CA168 - Digital World

Chapter 2: Multimedia Storage

The Nature of Sound

- Conversion of energy into vibrations in the air (or some other elastic medium)
- Most sound sources vibrate in complex ways leading to sounds with components at several different frequencies - harmonics in music
- Frequency spectrum – relative amplitudes of the frequency components of a sound
- Range of human hearing: roughly 20Hz–20kHz, falling off with age - which is why older adults can't hear high pitch ringtones ;-)

Chapter 2 Agenda

- Storage formats for information - text, image, sound, video, animation - have evolved rather than being planned, and that's bad;
- Yet there are standards (several in some cases) and these are stable;
- Here, we briefly cover:
  - Audio
  - Image
  - Video

Waveforms

- Sounds change over time except tones
  - e.g. musical note has attack and decay, speech changes constantly
- Frequency spectrum alters as sound changes
- Waveform is a plot of amplitude against time
  - Provides a graphical view of characteristics of a changing sound
  - Can identify syllables of speech, rhythm of music, quiet and loud passages, etc
  - See Waveform gallery for illustration !

Audio

Digitization – Sampling

- Sampling Theorem (hard maths) implies minimum rate of 40kHz to reproduce sound up to limit of hearing
- … which is why … CD: 44.1kHz .. Gives perfect reproduction (to our (your!) ears)
- 16 bits, 65536 quantization levels, CD quality
- Sub-multiples often used for low bandwidth – e.g. 22.05kHz for Internet audio

Data Size

- Sampling rate r is the number of samples per second
- Sample size s bits
- Each second of digitized audio requires rs/8 bytes
- CD quality: r = 44100, s = 16, hence each second requires just over 86 kbytes (k=1024), each minute roughly 5Mbytes (mono)

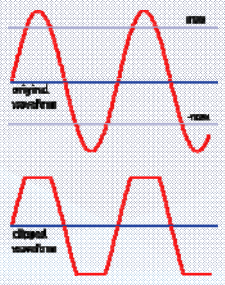
Compression

- Sampling audio and simply storing is a waste of space … we could save some.
- In general, lossy methods required because of complex and unpredictable nature of audio data
- CD quality, stereo, 3-minute song requires over 25 Mbytes
  - Data rate exceeds bandwidth of dial-up Internet connection so demand was for compression to allow streaming. Not so now, but we still get advantage of reduced data size anyway

Clipping

•If recording level is set too high, signal amplitude will exceed maximum that can be recorded, leading to unpleasant distortion

•But if level is set too low, dynamic range will be restricted



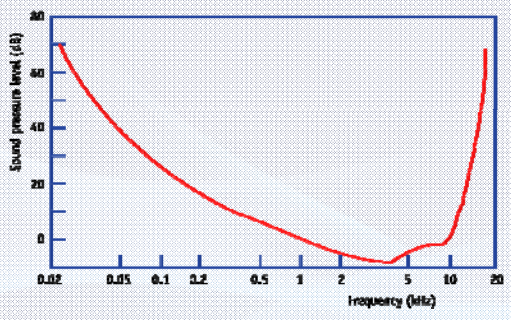
Perceptually-Based Compression

- Identify and discard data that doesn't affect our perception of the signal .. If its meant for hearing, then we as the hearers should be the judges of what to throw away
  - Needs a psycho-acoustical model, since ear and brain do not respond to sound waves in a simple way
  - Analogous to building a film or TV studio set - don't build the parts the viewer can't see !
  - Take advantage of …
    - Threshold of hearing – sounds too quiet to hear
    - Masking – sound obscured by some other sound

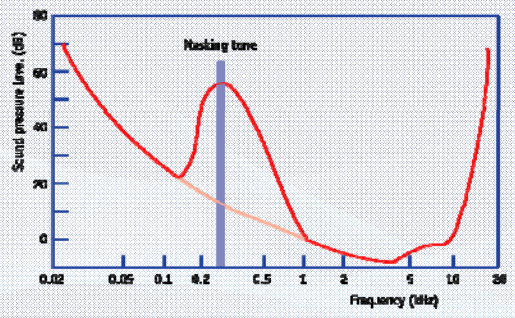
Effects and Filters

- Noise gate
- Low pass and high pass filters
- Notch filter
- De-esser
- Click repairer
- Reverb
- Graphic equalizer
- Envelope Shaping
- Pitch alteration and time stretching
- etc.
- … all are ways to vary the digitised waveform (which is just a set of numbers in a computer after all)

The Threshold of Hearing



Masking



AAC

- Advanced Audio Coding
- Defined in MPEG-2 standard, extended and incorporated into MPEG-4
- Not backward compatible with earlier standards
- Higher compression ratios and lower bit rates than MP3
- Subjectively better quality than MP3 at the same bit rate

Compression Algorithm

- Split signal into bands of frequencies using filters
  - Commonly use 32 bands
- Compute masking level for each band, based on its average value and a psycho-acoustical model
  - i.e. approximate masking curve by a single value for each band
- Discard signal if it is below masking level
- Otherwise quantize using the minimum number of bits that will mask quantization noise

Audio Formats

- Platform-specific file formats
  - AIFF, WAV
- Multimedia formats used as 'container formats' for sound compressed with different codecs
  - QuickTime, Windows Media, RealAudio
- MP3 has its own file format, but MP3 data can be included as audio tracks in QuickTime movies

MP3

- MPEG Audio, Layer 3
- Three layers of audio compression in MPEG-1
- Layer 1...Layer 3, encoding proces increases in complexity, data rate for same quality decreases
  - e.g. Same quality 192kbps at Layer 1, 128kbps at Layer 2, 64kbps at Layer 3
- 10:1 compression ratio at high quality
- Variable bit rate coding (VBR)

Computer Graphics

Computer Graphics

- Production and display of still images stored in digital form
  - Digitize printed image with a scanner
  - Capture image from digital camera
  - Grab frame from video camera
  - Create in digital form using graphics package
  - Generate visual representation of data

Memory Requirements

- Bitmapped – any picture of wxh pixels, using c bytes per pixel occupies whc bytes
- Vector – space required depends on complexity of picture (how many shapes, segments of path, etc)
  - Usually vector graphics smaller than bitmapped

Rendering

- Image is displayed on monitor etc as array of pixels
  - Rectangular (usually square) dots of colour
- Program (e.g. Web browser) sets pixels to an appropriate colour to produce desired image
  - Pixels merge optically to produce effect of continuous tone
- Computer must maintain a model of the image
  - May be stored in a file and read by program

Memory Requirements

- 128 px square with 20px blue outline filled in red
  - Bitmap using 24 bits per pixel, 128x128x3 = 48kbytes
  - Vector specified in SVG:
  - <path fill="#F8130D" stroke="#1E338B" stroke-width="20" d="M118,118H10V10h108V118z"/>
  - 86 bytes (plus 198 bytes SVG boilerplate)

Bitmapped and Vector Graphics

- Bitmapped graphics – image is modelled as an array of pixel values
  - Render by direct mapping of logical pixels to physical pixels of screen
    - May need scaling and clipping
- Vector graphics – image is modelled as mathematical description of curves, shapes
  - Render by computing pixels from description
- Big differences in resulting quality of image (sampling and quantization issues) and also of computational cost to calculate;

Image Editing

- Vectors – drawing programs
  - Select individual graphic objects (shapes, paths, &c)
  - Transform size, position, angle, &c
  - Change attributes: stroke and fill &c
- Bitmaps – painting programs
  - Select areas of pixels
  - Apply effects and filters

Scaling

- Vectors
  - Scaling is a simple mathematical operation on stored description (before rendering)
  - Curves and lines remain smooth at all sizes
- Bitmaps
  - Interpolate pixel values
    - More or less sophisticated algorithm
  - Usually produces loss of quality, blurring, jaggedness etc.

Vector Formats

- SVG (Scalable Vector Graphics)
  - W3C standard, not presently widely used
- SWF (Flash)
  - Primarily for vector animation, but can be used for still vector graphics; de facto standard
- EPS (Encapsulated PostScript)
  - Primarily print, use declining, superseded by PDF

File Formats

- Many different graphics file formats in existence
- Different ways of encoding image data
- Different amounts/form of supplementary data
- (Bitmaps) Different compression methods
  - Lossless – image can be reconstructed exactly from compressed version
  - Lossy – some information discarded, image can only be reconstructed approximately

Digital Video

WWW Bitmapped Formats

- GIF (Compuserve Graphics Interchange Format)
  - Lossless, 256 colours (indexed), transparency
- JPEG (Joint Photographic Experts Group)
  - Lossy (variable quality), millions of colours
- PNG (Portable Network Graphics)
  - Lossless, variable number of colours, W3C standard

Analogue Standards

- PAL (Phase Alternating Line)
  - Western Europe, Australia & New Zealand, China,…
- NTSC (National Television Standards Committee)
  - North America, Japan, Taiwan, parts of South America,…

Analogue Standards

- SECAM (Séquential Couleur avec Mémoire)
  - France and former Soviet Union
  - Standard only used for transmission
  - Uses PAL cameras etc

Digitization

- In the camera – DV + Firewire
- In the computer – video capture card
- Digitization in camera (DV), at time of capture, means less noise
- Less noise allows better compression

PAL and NTSC

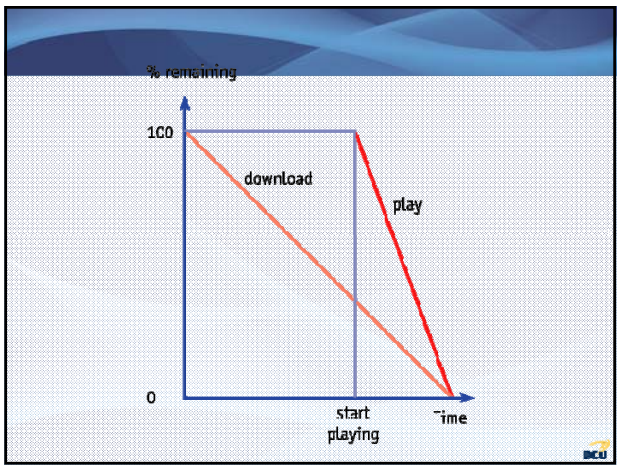
- PAL
  - Frame has 625 lines, 576 are picture
  - 25 frames (50 fields) per second
- NTSC
  - Frame has 525 lines, 480 are picture
  - 29.97 frames (59.94 fields) per second
    - (Often quoted as 30 frames per second)

Streamed Video

- Play back a video stream as it arrives over a network (like broadcast TV), instead of downloading an entire video clip and playing it from disk (like renting a DVD)
- Start playing a downloaded clip as soon as enough of it has arrived
- Starts when the (estimated) time to download the rest is equal to the duration of the clip

Video Data Size

- PAL uncompressed
  - 768x576 pixels per frame
    - x 3 bytes per pixel (24 bit colour)
    - x 25 frames per second
  - ≈ 31 MB per second
  - ≈ 1.85 GB per minute
- NTSC uncompressed
  - 640x480 pixels per frame
    - x 3 bytes per pixel (24 bit colour)
    - x 30 frames per second (approx)
  - ≈ 26 MB per second
  - ≈ 1.6 GB per minute



Video Standards

- Digital video devices must conform to standards
- Digital standards must maintain compatibility with older analogue standards for broadcast TV

MPEG-4

- Designed to support a range of multimedia data at bit rates from 10kbps to >1.8Mbps
- Applications from mobile phones to HDTV
- Video codec becoming popular for Internet use, is incorporated in QuickTime, RealMedia and DivX

Interlacing

- Required for analogue TV, so encountered in captured footage
- Each frame is divided into two fields
- Field 1: odd lines; Field 2: even lines
- Fields are transmitted one after the other
- Frame is built out of the interlaced fields

Video Compression

- Spatial (intra-frame) compression
  - Compress each frame in isolation, treating it as a bitmapped image
- Temporal (inter-frame) compression
  - Compress sequences of frames by only storing differences between them
- Always some compression because of sub-sampling

MPEG

- ISO/IEC Motion Picture Experts Group
- Series of standards including
  - MPEG-1 intended for video CD
  - MPEG-2 used in DVD and broadcast
  - MPEG-4 for low bitrate multimedia

Spatial Compression

- Image compression applied to each frame
- Can therefore be lossless or lossy, but lossless rarely produces sufficiently high compression ratios for volume of data
- Lossless compression implies a loss of quality if decompressed then recompressed
- Ideally, work with uncompressed video during post-production

Temporal Compression

- Key frames are spatially compressed only
  - Key frames often regularly spaced (e.g. every 12 frames)
- Difference frames only store the differences between the frame and the preceding frame or most recent key frame
- Difference frames can be efficiently spatially compressed

Post-Production

- Changing or adding to the material
  - Most changes are generalizations of image manipulation operations (e.g. colour correction, blurring and sharpening,…)
- Compositing – combining elements from different shots into a composite sequence
- Animating elements and combining animation with live action

Motion JPEG

- Purely spatial compression
- Apply JPEG to each frame
- Used by most analogue capture cards
- No standard, but MJPEG-A format widely supported

Preparing for Delivery

- Compromises required to bring resource requirements of video within capabilities of delivery media (e.g. networks) and low-end machines
  - Reduce frame size (e.g. downsample to quarter frame)
  - Reduce frame rate (12fps is OK for smooth motion, flicker not a problem on computer)
  - Reduce colour depth

Editing

- Making a constructed whole from a collection of parts
  - Selection, trimming and organization of raw footage
  - Apply transitions (e.g. dissolves) between shots
  - Combination of picture with sound
- No changes made to the footage itself

- That's all on MM formats.
- Next week we look at Search Engines