Fundamentals

Digital Multimedia, 3rd edition

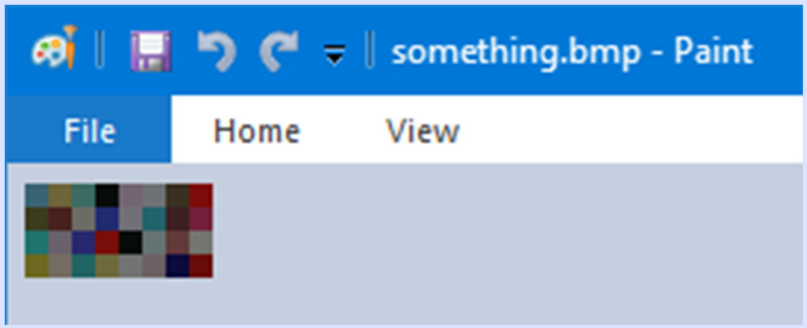
Chapter 2

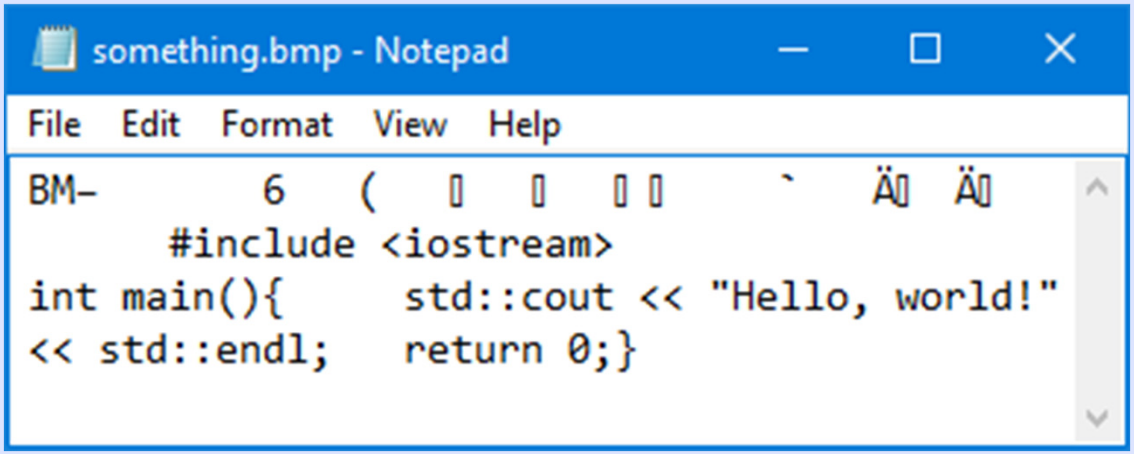
Digital Multimedia

“Digital multimedia: any combination of two or more media, represented in a digital form, sufficiently well integrated to be presented via a single interface, or manipulated by a single computer program”

Interpretation of bits

- Any digital media, signal, etc. will be stored as bits, 0 or 1
- Groups of bits can be interpreted as numbers or as pretty much anything depending on context
  - Binary numbers
    - 01100001 = 97 decimal
  - Characters - associate bit patterns with characters via a character set
    - 01100001 = a in ASCII
  - Brightness or color of a pixel in an image, amplitude of a sound wave, etc.





Digitization

- The world is analog
  - Often we need to convert analog signals to digital form so they can be manipulated by computer programs
  - Analog signals can vary continuously, digital is restricted to discrete values
- Two-stage process
  - Sampling - measure the value at discrete intervals
  - Quantization - restrict the value to a fixed set of quantization levels

Analog Vs. digital data

Analog

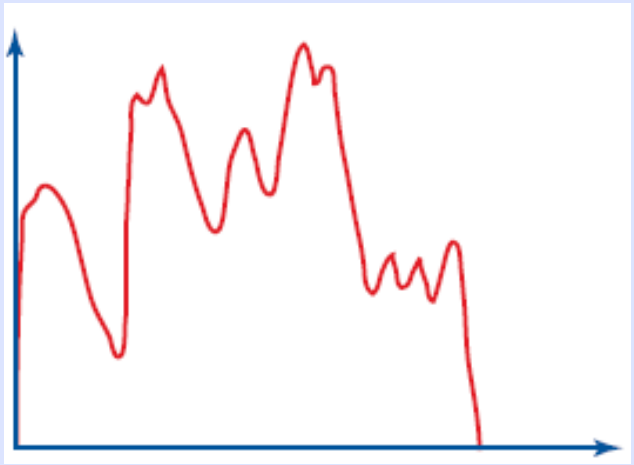
Clock hands can be in an unlimited number of positions

16:07:02

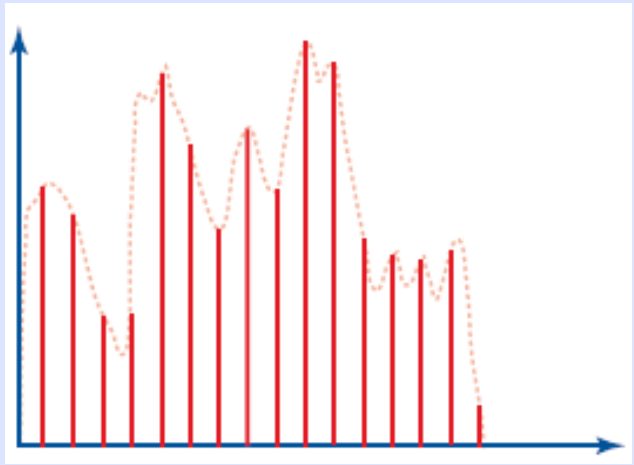
Digital

Clock limited by the chosen resolution (quantization levels)

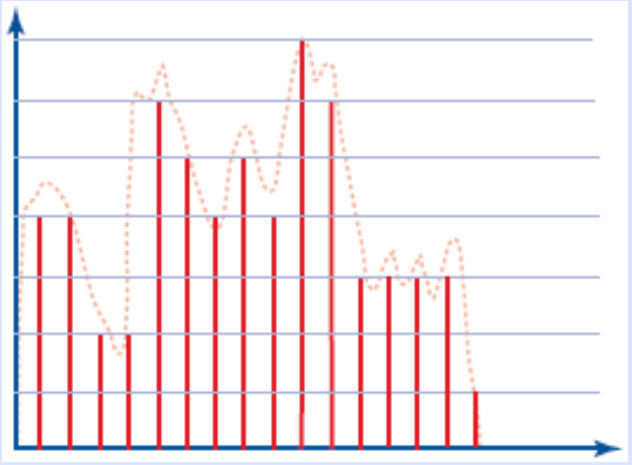
Sampling and quantization



Analog signal



Sampling

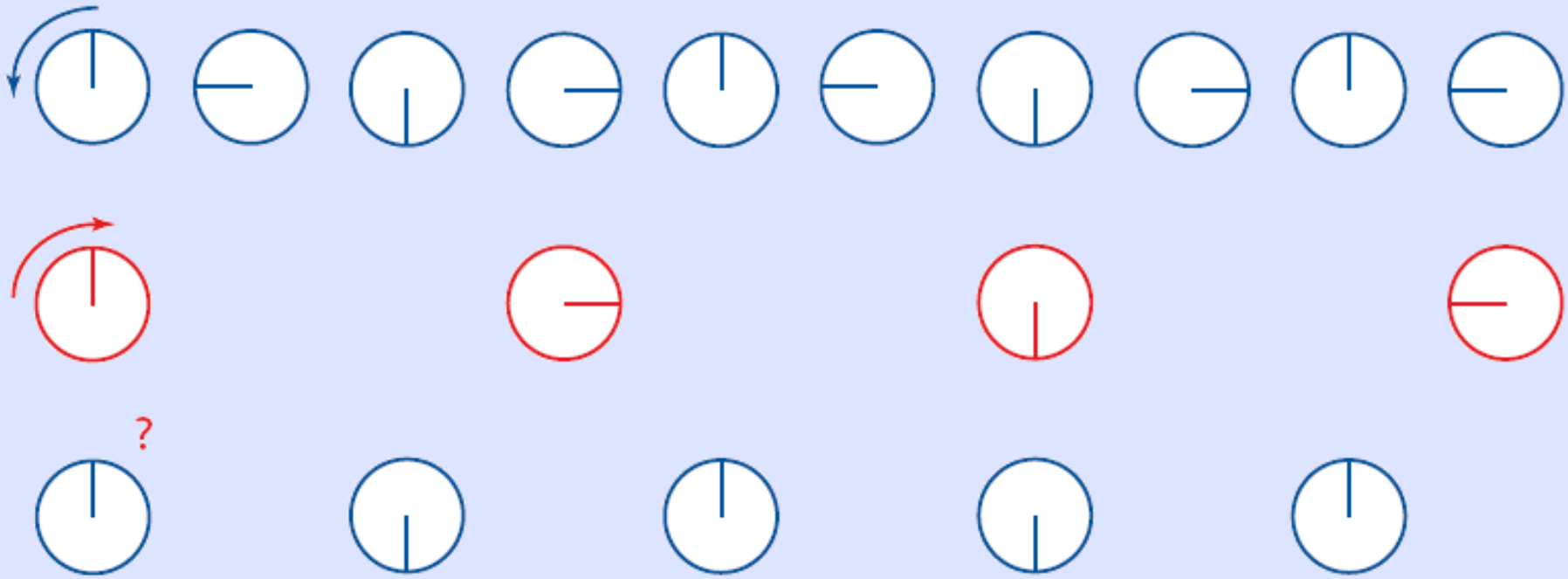


Quantization

Undersampling

- A low sample rate means few read values, and less data to store
  - Too low sample rate however leads to *undersampling*
  - Undersampling results in *aliasing*
    - Not possible to properly recreate the analog signal, resulting in sound distortion, image 'jaggies‘, jumpy motion in videos, etc.
- Sampling theorem
  - If the highest frequency component of a signal is at f_h the signal can be properly reconstructed if it has been sampled at a frequency > $2f_h$
    - Known as the Nyquist rate

Undersampling

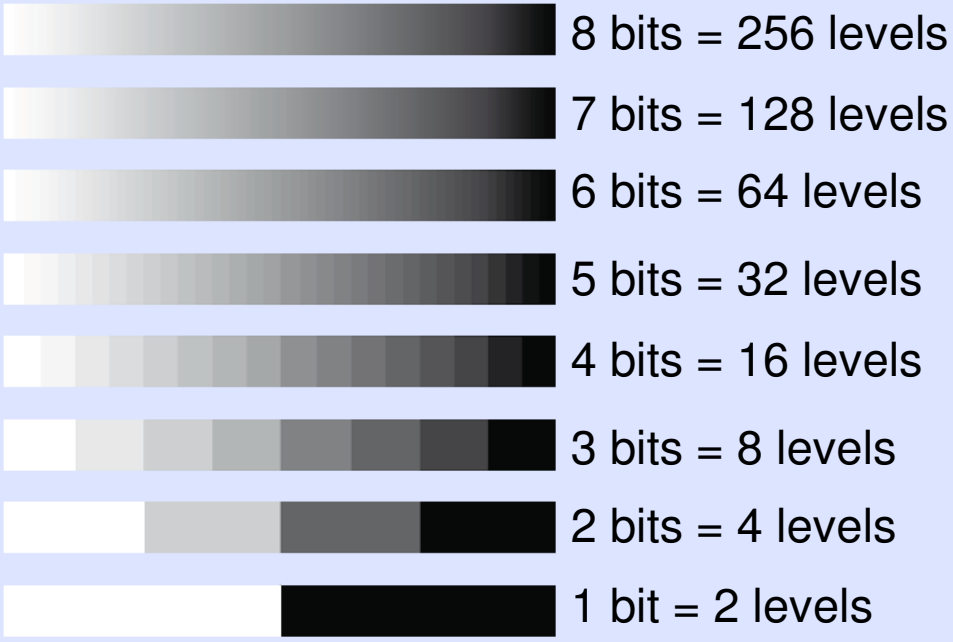


In movies tires on cars sometime appear to be rotating in the opposite direction. Sample rate (frame rate) too low compared to speed of the car

Too few quantization levels

- Reducing storage requirements can also be done by using fewer bits for each sampled value
  - Fewer quantization levels are available
- Too few quantization levels mean we cannot distinguish between values that fall between levels
  - Images: problems with banding and posterization
  - Sound: problems with hissing, loss of quiet passages, general fuzziness (*quantization noise*)

Banding and Moiré patterns



Banding:

Quantization levels are very noticeable as this gradient is stored in fewer and fewer bits

Moiré patterns:

Very noticeable distortions in fine patterns (like the one formed by bricks in this wall)

Posterization

Original picture

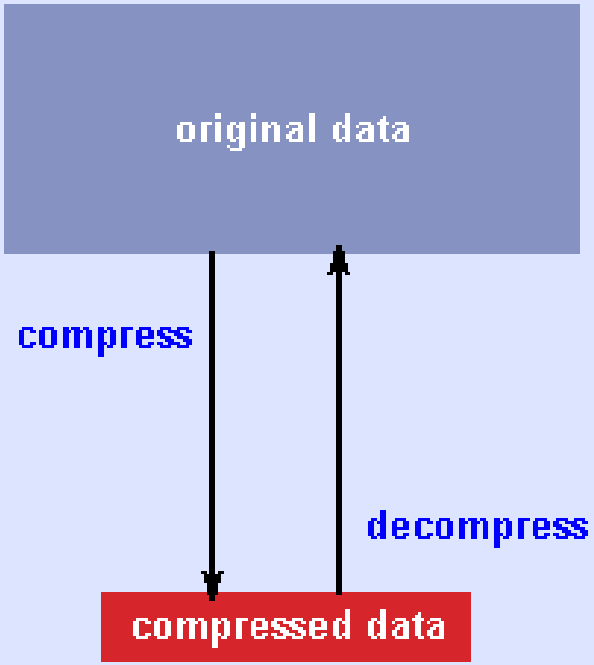
Few quantization levels (colors)

Compression

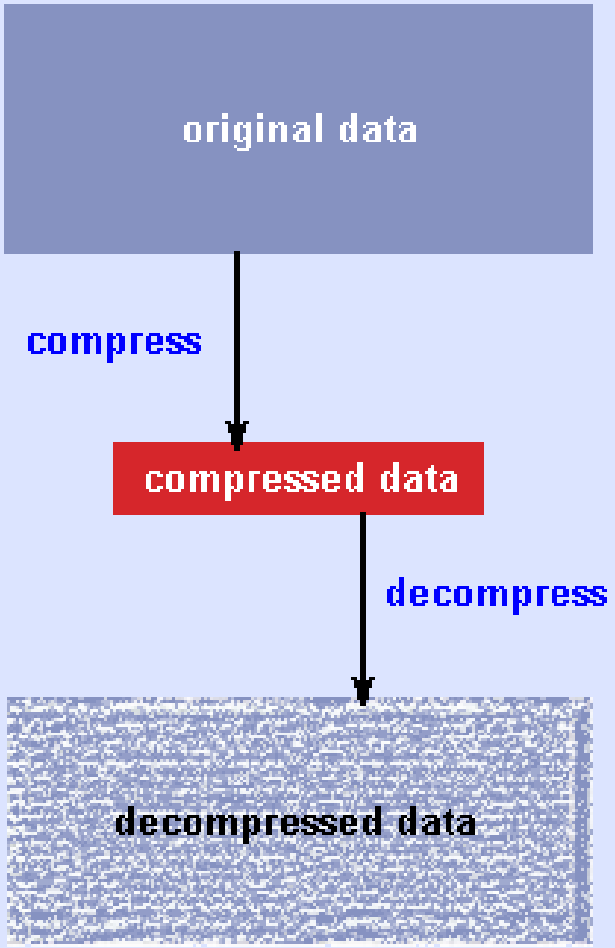
- Due to large file sizes, compression must often be applied to media data
- Compression may be lossless or lossy
  - Lossless = Always possible to decompress compressed data and obtain an exact copy of the original data
  - Lossy = Some (mostly unnoticeable) data is removed
    - Lossy compression is usually much more effective than lossless compression
  - Different compression algorithms are used for different types of media

Lossless / lossy compression

Lossless compression



Lossy compression



Digital representation of media

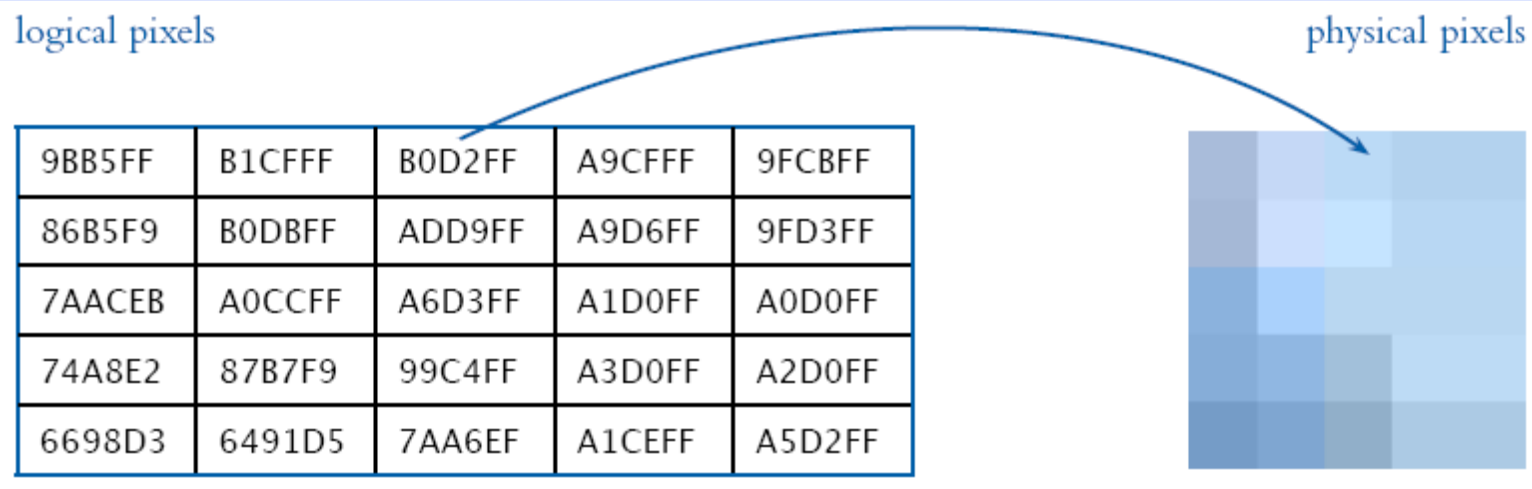
- There are established ways of representing images, video, animation, sound and text in bits
- The media data may be represented as plain text in a suitable language (e.g. XML), or as binary data in a specific format

Digital images

- Images are represented using an internal model, and displayed on screen as arrays of pixels
- Generating the pixels from the model is called rendering
- Images may be modeled as bitmapped graphics or as vector graphics

Bitmap images

- A bitmap is simply an array of pixels (stored color values) that can be mapped directly to the physical pixels on the screen



Vector graphics

- In vector graphics, the image is stored as a mathematical description of the individual lines, curves and shapes making up the image
- Rendered by calculating the appearance of the image

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE svg PUBLIC "-//W3C//DTD SVG 1.0//EN"
  "http://www.w3.org/TR/2001/REC-SVG-20010904/DTD/svg10.dtd">
<svg  xmlns="http://www.w3.org/2000/svg">
   <path fill="#F8130D" stroke="#1E338B" stroke-width="20"
    d="M118,118H10V10h108V118z"/>
</svg>
```

Vector graphics stored as SVG file

Vector graphics

- Mostly usable for generated/computer drawn images; photos should be stored as bitmaps
- Vector graphics are often smaller than bitmaps, but the size varies depending on the complexity of the graphics

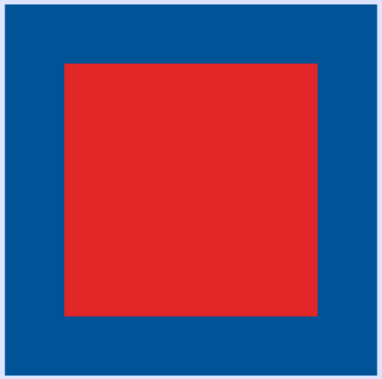


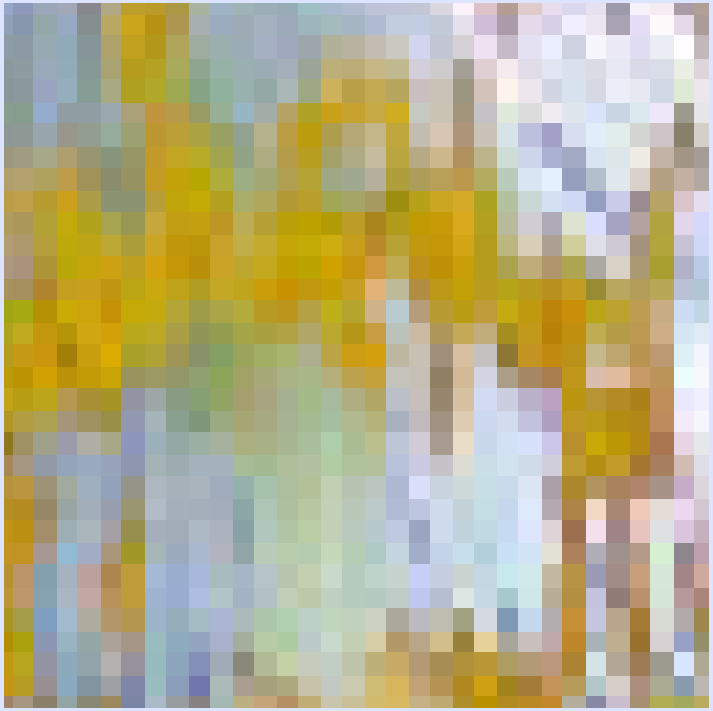
Image size: 128x128 pixels
Bitmap = 48000B
SVG = 284B

Scaling of vector graphics

- Resolution-independent and can be scaled without any loss of quality

Scaling of bitmap graphics

- Limited by a fixed number of stored pixels, bitmaps scale poorly



Animation and video

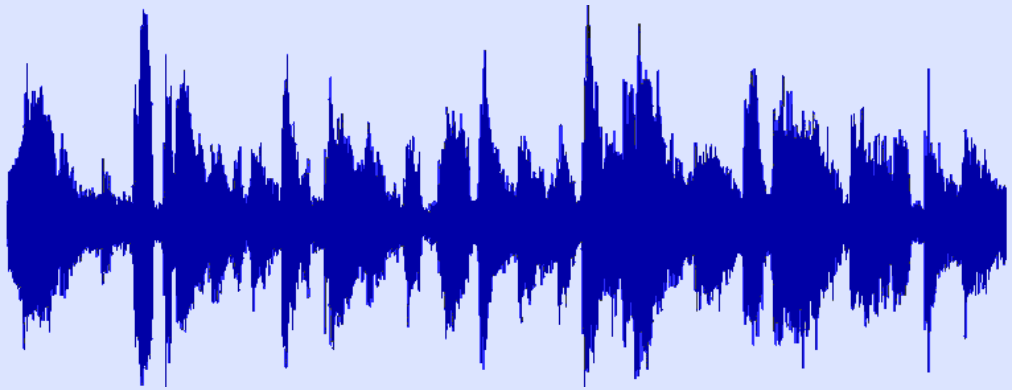
- A sequence of images displayed at a fast enough rate will appear as a continuous stream rather than as individual images
  - Usually around 40 images per second is needed
    - Known as the fusion frequency
  - Small changes between images will be perceived as movement
- Video
  - Recording of actual real-life motion
- Animation
  - Artificially created motion

Animation and video

- In video each frame is essentially stored as a bitmap image
  - Uncompressed video result in huge data sizes so effective compression is needed
    - Often not complete frames, but only difference between frames are stored
- In animation frames might be bitmap images or vector graphics
  - In some cases (e.g. used in Flash when that was still a thing) only a few key frames are stored, while everything else is calculated in runtime

Sound

- A sound often contains many different and complex frequencies with varying amplitudes
- After sampling and quantization, a sound can be represented as samples at a given bit rate
- Typically displayed as a waveform



Waveform

Sound

- CD audio is sampled at 44.1 kHz, but higher sampling rates are sometimes used
  - Humans can detect frequencies up to ~20 kHz, so a sample rate of 44.1 kHz fits the sampling theorem
- CD audio result in fairly large file sizes, so audio delivered over the Internet is usually compressed – e.g. using the MP3 codec
  - Other formats, like AAC, are taking over from MP3 these days

Text

- All the various characters used in text are stored as character codes and mapped to the actual characters by a character set
- Characters are displayed in a given font
  - Fonts consist of a number of character shapes (images) called glyphs

Text

- In multimedia text is used in various contexts
  - Directly as content on web pages and in text documents
    - Many aspects must be considered regarding the layout: Font size, font type, font color, text alignment, etc.
  - As the basis of other media types
    - SVG vector graphics is written in an XML based language
  - As meta-data in web sites or in various media files