Data Representation

Interpreting bits to give them meaning

Part 2: Media

Text, Web Pages, Pictures, Sound/Music, Video

Notes for CSC 100 - The Beauty and Joy of Computing
The University of North Carolina at Greensboro

Data is more than just numbers!

Data is stored using bits but represents many things:

- Documents
- Pictures
- Sound/music
- Video
- ...

How does this work?

- *File formats*: Structure bits in such a way that mapping between bits and what they represent is unambiguous
  - Standardized or open file formats
    - Specified so that anyone can write programs for them (JPEG, MPEG (and MP3), OpenDocument, HTML, ...)
    - "Open" and "standardized" doesn't mean "free" (MP3, GIF, ...)
- A *data capture* or creation program builds the file in the appropriate format
- A *rendering* program converts the file format to a recognizable form (image viewer, web browser, video player, ...)

Representations of Text

ASCII

When everything is 0's and 1's, how do you store or transmit something like "Hello World"?

Answer: Encode characters as binary strings

In early days there were several "encodings"

Most common for basic US/English use is *ASCII*

- **A**merican **S**tandard **C**ode for **I**nformation **I**nterchange
- Uses 7 bits per character
- Typically embedded in 8-bit bytes
- Hexadecimal bytes -> ASCII examples to the right

Less U.S.-centric encoding: Unicode

Some Special Characters

07 Bell	0C Form Feed
08 Backspace	0D Carriage Ret
0A New line	27 ESC

Punctuation Samples

20 Space	24 $	2E .
21 !	2B +	3A :
22 "	2C ,	3F ?

Digits

30 0 ... 39 9

Letters

41 A	4E N	61 a	6e n
42 B	4F O	62 b	6f o
43 C	50 P	63 c	70 p
44 D	51 Q	64 d	71 q
45 E	52 R	65 e	72 r
46 F	53 S	66 f	73 s
47 G	54 T	67 g	74 t
48 H	55 U	68 h	75 u
49 I	56 V	69 i	76 v
4A J	57 W	6A j	77 w
4B K	58 X	6B k	78 x
4C L	59 Y	6C l	79 y
4D M	5A Z	6D m	7A z

Representations of Text

ASCII - What does the highlighted part say?

```
0000000: 4c65 7420 7573 206e 6f74 2077 616c 6c6f
0000010: 7720 696e 2074 6865 2076 616c 6c65 7920
0000020: 6f66 2064 6573 7061 6972 2e20 4920 7361
0000030: 7920 746f 2079 6f75 2074 6f64 6179 206d
0000040: 7920 6672 6965 6e64 7320 2d2d 2073 6f20
0000050: 6576 656e 2074 686f 7567 6820 7765 2066
0000060: 6163 6520 7468 6520 6469 6666 6963 756c
0000070: 7469 6573 206f 6620 746f 6461 7920 616e
0000080: 6420 746f 6d6f 7272 6f77 2c20 4920 7374
0000090: 696c 6c20 6861 7665 2061 2064 7265 616d
00000a0: 2e20 4974 2069 7320 6120 6472 6561 6d20
00000b0: 6465 6570 6c79 2072 6f6f 7465 6420 696e
00000c0: 2074 6865 2041 6d65 7269 6361 6e20 6472
00000d0: 6561 6d2e 0a0a 4920 6861 7665 2061 2064
00000e0: 7265 616d 2074 6861 7420 6f6e 6520 6461
00000f0: 7920 7468 6973 206e 6174 696f 6e20 7769
0000100: 6c6c 2072 6973 6520 7570 2061 6e64 206c
0000110: 6976 6520 6f75 7420 7468 6520 7472 7565
0000120: 206d 6561 6e69 6e67 206f 6620 6974 7320
0000130: 6372 6565 643a 2022 5765 2068 6f6c 6420
0000140: 7468 6573 6520 7472 7574 6873 2074 6f20
0000150: 6265 2073 656c 662d 6576 6964 656e 742c
0000160: 2074 6861 7420 616c 6c20 6d65 6e20 6172
0000170: 6520 6372 6561 7465 6420 6571 7561 6c2e
```

Some Special Characters

```
07 Bell        0C Form Feed
08 Backspace   0D Carriage Ret
0A New line    27 ESC
```

Punctuation Samples

```
20 Space   24 $    2E .
21 !       2B +    3A :
22 "       2C ,    3F ?
```

Digits

```
30 0   ...   39 9
```

Letters

```
41 A   4E N      61 a   6e n
42 B   4F O      62 b   6f o
43 C   50 P      63 c   70 p
44 D   51 Q      64 d   71 q
45 E   52 R      65 e   72 r
46 F   53 S      66 f   73 s
47 G   54 T      67 g   74 t
48 H   55 U      68 h   75 u
49 I   56 V      69 i   76 v
4A J   57 W      6A j   77 w
4B K   58 X      6B k   78 x
4C L   59 Y      6C l   79 y
4D M   5A Z      6D m   7A z
```

Representations of Text

ASCII - The full hex dump!

```
0000000: 4c65 7420 7573 206e 6f74 2077 616c 6c6f  Let us not wallo
0000010: 7720 696e 2074 6865 2076 616c 6c65 7920  w in the valley
0000020: 6f66 2064 6573 7061 6972 2e20 4920 7361  of despair. I sa
0000030: 7920 746f 2079 6f75 2074 6f64 6179 206d  y to you today m
0000040: 7920 6672 6965 6e64 7320 2d2d 2073 6f20  y friends -- so
0000050: 6576 656e 2074 686f 7567 6820 7765 2066  even though we f
0000060: 6163 6520 7468 6520 6469 6666 6963 756c  ace the difficul
0000070: 7469 6573 206f 6620 746f 6461 7920 616e  ties of today an
0000080: 6420 746f 6d6f 7272 6f77 2c20 4920 7374  d tomorrow, I st
0000090: 696c 6c20 6861 7665 2061 2064 7265 616d  ill have a dream
00000a0: 2e20 4974 2069 7320 6120 6472 6561 6d20  . It is a dream
00000b0: 6465 6570 6c79 2072 6f6f 7465 6420 696e  deeply rooted in
00000c0: 2074 6865 2041 6d65 7269 6361 6e20 6472   the American dr
00000d0: 6561 6d2e 0a0a 4920 6861 7665 2061 2064  eam...I have a d
00000e0: 7265 616d 2074 6861 7420 6f6e 6520 6461  ream that one da
00000f0: 7920 7468 6973 206e 6174 696f 6e20 7769  y this nation wi
0000100: 6c6c 2072 6973 6520 7570 2061 6e64 206c  ll rise up and l
0000110: 6976 6520 6f75 7420 7468 6520 7472 7565  ive out the true
0000120: 206d 6561 6e69 6e67 206f 6620 6974 7320   meaning of its
0000130: 6372 6565 643a 2022 5765 2068 6f6c 6420  creed: "We hold
0000140: 7468 6573 6520 7472 7574 6873 2074 6f20  these truths to
0000150: 6265 2073 656c 662d 6576 6964 656e 742c  be self-evident,
0000160: 2074 6861 7420 616c 6c20 6d65 6e20 6172   that all men ar
0000170: 6520 6372 6561 7465 6420 6571 7561 6c2e  e created equal.
```

Formatted Text

HTML

ASCII provides letters - what about fonts, sizes, etc?

One option: HTML - **H**yper**T**ext **M**arkup **L**anguage

- The "language of web pages"
- "Markup" indicates formatting/style
- All characters are just regular character set (like ASCII) - including markup
- Must be *rendered* to convert character-based markup to formatted text
- A lot of formatting is now in CSS - **C**ascading **S**tyle **S**heets
- Much more involved than these examples!

HTML Source

```
This is formatted text, which can
be <b>bold</b> or <i>italic</i> or
<u>underlined</u> or <span style="
font-size: 150%">big</span>
or <span style="font-size: 50%"
>small</span> or ...
```

Rendered Text

This is formatted text, which can be **bold** or *italic* or underlined or big or small or ...

Pictures

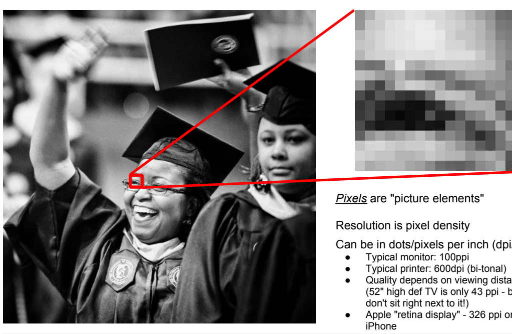
Grayscale

Grayscale images have levels of intensity, but no color

- More information than bi-tonal black and white (like fax machines or most printers)
- Less information than color

Pictures

Grayscale - Pixels



Pixels are "picture elements"

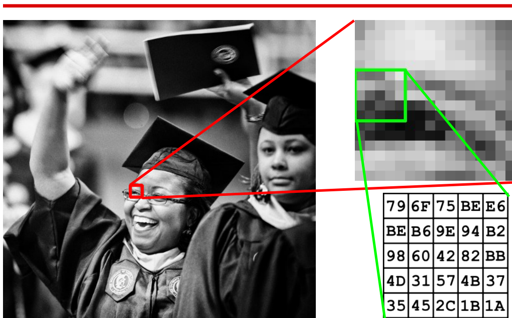
Resolution is pixel density

Can be in dots/pixels per inch (dpi/ppi)

- Typical monitor: 100ppi
- Typical printer: 600dpi (bi-tonal)
- Quality depends on viewing distance (52" high def TV is only 43 ppi - but you don't sit right next to it!)
- Apple "retina display" - 326 ppi on iPhone

Pictures

Grayscale - Pixels as numbers

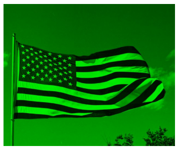


79	6F	75	BE	E6
BE	B6	9E	94	B2
98	60	42	82	BB
4D	31	57	4B	37
35	45	2C	1B	1A

Number of levels *typically* one byte

Pictures

Color - Three "color planes" (red, green, blue)



Each color plane is just like a grayscale image, with same issues:

- Resolution (ppi)
- Depth (bpp)

"24 bit color" means 8 bits per pixel in each of the 3 colors

Pictures

Why does this work?

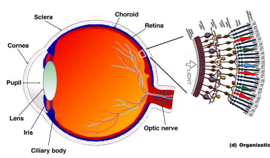


Fig. 1.1. A drawing of a section through the human eye with a schematic enlargement of the retina.

"Rods" and "cones" signal our brain about light we receive in our eye

Rods: monochrome only

Cones: Color - in three kinds, red, green, and blue

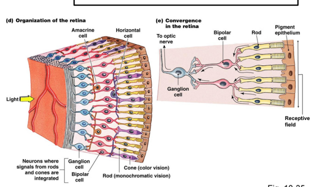
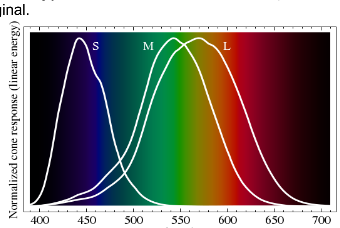


Fig. 10-35

Pictures

Why does this work?

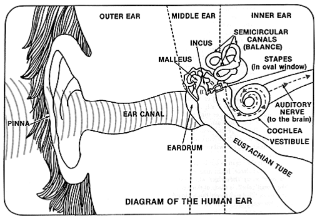
Bottom line: If humans can only perceive three colors (red, green, and blue) then reconstructing just those three colors allow us to *perceive* everything just as in an original.



Interesting question: What if someone were born with a mutation that gave them purple and yellow receptors?

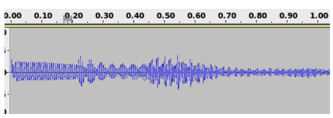
Sound
What is sound?

Sound is just rapid fluctuations in air pressure, detected by the (somewhat delicate!) organs in our ears



Sound
Sound waveforms

We can plot changes in pressure over time:

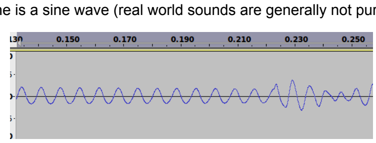


Main components:

- *Intensity* (how much pressure changes): We perceive this as "loudness" and in graph would be reflected in larger fluctuations
- *Frequency*: How rapid are the fluctuations?

Sound
Sound waveforms - Zooming in!

"Pure" tone is a sine wave (real world sounds are generally not pure!)



One *cycle* here is approximately 0.150 seconds to 0.157 seconds:

- *Period* is 0.007 seconds
- *Frequency* is 1/0.007 = 142.857... Hz (for "Hertz")
  - For reference, "middle C" is around 261.626 Hz
  - An octave doubles/halves frequency, so this note is a probably something like a "D below middle C" (which is 146.8 Hz)

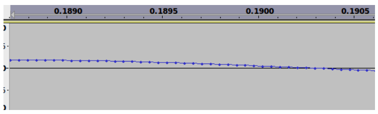
Question: How do we make this digital?

Sound

Sound waveforms - Zooming in even more!

Answer: We ***sample*** the waveform many times per second.

This is zoomed in enough where you can see actual samples:



Quality of sound reproduction depends on sample rate (samples per second):

- In this example, 22 samples between 0.1890 and 0.1900
  - *So 22/(0.190-0.189) = 22,000 samples per second*
- CD sound: 44,100 samples/second
- Typical DVD sound: 48,000 samples/second

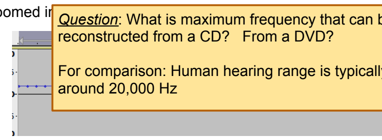
Nyquist Theorem: Perfect reconstruction of signals with frequency <= F if you sample at (2/F) samples/second

Sound

Sound waveforms - Zooming in even more!

Answer: We ***sample*** the waveform many times per second.

This is zoomed i



Question: What is maximum frequency that can be reconstructed from a CD?   From a DVD?

For comparison: Human hearing range is typically 20 Hz to around 20,000 Hz

Quality of sound reproduction depends on sample rate (samples per second):

- In this example, 22 samples between 0.1890 and 0.1900
  - *So 22/(0.190-0.189) = 22,000 samples per second*
- CD sound: 44,100 samples/second
- Typical DVD sound: 48,000 samples/second

Nyquist Theorem: Perfect reconstruction of signals with frequency <= F if you sample at (2/F) samples/second

Video

Basics

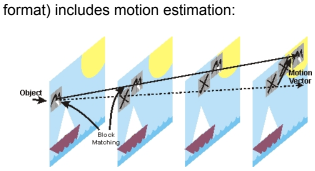
Can be viewed as a series of still images

- 24 frames per second (fps) in movies
- 30 fps in US television

Motion-JPEG (M-JPEG) is exactly this: JPEG image for each frame

- Benefit: Very simple format to work with and edit
- Drawback: Doesn't take advantage of temporal similarities between frames

MPEG (DVD format) includes motion estimation:



Video

A few more details...

Frames are no longer independent!

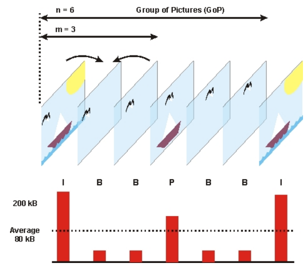
MPEG has three frame types:

- I-frames (intra-coded - independent)
- P-frames (predicted)
- B-frames (bi-predictive)

Must buffer B-frames until the next P-frame

Can only "enter" a video stream at an I-frame (or you see very blocky artifacts).

Video editors need to be very careful about this (splicing at non-I frames can be tricky!)



Video and Sound

A movie typically has multiple "streams" multiplexed together:

- Video stream
- Audio stream (maybe multiple for multi-language)
- Subtitles

Rendering software must synchronize streams - otherwise sound and video may be off (probably everyone has seen this happen!)

Compression

Taking advantage of redundancies and other structure to give smaller file sizes.

Two main types:

- Lossless: Allows perfect reconstruction of original data
  - Zip, RAR, FLAC, ... (JPEG has a lossless mode too!)
- Lossy: Reconstruction is an approximation of original
  - Most media formats: JPEG, MPEG, MP3, ...
  - Can usually trade off quality for compression

Note that digital sampling/capture is already a lossy process

(Remember taking advantage of human color vision?)

Compression

Examples, and what you can expect

Text: "Pride and Prejudice"

Original (uncompressed)	685 kB
Zip	250 kB
GZip	250 kB
RAR	217 kB
7Zip	204 kB
BZip	176 kB

Audio: "London Calling" (3:19 long)

CD audio (uncompressed)	35.2 MB
Zip (lossless, general)	33.9 MB
FLAC (lossless, audio)	25.4 MB
MP3 (lossy, 128 kbps)	3.2 MB
Ogg (lossy, quality 3)	3.1 MB

Notes:

- Zip is not designed for audio
- Both MP3 and Ogg sound good at this rate
- MP3 plays on almost all players
- MP3 encoding (using LAME) took 11.2 sec
- Ogg encoding took 6.1 sec

Compression

Examples, and what you can expect - cont'd

Picture: 3648 x 2736 (9.98 MPixel)

Raw	29.9 MB
Zip (lossless)	17.0 MB
BZip (lossless)	10.9 MB
PNG (lossless)	9.8 MB
JPEG (lossy - Q=95)	2.1 MB
JPEG (lossy - Q=85)	1.1 MB

Video: "Wizard of Oz" (1:41:42)
480x720 @30fps

Raw	190 GB
HQ DVD	3.6 GB

Notes:

- DVD compression is over 50:1
- DiVX / MP4 can give 200:1 or more

Summary

There's a ***lot*** more we could talk about

- Logarithmic scale of human perception (intensities, frequencies, etc.)
- Image formats: bitmapped vs vector formats
- Compression techniques
- Other imagery formats (multispectral images)
- ...

Explore this if it interests you! Following your curiosity is a great way to learn...