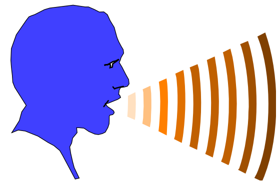
   

Video Processing & Communications

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

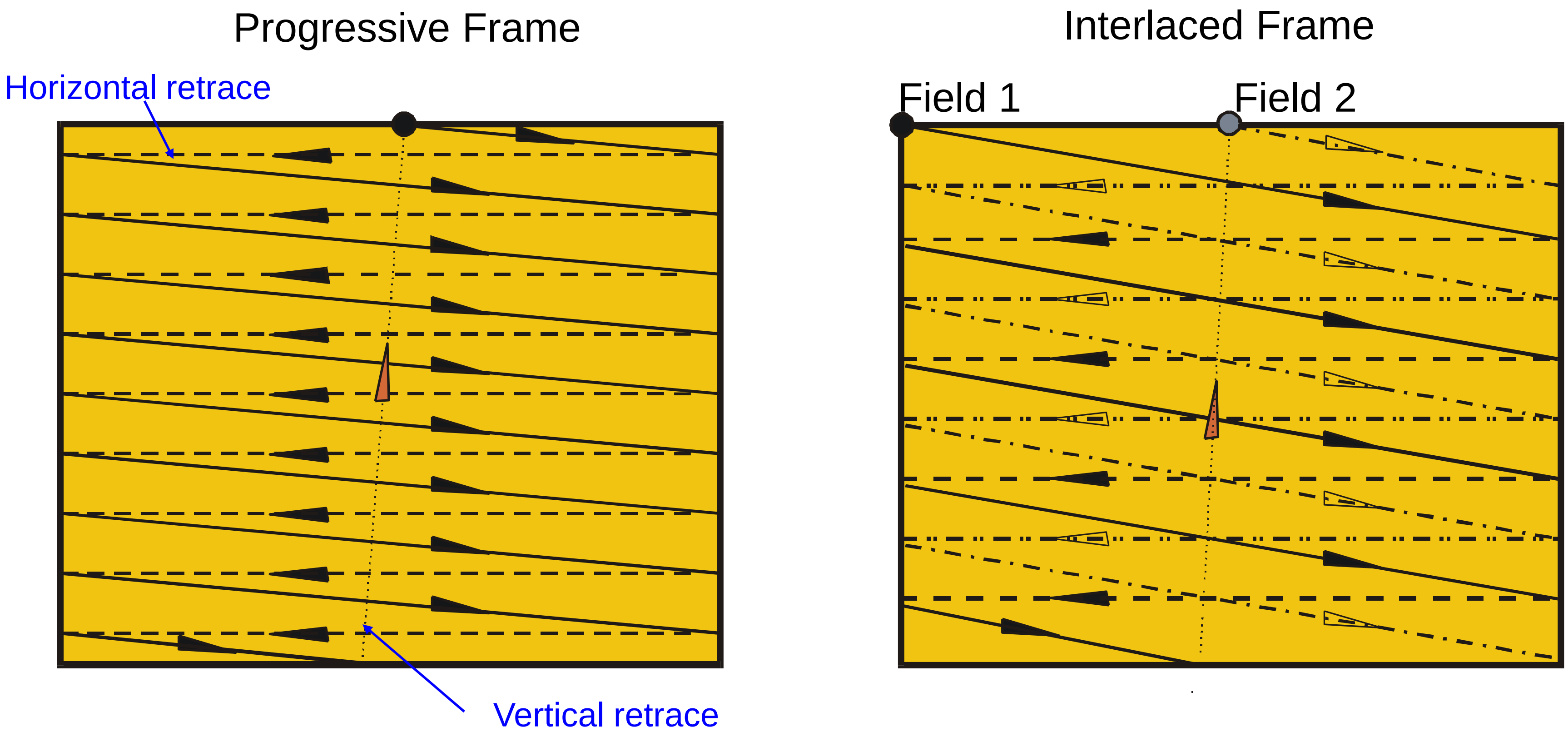
Analog Video

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

Raster Scan

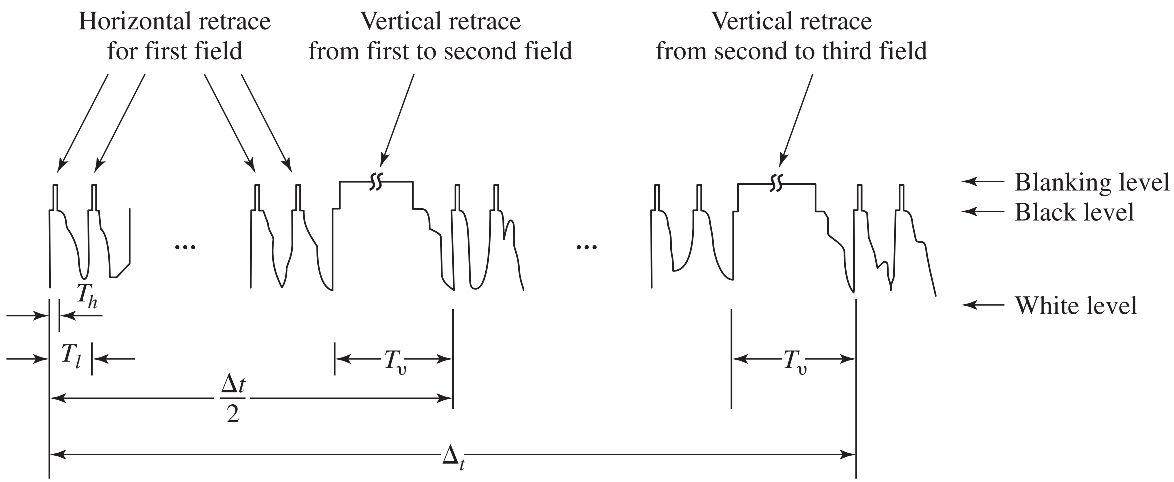
- Real-world scene is a continuous 3-D signal (temporal, horizontal, vertical)
- Analog video is stored in the raster format
  - Sampling in time: consecutive sets of frames
    - To render motion properly, >=30 frame/s is needed
  - Sampling in vertical direction: a frame is represented by a set of 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

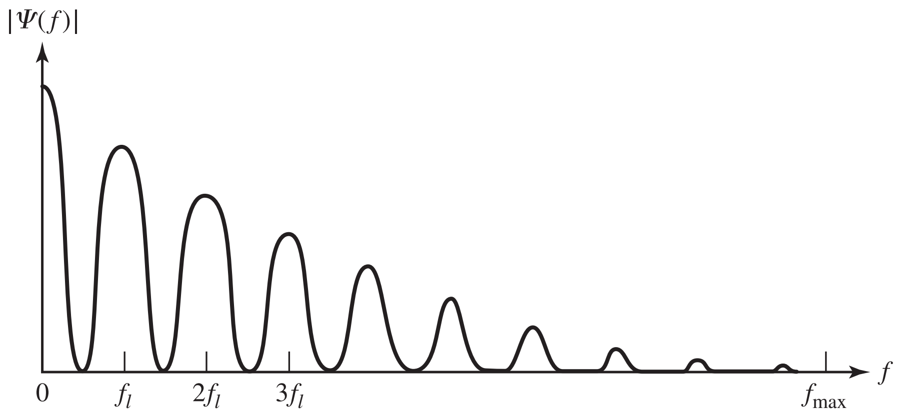


Interlaced scan is developed to provide a trade-off between temporal and vertical resolution, for a given, fixed data rate (number of line/sec).

Waveform and Spectrum of an Interlaced Raster

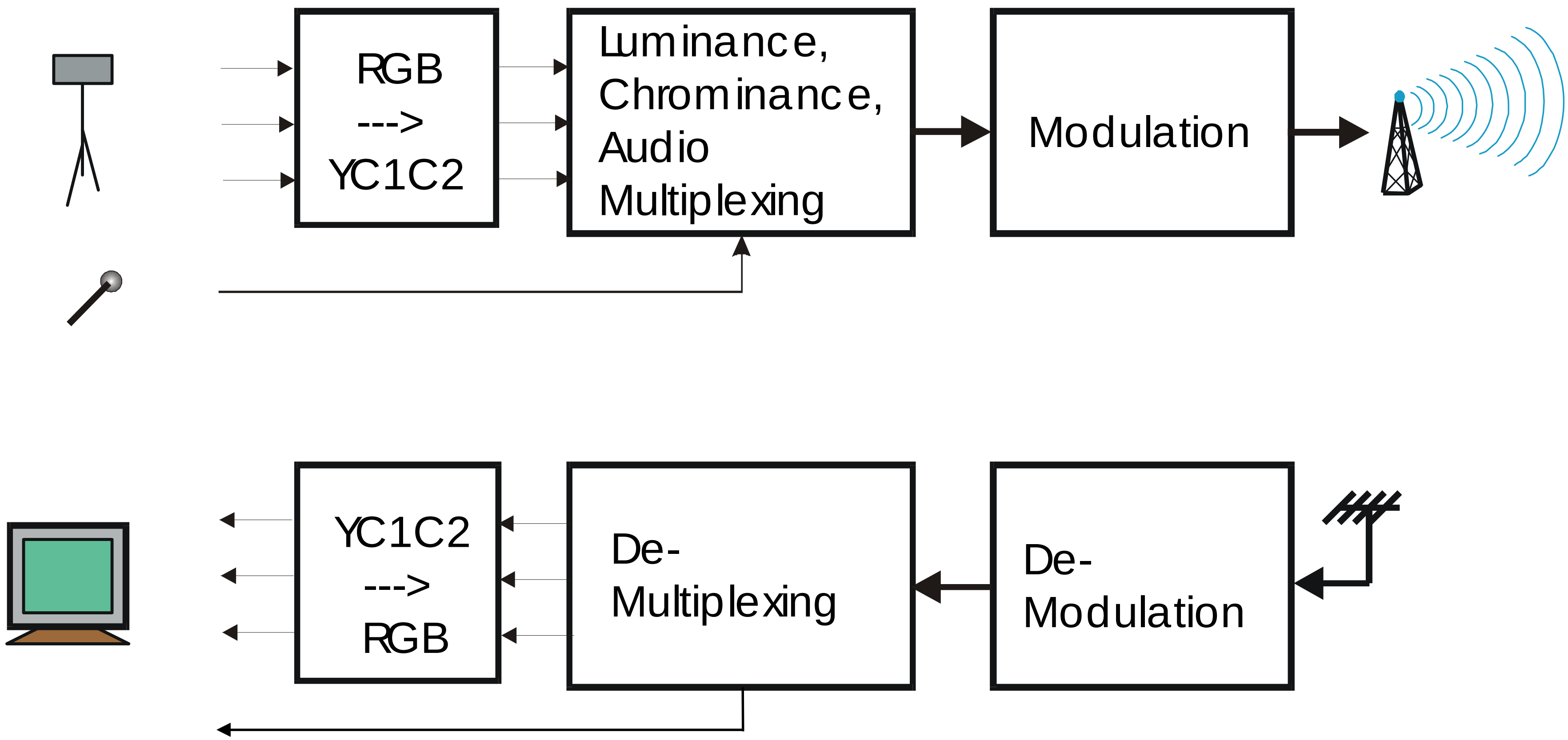


(a)



(b)

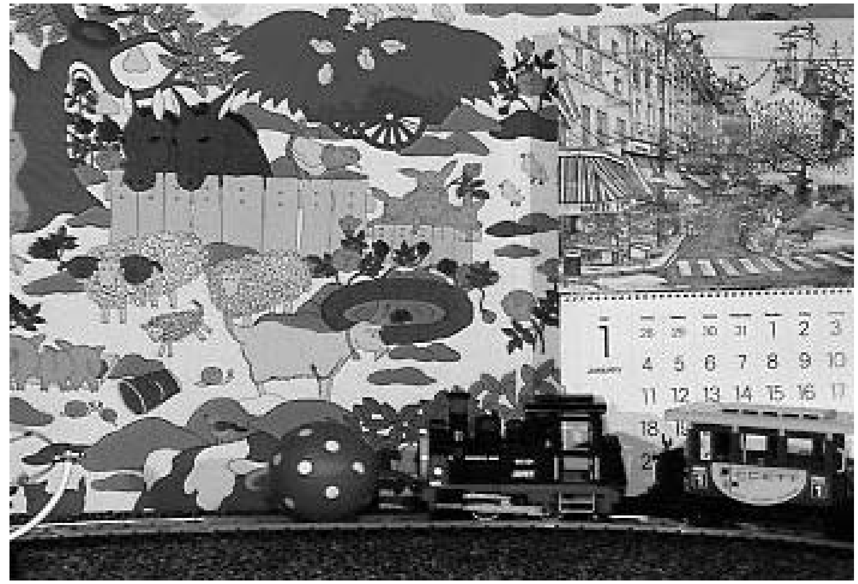
Color TV Broadcasting and Receiving



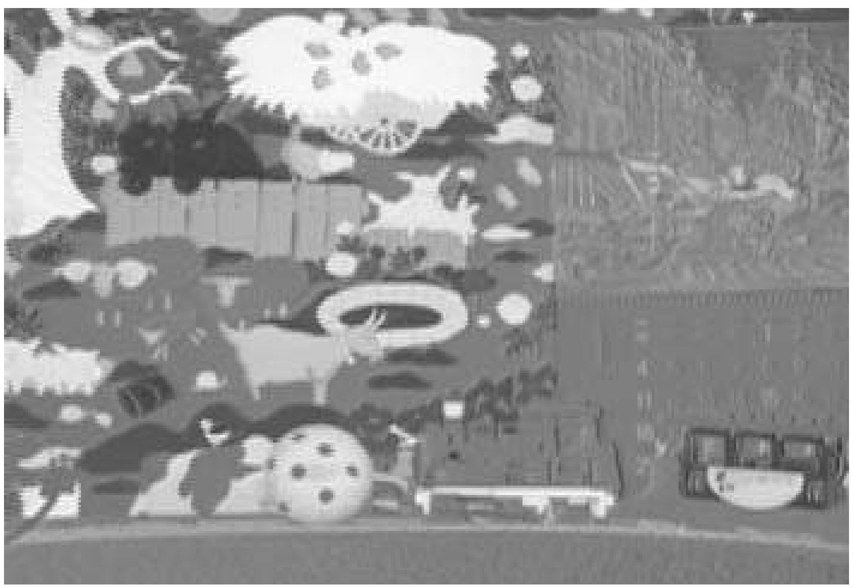
Why not using RGB directly?

- 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

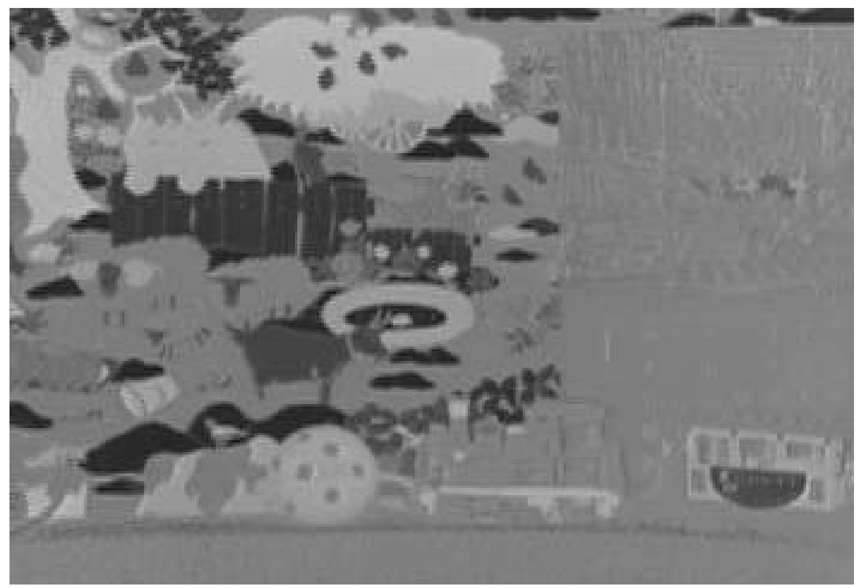
Color Image



Y image

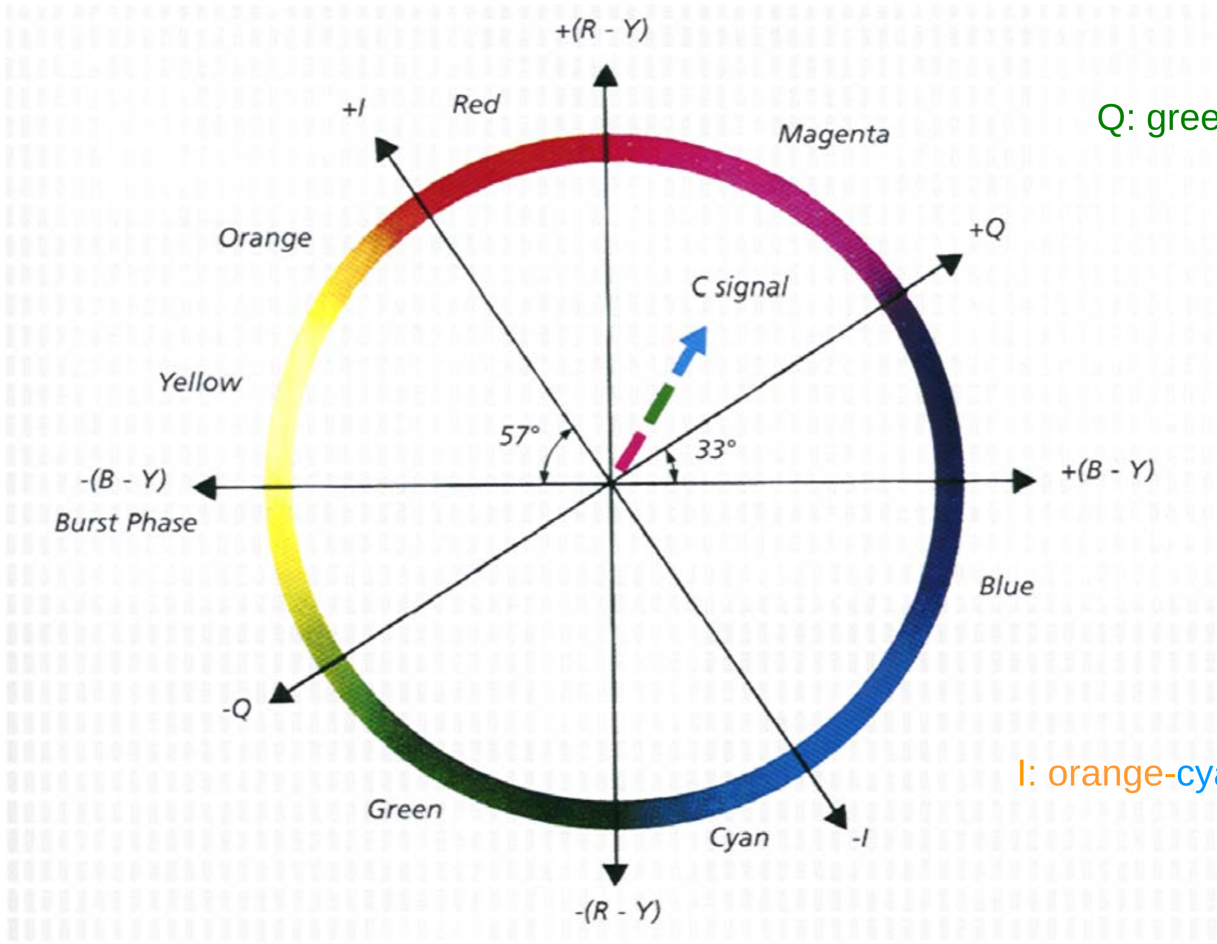


I image (orange-cyan)



Q image (green-purple)

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

    $$f_{v,\max} = Kf'_{s,y}/2$$

  - Maximum horizontal frequency (cycles/picture-width)

    $$f_{h,\max} = f_{v,\max} \cdot \mathrm{IAR}$$

  - Corresponding temporal frequency (cycles/second or Hz)

    $$f_{\max} = f_{h,\max}/T'_l = \mathrm{IAR} \cdot Kf'_{s,y}/2T'_l$$

  - For NTSC, $f_{\max} = 4.2\ \mathrm{MHz}$

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

Bandwidth of Chrominance Signals

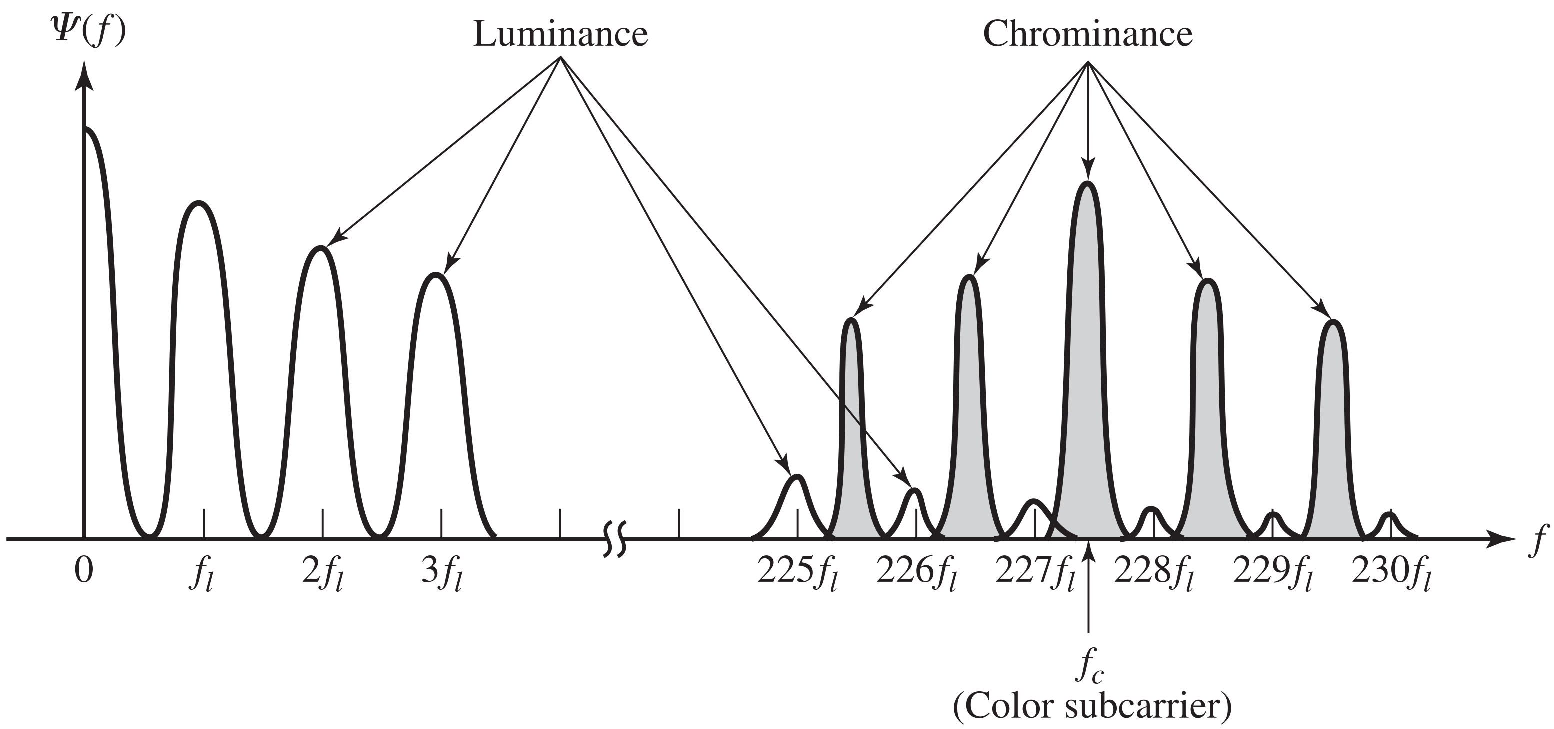
- Theoretically, for the same line rate, the chromiance signal can have as high frequency as the luminance signal
- However, with real video signals, the chrominance component typically changes much slower than luminance
- Furthermore, the human eye is less sensitive to changes in chrominance than to 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: bandlimitted to 1.5 MHz
  - Q: bandlimitted to 0.5 MHz

Multiplexing of Luminance and Chrominance

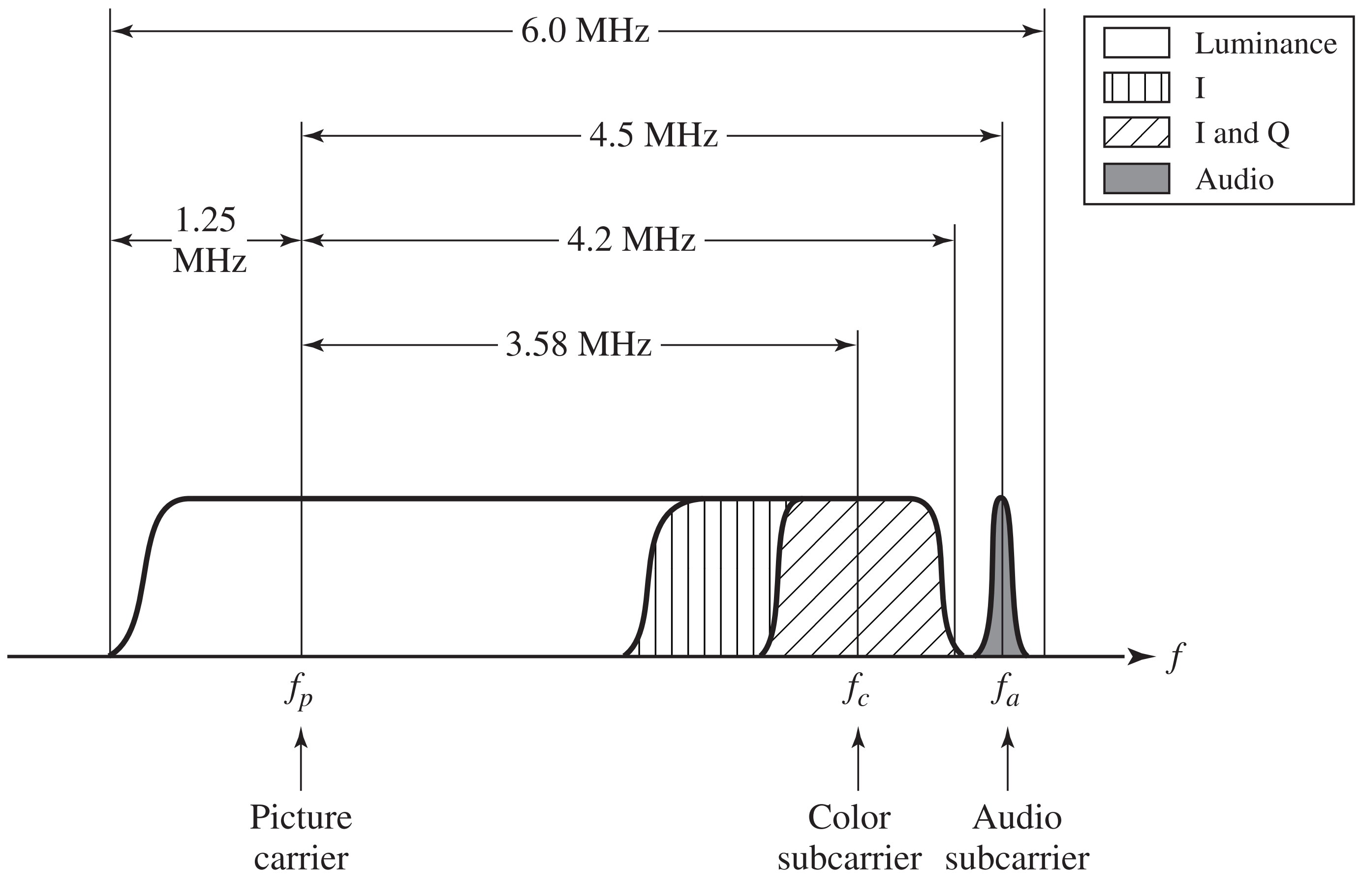
- Chrominance signal can be bandlimited
  - it usually has a narrower frequency span than the luminance and the human eye is less sensitive to high frequencies in chrominance
- The two chrominance components (I and Q) are multiplexed onto the same sub-carrier using QAM
  - The upper band of I is limited 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_l)
- The actual position should be such that the peaks of chrominance spectrum interlace with those of the luminance

$$f_c = 455 f_l / 2 \ \ (= 3.58 \ \text{Hz for NTSC})$$

Spectrum Illustration

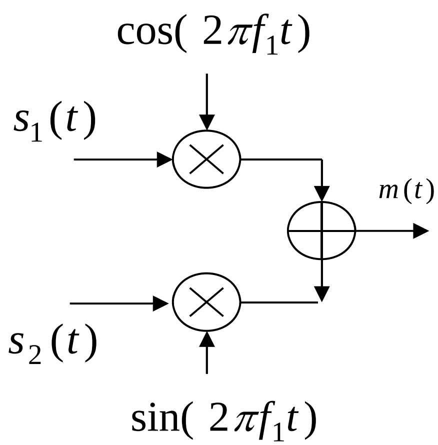


Multiplexing of luminance, chrominance and audio (Composite Video Spectrum)

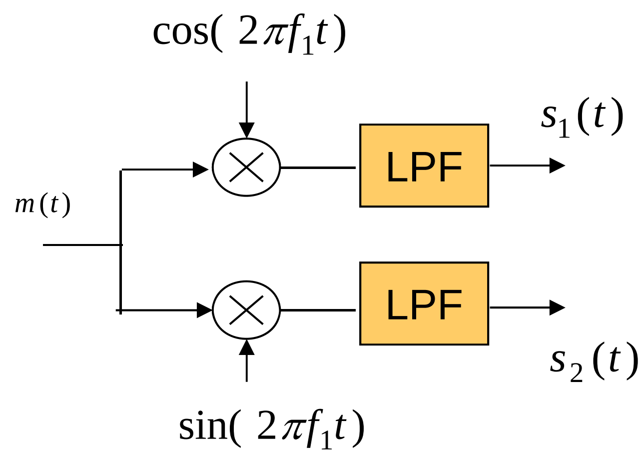


Quadrature Amplitude Modulation (QAM)

- A method to modulate two signals onto the same carrier frequency, but with 90º phase shift

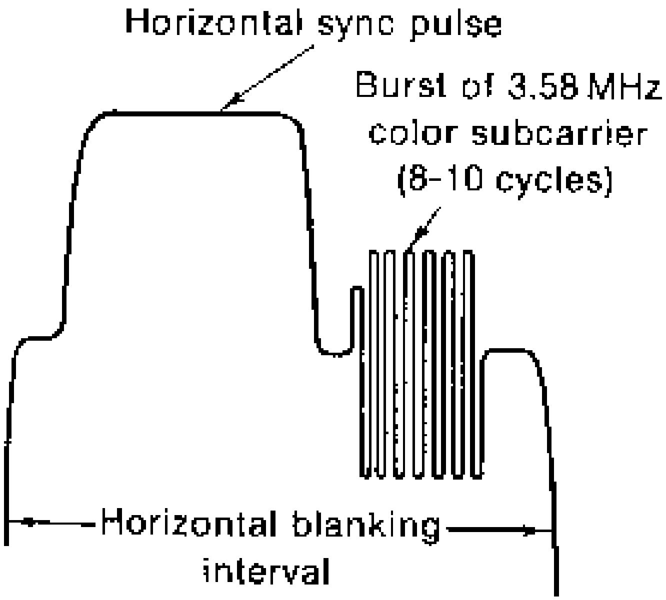


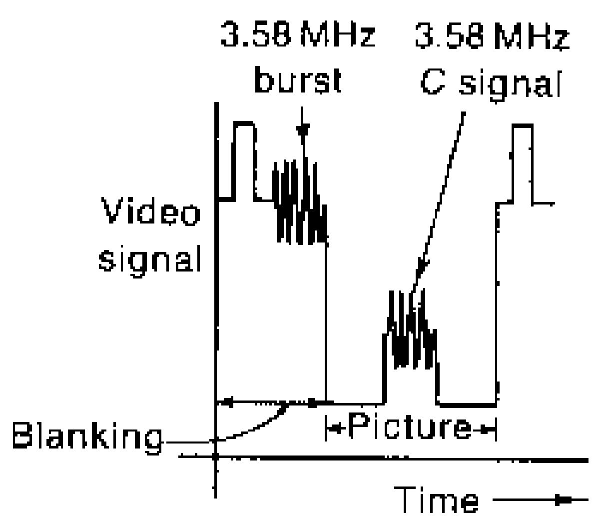
QAM modulator



QAM demodulator

Adding Color Bursts for Synchronization

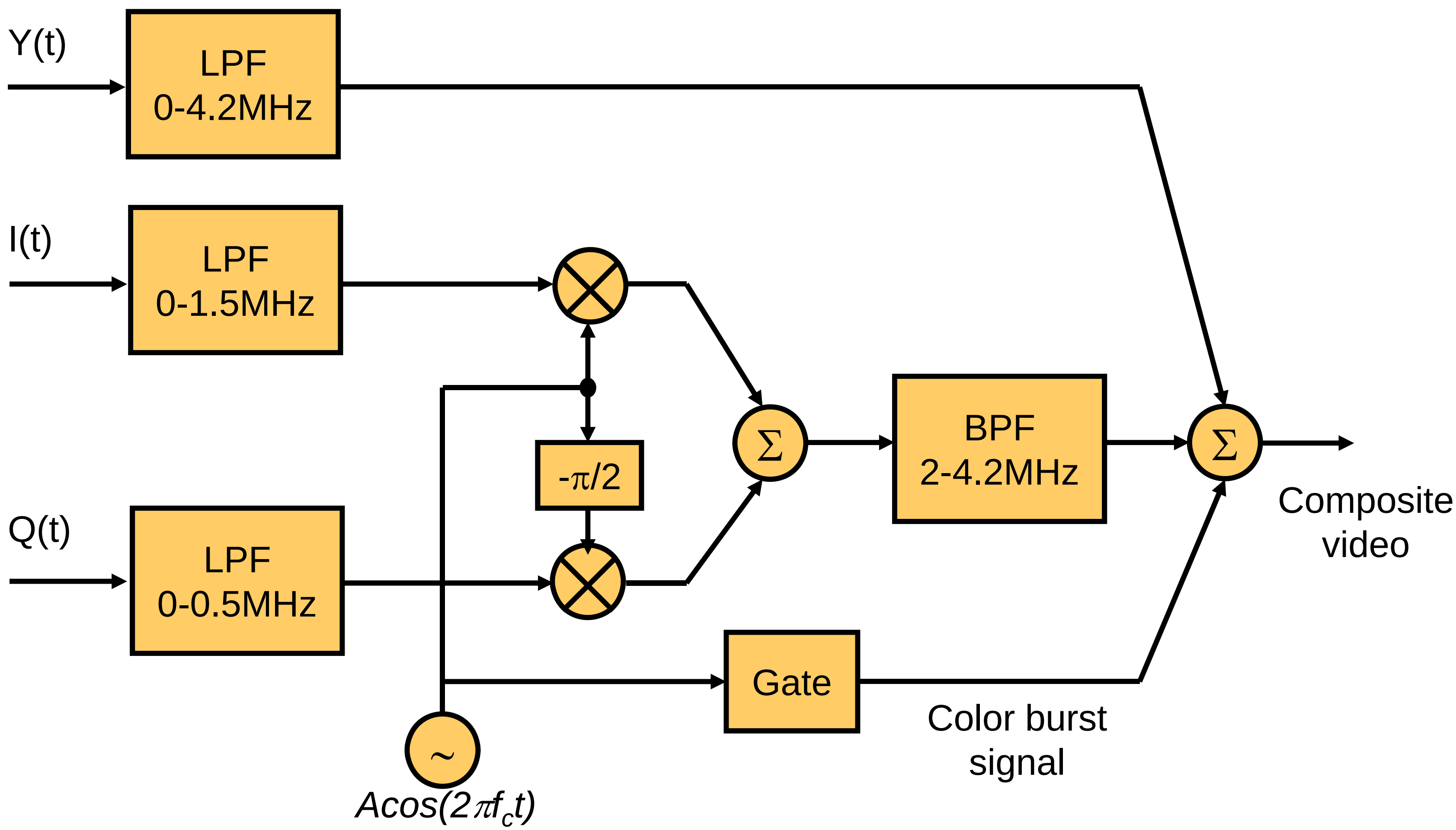




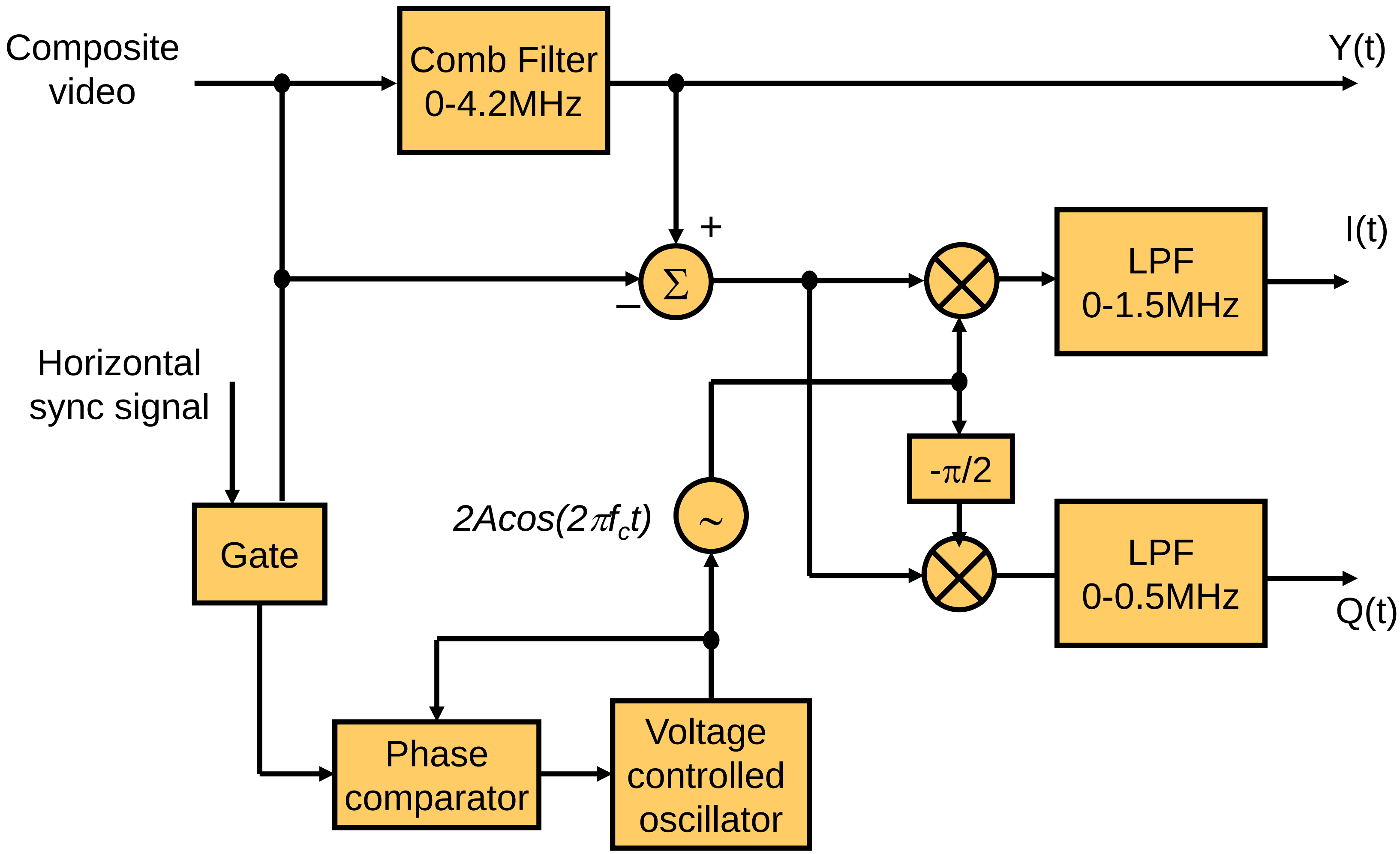
For accurate regeneration of the color sub-carrier signal at the receiver, a color burst signal is added during the horizontal retrace period

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



DeMultiplexing of Luminance and Chrominance



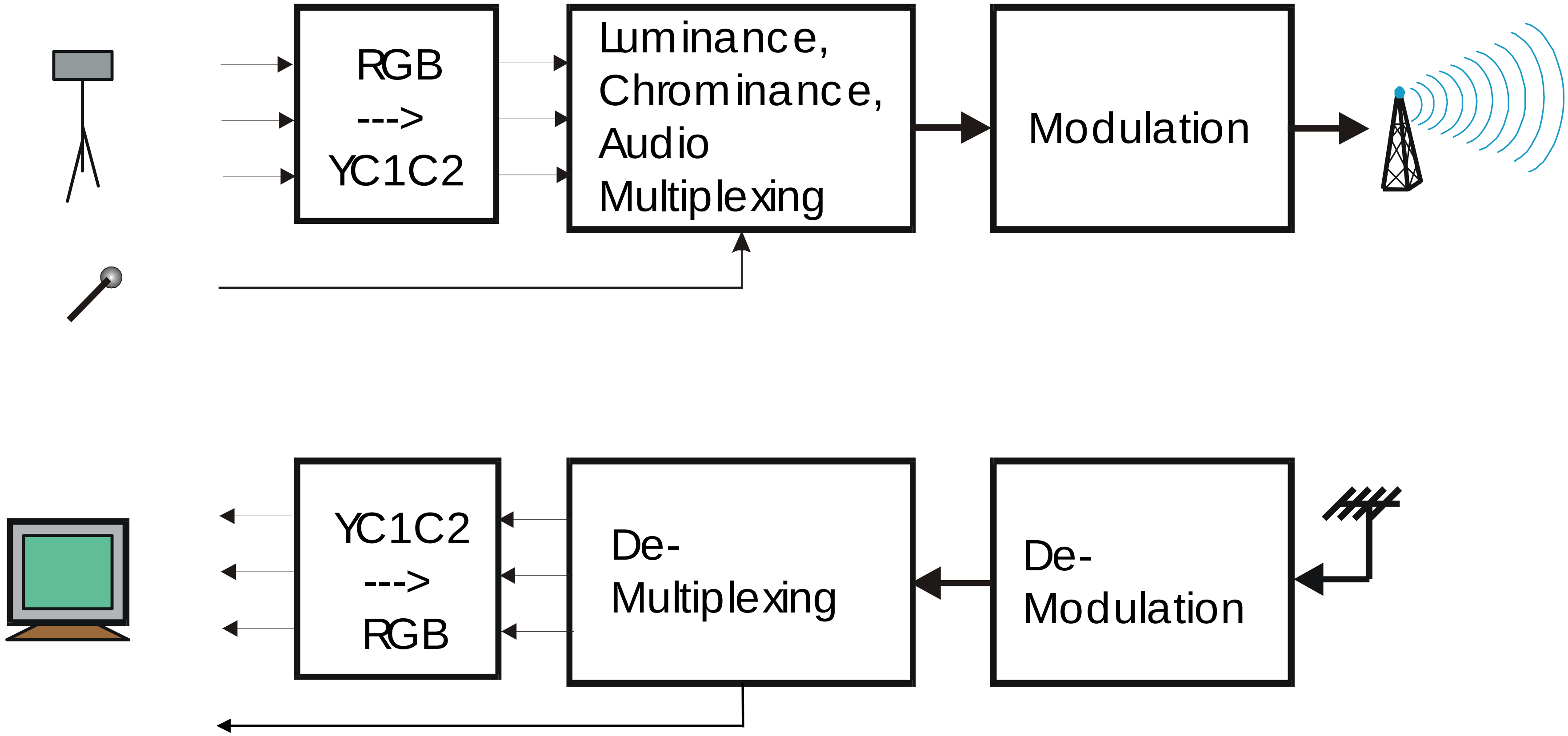
Luminance/Chrominance Separation

- 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

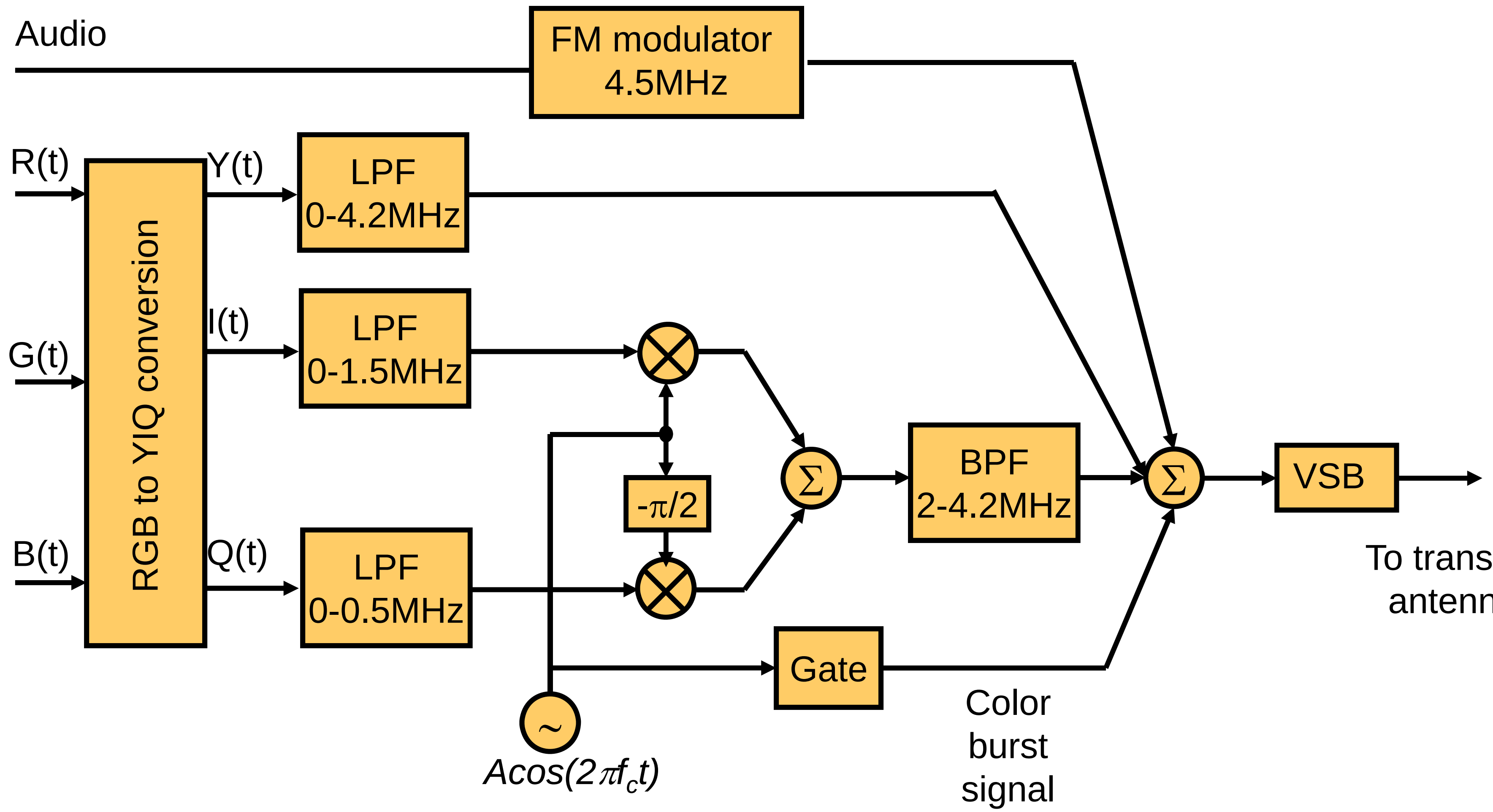
What will a Monochrome TV see?

- 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.

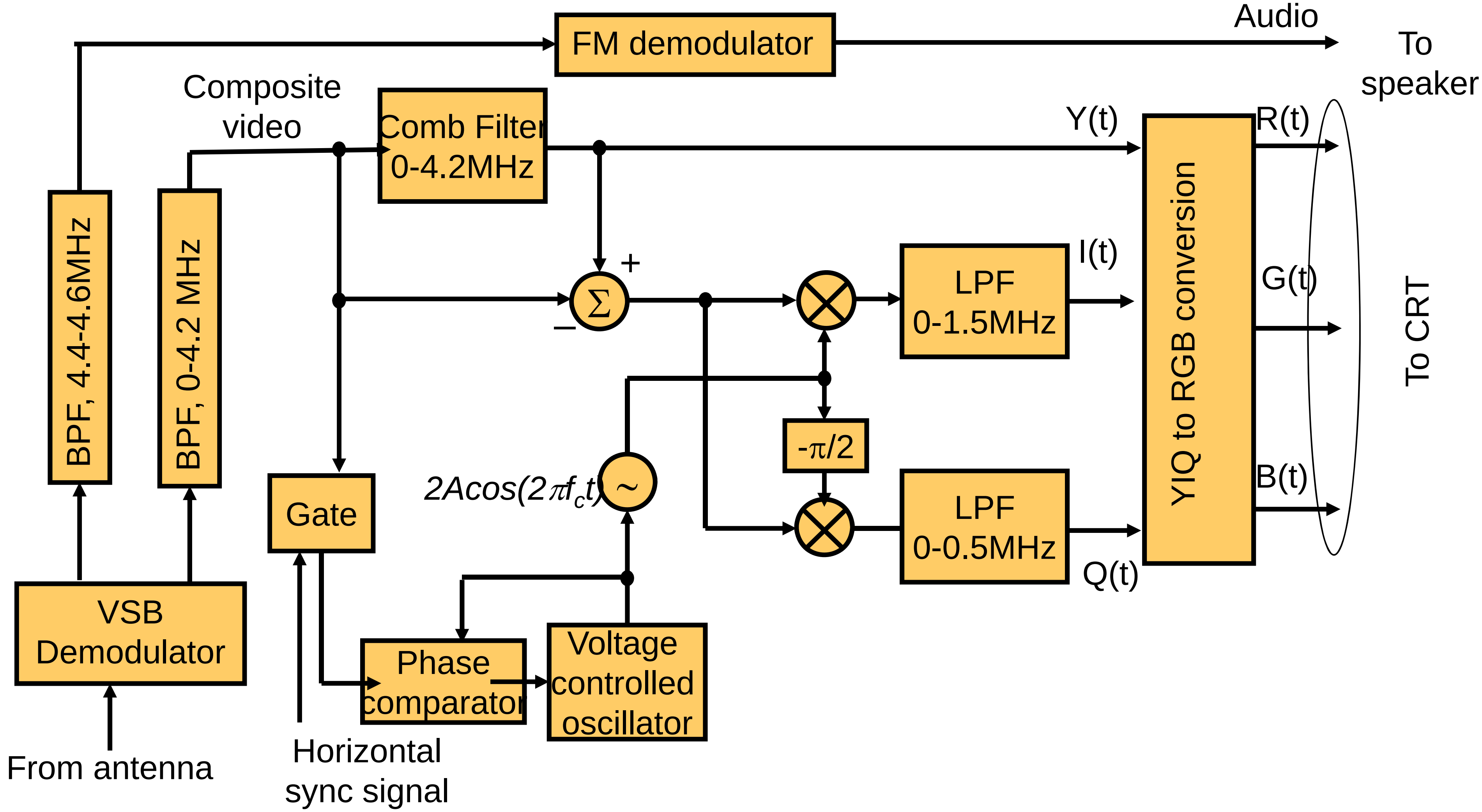
Color TV Broadcasting and Receiving



Transmitter in More Details



Receiver in More Details



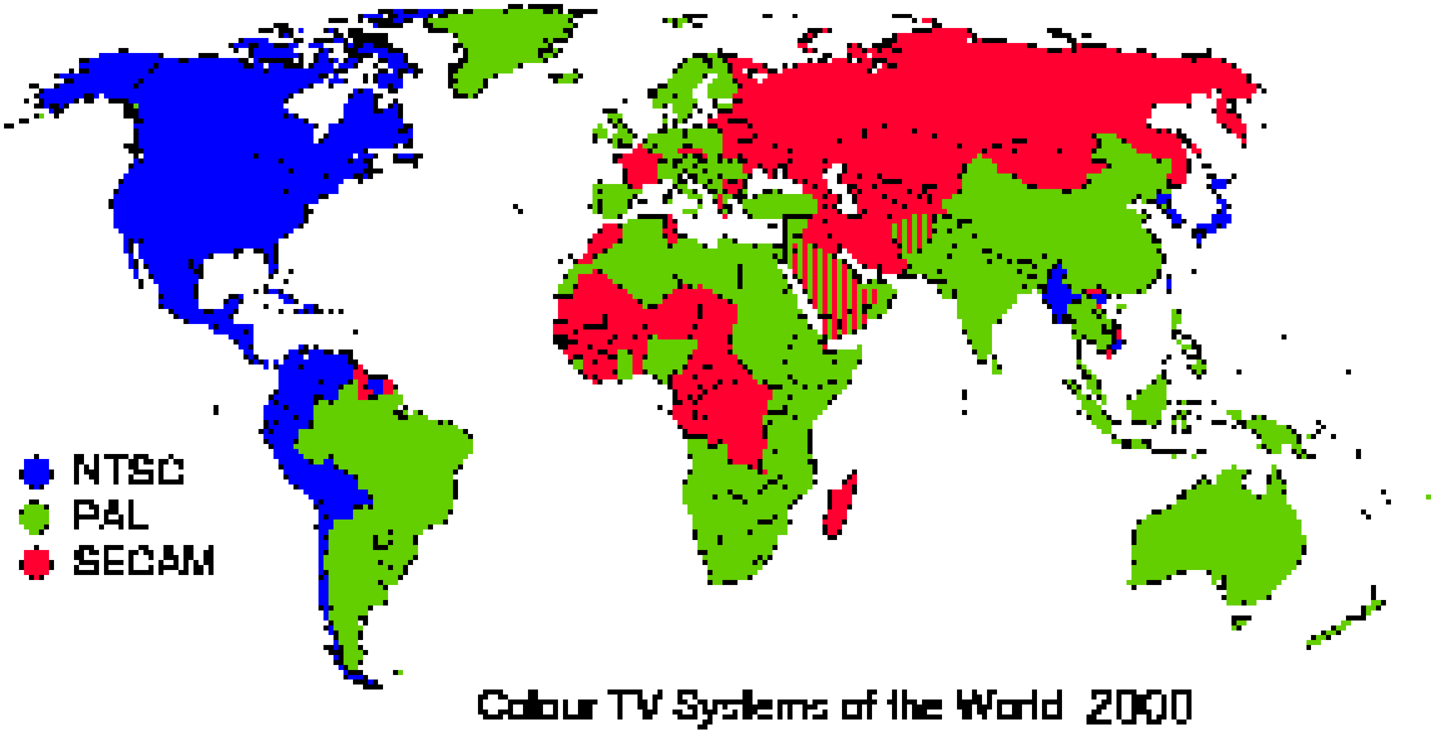
Original color frame

Recovered color frame

Different Color TV Systems

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

Who uses what?



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

Digital Video

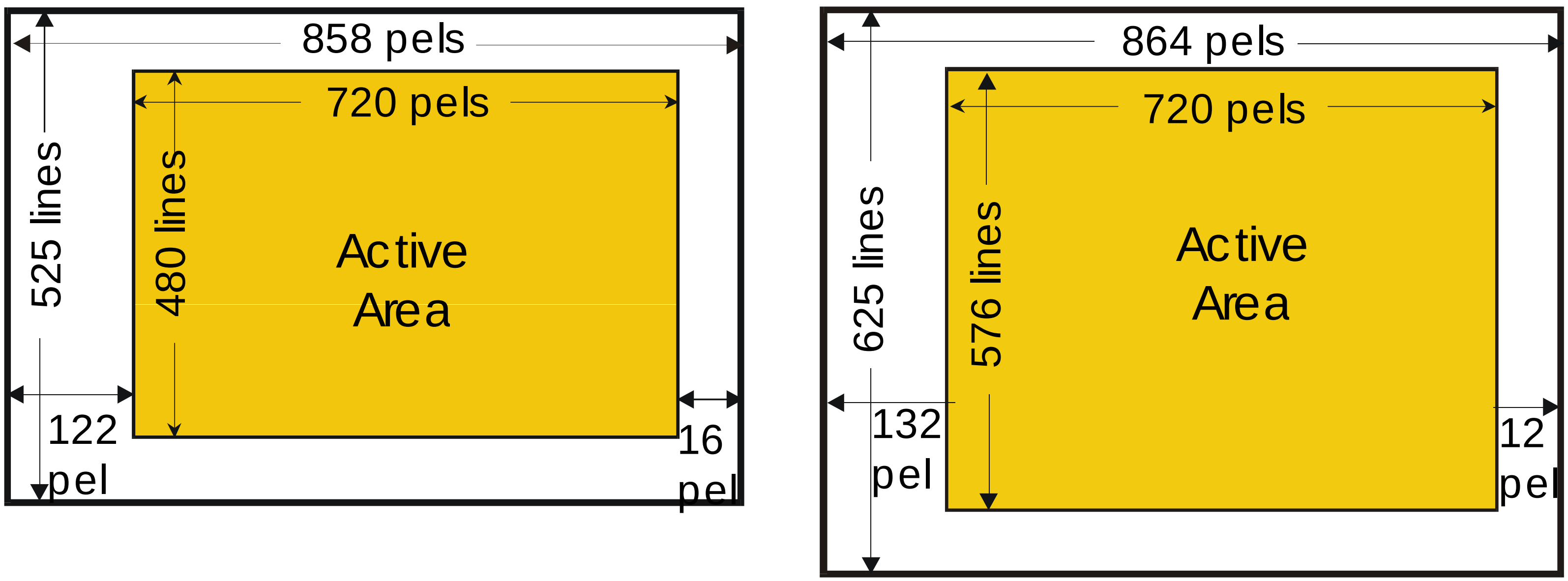
- Digital video by sampling/quantizing analog video raster → BT.601 video
- Directly capture digital video using digital cameras
- A digital video (including all color components) is compressed into a bit stream, which can be stored on a disk or transmitted over the air or through wires
- Transmission 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

$$f_s = 858 f_l (\text{NTSC}) = 864 f_l (\text{PAL/SECAM}) = 13.5 \text{ MHz}$$

BT.601* Video Format (commonly known as SDTV)



525/60: 60 field/s

625/50: 50 field/s

* BT.601 is formerly known as CCIR601

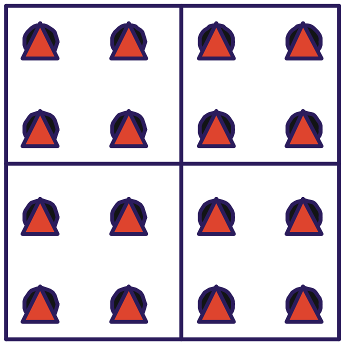
RGB <--> YCbCr

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

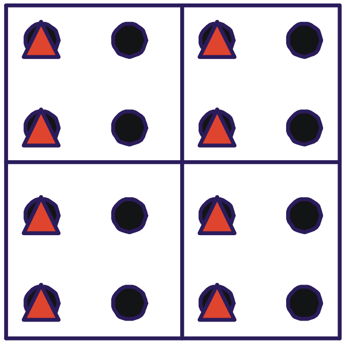
$$R_d = 1.164 Y_d' + 0.0 C_b' + 1.596 C_r',$$
$$G_d = 1.164 Y_d' - 0.392 C_b' - 0.813 C_r',$$
$$B_d = 1.164 Y_d' + 2.017 C_b' + 0.0 C_r',$$

$$Y_d' = Y_d - 16, C_b' = C_b - 128, C_r' = C_r - 128$$

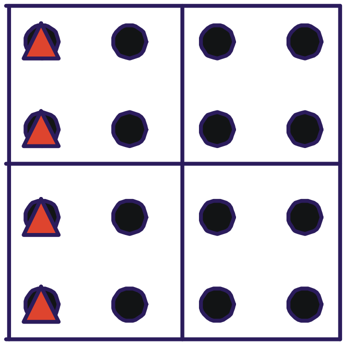
Chrominance Subsampling Formats



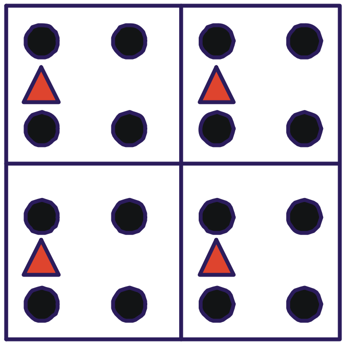
4:4:4
For every 2x2 Y Pixels
4 Cb & 4 Cr Pixel
(No subsampling)



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

Digital Video Formats

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

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