Course Code 005636 (Fall 2017)

Multimedia

Fundamental Concepts in Video

Prof. S. M. Riazul Islam, Dept. of Computer Engineering, Sejong University, Korea

E-mail: riaz@sejong.ac.kr

Outline

- Types of Video Signals
- Analog Video
- Digital Video

Types of Video Signals

- Component Video
- Composite Video
- S-Video

Component Video

Higher-end video systems make use of three separate video signals for the red, green, and blue image planes. Each color channel is sent as a separate video signal.

- gives the best color reproduction since there is no “crosstalk” between the three channels.
- requires more bandwidth and good synchronization of the three components.

Composite Video

Color ('chrominance', - I and Q, or U and V) and intensity ('luminance') signals are mixed into a *single* carrier wave.

- In NTSC TV, e.g., I and Q are combined into a chroma signal, and a color subcarrier is then employed to put the chroma signal at the high-frequency end of the signal shared with the luminance signal.

- The chrominance and luminance components can be separated at the receiver end and then the two color components can be further recovered.

- When connecting to TVs or VCRs, Composite Video uses only one wire and video color signals are mixed, not sent separately. The audio and *sync* signals are additions to this one signal.

- Since color and intensity are wrapped into the same signal, some interference between the luminance and chrominance signals is inevitable.

S-Video (Separate Video)

As a compromise, uses two wires, one for luminance and another for a composite chrominance signal,

- less crosstalk between the color and the crucial gray-scale information.
  - Humans are able to differentiate spatial resolution in grayscale images with a much more better than for the color part of color images.
  - As a result, we can send less accurate color information than must be sent for intensity information — we can only see fairly large blobs of color, so it makes sense to send less color detail.

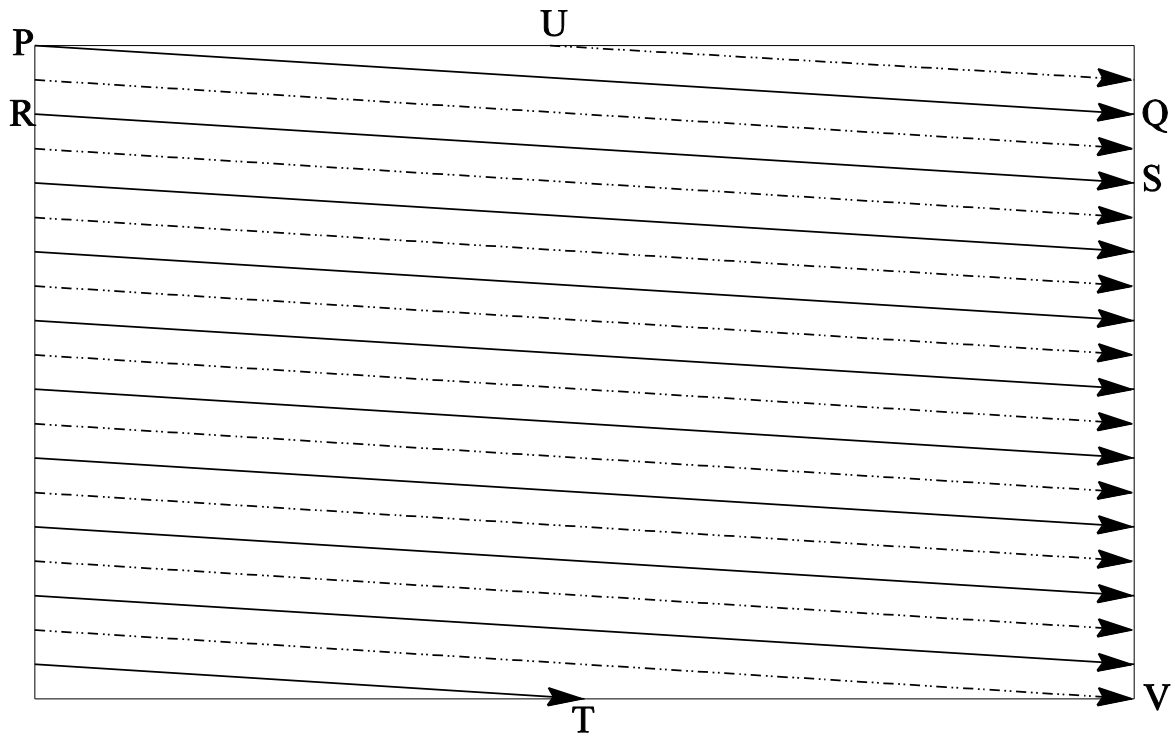
Analog Video

- An analog signal $f(t)$ samples a time-varying image.
- ‘Progressive’ scanning traces through a complete frame row-wise for each time interval.

- ‘Interlaced' scanning is used in TV, and in some monitors and multimedia standards as well.

  - The odd-numbered lines are traced first, and then the even-numbered lines. This results in “odd” and “even” fields — two fields make up one frame.

  - In fact, the odd lines (starting from 1) end up at the middle of a line at the end of the odd field, and the even scan starts at a half-way point.



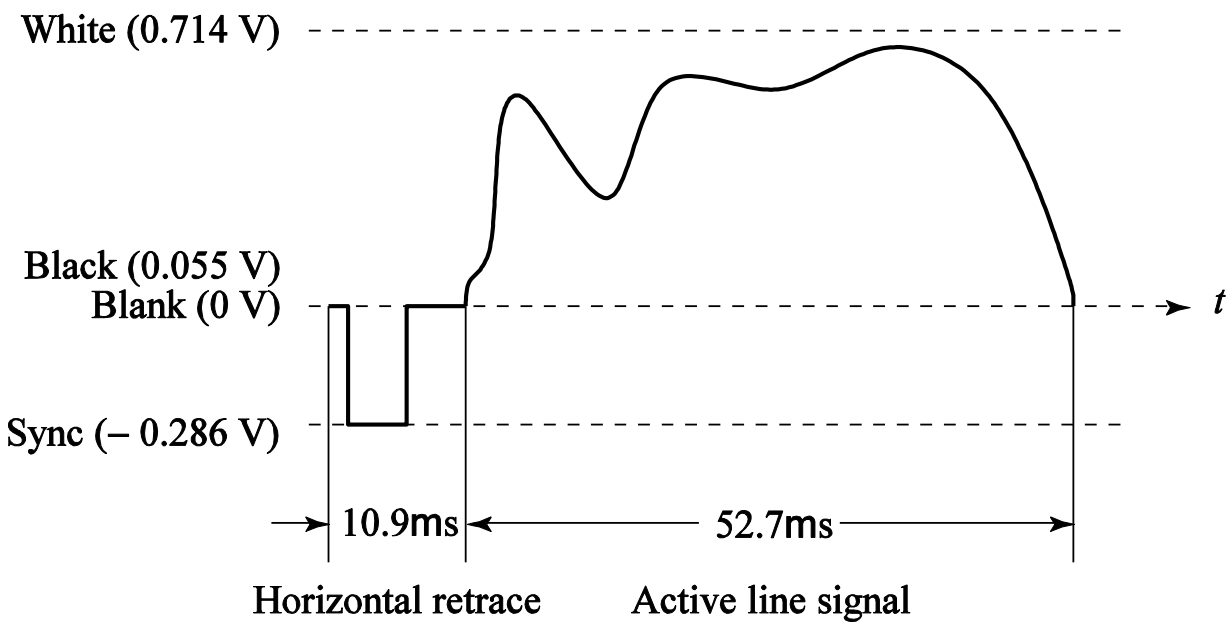
- First the solid (odd) lines are traced, P to Q, then R to S, etc., ending at T; then the even field starts at U and ends at V.

- The jump from Q to R, etc., is called the **horizontal** retrace, during which the electronic beam in the CRT is blanked. The jump from T to U or V to P is called the **vertical retrace**.

- The double number of fields presented to the eye reduces perceived flicker.

- The odd and even lines are displaced in time from each other — generally not noticeable except when very fast action is taking place on screen, when blurring may occur.

- 'De-interlace'
  - The simplest method consists of discarding one field and duplicating the scan lines of the other field.
- Analog video use a small voltage offset from zero to indicate 'black', and another value (zero) to indicate the start of a line.
  - A electronic signal for one scan line of NTSC composite video



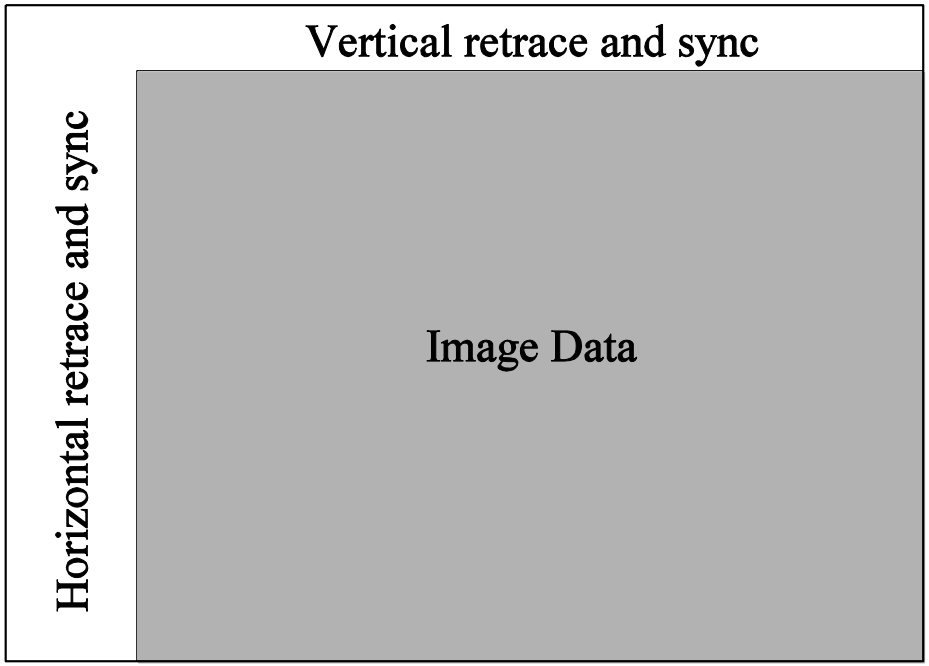
NTSC Video

- NTSC (National Television System Committee) TV standard is mostly used in North America and Japan. It uses the 4:3 **aspect ratio** (the ratio of picture width to its height) and 525 scan lines per frame at 30 (actually 29.97) frames per second (fps).
  - NTSC follows the interlaced scanning system, and each frame is divided into two fields, with 262.5 lines/field.
  - Thus the horizontal sweep frequency is $525 \times 29.97 \approx 15,734$ lines/sec, so that each line is swept out in $1/15.734 \approx 63.6\mu$sec.
  - Since the horizontal retrace takes 10.9 μsec, this leaves 52.7 µsec for the active line signal during which image data is displayed.

- Vertical retrace takes 20 lines for control information at the beginning of each field. Hence, the number of active *video lines* per frame is 485.

- Almost 1/6 of the raster at left is blanked for horizontal retrace and sync. The non-blanking pixels are called *active pixels*.

- It is known that pixels often fall in-between the scan lines. Therefore, even with non-interlaced scan, NTSC TV is only capable of showing about 340 (visually distinct) lines, i.e., about 70% of the 485 active lines. With interlaced scan, this could be 50%.



- NTSC video is an analog signal with no fixed horizontal resolution. Therefore one must decide how many times to sample the signal for display: each sample corresponds to one pixel output.
  - A 'pixel clock' is used to divide each horizontal line of video into samples. The higher the frequency of the pixel clock, the more samples per line there are.
  - Different video formats provide different numbers of samples per line.

Format	Samples per line
VHS	240
S-VHS	400-425
Betamax	500
Standard 8 m	300
Hi-8 mm	425

Color Model and Modulation of NTSC

- NTSC uses the YIQ color model, and the technique of **quadrature modulation** is employed to combine (the spectrally overlapped part of) *I* (in-phase) and *Q* (quadrature) signals into a single Chroma signal *C (***color subcarrier***)*:

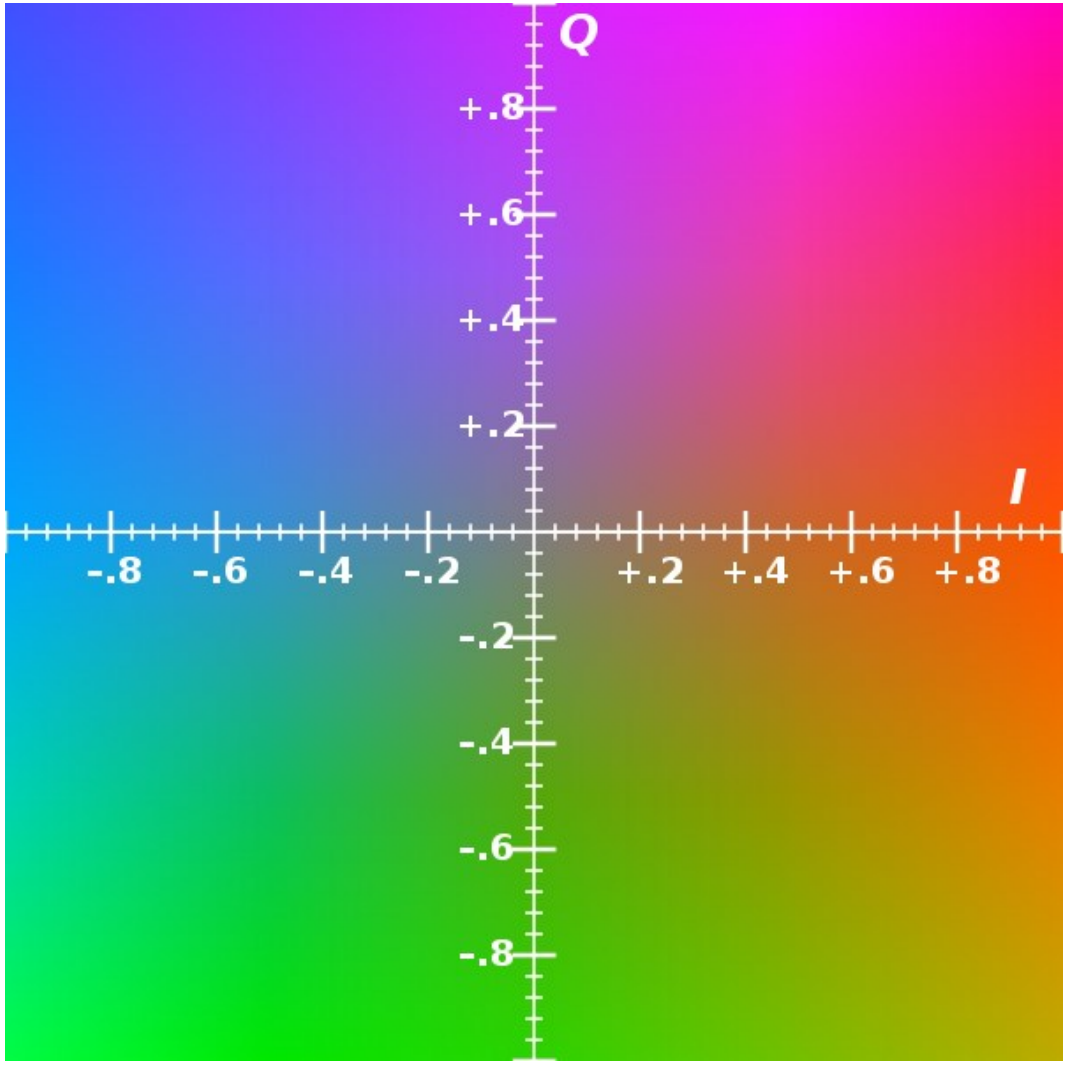
$$C = I\cos(F_{sc}t) + Q\sin(F_{sc}t)$$

- Its magnitude is $\sqrt{I^2+Q^2}$, and phase is $\tan^{-1}(Q/I)$. The frequency of C is $F_{sc} \approx$ 3.58 MHz.

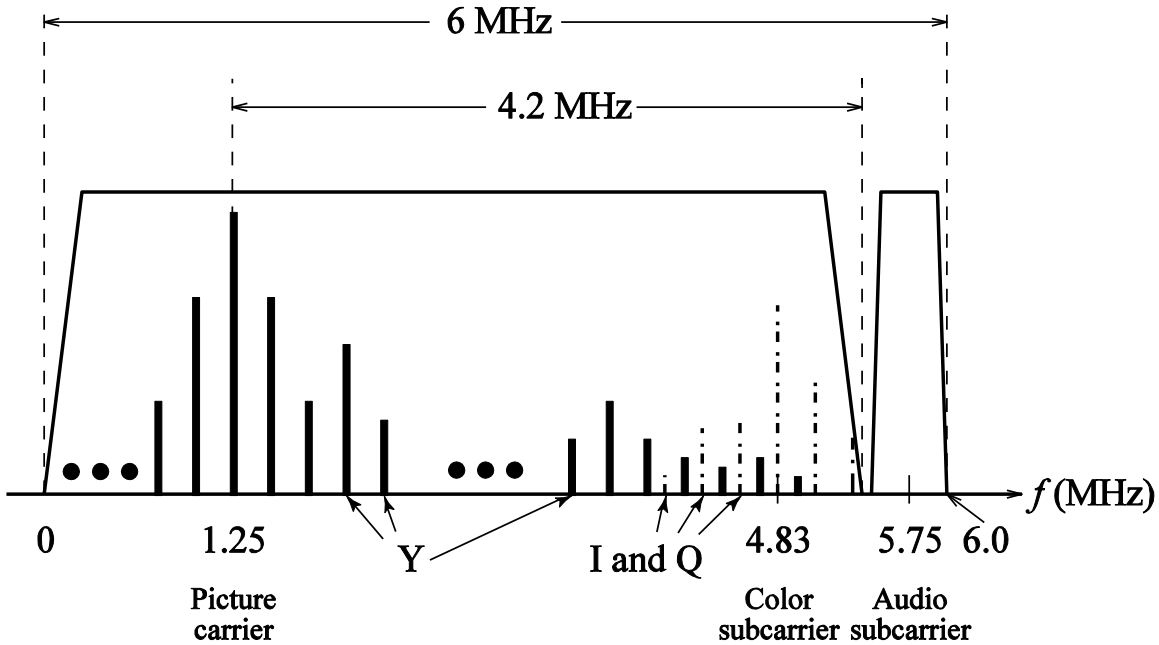
- The NTSC composite signal is a further composition of the luminance signal *Y* and the Chroma signal as defined below:

- $$\text{Composite} = Y + C = Y + I\cos(F_{sc}t) + Q\sin(F_{sc}t)$$

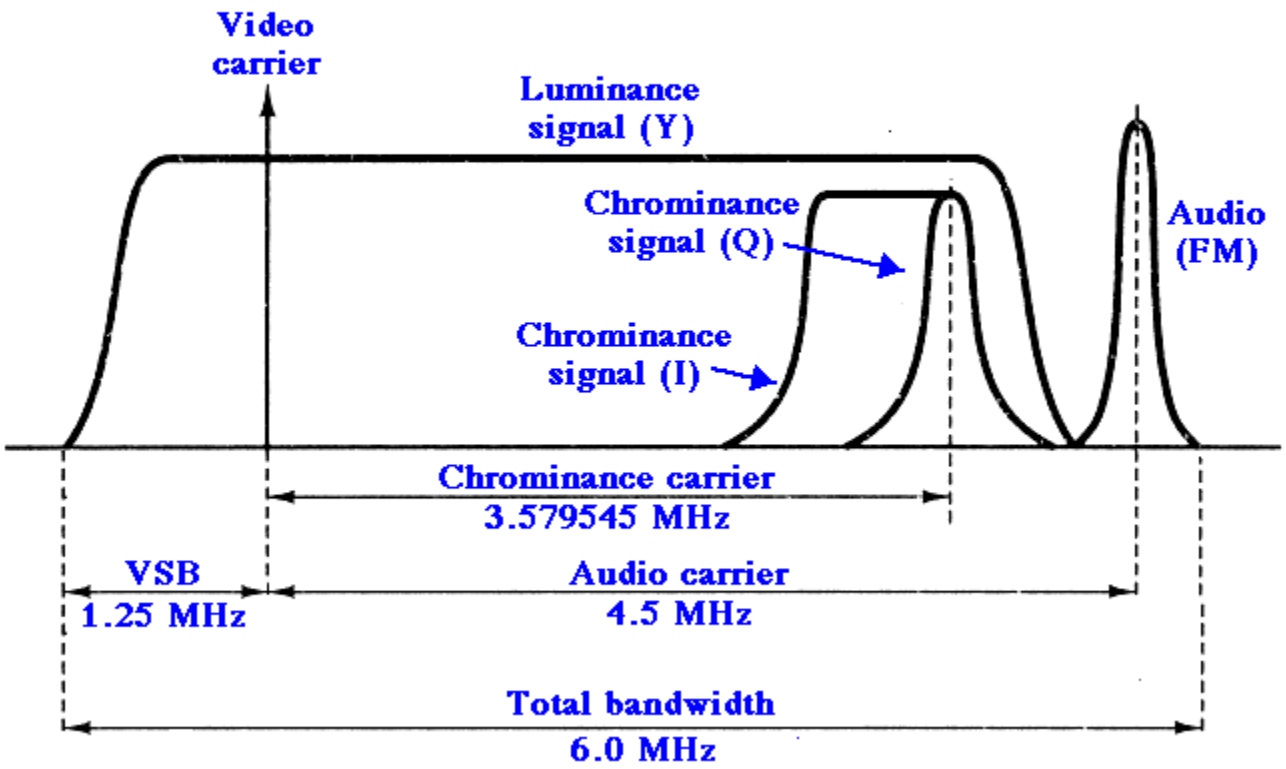
YIQ Color Space



- Interleaving Y and C signals in the NTSC spectrum
  - NTSC assigns a bandwidth of 4.2 MHz to *Y*, and only 1.6 MHz to *I* and 0.6 MHz to *Q* due to human insensitivity to color details (high frequency color changes).



Spectrum of Video Signal



Decoding NTSC Signals

- The first step is the separation of *Y* using a low-pass filter, which is located at the lower end of the channel.

- The chroma signal *C* can be demodulated to extract the components *I* and *Q* separately. To extract *I*:
  - Multiply the signal *C* by 2 cos($F_{sc}t$),

$$
\begin{aligned}
C \cdot 2\cos(F_{sc}t) &= I \cdot 2\cos^2(F_{sc}t) + Q \cdot 2\sin(F_{sc}t)\cos(F_{sc}t) \\
&= I \cdot (1 + \cos(2F_{sc}t)) + Q \cdot 2\sin(F_{sc}t)\cos(F_{sc}t) \\
&= I + I \cdot \cos(2F_{sc}t)) + Q \cdot 2\sin(2F_{sc}t)
\end{aligned}
$$

  - Apply a low-pass filter to obtain I and discard the two higher frequency ($2F_{sc}$) terms.

- Similarly, *Q* can be extracted by first multiplying *C* by 2sin($F_{sc}t$) and then low-pass filtering.

- The NTSC audio subcarrier frequency is 4.5 MHz. The Picture carrier is at 1.25 MHz, which places the center of the audio band at 1.25+4.5 = 5.75 MHz in the channel. The color is placed at 1.25+3.58 = 4.83 MHz.

- The audio is a bit too close to the color subcarrier — potential interference between the audio and color signals. Hence the NTSC color TV slowed down its frame rate to $30 \times 1,000/1,001 \approx 29.97$ fps.

- As a result, the adopted NTSC color subcarrier frequency is slightly lowered to

$$f_{sc} = 30 \times 1,000/1,001 \times 525 \times 227.5 \approx 3.579545 \text{ MHz},$$

where 227.5 is the number of color samples per scan line in NTSC broadcast TV.

PAL Video

- **PAL (Phase Alternating Line)** is a TV standard widely used in Western Europe, China, India, and many other parts of the world.

- PAL uses 625 scan lines per frame, at 25 frames/second, with a 4:3 aspect ratio and interlaced fields.
  - PAL uses the YUV color model. It uses an 8 MHz channel and allocates a bandwidth of 5.5 MHz to Y, and 1.8 MHz each to U and V. The color subcarrier frequency is $f_{sc} \approx 4.43$ MHz.

  - In order to improve picture quality, chroma signals have alternate signs (e.g., +U and -U) in successive scan lines, hence the name 'Phase Alternating Line'.

  - The signals in consecutive lines are averaged at the receiver so as to cancel the chroma signals (that always carry opposite signs) for separating Y and C and obtaining high quality Y signals.

SECAM Video

- **SECAM** (*Système Electronique Couleur Avec Mémoire*) is the third major broadcast TV standard.
  - uses 625 scan lines per frame, at 25 frames per second, with a 4:3 aspect ratio and interlaced fields.

- SECAM and PAL are very similar. They differ slightly in their color coding scheme:
  - In SECAM, U and V signals are modulated using separate color subcarriers at 4.25 MHz and 4.41 MHz respectively.

  - They are sent in alternate lines, i.e., only one of the U or V signals will be sent on each scan line.

- Comparison of Analog Broadcast TV Systems

TV System	Frame Rate (fps)	# of Scan Lines	Total Channel Width (MHz)	Bandwidth Allocation (MHz)		
				Y	I or U	Q or V
NTSC	29.97	525	6.0	4.2	1.6	0.6
PAL	25	625	8.0	5.5	1.8	1.8
SECAM	25	625	8.0	6.0	2.0	2.0

Digital Video

- The advantages of digital video:
  - Video can be stored on digital devices or in memory, ready to be processed (noise removal, cut and paste, etc.), and integrated to various multimedia applications;
  - Direct access is possible, which makes nonlinear video editing achievable as a simple task;
  - Repeated recording does not degrade image quality;
  - Ease of encryption and better tolerance to channel noise.

Chroma Subsampling

- Since humans see color with much less spatial resolution than they see black and white, it makes sense to 'decimate' the chrominance signal.

- Interesting names have arisen to label the different schemes used.
  - Numbers are given stating how many pixel values, per four original pixels, are actually sent.

  - The chroma subsampling scheme '4:4:4' indicates that no chroma subsampling is used: each pixel's Y, Cb and Cr values are transmitted, 4 for each of Y, Cb, Cr.

Chroma Subsampling

- The scheme '4:2:2' indicates horizontal subsampling of the Cb, Cr signals by a factor of 2. That is, of four pixels horizontally labelled as 0 to 3, all four Ys are sent, and every two Cb's and two Cr's are sent.
- The scheme '4:1:1' subsamples *horizontally* by a factor of 4.
- The scheme '4:2:0' subsamples in both the *horizontal* and *vertical* dimensions by a factor of 2. Theoretically, an average chroma pixel is positioned between the rows and columns.
- Scheme 4:2:0 along with other schemes is commonly used in JPEG and MPEG.

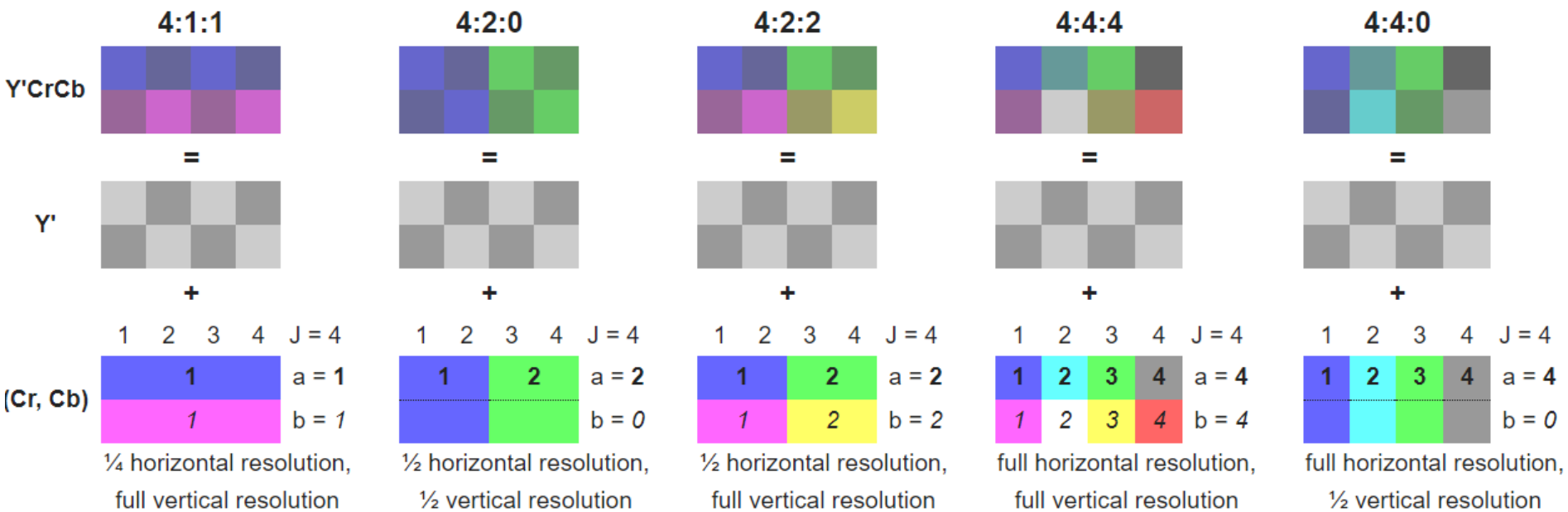
Chroma Subsampling

Original without color subsampling. 200% zoom.

Image after color subsampling (compressed with Sony Vegas DV codec, box filtering applied.)

http://en.wikipedia.org/wiki/Chroma_subsampling

Chroma Subsampling



CCIR Standards for Digital Video

- **CCIR** is the Consultative Committee for International Radio, and one of the most important standards it has produced is CCIR-601, for component digital video.
  - This standard has since become standard ITU-R-601, an international standard for professional video applications
    - adopted by certain digital video formats including the popular DV video.

- CIF stands for Common Intermediate Format specified by the CCITT.
  - The idea is to specify a format for lower bitrate.
  - CIF is about the same as Video Home System (VHS) quality. It uses a progressive (non-interlaced) scan.
  - QCIF stands for 'Quarter-CIF'. All the CIF/QCIF resolutions are evenly divisible by 8, and all except 88 are divisible by 16; this provides convenience for block-based video coding in H.261 and H.263.

- CIF is a compromise of NTSC and PAL in that it adopts the NTSC frame rate and half of the number of active lines as in PAL.

	CCIR 601 525/60 NTSC	CCIR 601 625/50 PAL/SECAM	CIF	QCIF
Luminance resolution	720 x 480	720 x 576	352 x 288	176 x 144
Chrominance resolution	360 x 480	360 x 576	176 x 144	88 x 72
Colour Subsampling	4:2:2	4:2:2	4:2:0	4:2:0
Fields/sec	60	50	30	30
Interlaced	Yes	Yes	No	No

NTSC

$$525 \times 858 \times 30 \times 2 \text{ bytes} \times 8 \frac{\text{bits}}{\text{byte}} \approx 216 \text{ Mbps}$$

HDTV (High Definition TV)

- The main thrust of **HDTV** (High Definition TV) is to increase the visual field especially in its width.
  - The first generation was based on an analog technology developed by Sony and NHK in Japan in the late 1970s.
  - MUSE (MUltiple sub-Nyquist Sampling Encoding) was an improved NHK HDTV with hybrid analog/digital technologies that was put in use in the 1990s. It has 1,125 scan lines, interlaced (60 fields per second), and 16:9 aspect ratio.
  - Since uncompressed HDTV will easily demand more than 20 MHz bandwidth, which will not fit in the current 6 MHz or 8 MHz channels, various compression techniques are being investigated.
  - High quality HDTV signals will be transmitted using more than one channel even after compression.

- Advanced Digital TV formats supported by ATSC
  - 'I' mean interlaced scan and 'P' means progressive (non-interlaced) scan.

# of Active Pixels per line	# of Active Lines	Aspect Ratio	Picture Rate
1,920	1,080	16:9	60I 30P 24P
1,280	720	16:9	60P 30P 24P
704	480	16:9 & 4:3	60I 60P 30P 24P
640	480	4:3	60I 60P 30P 24P

- For video, MPEG-2 is the compression standard. For audio, AC-3 is the standard. It supports the 5:1 channel Dolby surround sound, i.e., five surround channels plus a subwoofer channel.

- The salient difference between conventional TV and HDTV:
  - HDTV has a much wider aspect ratio of 16:9 instead of 4:3.

  - HDTV moves toward progressive (non-interlaced) scan. The rationale is that interlacing introduces serrated edges to moving objects and flickers along horizontal edges.

Q&A