CM3106 Chapter 3: Multimedia Data Basics

Prof David Marshall
dave.marshall@cs.cardiff.ac.uk
and
Dr Kirill Sidorov
K.Sidorov@cs.cf.ac.uk
www.facebook.com/kirill.sidorov

School of Computer Science & Informatics
Cardiff University, UK

Multimedia Data Basics

Multimedia systems/applications have to deal with the

- Generation of data,
- Manipulation of data,
- Storage of data,
- Presentation of data, and
- Communication of information/data

Lets consider some broad implications of the above

Discrete v Continuous Media

RECALL: Our Definition of Multimedia

- All data must be in the form of digital information.
- The data may be in a variety of formats:
  - text,
  - graphics,
  - images,
  - audio,
  - video.

Synchronisation

A majority of this data is large and the different media may need **synchronisation**:

- The data will usually have temporal relationships as an integral property.

Click here or image above to run movies

Static and Continuous Media

Static or **Discrete** Media :

Some media is time **independent**:

- Normal data, text, single images, graphics are examples.

Continuous Media :

Time **dependent** Media:

- Video, animation and audio are examples.

Analog and Digital Signals

- Some basic definitions – **Studied HERE**
- Overviewing of technology — **Studied HERE**
  - **Recap from CM2208/CM2202**
- **More in depth study later.**

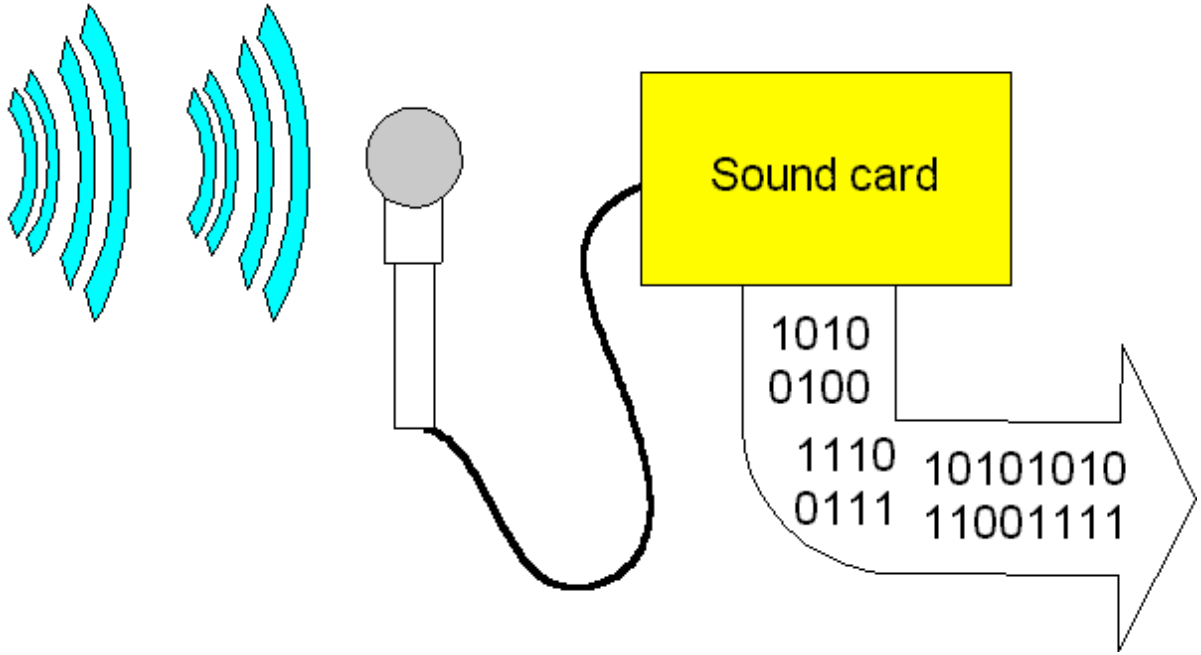
Analog and Digital Signal Conversion

The world we sense is full of analog signals:

- Electrical sensors convert the medium they sense into electrical signals
  - *E.g.* transducers, thermocouples: temperature sensor, microphones: acoustic sensor
    Cameras (Still and Video): light sensor.
  - (usually) **continuous Analog** signals (e.g. Sound and Light)
- *Analog*: continuous signals must be converted or **digitised** for computer processing.
- *Digital*: discrete digital signals that computer can readily deal with.

Analog-to-Digital Converter (ADC)

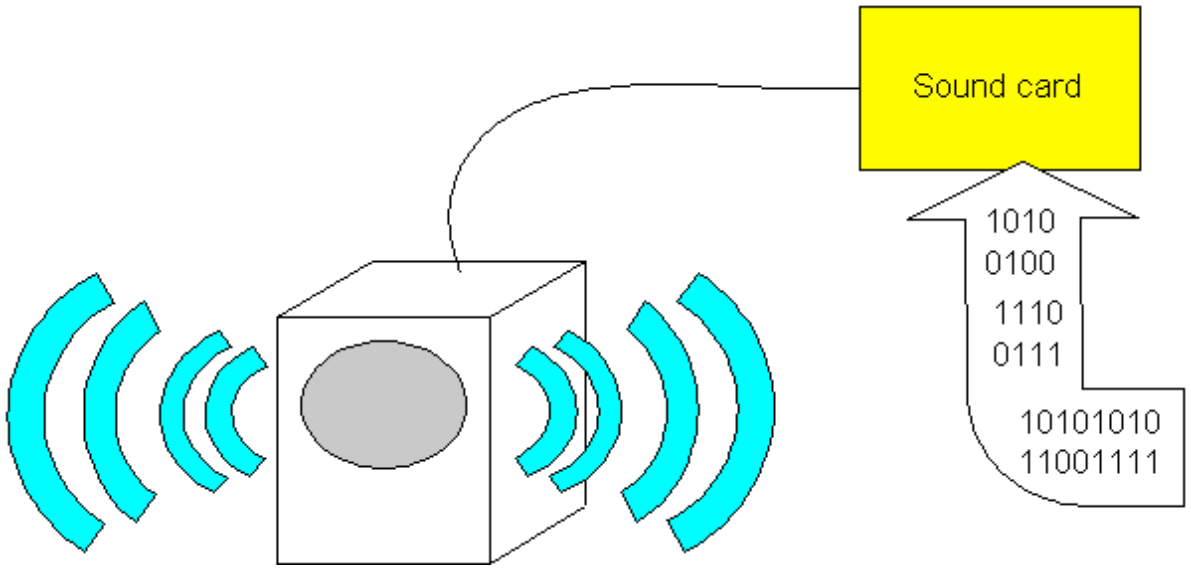
- Special hardware devices : **Analog-to-Digital** converters.
  E.g. Audio:



Take analog signals from analog sensor (*e.g.* microphone) and digitally sample data
(More details later)

Digital-to-Analog Converter (DAC)

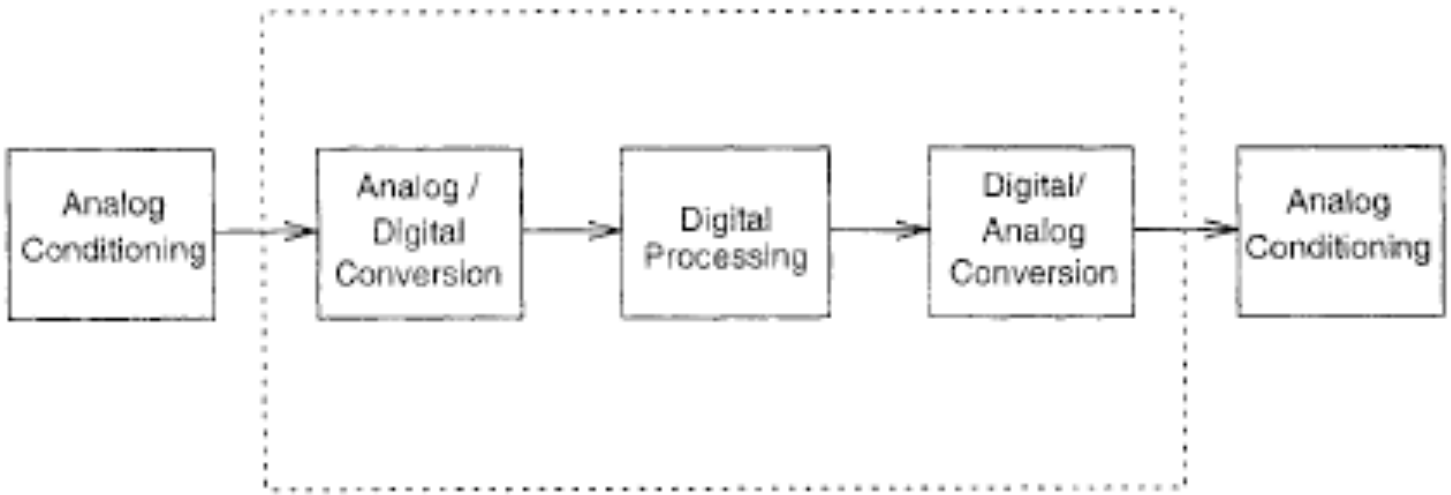
- Playback – a converse operation to *Analog-to-Digital*



- Takes digital signal, possible after modification by computer (*e.g.* volume change, equalisation)
- Outputs an analog signal that may be played by analog output device (*e.g.* loudspeaker, RGB monitor/display)

Analog-to-Digital-to-Analog Pipeline (1)

- Begins at the conversion from the analog input and ends at the conversion from the output of the processing system to the analog output as shown:



Analog-to-Digital-to-Analog Pipeline (2)

- Anti-aliasing filters (major part of *Analog Conditioning*) are needed at the input to remove frequencies above the sampling limit that would result in *aliasing*. **More later**
  The anti-aliasing filter at the output removes the aliases that result from the sampling (**see sampling theorem**).
- After the anti-aliasing filter, the analog/digital converter (ADC) **quantises** the continuous input into discrete levels.
- After digital processing, the output of the system is given to a digital/analog converter (DAC) which converts the discrete levels into continuous voltages or currents.
- This output must also be filtered with a low pass filter to remove the aliases from the sampling.
  Subsequent processing can include further filtering, mixing, or other operations.
  However, these will not be discussed further in this course.

Multimedia Data: Input and format

How to capture and store each Media format?

Note that text and graphics (and some images) are mainly generated directly by computer/device (e.g. drawing/painting programs) and *do not* require digitising:

They are generated directly in some (usually binary) format.

- Printed text and some handwritten text can be scanned via Optical Character Recognition
- Handwritten text could also be digitised by electronic pen sensing
- Printed imagery/graphics can be flatbed scanned directly to image formats.

Text and Static Data

- Source: keyboard, speech input, optical character recognition, data stored on disk.

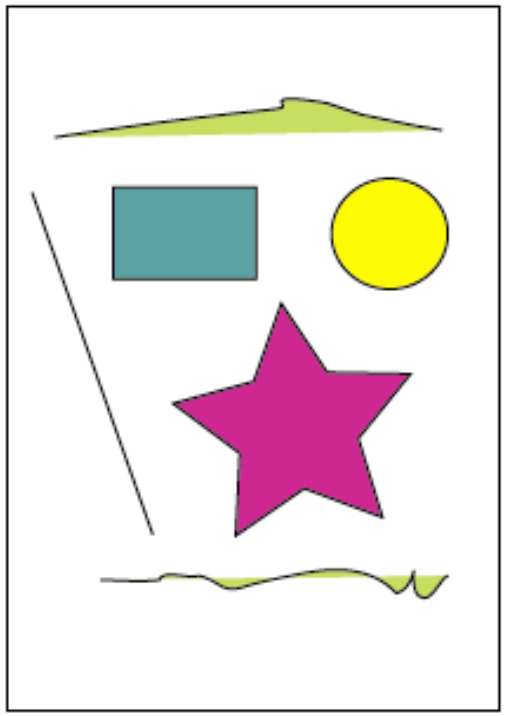
- Stored and input character by character:
  - Storage: 1 byte per character (text or format character), *e.g.* ASCII; more bytes for Unicode.
  - For other forms of data (*e.g.* Spreadsheet files). May store as text (with formatting, e.g. CSV – Comma-Separated Values) or may use binary encoding.

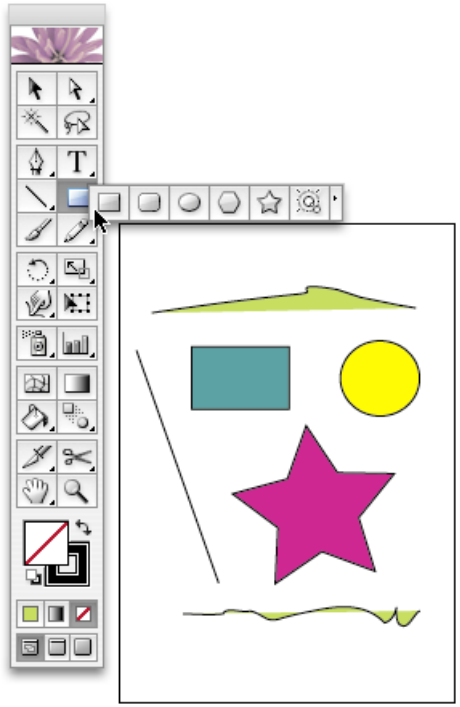
Text and Static Data (cont.)

- Formatted Text: Raw text or formatted text *e.g* HTML, Rich Text Format (RTF), Word or a program language source (Java, Python, MATLAB *etc.*)
- Data **Not temporal** — **BUT** may have natural implied sequence *e.g.* HTML format sequence, Sequence of Java program statements.
- Size Not significant w.r.t. other Multimedia data formats.
- Compression: convenient to bundle files for archiving and transmission of larger files. *E.g. Zip, RAR, 7-zip*. General purpose compression programs may not work well for other media types: audio, image, video etc.

Graphics

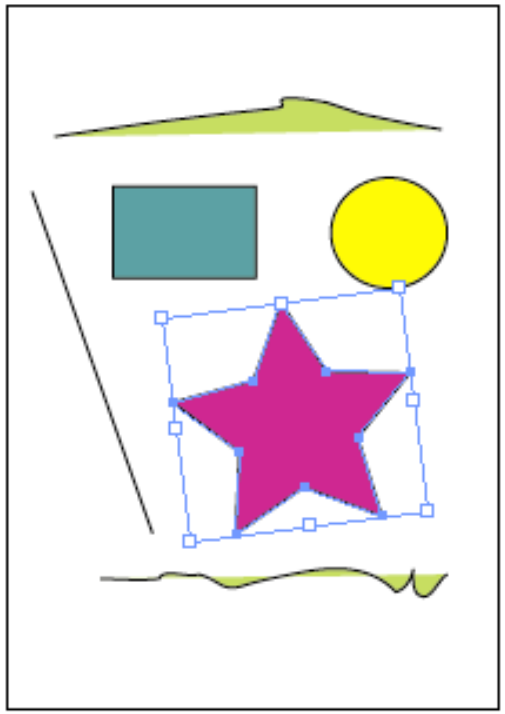
- Format: constructed by the composition of primitive objects such as lines, polygons, circles, curves and arcs.
- Input: Graphics are usually generated by a graphics editor program (*e.g.* Illustrator, Freehand) or automatically by a program (*e.g.* Postscript).





Graphics (cont.)

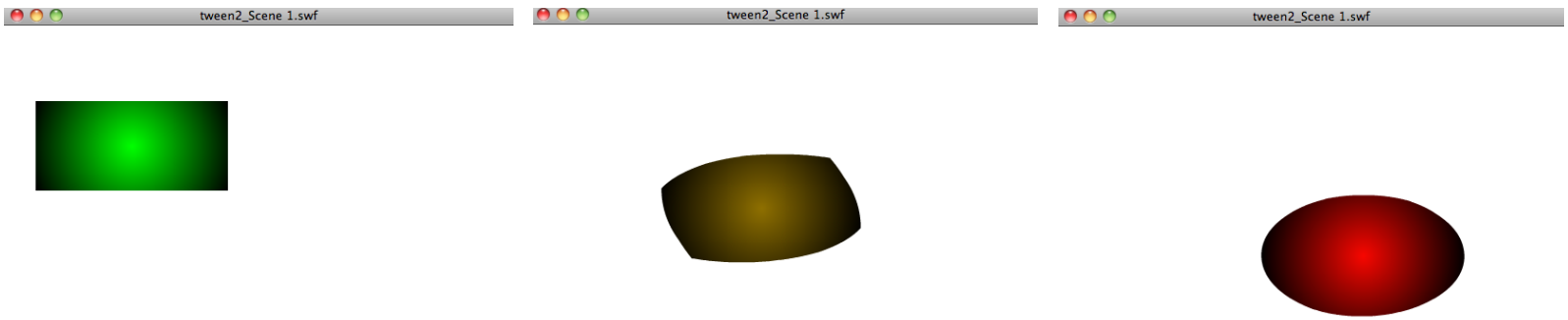
- Graphics input devices: keyboard (for text and cursor control), mouse, trackball or graphics tablet.
- Graphics are usually selectable and editable or revisable (unlike images).



- Graphics files usually store the primitive assembly
- Do not take up a **very high** storage overhead.

Graphics (cont.)

- Graphics standards : OpenGL - Open Graphics Library, a standard specification defining a *cross-language*, *cross-platform* API for writing applications that produce 2D/3D graphics.
- Animation: can be generated via a sequence of slightly changed graphics
  - 2D animation: e.g. Flash — Key frame interpolation: *tweening*: motion & shape



Simple Flash Demo: CM3106 Web Page Splash Screen

Module No: CM3106

© David Marshall

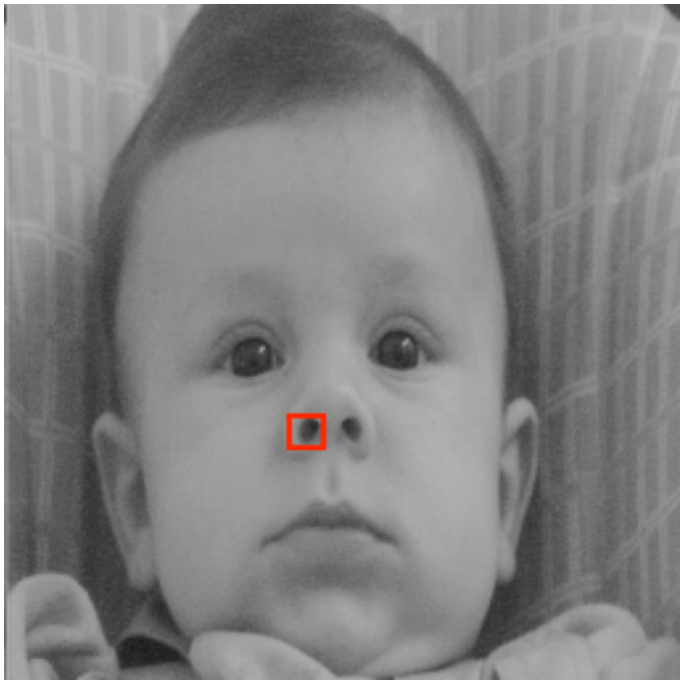
Click here or image above to run movies

Graphics (cont.)

- Animation (cont.)
  - 3D animation: e.g. Maya.
    Change of *shape/texture/position*, *lighting*, *camera*
- Graphics animation is *compact*
  – suitable for network transmission (e.g. Flash).

Images

- Still pictures which (uncompressed) are represented as a bitmap (a grid of pixels).



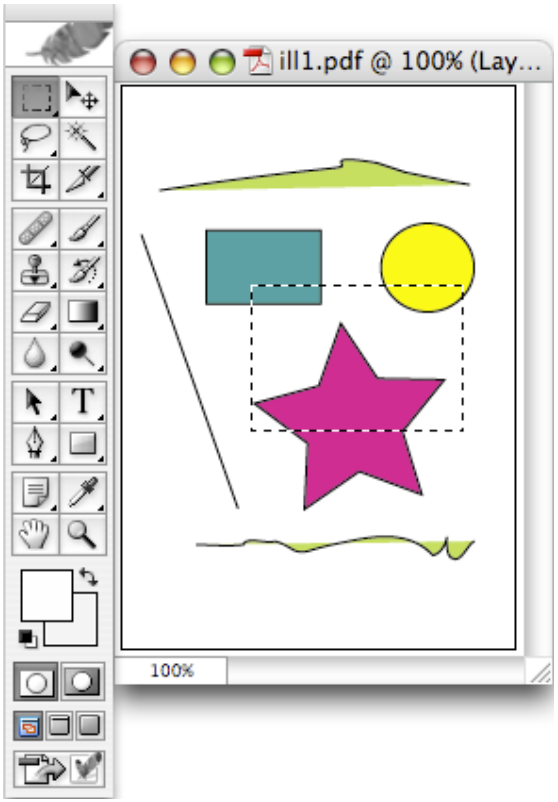
99	71	61	51	49	40	35	53	86	99
93	74	53	56	48	46	48	72	85	102
101	69	57	53	54	52	64	82	88	101
107	82	64	63	59	60	81	90	93	100
114	93	76	69	72	85	94	99	95	99
117	108	94	92	97	101	100	108	105	99
116	114	109	106	105	108	108	102	107	110
115	113	109	114	111	111	113	108	111	115
110	113	111	109	106	108	110	115	120	122
103	107	106	108	109	114	120	124	124	132

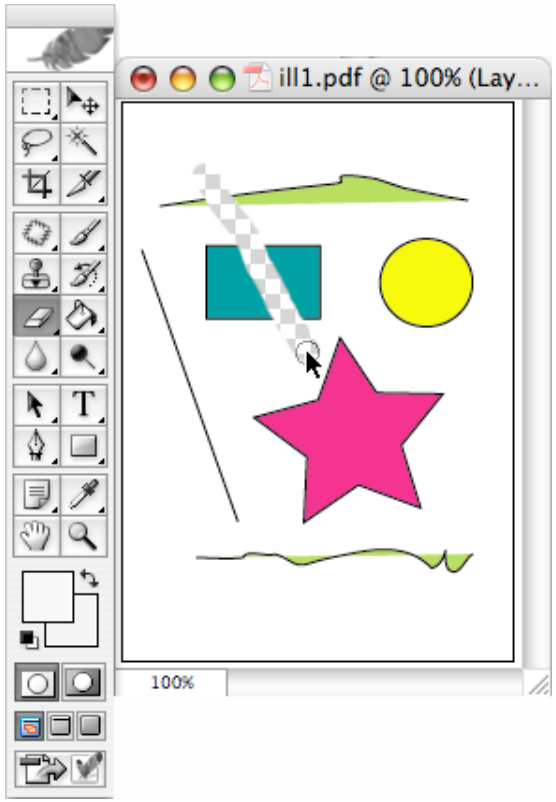
Images (cont.)

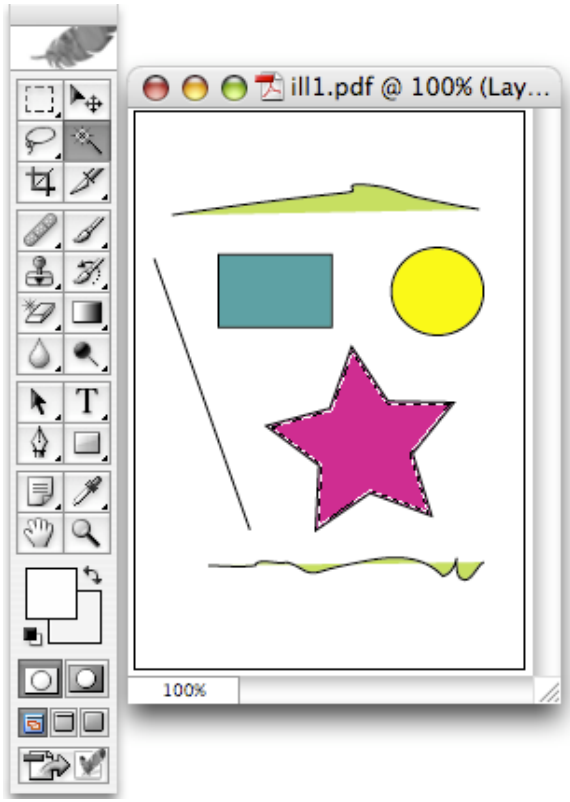
- Input: scanned for photographs or pictures using a digital scanner or from a digital camera.
- Input: May also be generated by programs similar to graphics or animation programs.
- Analog sources will require digitising.
- Stored at 1 bit per pixel (Black and White), 8 Bits per pixel (Grey Scale, Colour Map) or 24 Bits per pixel (True Colour)
- Size: a 512x512 Grey scale image takes up 1/4 MB, a 512x512 24 bit image takes 3/4 MB with no compression.
- This overhead soon increases with image size — modern high digital camera 10+ Megapixels $\approx$ 29MB uncompressed!
- Compression is commonly applied.

Images (cont.)

- Can usually only edit individual or groups of pixels in an image editing application, *e.g.* photoshop.

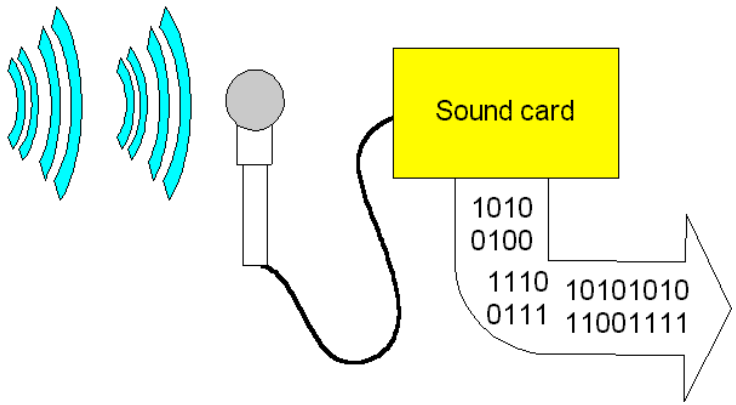






Audio

- Audio signals are continuous analog signals.
- Input: microphones and then digitised and stored



- CD Quality Audio requires 16-bit sampling at 44.1 KHz: Even higher audiophile rates (e.g. 24-bit, 96 KHz)
- 1 Minute of Mono CD quality (uncompressed) audio = 5 MB.
  Stereo CD quality (uncompressed) audio = 10 MB.
- Usually compressed (E.g. MP3, AAC, Flac, Ogg Vorbis)

Video

- Input: Analog Video is usually captured by a video camera and then digitised, although digital video cameras now essentially perform both tasks.
- There are a variety of video (analog and digital) formats (**more later**)
- Raw video can be regarded as being a series of single images. There are typically 25, 30 or 50 frames per second.

Click here or image above to run movie

Video (cont)

Video Size:

- A 512x512 size monochrome video images take 25*0.25 = 6.25MB for a second to store uncompressed.
- Typical PAL digital video (720×576 pixels per colour frame) $\approx 1.2 \times 25 = 30$MB for a second to store uncompressed.
- High Definition video on Blu-ray (up to $1920 \times 1080 = 2$ Megapixels per frame) $\approx 6 \times 25 = 150$MB for a second to store uncompressed, i.e. 9GB for a minute to store uncompressed. (There are higher possible frame rates!)
- Digital video clearly needs to be compressed.

Multimedia Data Representation

Issues to be covered (Over coming lectures):

- Digital Audio
  - Digital Audio Synthesis
  - MIDI — Synthesis and Compression Control
  - Digital Audio Signal Processing/Audio Effects
- Graphics/Image Formats
  - Colour Representation/Human Colour Perception
- Digital Video
  - Chroma Subsampling

General Themes across all above

- Sampling/Digitisation
  - Sampling Artifacts — *Aliasing*
- Compression requirements
  - Data formats especially size
  - Human Perception $\longrightarrow$ compression ideas

Building up to full Multimedia Compression Algorithms — following lectures