Multimedia Information Representation

Gail Reynard

In this lecture …

- How multimedia information is converted from analogue to digital
- How it is represented digitally
- Basic principles needed for later material

Introduction

- All types of multimedia are stored on computer in a digital form
  - Text
    - string of codewords
  - Graphics
    - line - start and end coordinates
  - Audio and video
    - analogue signals need to be converted to digital
    - signal encoder samples the amplitude
    - signal decoder converts it back for playing

Digitization

- Analogue signals
  - amplitude varies with time
  - range of frequencies of the sinusoidal components that make up a signal is called the signal bandwidth
  - for transmission:
    - bandwidth of transmission channel should be \( \geq \) bandwidth of signal
    - otherwise low and/or high frequencies will be lost

Encoder Design

- Consists of:
  - bandlimiting filter
  - analogue-to-digital converter (ADC), consisting of:
    - a sample-and-hold
      - samples the amplitude at regular time intervals
    - a quantizer
      - converts each sample amplitude into a codeword

Signal Encoding
Nyquist Sampling Theorem

- “in order to obtain an accurate representation of a time-varying analogue signal, its amplitude must be sampled at a minimum rate that is equal to or greater than twice the highest sinusoidal frequency component that is present in the signal”
- known as the Nyquist rate
- represented as Hz or samples per second (sps)

Sampling Rate

- Sampling at a rate lower than the Nyquist rate:
  - causes additional frequency components to be generated
  - original signal becomes distorted
  - replaced by an alias signal of lower frequency
- known as the Nyquist rate
- represented as Hz or samples per second (sps)

Sampling Rate (2)

- To prevent this distortion:
  - bandlimiting filter is used:
    - removes frequency components higher than half the sampling frequency
    - therefore also known as an antialiasing filter

Quantization Intervals

- Each sample can only be represented by a number of discrete levels
- Quantization interval
  - Determined by:
    - maximum positive and negative signal amplitude
    - number of binary bits used to represent values
    - ALSO influenced by:
      - smallest amplitude relative to peak amplitude (dynamic range)
- Quantization error
  - the difference between the actual signal amplitude and the corresponding nominal amplitude
  - a.k.a. Quantization noise

Decoder Design

- Each digital codeword is converted to an equivalent analogue sample using a digital-to-analogue converter (DAC)
- Output of DAC is passed through a low-pass filter
- Most audio/video encoders and decoders are combined into a single unit - audio/video codec
Digitized Pictures

- Colour principles
- Raster-scan principles
- Pixel depth
- Aspect ratio

Colour Principles

- A whole spectrum of colours (colour gamut) can be produced by mixing Red (R), Green (G) and Blue (B)
- Additive colour mixing
  - black is produced when R=0, G=0, B=0
  - good for producing a colour image on a black surface
  - good for TVs and computer monitors

Colour Principles (2)

- Subtractive colour mixing
  - produces a similar range of colours, using Cyan (C), Magenta (M) and Yellow (Y)
  - white is produced when C=0, M=0, Y=0
  - good for producing a colour image on a white surface
  - used for printing applications
- Digitization yields a colour image that can be displayed directly on computer monitors (i.e. RGB)

Raster-scan Principles

- In most TV sets, the picture tubes use a raster scan
- Progressive scanning
  - top left to bottom right
- Frame
  - a complete set of horizontal lines
    - 525 (North and South America) or 625 (Europe)
- Frame rate
  - number of frames per second (Hz)

Raster-scan Principles (2)

- Most current picture tubes operate in analogue mode
- For digital TV and digitized pictures
  - colour signals comprise a string of pixels (fixed number of pixels per scan line)
  - each line is read from memory during the scanning process
  - each is converted to analogue using a digital-to-analogue converter (DAC)
  - normally a separate block of memory (video RAM) is used

Pixel Depth

- No. of bits per pixel
  - determines the range of different colours that can be produced
  - e.g.
    - 12 bits - 4 bits per primary colour - 4096 colours
    - 24 bits - 8 bits per primary colour - 2^24 colours
  - In practice, only a subset of colours are used
    - human eye can’t discriminate between all of them
    - a colour look-up table (CLUT) is used to store them
Aspect Ratio

- Ratio of screen width to screen height
- Current older TVs - 4/3
- Widescreen TVs - 16/9
- National Television Standards Committee (NTSC) uses 525 scan lines per frame
- PAL (UK), CCIR (Germany) and SECAM (France) use 625 scan lines per frame
- Not all lines are displayed on screen
- Some lines carry control information

Aspect Ratio (2)

<table>
<thead>
<tr>
<th>Standard</th>
<th>Scan lines per frame</th>
<th>Visible lines per frame</th>
<th>Pixels per line with 4/3 aspect ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTSC</td>
<td>525</td>
<td>480</td>
<td>640</td>
</tr>
<tr>
<td>PAL, CCIR, SECAM</td>
<td>625</td>
<td>576</td>
<td>768</td>
</tr>
</tbody>
</table>

Video

- Broadcast TV
- Digital TV
- PC video

Broadcast Television

- Scanning sequence
  - minimum refresh rate of 50 times per second is needed to avoid flicker
  - refresh rate of 25 frames per second is enough to produce smooth motion
  - therefore, to ease transmission bandwidth, the image associated with each frame is transmitted in two halves (fields)
    - odd scan lines, then even scan lines
    - interlaced scanning - integration of the two fields

Broadcast Television (2)

- Colour signals
  - different set of colour signals from R, G, B are used:
    - brightness
      - amount of energy that stimulates the eye
    - hue
      - actual colour of the source
    - saturation
      - strength or vividness of the colour
  - luminance
  - chrominance

Broadcast Television (3)

- A range of colours can be produced by varying the magnitude of the three electrical signals that energise the red, green and blue phosphors
  - \( 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B = \text{white} \)
  - luminance is only a function of amount of white light in a source
    - therefore, can be determined:
      - \( Y_s = 0.299 \cdot R_s + 0.587 \cdot G_s + 0.114 \cdot B_s \)
  - two other signals used to represent colouration
    - blue chrominance \( C_b = \text{hue} \)
    - red chrominance \( C_r = \text{saturation} \)
Broadcast Television (4)
\[ \Delta C_b \text{ and } C_r \text{ are obtained from the two colour difference signals} \]
\[ C_b = B_s - Y_s \text{ and } C_r = R_s - Y_s \]
\[ \Delta \text{ since } Y \text{ is a function of all three colours, } G \text{ can be computed from these two signals} \]
\[ \Delta \text{ Therefore, the combination of three signals } Y, C_b \text{ and } C_r: \]
\[ \cdot \text{ contains all the information needed to describe a colour signal} \]
\[ \cdot \text{ also compatible with monochrome TV sets} \]
\[ \cdot \text{ to fit these into the same bandwidth as used by monochrome TVs, they are combined to produce a composite video signal} \]

Broadcast Television (5)
\[ \Delta \text{ In practice the magnitude of the two colour difference signals are scaled down} \]
\[ \cdot \text{ otherwise the amplitude of the luminance signal can become greater than that of the equivalent monochrome signal} \]
\[ - \text{ degradation of monochrome picture} \]
\[ \cdot \text{ PAL} \]
\[ - C_b \text{ and } C_r \text{ are referred to as } U \text{ and } V \]
\[ - Y = 0.299R + 0.587G + 0.114B \]
\[ - U = 0.493(B-Y) \]
\[ - V = 0.877(R-Y) \]
\[ \cdot \text{ NTSC} \]
\[ - C_b \text{ and } C_r \text{ are combined to give two different signals, } I \text{ and } Q \]

Digital Video
\[ \Delta \text{ The resolution of the eye is less sensitive to colour than to luminance} \]
\[ \Delta \text{ the two chrominance signals can tolerate a reduced resolution relative to that of the luminance signal} \]
\[ \cdot \text{ so use luminance and the two colour difference signals rather than RGB} \]
\[ \Delta \text{ various digitization formats exploit this:} \]
\[ \cdot \text{ 4:2:2 format, 4:2:0 format} \]
\[ \cdot \text{ HDTV formats (SIF, CIF, QCIF)} \]

Digital Video (2)
\[ \Delta \text{ 4:2:2 format} \]
\[ \Delta \text{ original digitization format used for TV studios} \]
\[ \Delta \text{ 4Y samples for every 2C}_b \text{ and } 2C_r \text{ samples} \]
\[ \cdot \text{ i.e. reduced resolution of } C_b \text{ and } C_r \]
\[ \Delta \text{ 4:2:0 format} \]
\[ \Delta \text{ a derivative of 4:2:2 format} \]
\[ \Delta \text{ used in digital video broadcast applications} \]

HDTV Formats - SIF
\[ \Delta \text{ Source intermediate format} \]
\[ \Delta \text{ picture quality comparable to a VCR} \]
\[ \Delta \text{ uses half the spatial resolution horizontally and vertically as 4:2:0 format} \]
\[ \cdot \text{ subsampling} \]
\[ \Delta \text{ uses half the refresh rate (temporal resolution)} \]

HDTV Formats - CIF
\[ \Delta \text{ Common intermediate format} \]
\[ \Delta \text{ for use in videoconferencing applications} \]
\[ \Delta \text{ derived from SIF} \]
\[ \Delta \text{ uses a combination of:} \]
\[ \cdot \text{ the spatial resolution used for SIF in the 625-line system} \]
\[ \cdot \text{ the temporal resolution used by SIF in the 525-line system} \]
HDTV Formats - QCIF

- **Quarter CIF**
  - For use in video telephony applications
  - Derived from CIF
  - Uses half the spatial resolution of CIF horizontally and vertically
  - Uses half or a quarter of the temporal resolution of CIF

PC Video

- Digitization formats described are used with PC based video
- But resolutions differ in order to avoid distortion on the computer monitor

Summary

- Multimedia formats
- Digitization
- Digitized pictures
- Broadcast video
- Video digitization formats