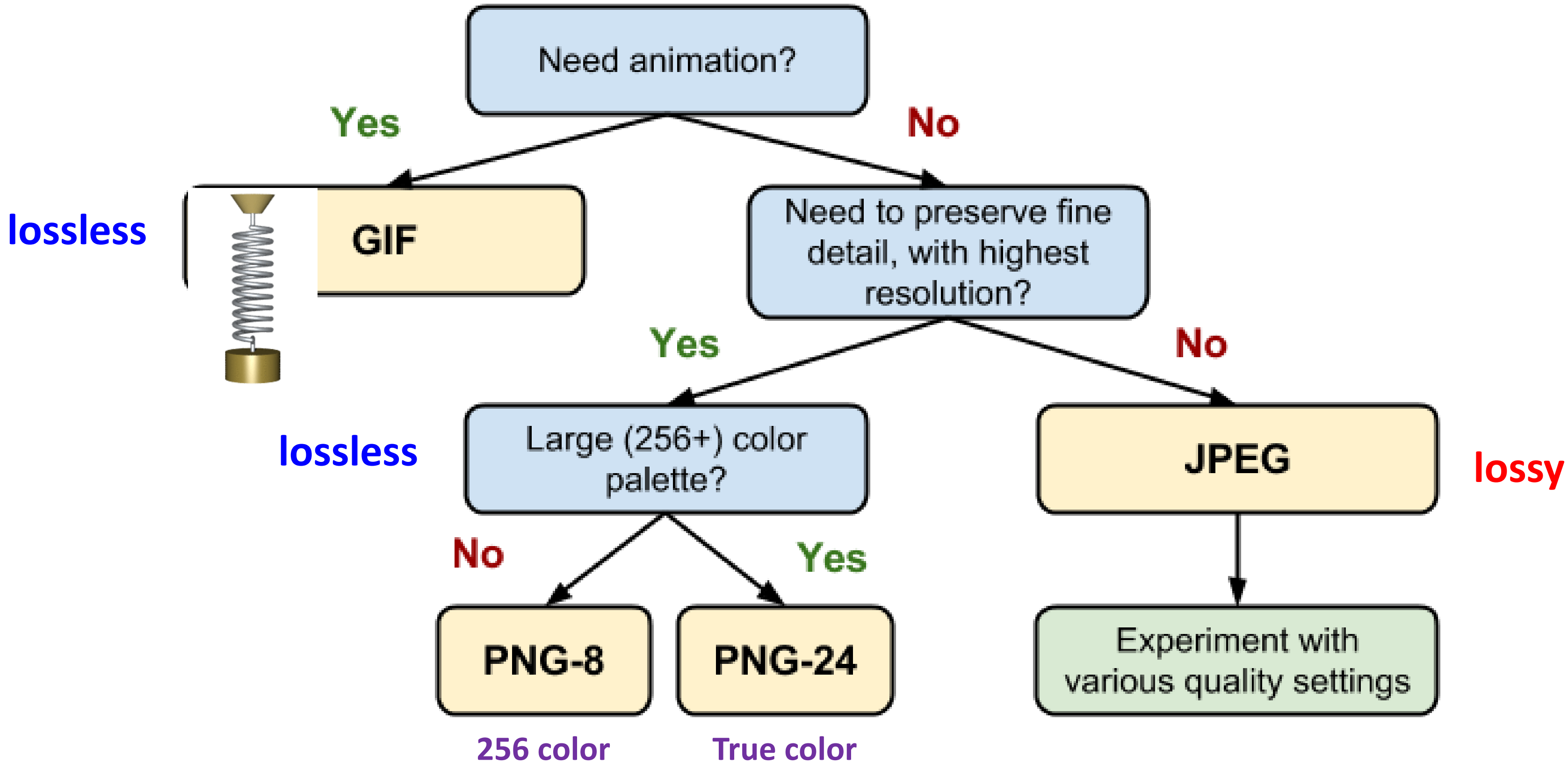


PNG used for a text screenshot

JPEG used for a text screenshot

Guide to Choosing Image Format

- Google has an [excellent guide]
  - Aimed at developers
  - “TL;DR” summaries for each section



File Format	Number of Colors	Lossless
PNG	unlimited	✓
JPG	unlimited	✗
GIF	256	✓

summary

- RGB: RRGGBB 3 bytes
- Pixel
- Vector
- Raster
  - Indexed
- formats
  - **Lossless**: GIF (animation), PNG
  - **Lossy**: JPG

RGB = (178,102,255) #B266FF

24 bits

8bits R 256
8bits G 256 256x256x256=16M
8bits B 256

Digital Data Representation

Topics for the next few weeks

Representing Numbers

Representing Text

Representing Audio

Representing Images

Representing Video

Video

Actions speak louder than words
(Hmm… Perhaps I should make this into a video)

Analog TV of the past century

- **NTSC :** National Television Standards Committee (1952)
- **CRT:** Cathode Ray Tube
  - **Electron Gun** sweeps the lines with phosphor dots
  - Visibly flickers when you refresh it too slowly
- **Resolution:** 525 lines *fully* refreshed at 30 Hz (every 2 fields)
  - **Over-scanning:** 525 lines reduced to 484 *visible* lines
  - **Aspect Ratio:** 4:3 (screen width-to-height ratio)
- **Frame/Field Rate:** 30 [fully refreshed] frames per second (60 fields per second)
  - **Interlaced scan:** scan 1 field at a time (for faster *partial* refresh)
  - **Fields:** frames divided into 2 fields (e.g., odd # vs even #)
- **Colour TV:** composite signal of luminance (the black-and-white part) & chrominance (how different the colour is from gray) (Critics: NTSC = "Never Twice the Same Colour"; PAL was an improvement, so was SECAM)

For your information

Today

- CRTs are not used anymore
  - Interlaced scan is not needed (but remains in some surviving standards: 1080i vs 1080p)
  - Progressive scan is the norm
  - No hidden lines are needed, since there is no beam to move
- Transmission and storage are digital
  - Compression can be used
  - Resolution is much higher
  - Colour was present in the new standards from the beginning
- New output devices have much higher accuracy
  - High resolution requires more careful preparation
  - More research on human perception exists
  - High Dynamic Range techniques

Virtual Reality

- “Immersive” technologies
  - External: CAVE
  - Head-mounted displays
- Images change with head movements/rotations to always show proper perspective
- Challenges
  - Processing power
  - More complex hardware
  - Images need to be generated much faster than for movies
    - Delays cause motion sickness
  - Cannot be used for too long

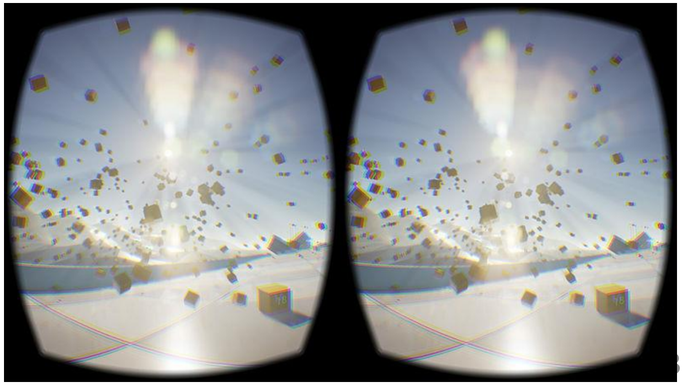
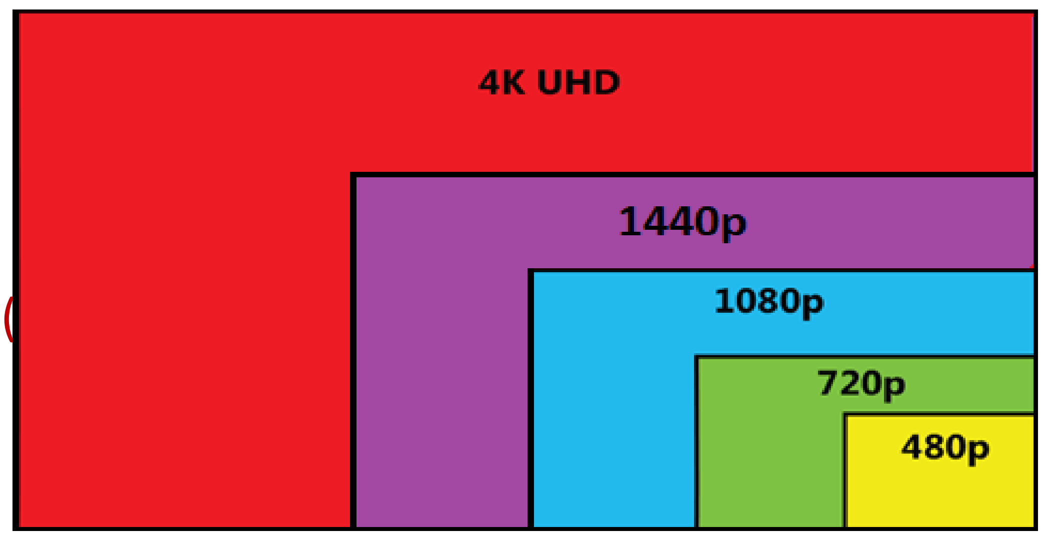


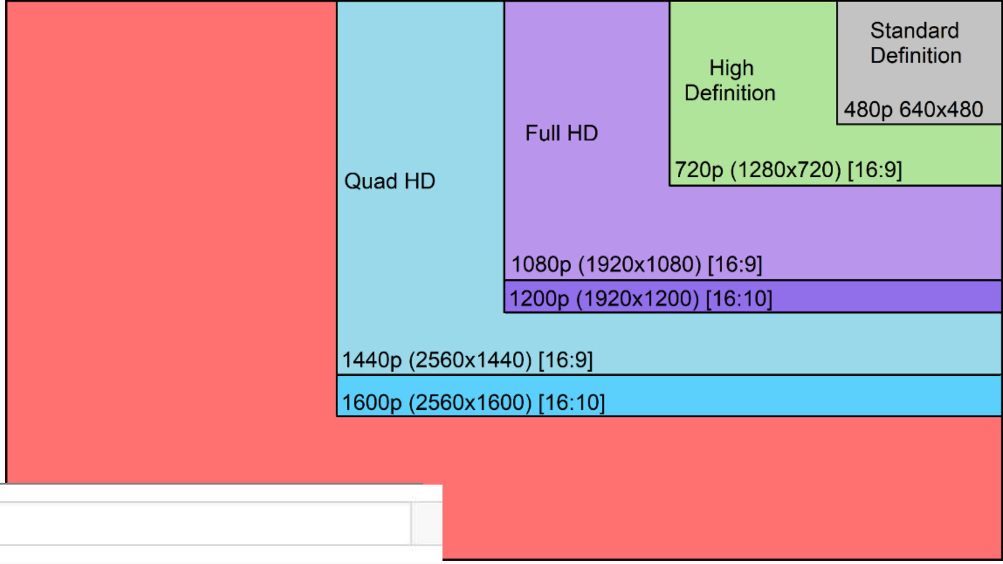
Image by Ats Kurvet, via Wikipedia

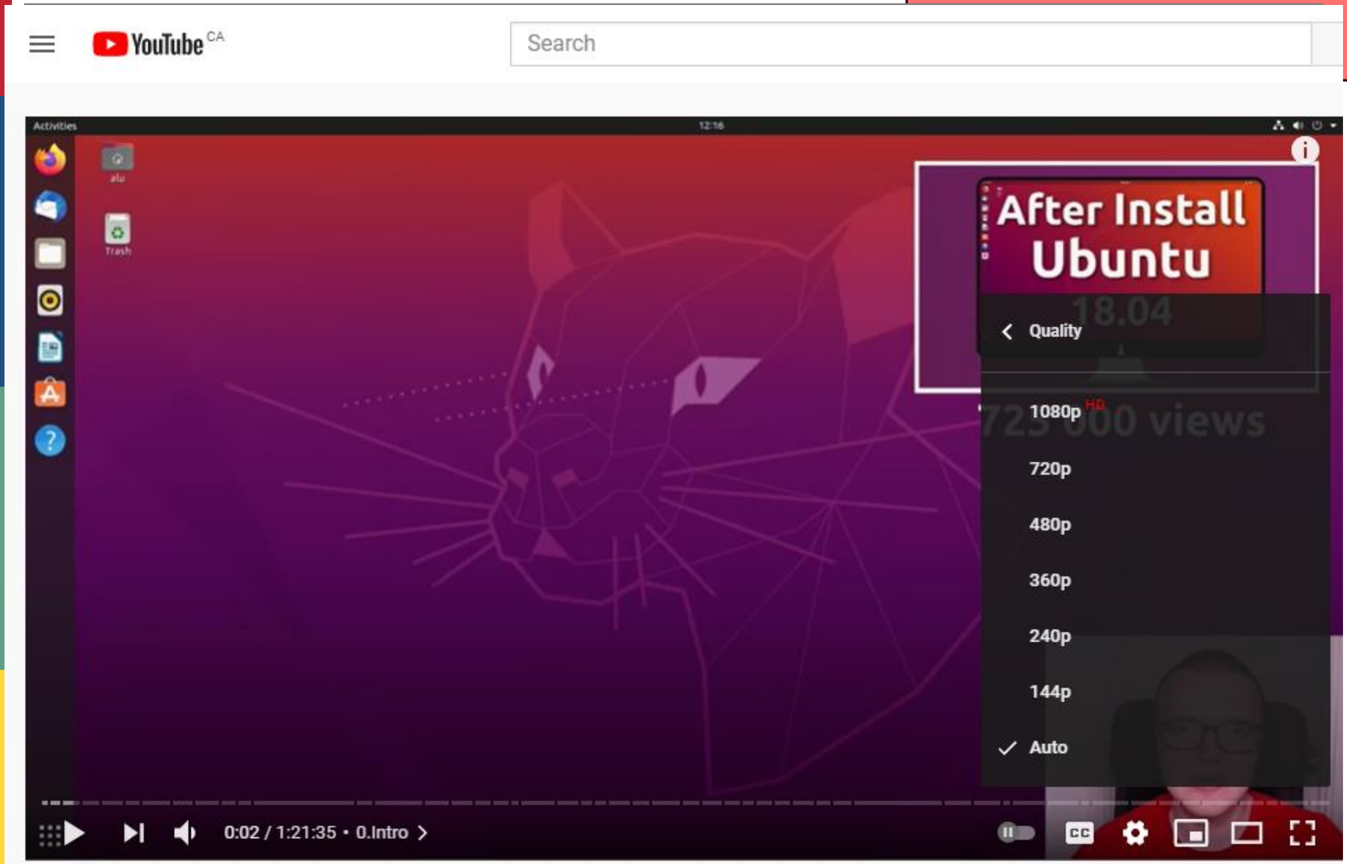
Video

- Comprised of frames of still images (combined with audio)
- When viewed in rapid succession, the images give the appearance of motion
  - Typical **frame rates** are 24~30 **frames per second** to 60 **fps**
- **Aspect Ratio:** 16:9 typically
- **Resolution:**
  - **4096**x2160 pixels, **4K**
  - 1920x**1080**, **Full HD**
  - 1280x**720**, **HD**
  - 853x**480**, **SD**



- Uncompressed video occupies vast storage space (more than 100 MB per second for Full HD) 1920*1080*3*30/sec
  - Compression techniques are usually used 186M bytes/sec


Standard Definition
480p 640x480
High Definition
720p (1280x720) [16:9]
Full HD
1080p (1920x1080) [16:9]
1200p (1920x1200) [16:10]
Quad HD
1440p (2560x1440) [16:9]
1600p (2560x1600) [16:10]

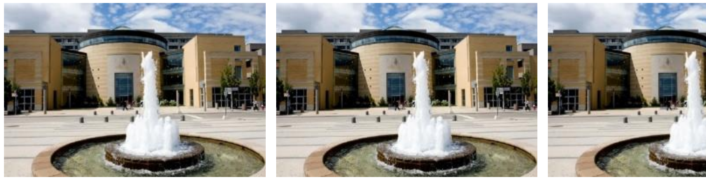

YouTube CA
Search
After Install Ubuntu
18.04
Quality
1080p HD
720p
480p
360p
240p
144p
Auto
0:02 / 1:21:35 • 0.Intro

Techniques for Video Compression

1 Each frame may contain duplicated info (e.g., block of blue sky)

2 Typically little change between successive video frames

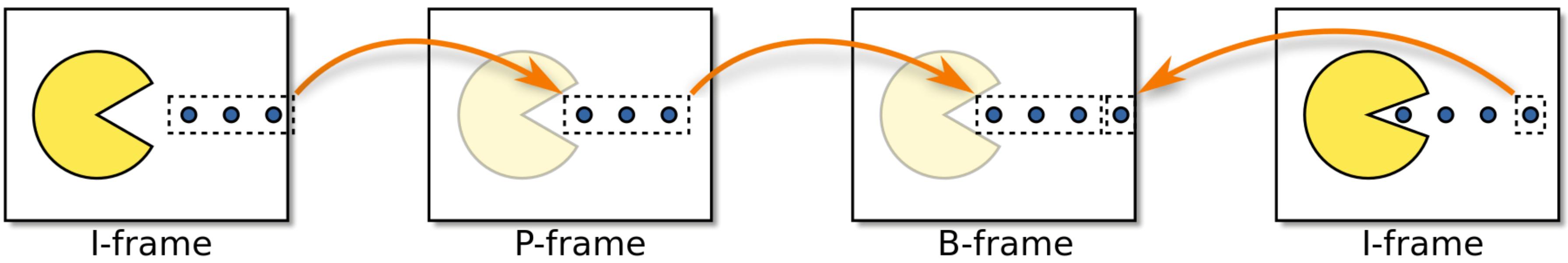
- E.g., scene with little motion.
- Use techniques to remove that redundant info



- **Spatial** (Intra-frame) techniques (e.g., M-JPEG)
  - Remove duplicate info from within the same frame to reduce file size
  - Block of clear sky, store color and coordinates (similar to run-length)

- **Temporal** (Inter-frame) techniques (e.g., MPEG)
  - Use data from nearby frames (before or after) to reduce the file size
  - **Key-frames** (I-frames):
    - typically compressed (separately from P B), using only spatial techniques
    - can be reproduced independently
    - inserted automatically at scene changes and/or at regular intervals to preserve quality of streaming or playback
  - **Delta frames: P-frames & B-frames:** Other frames encoded by saving differences between it and previous key-frame (P=Predictive, B=Bidirectional)
  - If a key-frame is lost (e.g., during streaming) or skipped (e.g., when skipping forwards or backwards) the displayed video can be distorted
  - Suitable for video distribution, but not video editing

Temporal (Inter-frame) techniques



Three types of *pictures* (or frames) are used in video compression: I, P, and B frames.

- An **I-frame** (Intra-coded picture) is a complete image, like a JPG or BMP image file.

- A **P-frame** (Predicted picture) holds only the changes in the image from the previous frame. For example, in a scene where a car moves across a stationary background, only the car's movements need to be encoded. The encoder does not need to store the unchanging background pixels in the P-frame, thus saving space. P-frames are also known as *delta-frames*.

- A **B-frame** (Bidirectional predicted picture) saves even more space by using differences between the current frame and both the preceding and following frames to specify its content.

Bitrate vs Size

- If the bitrate is too low, the video becomes pixelated

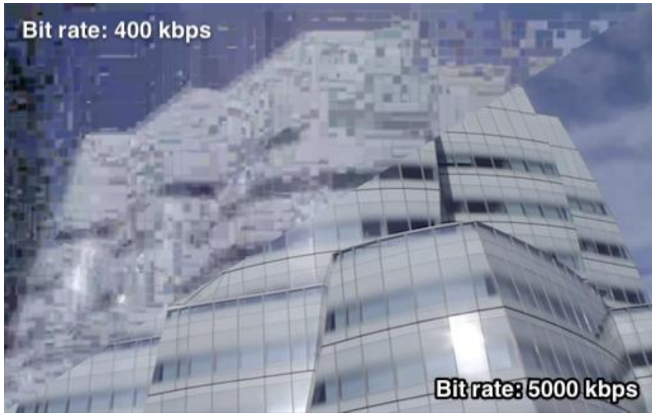


Image courtesy of vimeo.com

Constant bitrate (CBR): same bit rate throughout the video

Variable bitrate (VBR): can vary (up to a maximum)

  - Allows higher bitrates during fast motion scenes, and lower bitrates when there is little motion

- Examples
  - EECS 1520 Zoom Lecture recordings: HD(1920x968x25fps) approx. 0.2 Mbps VBR, static slides are very easy to compress!
  - Netflix: 5 Mbps at 1080p, up to 16 Mbps for 4K
  - Blu-ray: about 40 Mbits/s for **1080p** (older, less efficient scheme)
  - *"Do you need a Gigabit fiber?"*

Effect of Bitrate

To further reduce file size, video files (like audio files) can limit their bitrate

- If the bitrate is too low the video becomes pixelated

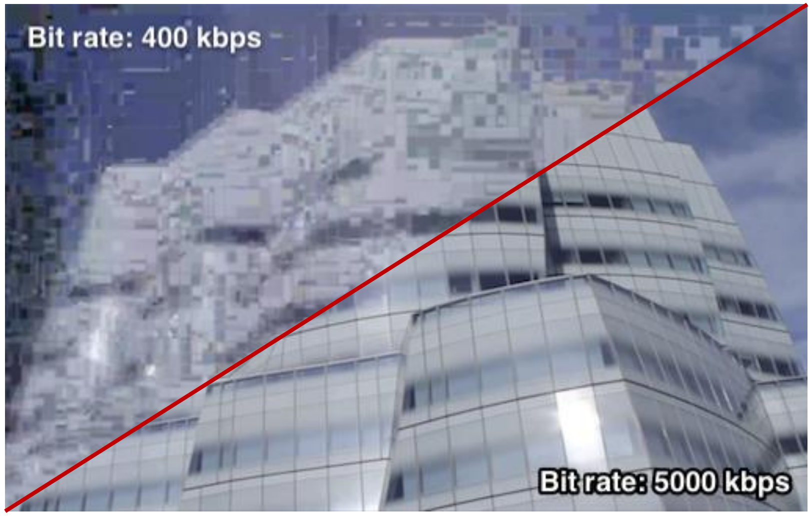


Image courtesy of vimeo.com

Constant bitrate (CBR): same bit rate throughout the video

Variable bitrate (VBR): can vary (up to a maximum)

- Allows higher bitrates during fast motion scenes, and lower bitrates when there is little motion

Better than CBR, but complicated

Example Codecs

- **HuffYUV** (**Huff**man, **YUV** colour space)
  - **Lossless**, Huffman encoding, compression ratios around 47% [reference]
- **MPEG-2**
  - **Lossy**, temporal and spatial compression, Huffman encoding
  - Compression ratios around 3.3%
    - Highly dependent on bitrate and other encoding options
- **H.264**
  - **Lossy**, temporal and spatial compression, frame prediction
    - Predicts the next frame using spatial and temporal information, compares with the actual frame, saves only the difference
  - Compression ratios around 1.6%
    - Highly dependent on bitrate and other encoding options
  - Perceived quality higher than MPEG-2 at same bitrate [reference]
- **H.265 (HEVC)**
  - successor of H.264: further 50 % reduction of size, much more complex!
  - Often drains battery faster

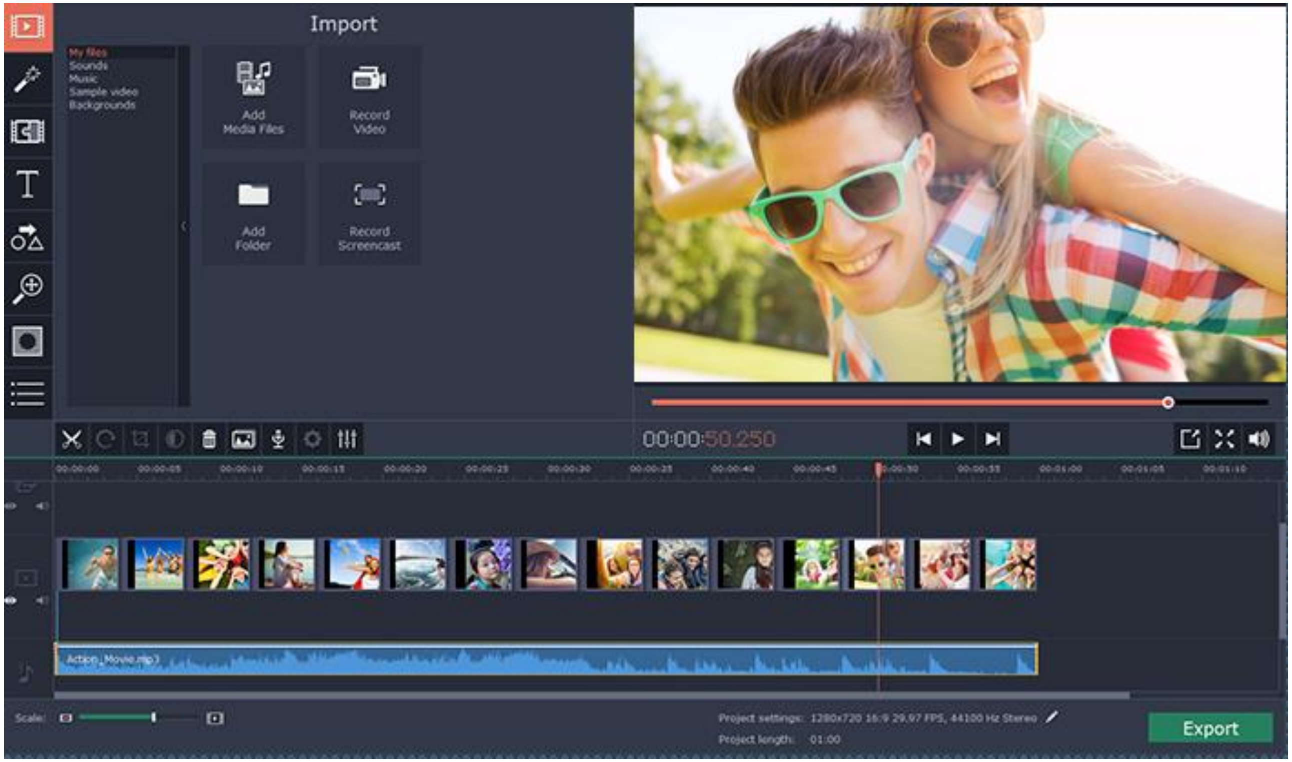
File Containers

- Combine a video stream and an audio stream into one file
- Does not necessarily indicate the codecs used
- Some allow for additional data:
  - Additional video stream representing different viewing angles
  - Multiple language audio tracks
  - Optional subtitles
- Containers are indicated by the file extension
  - **.avi** (Audio Video Interleave): not good for streaming
  - **.mp4** (MPEG-4): supports multiple audio/video streams and subtitles, but limited codecs, is popular on mobile devices
  - **.mkv** (Matroska): like mp4, but supports unlimited streams and unlimited codecs, poor support on mobile devices

Example Q: which of the following is container, is codec HuffYUV, mp4

Video Editors

- iMovie (macOS, iOS)
- OpenShot (Linux, macOS, Windows)
- VSDC Free Video Editor (Windows)
- Adobe Premiere Pro (macOS, Windows; professional app)



summary

- Techniques for Video Compression
  - spatial
  - temporal   key frame, delta frame

- Bitrate   CBR vs VBR
- Codecs
  - Huffyuv   **lossless**
  - MPEG
  - H.264
  - H.265

lossy

- File containers
  - .avi
  - .mp4
  - .mkv