CMPT 365 Multimedia Systems

Media Representations
- Video

Spring 2017

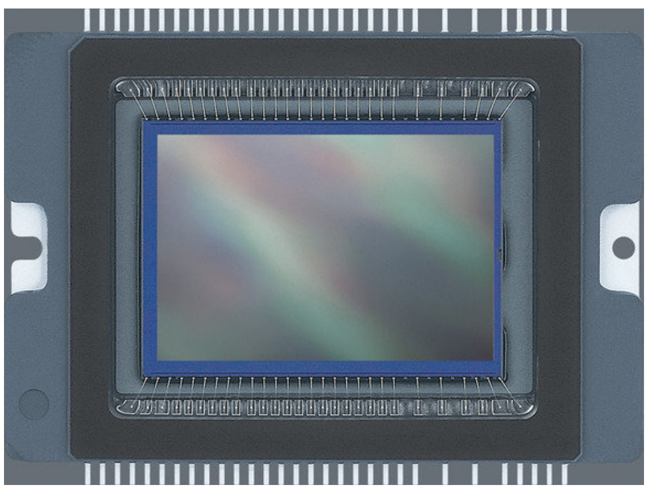
Edited from slides by Dr. Jiangchuan Liu

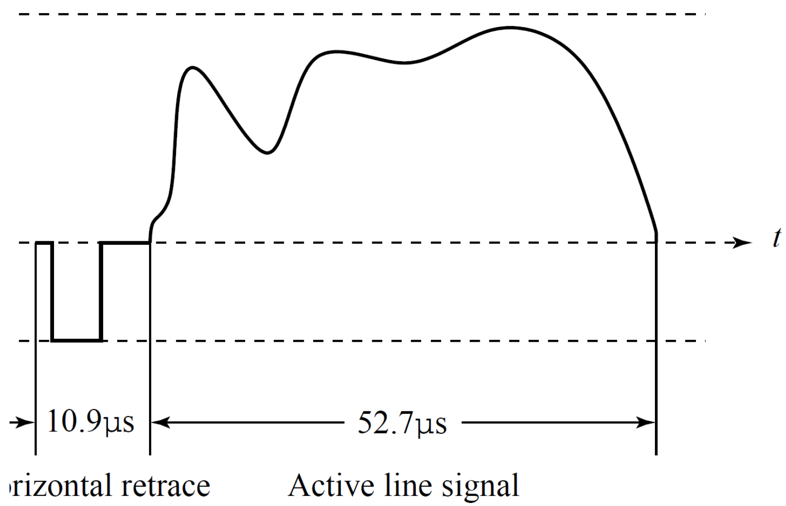
Outline

- Analog Video
- Digital Video
- Video Interfaces
- HDTV
- Further Exploration

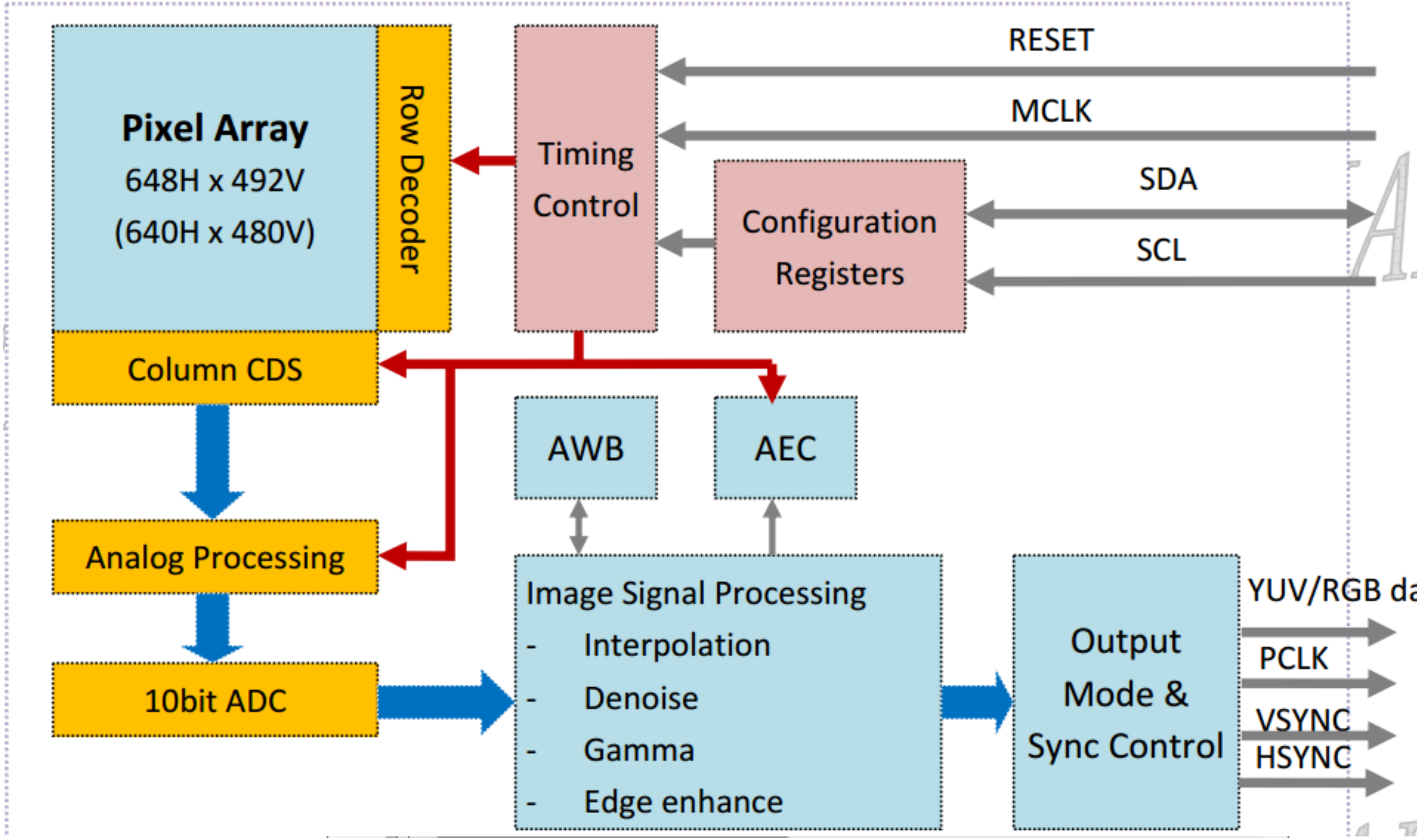
Review - Camera

- CCD/CMOS are just sensor type
- Usually 2d matrix array
- Scan through each point to produce electronical signal
- Digital camera has Analog to Digital conversion





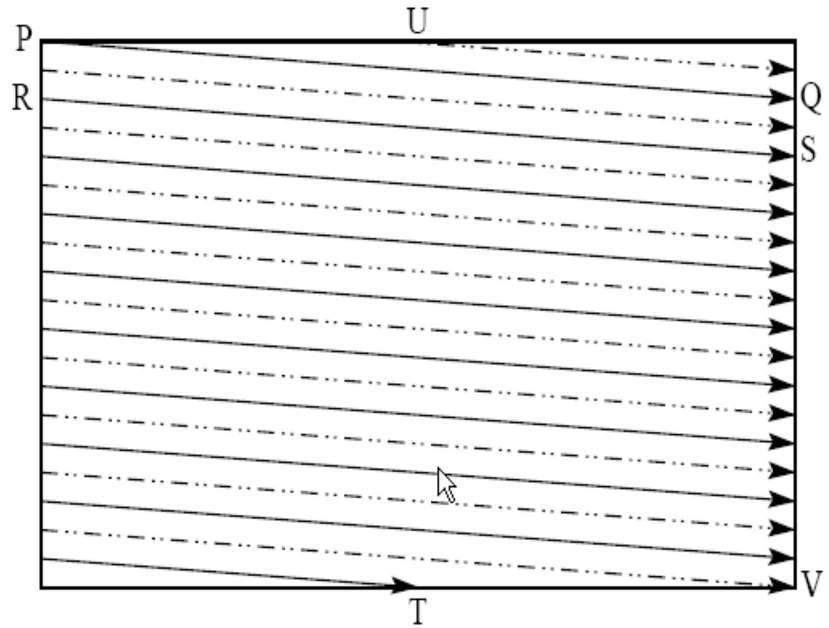
Review - Camera



Analog Video

- An analog signal $f(t)$ samples a time-varying image
- Progressive scanning
  - traces through a complete picture (a frame) row-wise for each time interval.
- Interlaced scanning
  - Odd-numbered lines traced first, and then the even-numbered lines.
  - "odd" and "even" fields - two fields make up one frame
  - Widely used in traditional (non-digital) TV

Interlaced Scan



- 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**

Interlaced Scan

- Because of interlacing, 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.
- For example, in the video in Fig. 5.2, the moving helicopter is blurred more than is the still background.

Example of Interlaced Scan

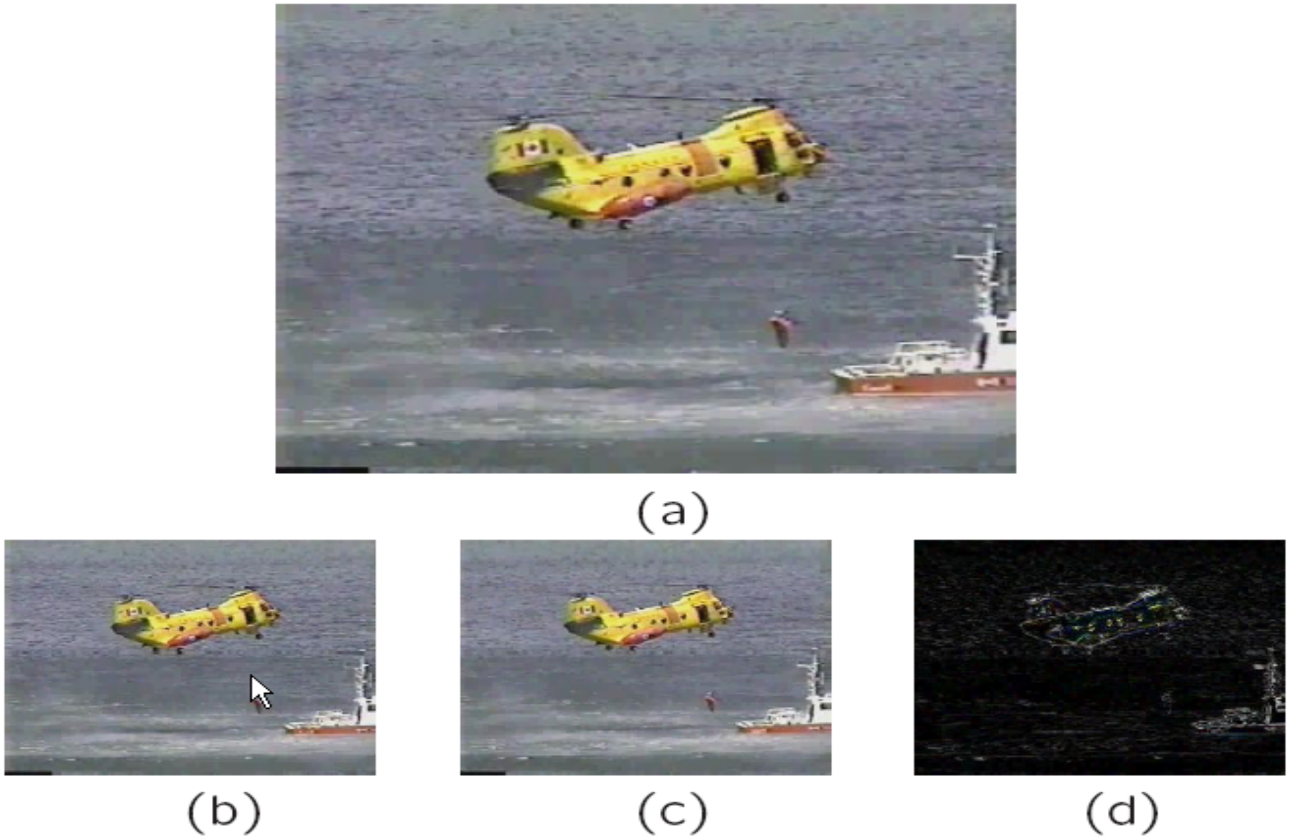
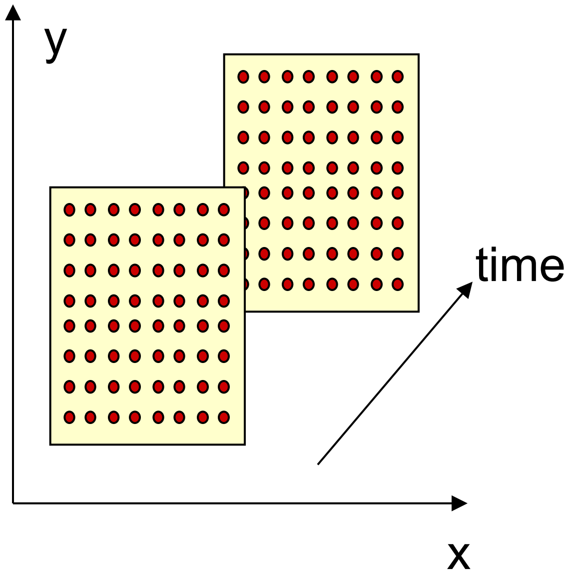


Fig. 5.2: Interlaced scan produces two fields for each frame. (a) The video frame, (b) Field 1, (c) Field 2, (d) Difference of Fields

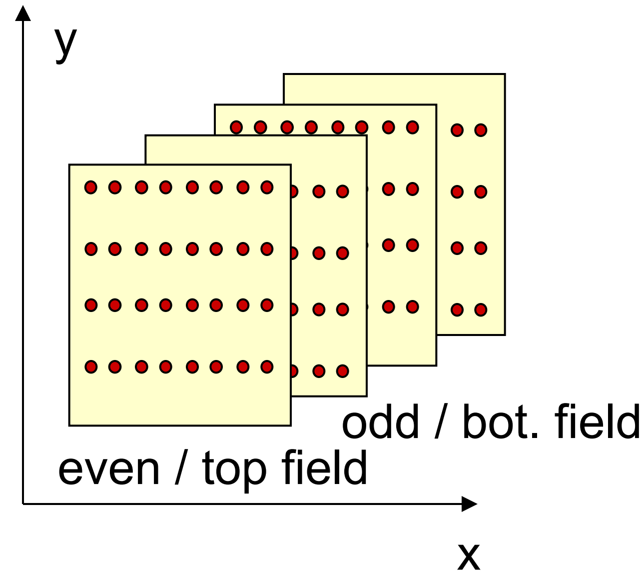
- Since it is sometimes necessary to change the frame rate, resize, or even produce stills from an interlaced source video, various schemes are used to "de-interlace" it.
  a) The simplest de-interlacing method consists of discarding one field and duplicating the scan lines of the other field. The information in one field is lost completely using this simple technique.
  b) Other more complicated methods that retain information from both fields are also possible.

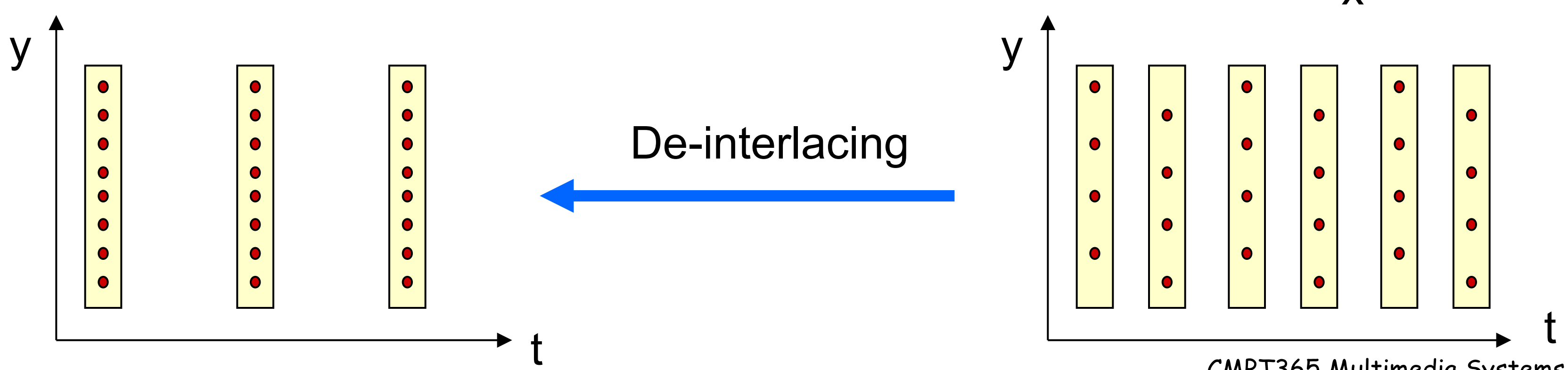
Digital vs Analog TV Signal

Digital video:
Progressive Frame
All lines



TV signal:
Interlaced Fields: even lines or odd lines only
Tradeoff between frame rate and bandwidth





Analog video use a small voltage offset from zero to indicate "black", and another value such as zero to indicate the start of a line. For example, we could use a "blacker-than-black" zero signal to indicate the beginning of a line.

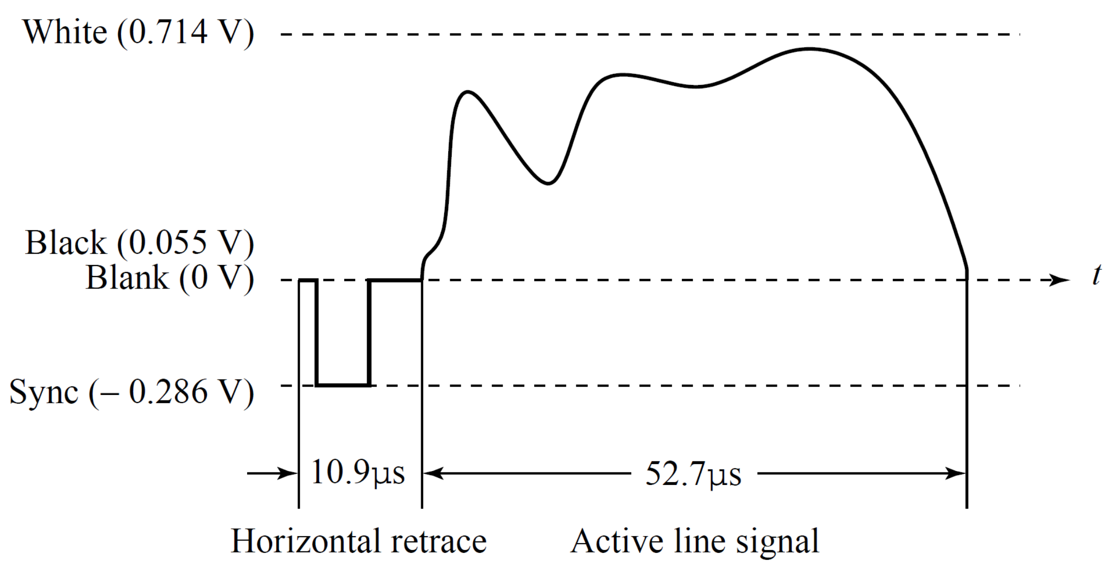


Fig. 5.3: Electronic signal for one NTSC scan line.

5.1.1 NTSC Video

- NTSC (National Television System Committee) TV standard is mostly used in North America and Japan. It uses the familiar 4:3 **aspect ratio** (i.e., the ratio of picture width to its height) and uses 525 scan lines per frame at 30 frames per second (fps).

  a) NTSC follows the interlaced scanning system, and each frame is divided into two fields, with 262.5 lines/field.

  b) Thus the horizontal sweep frequency is $525 \times 29.97 \approx 15,734$ lines/sec, so that each line is swept out in $1/15.734 \times 103$ sec $\approx 63.6\mu$sec.

  c) Since the horizontal retrace takes 10.9 μsec, this leaves 52.7 μsec for the active line signal during which image data is displayed (see Fig.5.3).

- Fig. 5.4 shows the effect of "vertical retrace & sync" and "horizontal retrace & sync" on the NTSC video raster.

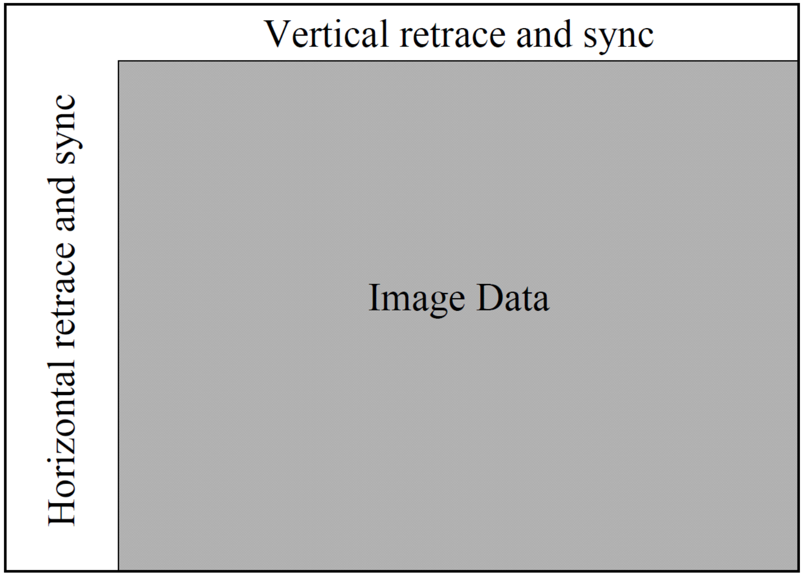


Fig. 5.4: Video raster, including retrace and sync data

a) Vertical retrace takes place during 20 lines reserved for control information at the beginning of each field. Hence, the number of active *video lines* per frame is only 485.

b) Similarly, almost 1/6 of the raster at the left side is blanked for horizontal retrace and sync. The non-blanking pixels are called *active pixels*.

c) Since the horizontal retrace takes 10.9 μsec, this leaves 52.7 μsec for the active line signal during which image data is displayed (see Fig.5.3).

- 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, as listed in Table 5.1.

Table 5.1: Samples per line for various video formats

Format	Samples per line
VHS	240
S-VHS	400-425
Betamax	500
Standard 8 m	300
Hi-8 mm15	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:

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

(5.1)

- This modulated chroma signal is also known as the **color subcarrier**, whose 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:

$$\text{composite} = Y + C = Y + I\cos(F_{sc}t) + Q\sin(F_{sc}t) \qquad (5.2)$$

- Fig. 5.5: 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).

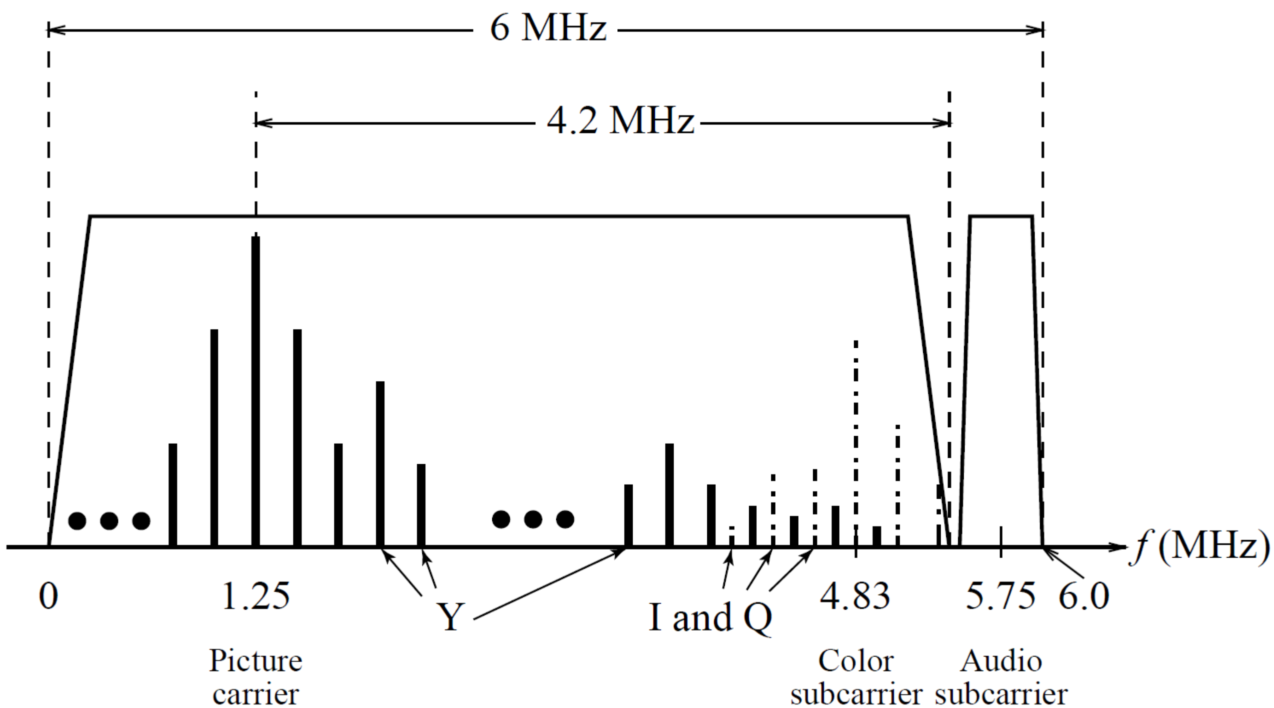


Fig. 5.5: Interleaving Y and C signals in the NTSC spectrum.

Decoding NTSC Signals

- The first step in decoding the composite signal at the receiver side is the separation of Y and C.
- After the separation of Y using a low-pass filter, the chroma signal C can be demodulated to extract the components I and Q separately. To extract I:

1. Multiply the signal C by $2\cos(F_{sc}t)$, i.e.,

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

2) 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 $2\sin(F_{sc}t)$ and then low-pass filtering.

- The NTSC bandwidth of 6 MHz is tight. Its 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 (Fig. 5.5). But notice that the color is placed at 1.25+3.58 = 4.83 MHz.

- So the audio is a bit too close to the color subcarrier — a cause for potential interference between the audio and color signals. It was largely due to this reason that the NTSC color TV actually 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
  - $f_{sc} = 30 \times 1{,}000/1{,}001 \times 525 \times 227.5 \approx 3.579545$ MHz,
- ** 227.5 is the # of color samples per scan line in NTSC broadcast TV.

5.1.2 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.
  (a) 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.

5.1.3 SECAM Video

- **SECAM** stands for *Système Electronique Couleur Avec Mémoire*, the third major broadcast TV standard.
- SECAM also 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:
  - (a) In SECAM, U and V signals are modulated using separate color subcarriers at 4.25 MHz and 4.41 MHz respectively.
  - (b) They are sent in alternate lines, i.e., only one of the U or V signals will be sent on each scan line.

- Table 5.2 gives a comparison of the three major analog broadcast TV systems.

❒ **Table 5.2:** Comparison of Analog Broadcast TV Systems

TV System	Fram e 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

NTSC Video

- **NTSC** (National Television System Committee) TV standard is mostly used in North America and Japan
  - YIQ color model
  - 4:3 **aspect ratio** (i.e., the ratio of picture width to its height)
  - 525 scan lines per frame at 30 frames per second (fps).
- Interlaced scanning, and each frame is divided into two fields, with 262.5 lines/field
  - horizontal sweep frequency is 525x29.97 = 15,734 lines/sec,
  - each line is swept out in 1/15,734 = 63.6 us
  - the horizontal retrace takes 10.9 sec, this leaves 52.7 sec for the active line signal during which image data is displayed
- PAL in Asia/Europe, SECAM in Europe
- All faded out (Canada, Aug 31, 2011)

Outline

- Analog Video
- Digital Video
- Video Interfaces
- HDTV
- Further Exploration

Digital Video

- Why digital video ?
- Advantages
  - Stored on digital device or in memory
  - Faithful duplication in digital domain
    - Good or bad ?
  - Direct (random) access,
    - nonlinear video editing achievable as a simple, rather than a complex task
  - Ease of manipulation (noise removal, cut and paste, etc.)
  - Ease of encryption and better tolerance to channel noise
    - Multimedia communications
  - Integration to various multimedia applications

ITU-R digital video specifications

	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

Note, 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

- CCIR-601 for component digital video
  - specified by Consultative Committee for International Radio (CCIR)
  - aspect ratio of 4:3
  - interlaced scan, so each field has only half as much vertical resolution
  - Now become standard ITU-R-601, adopted by many digital video formats including the popular DV video.

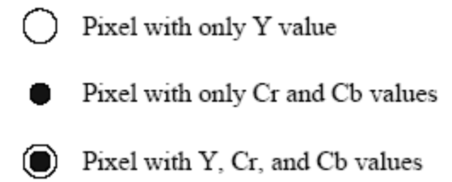
CIF

- CIF: Common Intermediate Format
  - Specified by CCITT (***C**omité **C**onsultatif **I**nternational **T**éléphonique et **T**élégraphique*).
- A format for lower bitrate
  - CIF is about the same as VHS quality.
  - Progressive (non-interlaced) scan.
- QCIF: "Quarter-CIF"
- 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, discussed later

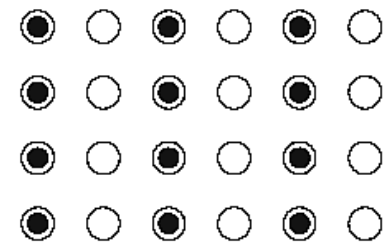
Chroma Subsampling

- Since humans see color with much less spatial resolution than they see black and white, it makes sense to subsample chrominance signal
- Interesting (but not necessarily informative!) names have arisen to label the different schemes used.
  - 4:4:4
  - 4:2:2
  - 4:1:1
  - 4:2:0
- To begin with, 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.

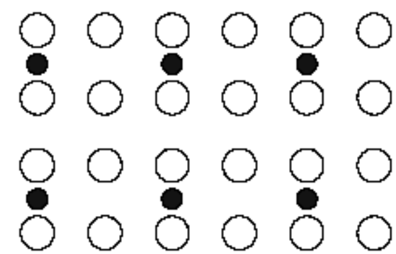
Chroma Subsampling cont'd



- 4:2:2: *horizontal subsampling* of the Cb, Cr signals by a factor of 2.
  - 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, as (Cb0, Y0)(Cr0, Y1)(Cb2, Y2)(Cr2, Y3)(Cb4, Y4), and so on (or averaging is used).



- 4:2:0: subsamples in *both the horizontal and vertical* dimensions by a factor of 2.
  - 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 (more later).

Outline

- Analog Video
- Digital Video
- Video Interfaces
- HDTV
- Further Exploration

Analog Video Display Interfaces

Component video, Composite video, S-video, VGA

Component video

- **Component video**: three separate video signals for the red, green, and blue image planes. Each color channel is sent as a separate video signal.
  - For higher-end video systems
  - Supported by most computer systems
  - Best color reproduction
    - no "crosstalk" between the three channels.
  - But more bandwidth and good synchronization

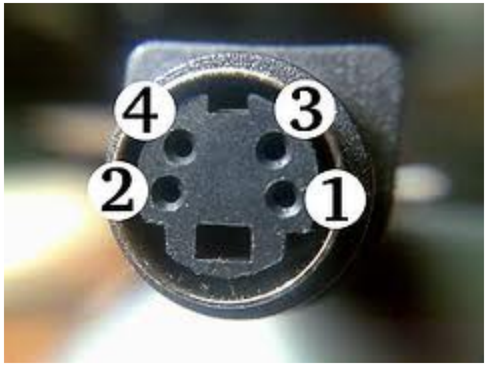
Composite Video - 1 Signal

- **Composite video**: color ("chrominance") and brightness ("luminance") signals are mixed into a *single* wire
  - **Chrominance** (I and Q, or U and V).
    - Combined into a chroma signal, and then put at the high-frequency end of the signal shared with the luminance signal Y.
  - Chrominance and luminance components separated at the receiver end and then two color components be further recovered.
  - Only one wire for video signal
    - Audio signals added through separate wires

- Interference is inevitable.

S-Video - 2 Signals

- **S-Video**: Separated video, or Super-video
  - a compromise, with two wires
  - one for luminance and another for a composite chrominance signal.
  - Less crosstalk between color and the crucial gray-scale information.
- Reason for placing luminance into its own part
  - black-and-white is most crucial for visual perception.
    - Both in terms of brightness and spatial resolution
    - Less information for color is fine



VGA (Video Graphics Array)

- Analog only
- Introduced with IBM x86 machines (1987), but became a universal analog display interface
- R, G, B, plus power, syn, control etc
- A VGA (D-sub) connector

Digital Display Interfaces

Digital interfaces emerged in 1980s (e.g., Color Graphics Adapter (CGA), and evolved rapidly.

- Connectors of different digital display interfaces: DVI, HDMI, DisplayPort.

DVI (Digital Visual Interface)

- Uncompressed digital video
- Almost a ubiquitous computer display link replacing VGA (since 1999)
- Uncompressed video only
  - R, G, B (both digital and analog)
  - plus clock, syn, power, control etc.

- Single link: 1920x1080 60Hz
- Dual link: 2560x1600 60Hz

HDMI (High-Definition Multimedia Interface)

- (2002) backward-compatible with DVI
- RGB or YCbCr + digital audio
  - + bidirectional audio, ethernet
- High bandwidth digital content protection (HDCP)
- HDMI 1.3 2560x1600

Display Port

- (2006) Packetized transmission (like Internet)
  - 4K video support
- Royalty-free (HDMI is not!)

- Variation/enhancement: Thunderbolt

Example: DVD Player

Example: Home Theatre Receiver

Example: Xbox 360

- Simple package vs premium package

Example: PS4


PS4
Designed and Engineered by Sony Computer Entertainment in JAPAN
(OPTICAL)
LAN
AUX
TECHINSIGHTS
teardown.com

Outline

- Analog Video
- Digital Video
- Video Interfaces
- HDTV
- Further Exploration

HDTV

- Main objective of **HDTV** (High Definition TV)
  - not necessary to increase the "definition" in each unit area
  - but rather to increase the visual field especially in its width.
- First generation of HDTV
  - an analog technology developed by Sony and NHK in Japan in the late 1970s.
- MUSE (MUltiple sub-Nyquist Sampling Encoding)
  - an improved NHK HDTV with hybrid analog/digital technologies in the 1990s.
  - 1,125 scan lines, interlaced (60 fields per second), and 16:9 aspect ratio.
- Need for compressions
  - uncompressed HDTV will easily demand more than 20 MHz bandwidth, which will not t in the current 6 MHz or 8 MHz channels
  - high quality HDTV signals would be transmitted using more than one channel even after compression.

HDTV in North America

- 1987: FCC (Federal Communications Commission) decided that HDTV standards must be compatible with existing NTSC and be confined to the existing VHF (Very High Frequency) and UHF (Ultra High Frequency) bands.
- 1990: FCC announced a very different initiative
  - preference for a full-resolution HDTV
  - HDTV would be simultaneously broadcast with the existing NTSC TV and eventually replace it.
- 1993: after a boom of proposals for digital HDTV, the FCC made a key decision to go all-digital.
  - A "grand alliance" was formed that included four main proposals, by General Instruments, MIT, Zenith, and AT&T, and by Thomson, Philips, Sarnoff and others.
  - This eventually led to the formation of the ATSC (Advanced Television Systems Committee) -- responsible for the standard for TV broadcasting of HDTV.
- 1995: U.S. FCC Advisory Committee on Advanced Television Service recommended that the ATSC Digital Television Standard be adopted.

Advanced Digital Formats by ATSC

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

"I": interlaced scan
"P": progressive (non-interlaced) scan

More about HDTV

- For video, MPEG-2 is chosen as the compression standard.
- For audio, AC-3 is the standard
  - supports 5.1 channel Dolby surround sound -- 5 surround channels plus a subwoofer channel
- Difference between conventional TV and HDTV:
  - Much wider aspect ratio of 16:9 instead of 4:3.
  - Move towards progressive (non-interlaced) scan
    - interlacing introduces serrated edges to moving objects and flickers along horizontal edges

Recent Advances

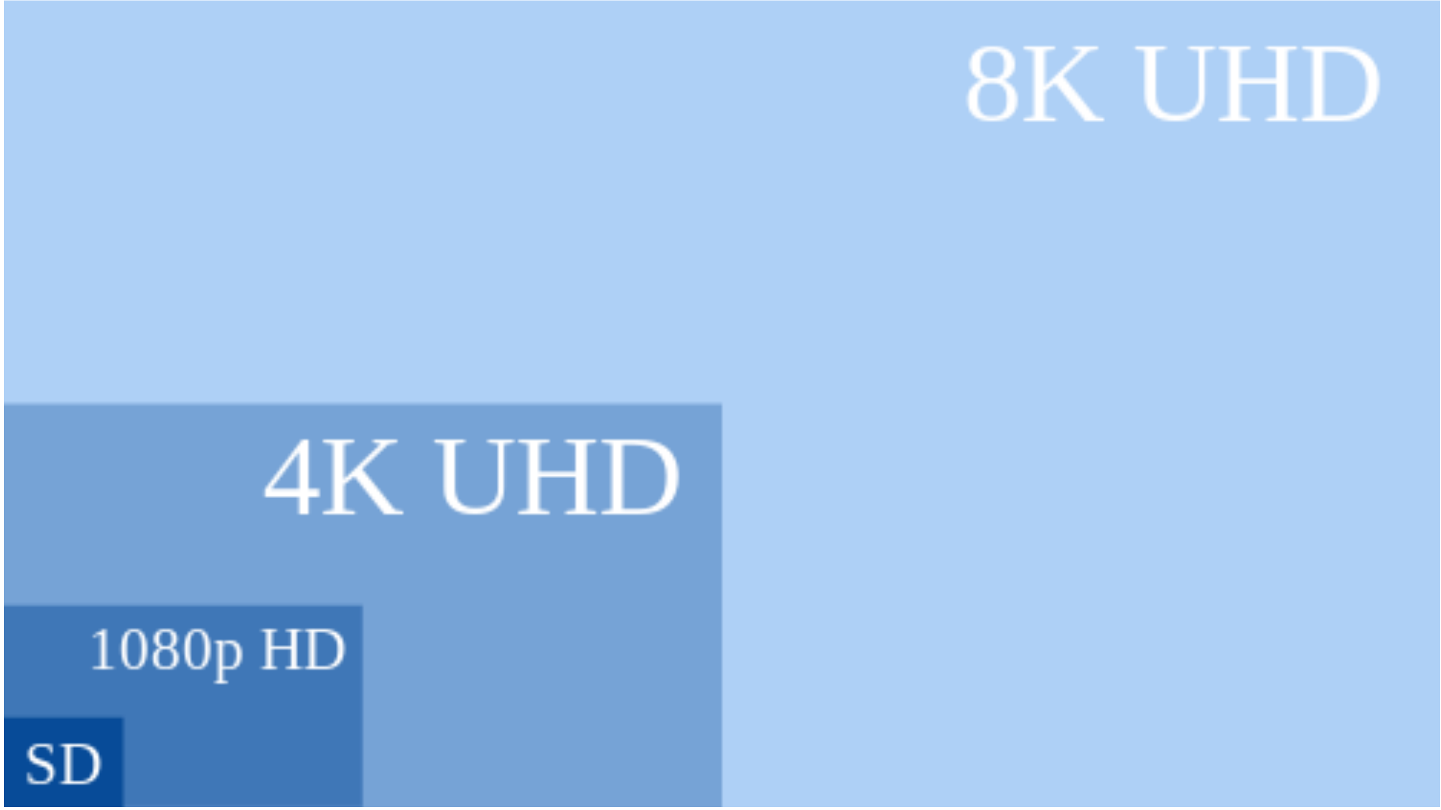
- The FCC (Federal Communications Commission) has planned to replace all analog broadcast services with digital TV broadcasting by the year 2006.
  - later delayed to June 12,2009 in US
  - Canada: August 31, 2011 (one year extension for some CBC transmitters)
- The services provided will include:
  - **SDTV (Standard Definition TV)**: the current NTSC TV or higher
  - **EDTV (Enhanced Definition TV)**: 480 active lines or higher, i.e., the third and fourth rows in the Table
  - **HDTV (High Definition TV)**: 720 active lines or higher

Ultra High Definition TV (UHD, or 4K)

- Announced in 2012
  - 4K UHDTV: 2160P (3,840×2,160, progressive scan) and 8K UHDTV: 4320P (7,680×4,320, progressive scan).
- Aspect ratio is 16:9. Bit-depth can be up to 12 bits, and the chroma subsampling can be 4:2:0 or 4:2:2.
- Supported frame rate has been gradually increased to 120 fps.

UHDTV will provide superior picture quality, comparable to IMAX movies, but it will require a much higher bandwidth and bitrate – 40MBps encoded !

Ultra High Definition TV (UHD, or 4K)



CCD/CMOS Image Sensor Size

Common Sensor Sizes

Sensor Type	1/2.5"	1/1.8"	2/3"	4/3"	APS-C	35mm
Aspect Ratio	4:3	4:3	4:3	4:3	2:3	2:3
Diagonal (mm)	7.2	8.9	11	22.5	27.3	43.3
Width (mm)	5.8	7.2	8.8	18	22.7	36
Height (mm)	4.3	5.3	6.6	13.5	15.1	24

Outline

- Types of Video Signals
- Analog Video
- Digital Video
- HDTV
- Further Exploration

Further Exploration

- Chapter 5.1, 5.2, 5.3