Video Processing & Communications

Basics of Video

Yao Wang
Polytechnic University, Brooklyn, NY11201
yao@vision.poly.edu
Outline

- Color perception and specification (review on your own)
- Video capture and display (review on your own)
- Analog raster video
- Analog TV systems
- Digital video
Analog Video

- Video raster
- Progressive vs. interlaced raster
- Analog TV systems
Raster Scan

- Real-world scene is a continuous 3-D signal (temporal, horizontal, vertical)
- Analog video is stored in the raster format
  - Sampling in time: consecutive sets of frames
    - To render motion properly, >=30 frame/s is needed
  - Sampling in vertical direction: a frame is represented by a set of scan lines
    - Number of lines depends on maximum vertical frequency and viewing distance, 525 lines in the NTSC system
  - Video-raster = 1-D signal consisting of scan lines from successive frames
Progressive and Interlaced Scans

Interlaced scan is developed to provide a trade-off between temporal and vertical resolution, for a given, fixed data rate (number of line/sec).
Waveform and Spectrum of an Interlaced Raster

(a) Horizontal retrace for first field
Vertical retrace from first to second field
Vertical retrace from second to third field

Blanking level
Black level
White level

(b) $|\Psi(f)|$

0 $f_1$ $2f_1$ $3f_1$ $f_{max}$
Color TV Broadcasting and Receiving

Diagram:

- RGB ---> YC1C2
- Luminance, Chrominance, Audio Multiplexing ---> Modulation
- Modulation

- De-Multiplexing
- De-Modulation

Yao Wang, 2013 Video Basics
Why not using RGB directly?

- R, G, B components are correlated
  - Transmitting R, G, B components separately is redundant
  - More efficient use of bandwidth is desired
- RGB->YC1C2 transformation
  - Decorrelating: Y, C1, C2 are uncorrelated
  - C1 and C2 require lower bandwidth
  - Y (luminance) component can be received by B/W TV sets
- YIQ in NTSC
  - I: orange-to-cyan
  - Q: green-to-purple (human eye is less sensitive)
    - Q can be further bandlimited than I
  - Phase=Arctan(Q/I) = hue, Magnitude=sqrt (I^2+Q^2) = saturation
  - Hue is better retained than saturation
Color Image

Y image

I image (orange-cyan)

Q image (green-purple)

Yao Wang, 2013  Video Basics  9
I and Q on the color circle

- Q: green-purple
- I: orange-cyan
Conversion between RGB and YIQ

- **RGB -> YIQ**
  
  \[ Y = 0.299 \, R + 0.587 \, G + 0.114 \, B \]
  \[ I = 0.596 \, R - 0.275 \, G - 0.321 \, B \]
  \[ Q = 0.212 \, R - 0.523 \, G + 0.311 \, B \]

- **YIQ -> RGB**
  
  \[ R = 1.0 \, Y + 0.956 \, I + 0.620 \, Q \]
  \[ G = 1.0 \, Y - 0.272 \, I - 0.647 \, Q \]
  \[ B = 1.0 \, Y - 1.108 \, I + 1.700 \, Q \]
TV signal bandwidth

- Luminance
  - Maximum vertical frequency (cycles/picture-height) = black and white lines interlacing
    \[ f_{v,\text{max}} = \frac{Kf'_{s,y}}{2} \]
  - Maximum horizontal frequency (cycles/picture-width)
    \[ f_{h,\text{max}} = f_{v,\text{max}} \cdot \text{IAR} \]
  - Corresponding temporal frequency (cycles/second or Hz)
    \[ f_{\text{max}} = \frac{f_{h,\text{max}}}{T'l} = \text{IAR} \cdot \frac{Kf'_{s,y}}{2T'l} \]
  - For NTSC, \[ f_{\text{max}} = 4.2 \text{ MHz} \]

- Chrominance
  - Can be bandlimited significantly
    - I: 1.5 MHz, Q: 0.5 MHz.
Bandwidth of Chrominance Signals

- Theoretically, for the same line rate, the chrominance signal can have as high frequency as the luminance signal.
- However, with real video signals, the chrominance component typically changes much slower than luminance.
- Furthermore, the human eye is less sensitive to changes in chrominance than to changes in luminance.
- The eye is more sensitive to the orange-cyan range (I) (the color of face!) than to green-purple range (Q).
- The above factors lead to:
  - I: bandlimited to 1.5 MHz
  - Q: bandlimited to 0.5 MHz
Multiplexing of Luminance and Chrominance

- Chrominance signal can be bandlimited
  - it usually has a narrower frequency span than the luminance and the human eye is less sensitive to high frequencies in chrominance
- The two chrominance components (I and Q) are multiplexed onto the same sub-carrier using QAM
  - The upper band of I is limited to 0.5 MHz to avoid interference with audio
- Position the bandlimited chrominance at the high end spectrum of the luminance, where the luminance is weak, but still sufficiently lower than the audio (at 4.5 MHz=286 \( f_i \))
- The actual position should be such that the peaks of chrominance spectrum interlace with those of the luminance
  \[
  f_c = \frac{455 f_i}{2} \ (\text{= 3.58 Hz for NTSC})
  \]
Spectrum Illustration

Luminance

Chrominance

$\Psi(f)$

$0 \quad f_i \quad 2f_i \quad 3f_i \quad 225f_i \quad 226f_i \quad 227f_i \quad 228f_i \quad 229f_i \quad 230f_i$

$f_c$

(Color subcarrier)
Multiplexing of luminance, chrominance and audio (Composite Video Spectrum)

- Luminance
- I
- I and Q
- Audio

- Picture carrier: $f_p$
- Color subcarrier: $f_c$
- Audio subcarrier: $f_a$

- 6.0 MHz
- 4.5 MHz
- 1.25 MHz
- 4.2 MHz
- 3.58 MHz

Yao Wang, 2013

Video Basics
Quadrature Amplitude Modulation (QAM)

- A method to modulate two signals onto the same carrier frequency, but with $90^\circ$ phase shift

\[
\begin{align*}
\cos(2\pi f_1 t) \\
\sin(2\pi f_1 t)
\end{align*}
\]

QAM modulator

\[
\begin{align*}
\cos(2\pi f_1 t) \\
\sin(2\pi f_1 t)
\end{align*}
\]

QAM demodulator

Yao Wang, 2013 Video Basics 17
Adding Color Bursts for Synchronization

For accurate regeneration of the color sub-carrier signal at the receiver, a color burst signal is added during the horizontal retrace period.

Figure from From Grob, Basic Color Television Principles and Servicing, McGraw Hill, 1975
http://www.ee.washington.edu/conselec/CE/kuhn/ntsc/95x417.gif
Multiplexing of Luminance and Chrominance

The diagram illustrates the process of multiplexing luminance (Y) and chrominance (I, Q) signals. The luminance signal (Y(t)) is passed through a low-pass filter (LPF) with a frequency response of 0-4.2MHz. Similarly, the chrominance signals (I(t), Q(t)) are filtered by LPFs with frequency responses of 0-1.5MHz and 0-0.5MHz, respectively. The filtered luminance and chrominance signals are then combined using a summing junction (Σ). The combined signal is then passed through a band-pass filter (BPF) with a frequency response of 2-4.2MHz. The output of the BPF is further modified by a gate and an amplitude modulation (Acos(2πfₜ)), resulting in the composite video signal. The color burst signal is also generated and combined with the composite video signal.
DeMultiplexing of Luminance and Chrominance

Composite video

Comb Filter 0-4.2MHz

Y(t)

I(t)

LPF 0-1.5MHz

Q(t)

LPF 0-0.5MHz

Horizontal sync signal

Gate

2Acos(2πf_c t)

Phase comparator

Voltage controlled oscillator

-π/2

-π/2
Luminance/Chrominance Separation

• In low-end TV receivers, a low pass filter with cut-off frequency at 3MHz is typically used to separate the luminance and chrominance signal.
  – The high frequency part of the I component (2 to 3 Mhz) is still retained in the luminance signal.
  – The extracted chrominance components can contain significant luminance signal in a scene with very high frequency (luminance energy is not negligible near $f_c$)
  – These can lead to color bleeding artifacts
• For better quality, a comb filter can be used, which will filter out harmonic peaks correspond to chrominance signals.
• Show example of comb filter on board
What will a Monochrome TV see?

• The monochrome TV receiver uses a LPT with cut-off at 4.2 MHz, and thus will get the composite video (baseband luminance plus the I and Q signal modulated to $f_c = 3.58$ MHz)
  – Because the modulated chrominance signal is at very high frequency (227.5 cycles per line), the eye smoothes it out mostly, but there can be artifacts
  – The LPF in Practical TV receivers have wide transition bands, and the response is already quite low at $f_c$. 
Color TV Broadcasting and Receiving
Transmitter in More Details

Audio

R(t) → Y(t) → LPF 0-4.2MHz

G(t) → I(t) → LPF 0-1.5MHz

B(t) → Q(t) → LPF 0-0.5MHz

\[ A_{\cos}(2\pi f_c t) \]

\( -\pi/2 \)

Gate

BPF 2-4.2MHz

\( \sum \)

To transmit antenna

VSB

To transmit antenna
Receiver in More Details

From antenna

BPF, 4.4-4.6MHz

BPF, 0.4-2 MHz

VSB Demodulator

Composite video

Comb Filter
0-4.2MHz

Comb Filter
0-4.2MHz

2Acos(2πfct)

Gate

2Acos(2πfct)

Horizontal sync signal

Phase comparator

Voltage controlled oscillator

FM demodulator

Audio

To speaker

To CRT

Y(t)

R(t)

LPF
0-1.5MHz

LPF
0-0.5MHz

LPF
0-1.5MHz

L(t)

Q(t)

G(t)

B(t)

YIQ to RGB conversion

Horizontal sync signal

From antenna

Yao Wang, 2013

Video Basics
Original color frame

Recovered color frame
### Different Color TV Systems

<table>
<thead>
<tr>
<th>Parameters</th>
<th>NTSC</th>
<th>PAL</th>
<th>SECAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Rate (Hz)</td>
<td>59.95 (60)</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Line Number/Frame</td>
<td>525</td>
<td>625</td>
<td>625</td>
</tr>
<tr>
<td>Line Rate (Line/s)</td>
<td>15,750</td>
<td>15,625</td>
<td>15,625</td>
</tr>
<tr>
<td>Color Coordinate</td>
<td>YIQ</td>
<td>YUV</td>
<td>YDbDr</td>
</tr>
<tr>
<td>Luminance Bandwidth (MHz)</td>
<td>4.2</td>
<td>5.0/5.5</td>
<td>6.0</td>
</tr>
<tr>
<td>Chrominance Bandwidth (MHz)</td>
<td>1.5(I)/0.5(Q)</td>
<td>1.3(U,V)</td>
<td>1.0 (U,V)</td>
</tr>
<tr>
<td>Color Subcarrier (MHz)</td>
<td>3.58</td>
<td>4.43</td>
<td>4.25(Db),4.41(Dr)</td>
</tr>
<tr>
<td>Color Modulation</td>
<td>QAM</td>
<td>QAM</td>
<td>FM</td>
</tr>
<tr>
<td>Audio Subcarrier</td>
<td>4.5</td>
<td>5.5/6.0</td>
<td>6.5</td>
</tr>
<tr>
<td>Total Bandwidth (MHz)</td>
<td>6.0</td>
<td>7.0/8.0</td>
<td>8.0</td>
</tr>
</tbody>
</table>
Who uses what?

From http://www.stjarnhimlen.se/tv/tv.html#worldwide_0
Digital Video

- Digital video by sampling/quantizing analog video raster → BT.601 video
- Directly capture digital video using digital cameras
- A digital video (including all color components) is compressed into a bit stream, which can be stored on a disk or transmitted over the air or through wires
- Transmission is achieved through digital modulation
  - Converting each bit or a group of bits into a preset waveform
Digitizing A Raster Video

- Sample the raster waveform = Sample along the horizontal direction
- Sampling rate must be chosen properly
  - For the samples to be aligned vertically, the sampling rate should be multiples of the line rate
  - Horizontal sampling interval = vertical sampling interval
  - Total sampling rate equal among different systems

\[ f_s = 858 f_i (NTSC) = 864 f_i (PAL/SECAM) = 13.5 \text{ MHz} \]
BT.601* Video Format (commonly known as SDTV)

* BT.601 is formerly known as CCIR601
RGB $\leftrightarrow$ YCbCr

\[ Y_d = 0.257\, R_d + 0.504\, G_d + 0.098\, B_d + 16, \]
\[ C_b = -0.148\, R_d - 0.291\, G_d + 0.439\, B_d + 128, \]
\[ C_r = 0.439\, R_d - 0.368\, G_d - 0.071\, B_d + 128, \]

\[ R_d = 1.164\, Y_d' + 0.0\, C_b' + 1.596\, C_r', \]
\[ G_d = 1.164\, Y_d' - 0.392\, C_b' - 0.813\, C_r', \]
\[ B_d = 1.164\, Y_d' + 2.017\, C_b' + 0.0\, C_r', \]

\[ Y_d' = Y_d - 16, \quad C_b' = C_b - 128, \quad C_r' = C_r - 128 \]
Chrominance Subsampling Formats

4:4:4
For every 2x2 Y Pixels
4 Cb & 4 Cr Pixel
(No subsampling)

4:2:2
For every 2x2 Y Pixels
2 Cb & 2 Cr Pixel
(Subsampling by 2:1 horizontally only)

4:1:1
For every 4x1 Y Pixels
1 Cb & 1 Cr Pixel
(Subsampling by 4:1 horizontally only)

4:2:0
For every 2x2 Y Pixels
1 Cb & 1 Cr Pixel
(Subsampling by 2:1 both horizontally and vertically)

Y Pixel

Cb and Cr Pixel

Yao Wang, 2013

Video Basics
# Digital Video Formats

<table>
<thead>
<tr>
<th>Video Format</th>
<th>Y Size</th>
<th>Color Sampling</th>
<th>Frame Rate (Hz)</th>
<th>Raw Data Rate (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HDTV Over air. cable, satellite, MPEG2 video, 20-45 Mbps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMPTE296M</td>
<td>1280x720</td>
<td>4:2:0</td>
<td>24P/30P/60P</td>
<td>265/332/664</td>
</tr>
<tr>
<td>SMPTE295M</td>
<td>1920x1080</td>
<td>4:2:0</td>
<td>24P/30P/60I</td>
<td>597/746/746</td>
</tr>
<tr>
<td><strong>Video production, MPEG2, 15-50 Mbps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BT.601</td>
<td>720x480/576</td>
<td>4:4:4</td>
<td>60I/50I</td>
<td>249</td>
</tr>
<tr>
<td>BT.601</td>
<td>720x480/576</td>
<td>4:2:2</td>
<td>60I/50I</td>
<td>166</td>
</tr>
<tr>
<td><strong>High quality video distribution (DVD, SDTV), MPEG2, 4-10 Mbps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BT.601</td>
<td>720x480/576</td>
<td>4:2:0</td>
<td>60I/50I</td>
<td>124</td>
</tr>
<tr>
<td><strong>Intermediate quality video distribution (VCD, WWW), MPEG1, 1.5 Mbps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIF</td>
<td>352x240/288</td>
<td>4:2:0</td>
<td>30P/25P</td>
<td>30</td>
</tr>
<tr>
<td><strong>Video conferencing over ISDN/Internet, H.261/H.263, 128-384 Kbps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIF</td>
<td>352x288</td>
<td>4:2:0</td>
<td>30P</td>
<td>37</td>
</tr>
<tr>
<td><strong>Video telephony over wired/wireless modem, H.263, 20-64 Kbps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCIF</td>
<td>176x144</td>
<td>4:2:0</td>
<td>30P</td>
<td>9.1</td>
</tr>
</tbody>
</table>

Yao Wang, 2013
Video Terminology

- **Component video**
  - Three color components stored/transmitted separately
  - Use either RGB or YIQ (YUV) coordinate
  - New digital video format (YCrCb)
  - Betacam (professional tape recorder) use this format
- **Composite video**
  - Convert RGB to YIQ (YUV)
  - Multiplexing YIQ into a single signal
  - Used in most consumer analog video devices
- **S-video**
  - Y and C (QAM of I and Q) are stored separately
  - Used in high end consumer video devices
- **High end monitors can take input from all three**
Homework

• Reading assignment:
  – Chap. 1.
  – Specific technique for multiplexing / demultiplexing not required

• Problems:
  – Prob. 1.5.
  – Prob. 1.6.
  – Prob. 1.7.
  – Prob. 1.8.
  – Prob. 1.9.
  – Prob. 1.11