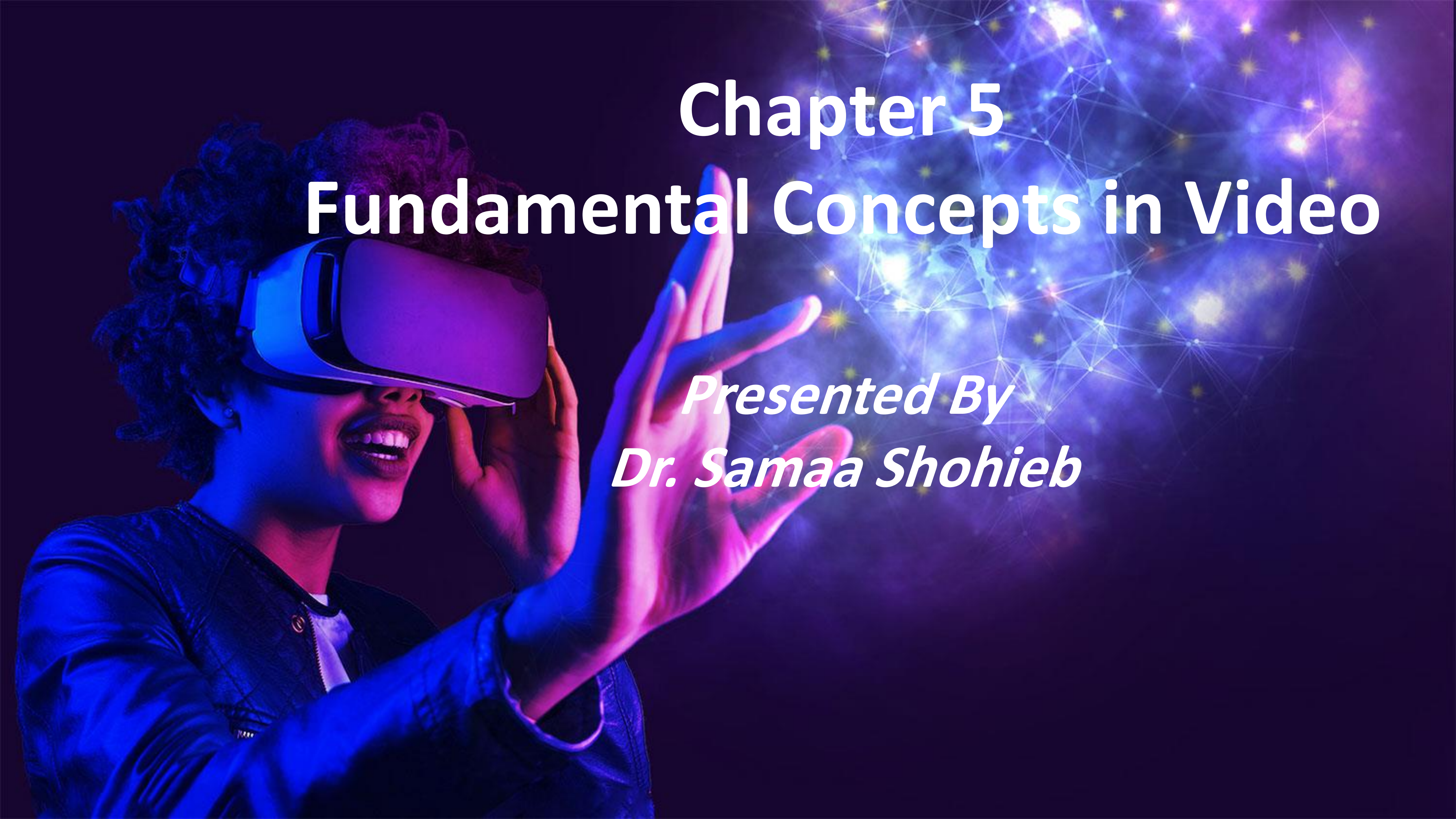
Chapter 5
Fundamental Concepts in Video



Presented By

Dr. Samaa Shohieb

Agenda

5.1 Analog Video
5.2 Digital Video
5.3 Video Display Interfaces
5.4 360o Video
5.5 3D Video and TV
5.6 Video Quality Assessment (VQA)

5.1 Analog Video

- An analog signal $f(t)$ samples **a time-varying image**.
- So-called ***"progressive" scanning*** traces through a **complete picture** (***a frame***) **row-wise** for **each time** interval.
- **In TV**, and in some monitors and multimedia standards as well, another system, ***called "interlaced" scanning*** is used:

a) The **odd-numbered lines** are **traced first**, and **then** the **even-numbered lines** are **trace**d.

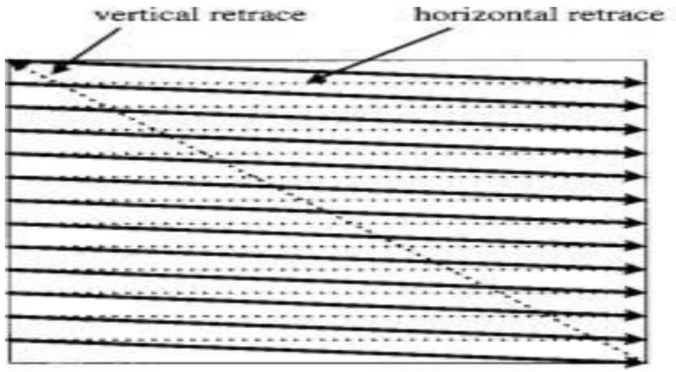
b) one frame = "odd" +"even" fields — (two fields make up).

c) 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.

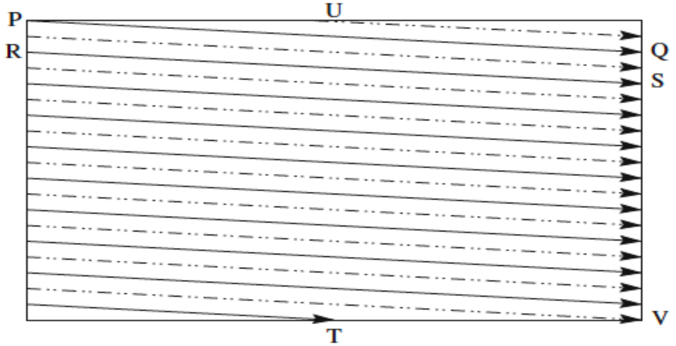
d) **the interlaced scanning**. First the solid (odd) lines are traced, P to Q, then R to S, etc., ending at T; then the even field starts at U and ends at V.

e) The **jump** from **Q to R,** etc. ➔ called the **horizontal retrace**, during which the electronic beam in the CRT(Cathode-ray tube) is blanked.

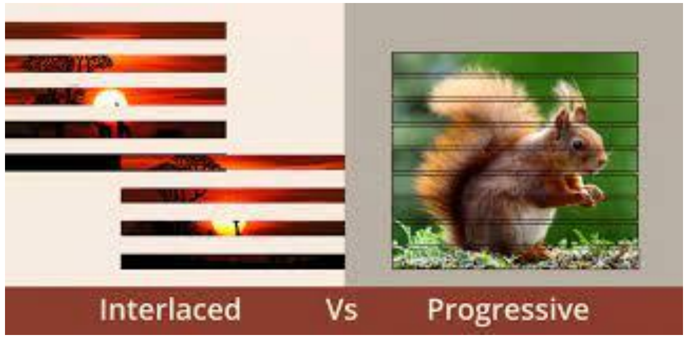
f) The **jump** from **T to U or V to P** ➔called the **vertical retrace**.



Progressive



Interlacing Scanning



5.1 Analog Video cont.

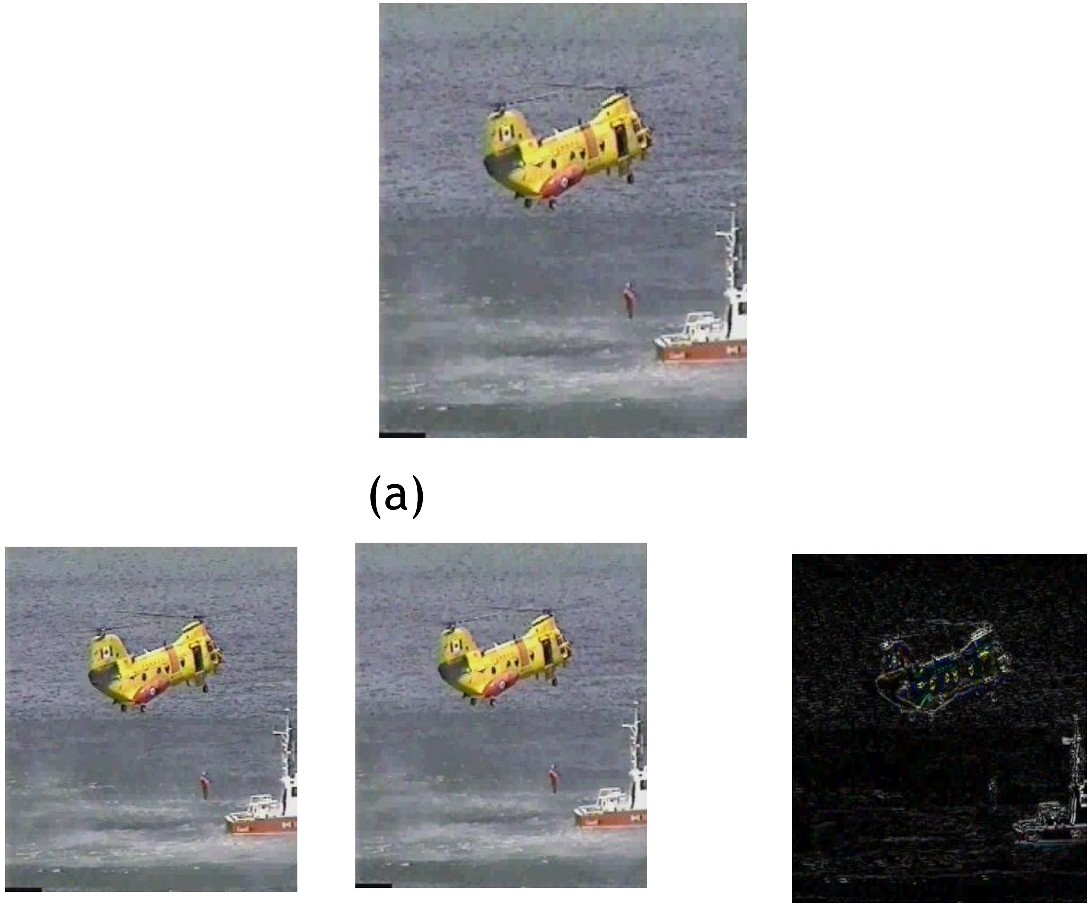
- **<u>Why Interlacing</u>** ? It was difficult to transmit the amount of information in a full frame quickly enough **to avoid *<u>Flicker</u>*→ Every ½ the T sending 1 field (like a new image) →Gives the feeling of continuity of the video→ reduced the *<u>Flicker</u>***

- 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.**

    - ***<u>Called Aliasing effect</u>***

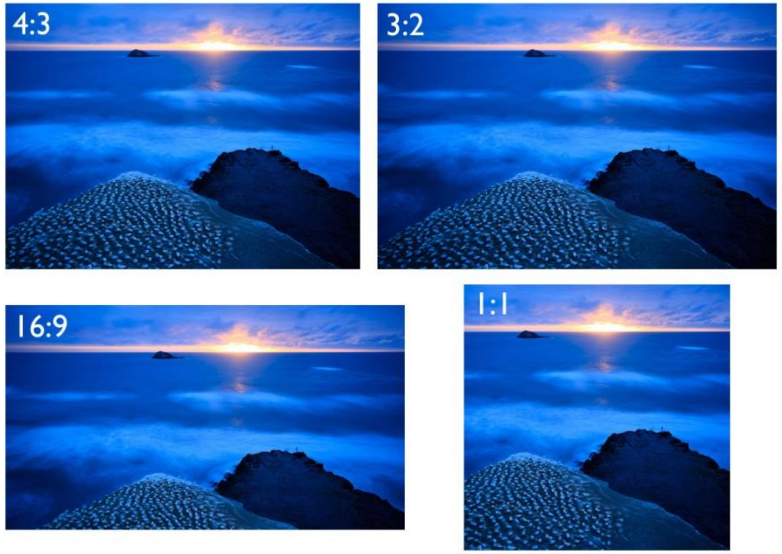


(a) (b) (c) (d)

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.

5.1 Analog Video cont.

- Sometimes necessary to change the **frame rate** (number of frames per second), **resize**, or even produce **stills from an interlaced source video**, various schemes are used **to "de-interlace"** it.
- 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.
  - Other more **complicated methods are also possible.**
- **Aspect Ratio→ The width : The height (of the Image)**



Aspect Ratio Examples

Color Models in Video

(a) YUV→ **Analog**
(b) YIQ→ **Analog**
(c) YCbCr→ **Digital**

- Y➔ codes are **luminance** (for gamma-corrected signals) the "**Luma**".➔ **Brightness or Lightness**
- **U & V/ I & Q** are **Chrominance ➔ Color Information (Chroma)**

(c) For gray, $R' = G' = B'$, the luminance

Y' equals to that gray, = 0.299R+0.587G+0.114 B= 1.0.

And for a gray ("black and white") image, the chrominance (U, V) is **zero**.

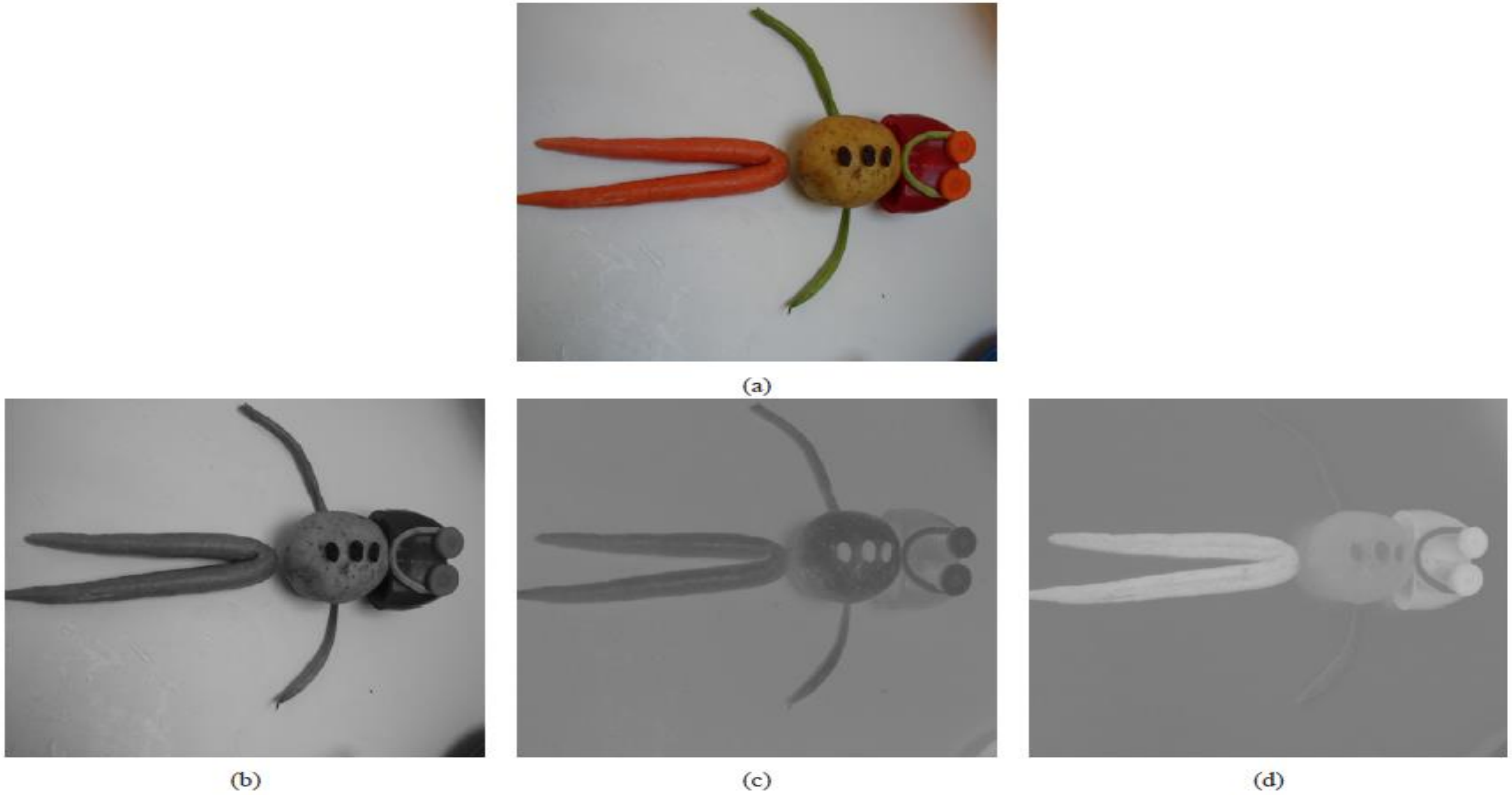


Fig. 4.18: Y ′UV decomposition of color image. Top image (a) is original color image; (b) is Y ′; (c,d) are (U, V)

YIQ Color Model

- YIQ is used in NTSC color TV broadcasting. Again, gray pixels generate zero (I, Q) chrominance signal.

(a) **I** and **Q** are ***a rotated version*** of **U** and **V** .

(b) Y' in YIQ is the same as in YUV; U and V are rotated by 33°:

$$I = 0.492111(R' - Y') \cos 33° - 0.877283(B' - Y') \sin 33°$$

$$Q = 0.492111(R' - Y') \sin 33° + 0.877283(B' - Y') \cos 33°$$

(c) This leads to the following matrix transform:

$$\begin{bmatrix} Y' \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595879 & -0.274133 & -0.321746 \\ 0.211205 & -0.523083 & 0.311878 \end{bmatrix} = \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{4.37}$$

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).

- **NTSC** follows the **interlaced scanning system**,
- frame **is divided into two fields**,
  - with **262.5** lines/field.
- Thus **the horizontal sweep frequency** is **525×29.97 msec**
- **≈15, 734 lines/sec**, so that each line is swept out in **1/15.734 ×10⁶sec ≈ 63.6μsec**.
- Since the **horizontal retrace** takes **10.9** μsec,
- this leaves **63.6 - 10.9 = 52.7 μsec** for the **active line signal** during which **image data is displayed.**

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

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.

  (b) It uses an **8 MHz** channel and allocates

  (a) a bandwidth of **5.5 MHz to Y**,

  (b) and 1.8 MHz each to **U and V**. (WHY LESS -→ **CHROMA SUBSAMPLING** Later)

  (c) The color subcarrier frequency is $f_{sc} \approx 4.43$ MHz.

  (c) In order to **improve picture quality,**

  (d) Chroma signals have alternate signs (e.g., +U and -U) in **successive scan lines**, hence the name "**Phase Alternating Line**".

  (e) This facilitates the use of a (line rate) comb filter at the receiver

  (f) The signals in consecutive lines are **averaged** so as to **cancel the Chroma signals** (that always **carry opposite signs (low quality)**) for separating Y and C and obtaining high quality Y signals.

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**.

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=~ 30	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

5.2 Digital Video

- The ***advantages of digital representation*** for video are many. For example:

  (a) **Video** can be **stored on digital devices** or in **memory**,

  (b) Ready to be **processed** (**noise removal**, **cut and paste**, etc.),

  (c) and **integrated to various multimedia applications**

  (d) Direct access is possible, which **makes nonlinear video editing** achievable as a simple, rather than a complex, task;

  (e) **Repeated recording** does not **degrade** image quality;

  (f) **Ease of encryption** and **better tolerance** to channel noise.

YCbCr color model

- **YCbCr color model**: uses in digital video model.
- Y′ is the Luma component and Cb and Cr are the Blue-difference and Red-difference Chroma components.
- Closely related to YUV: scaled and shifted YUV

$$\boldsymbol{Cb = ((B - Y)/2) + 0.5}$$
$$\boldsymbol{Cr = ((R - Y)/1.6) + 0.5}$$

Y C_B C_R

5.2.2 CCIR, ITU-R and CIF 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 application
  - Adopted by **certain digital video formats** including the popular DV video.

- **CIF** stands for **Common Intermediate Format** specified by the **CCITT**.
  - (a) The idea of CIF is to specify a format for lower bitrate.
  - (b) CIF is about the same as **VHS quality.**
  - (c) It uses a ***progressive*** **(non-interlaced)** scan.
  - (d) QCIF stands for "Quarter-CIF". All the CIF/QCIF resolutions = divisible by 8,
  - (e) 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.**

Table 5.3: 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

5.2.1 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 (but not necessarily informative!) names have arisen to label the different schemes used.
- To begin with, numbers are given stating how many pixel values, per four original pixels, are actually sent:

  a) **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.

  b) The scheme **"4:2:2"** indicates **horizontal subsampling of the Cb**, **Cr** signals **by a factor of 2**.

  c) 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**, as **(Cb0, Y0)(Cr0, Y1)(Cb2, Y2)(Cr2, Y3)(Cb4, Y4), and so on (or averaging is used).**

  d) The scheme "4:1:1" subsamples *horizontally* by a **factor of 4**.

  e) 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 as shown Fig.5.6.

  – Scheme 4:2:0 along with other schemes is commonly used in JPEG and MPEG (see later chapters in Part 2).

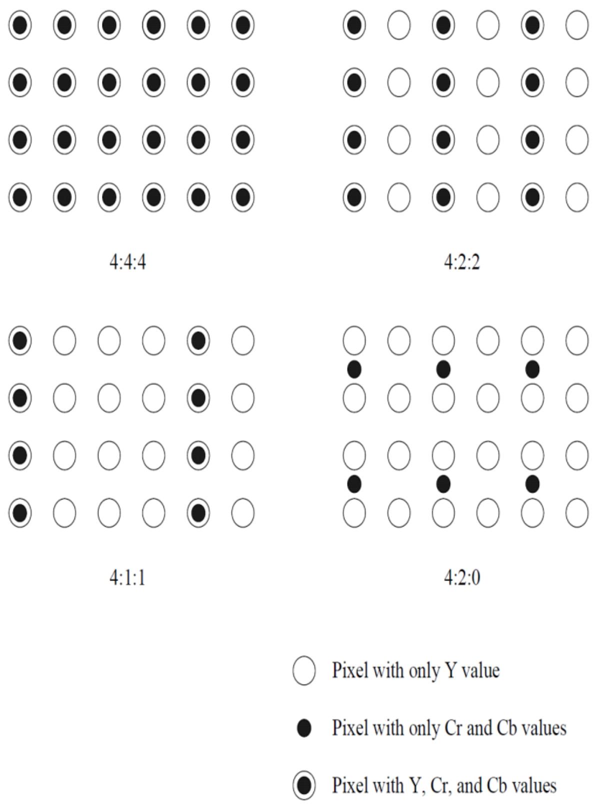


Fig. 5.6: Chroma subsampling

Bit Rate & Video Size

$$Bitrate\ (BR) = width(W) \times Hight\ (H) \times Color\ depth \times fps$$

$$Video\ size = width\ (W) \times Hight\ (H) \times Color\ depth \times fps \times duration$$

For example, video of duration of 1 hour (3600sec) and frame size of 640 x 480 at color depth of 24 bits and frame rate 25 fps. This video has the following properties:

$$\text{Pixels per frame} = 640 \times 480 = 307{,}200$$

$$\text{Bits per frame} = 307{,}200 \times 24 = 7{,}372{,}800 = 7.37\ Mbits$$

$$\text{Bitrate (BR)} = 7.37 \times 25 = 184{,}25\ Mbits/sec$$

$$\text{Video size} = Bitrate \times duration = 184{,}25 \times 3600 sec = 662{,}400\ Mbits$$

$$= 82{,}800\ Mbytes = 82.8\ Gbytes$$

Bit Rate & Video Size Example

Example 1: Resolution 720 × 485 frame rate 30 frames per sec (fps) using 4:4:4 and 4:2:2 sampling. **Calculate video bit rate**.

Solution:

Using 4:4:4 sampling:

Pixels per frame = 720 × 485 = 349,200 pixels/frame

4:4:4 sampling gives 720 × 485 × 3 =1,047,600 bytes/frame ≈ 1.05 M/frame

Video bit rate= $1.05 \times 30 = 31.5\ MBytes/sec \rightarrow 31.5M \times 8bits = 250\ Mbps$

Using 4:2:2 subsampling

Pixels per frame= $720 \times 485 \times 2 = 698{,}400\ bytes/frame \approx 0.698\ M/frame$

Video bit rate = $0.698 \times 30 = 21\ MB/sec = 21M \times 8 = 168Mbps$

Bit Rate & Video Size Example Cont.

Example 2: Resolution 1280 × 720 frame rate 30fps using 4:2:0 subsampling. **Find video bit rate**.

Example 3: Resolution 1080 × 1920 frame rate 60fps using 4:4:4 subsampling. **Find video bit rate**.

Bit Rate & Video Size Example Cont.

Example 2: Resolution 1280×720 frame rate 30fps using 4:2:0 subsampling. **Find video bit rate**.

Solution:

Pixels per frame = $1280 \times 720 = 921{,}600$ pixels/frame

4:2:0 $\rightarrow 921{,}600 \times 1.5 = 1{,}382{,}400$ bytes/frame ≈ 1.38 MB/frame

Video bit rate = $1.38M \times 30 = 41MB/sec \rightarrow 41 \times 8 = 328Mbps.$

Example 3: Resolution 1080×1920 frame rate 60fps using 4:4:4 subsampling. **Find video bit rate**.

Solution:

Pixels per frame = $1080 \times 1920 = 2{,}073{,}600$ pixels/frame

4:4:4 $\rightarrow 2{,}073{,}600 \times 3 = 6{,}220{,}800$ bytes/frame ≈ 6.22 MB/frame

Video bit rate = $6.22\ M \times 60 = 373.25\ MB/sec \rightarrow 373.25 \times 82985.98Mbps\ \approx 3\ Gbps.$

5.2.3 High Definition TV (HDTV)

- The main thrust of **HDTV** (High Definition TV) is not to increase the "definition" in each unit area, but rather **to increase the visual field especially in its width**.
  a) The first generation of HDTV was based on an analog technology developed by Sony and NHK in Japan in the late 1970s.
  b) 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.**
  c) Demand more **than 20 MHz bandwidth**
  d) Will not fit in the current **6 MHz or 8 MHz channels→** various compression techniques are being investigated.
  e) It is also anticipated that high quality HDTV signals will be transmitted using more than one channel even after compression.

5.2.3 High Definition TV (HDTV) cont.

- The services provided include:
  - **SDTV (Standard Definition TV)**: the NTSC TV or higher.
  - **EDTV (Enhanced Definition TV)**: 480 active lines or higher, i.e., the third and fourth rows in Table 5.4.
  - **HDTV (High Definition TV)**: 720 active lines or higher. So far, the popular choices are:
    - 720P (1,280 $\times$ 720, progressive scan, 30 fps)
    - 1080I (1,920 $\times$ 1,080, interlaced, 30 fps)
    - 1080P (1,920 $\times$ 1,080, progressive scan, 30 or 60 fps).

Table 5.4: Advanced Digital TV formats supported by

Advanced TV Systems Committee (ATSC)

# of Active Pixels per line	# of Active Lines	Aspect Ratio	Frame 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

5.2.4 Ultra High Definition TV (UHDTV)

- UHDTV is a new generation of HDTV. The standards initiated in 2012 support **4K UHDTV**: 2160P **(3,840×2,160, progressive scan**) and **8K UHDTV**: **4320P** (**7,680×4,320, progressive scan**).
- The aspect ratio is 16:9. The bit-depth is 10 or 12 bits per sample, and the **chroma subsampling can be 4:2:0, 4:2:2, or 4:4:4.**
- **The supported frame rate has been gradually increased to 120 fps.**
- **The UHDTV** will provide superior picture quality, comparable to IMAX movies, but it will require a much higher bandwidth and bitrate.
- **16K UHDTV** has been demonstrated in 2018, targeting applications such as Virtual Reality with true immersion. Its resolution is 15,360 × 8,640 for a total of **132.7 megapixels.**

Type of UHDTV	Resolution	Bit Depth	Aspect Ratio	Frame Rate
4K UHD (2160P)	3840 x 2160	10 or 12 bits	16:9	Up to 120P
8K UHD (4320P)	7680 x 4320	10 or 12 bits	16:9	Up to 120P
16K UHD (8640P)	15360 x 8640	10 or 12 bits	16:9	Up to 240P

5.3 Video Display Interfaces

5.3.1 Analog Display Interfaces

- Analog video signals are often transmitted in one of three different interfaces: **Component video, Composite video, and S-video.** Figure 5.7 shows the typical connectors for them.

Fig. 5.7: Connectors for typical analog display interfaces. From left to right: Component video, Composite video, S-video, and 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.
  a) **Most computer systems** use **Component Vide**o, with **separate** signals for **R**, **G**, and **B** signals.
  b) **➔ Component Video gives *<u>the best color reproduction</u>* since there is *<u>no “crosstalk</u>*” between the three channels.**
  c) Component video, **requires more bandwidth** and **good synchronization of the three components**.

Composite Video

- **Composite video**: **color (“chrominance**”) and **intensity (“luminance”)** signals are mixed into a ***single* carrier wave.**

  a) Chrominance is a composition of **two color components** (I and Q, or U and V).

  b) 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.**

  c) The **chrominance** and **luminance** components **can be separated at the receiver end**

  d) one wire and video color signals are mixed, not sent separately.

  e) 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

- **S-Video**: as a compromise, (**separated video, or Super-video,** e.g., in S-VHS)
- **Black-and-white information is most crucial for visual perception.**
  - **Humans are able to differentiate spatial resolution** in grayscale images with a much **higher acuity 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.

5.3.2 Digital Display Interfaces

- Digital interfaces emerged in 1980s (e.g., Color Graphics Adapter (CGA)),
- Today, the most widely used digital video interfaces include Digital Visual Interface (DVI),
- High-Definition Multimedia Interface (HDMI), and DisplayPort.

- **Fig. 5.8:** Connectors of different digital display interfaces. From left to right: **DVI**, **HDMI**, **DisplayPort**.

High-Definition Multimedia Interface (HDMI)

- HDMI is a newer digital audio/video interface developed to be backward compatible with **DVI.**

1. HDMI doesn't carry analog signal and hence is not compatible with VGA.
2. HDMI supports both RGB and **YCbCr <u>4:4:4 or 4:2:2</u>.**
3. [DVI is limited to the RGB color range (0-255).]
4. HDMI supports digital audio, in addition to digital video.

- The maximum pixel clock rate for HDMI 1.0 is 165 MHz, which is sufficient to support **1080P (1,920 × 1,200) at 60 Hz.**
- HDMI 2.0 was released in 2013, which supports 4K resolution at 60 frames per second.

5.4 360^{o} Video

- 360^{o} Video is also known **as *Omnidirectional Video*, *Spherical Video* or *Immersive Video*.**
- 360^{o} Video can span 360^{o} horizontally and 180^{o} vertically.
- It is captured by cameras at (almost) all possible viewing angles.
- In actual implementations, the video
- is usually captured by an ***omnidirectional camera*** or a collection of cameras with a wide field of view.
- 360^{o} Video can be ***monoscopic or stereoscopic***. The latter is especially suitable for **VR applications**.

5.5 3D Video and TV

- Three-dimensional (3D) pictures and movies have been in existence for decades.
- the success of the **2009** film **Avatar** have pushed 3D video to its peak.
- Increasingly, it is in movie theaters, broadcast TV (e.g., sporting events), PCs, and various hand-held devices.
- The main advantage of the 3D video is that it enables the experience of immersion — be there, and really Be there!

5.6 Video Quality Assessment (VQA)

- When **comparing the coding efficiency** of different video compression methods, a common practice is to **compare the bitrates** of the coded video **bitstreams at the same *quality***.

- The video quality assessment approaches can be ***objective or subjective***: the former is done automatically by computers and the latter requires human judgment.


THANK YOU