ICT 514
Multimedia Systems

Topic 3: Science of Colour & Video

Ref:
Li, Z. & Drew M., “Fundamentals of Multimedia”, Chapter 4 & 5
Science of Colour

• Light is an **electromagnetic** wave. Its color is characterized by the wavelength content of the light.
  
  (a) Laser light is a single wavelength.
  (b) Most light sources made up of many wavelengths.
  (c) Humans cannot detect all light, just those within the visible wavelengths.
  (d) Short wavelengths produce a blue sensation, long wavelengths produce a red one.

• **Spectrophotometer** is a device used to measure visible light.
Colour and the science behind

- Visible light is an electromagnetic wave in the range 400 nm to 700 nm (where 1 nm is $10^{-9}$ meter).
- A graph indicating the intensity of the light components is called a Spectral Power Distribution (SPD) or a spectrum.
- The symbol for wavelength is $\lambda$. This curve is called $E(\lambda)$. 
Fig. 4.2: Spectral power distribution of daylight.
Human Vision

• The eye works like a camera, with the lens focusing an image onto the retina (upside-down and left-right reversed).

• The retina consists of an array of rods and three kinds of cones. The rods come into play when light levels are low and produce a image in shades of gray.

• For higher light levels, the cones each produce a signal. Because of their differing pigments, the three kinds of cones are most sensitive to red ($R$), green ($G$), and blue ($B$) light.

• It seems likely that the brain makes use of differences $R-G$, $G-B$, and $B-R$, as well as combining all of $R$, $G$, and $B$ into a high-light-level achromatic channel.
Spectral Sensitivity of the Eye

- The eye is most sensitive to light in the middle of the visible spectrum.
- The sensitivity of our receptors is also a function of wave-length.
- Blue is a late addition in evolution and statistically, Blue is the favorite color of humans, regardless of nationality. Perhaps because Blue is a latecomer and thus is a bit surprising!
Fig. 4.3: R, G, and B cones, and Luminous Efficiency curve $V(\lambda)$. 

Li & Drew ©Prentice Hall 2003
• Fig. 4.3 shows the overall sensitivity as a dashed line. This important curve is called the luminous-efficiency function.

• The Blue receptor sensitivity is not shown to scale because it is much smaller than the curves for Red or Green.

• It is usually denoted $V(\lambda)$ and is formed as the sum of the response curves for Red, Green, and Blue.

• The rod sensitivity curve looks like the luminous-efficiency function $V(\lambda)$ but is shifted to the red end of the spectrum.

• The achromatic channel produced by the cones is approximately proportional to $2R+G+B/20$. 
• These spectral sensitivity functions are usually denoted by letters other than “R; G; B”. They can be represented by a vector function $\mathbf{q} (\lambda)$, with components

$$
\mathbf{q} (\lambda) = (q_R(\lambda); q_G(\lambda); q_B(\lambda))^T
$$
• Adding up the cone responses for all wavelengths, weighted by the eye's relative response at that wavelength

\[ R = \int E(\lambda) \, q_R(\lambda) \, d\lambda \]
\[ G = \int E(\lambda) \, q_G(\lambda) \, d\lambda \]
\[ B = \int E(\lambda) \, q_B(\lambda) \, d\lambda \]

• \( \int \) is the symbol of “integral”.

Reflection

• Surfaces reflect different amounts of light at different wavelengths, and dark surfaces reflect less energy than light surfaces.

• The reflectance function is denoted $S(\lambda)$. 
Fig. 4.4: Surface spectral reflectance functions $S(\lambda)$ for objects.
Image formation

• Light from the illuminant with SPD \( E(\lambda) \) impinges on a surface, with surface spectral reflectance function \( S(\lambda) \), is reflected, and then is filtered by the eye's cone functions \( q(\lambda) \).

• The function \( C(\lambda) \) is called the color signal and consists of the product of \( E(\lambda) \), the illuminant, times \( S(\lambda) \), the reflectance:

\[
C(\lambda) = E(\lambda) S(\lambda)
\]
Fig. 4.5: Image formation model.
Gamma Correction

- The light emitted is in fact roughly proportional to the voltage raised to a power; this power is called **gamma**, with symbol $\gamma$ (Greek Small letter Gamma).
  
  (a) If the value in the red channel is $R$, the screen emits light proportional to $R$, with SPD equal to that of the red phosphor paint on the screen that is the target of the red channel electron gun. The value of gamma is around 2.2.
  
  (b) It is customary to append a prime to signals that are **gamma-corrected** by raising to the power $(1/\gamma)$ before transmission in order to recover the linear signals.
Fig. 4.6: (a): Effect of CRT on light emitted from screen (voltage is normalized to range 0..1). (b): Gamma correction of signal.
Fig. 4.7: (a): Display of ramp from 0 to 255, with no gamma correction. (b): Image with gamma correction applied
Color-Matching Functions

• Even without knowing the eye-sensitivity curves of Fig.4.3, a technique evolved in psychology for matching a combination of basic $R$, $G$, and $B$ lights to a given shade.
• The particular set of three basic lights used in an experiment are called the set of color primaries.
• To match a given color, a subject is asked to separately adjust the brightness of the three primaries using a set of controls until the resulting spot of light most closely matches the desired color.
• The basic situation is shown in Fig.4.8. A device for carrying out such an experiment is called a colorimeter.
Fig. 4.8: Colorimeter experiment.
• The amounts of R, G, and B the subject selects to match each single-wavelength light forms the color-matching curves. These are denoted $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ and are shown in Fig. 4.9.

Fig. 4.9: CIE RGB color-matching functions $\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda)$. 
CIE Chromaticity Diagram

- Since the \( \tilde{r}(\lambda) \) color-matching curve has a negative lobe, a set of fictitious primaries were devised that lead to color-matching functions with only positives values.

(a) The resulting curves are shown in Fig. 4.10; these are usually referred to as *the* color-matching functions.

(b) They are a \( 3 \times 3 \) matrix away from \( \tilde{r}, \tilde{g}, \tilde{b} \) curves, and are denoted \( \tilde{x}(\lambda), \tilde{y}(\lambda), \tilde{z}(\lambda) \).

(c) The matrix is chosen such that the middle standard color-matching function \( \tilde{y}(\lambda) \) exactly equals the luminous-efficiency curve \( V(\lambda) \) shown in Fig. 4.3.
Fig. 4.10: CIE standard XYZ color-matching functions $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$. 
For a general SPD $E(\lambda)$, the essential “colorimetric” information required to characterize a color is the set of *tristimulus values* $X, Y, Z$ defined in analogy to (Eq. 4.2) as ($Y = \text{luminance}$):

\[
X = \int E(\lambda) \, \bar{x}(\lambda) \, d\lambda
\]

\[
Y = \int E(\lambda) \, \bar{y}(\lambda) \, d\lambda
\]

\[
Z = \int E(\lambda) \, \bar{z}(\lambda) \, d\lambda
\]

(4.6)

3D data is difficult to visualize, so the CIE devised a 2D diagram based on the values of $(X, Y, Z)$ triples implied by the curves in Fig. 4.10.
• We go to 2D by factoring out the magnitude of vectors \((X; Y; Z)\) by dividing by the sum \((X + Y + Z)\) to make the **chromaticity**:

\[
x = \frac{X}{X + Y + Z}
\]

\[
y = \frac{Y}{X + Y + Z}
\]

\[
z = \frac{Z}{X + Y + Z}
\]
• This effectively means that one value out of the set \((x; y; z)\) is redundant since we have
\[
x + y + z = \frac{(X + Y + Z)}{(X + Y + Z)} = 1
\]
• So,
\[
z = 1 - x - y
\]
- Effectively, we are projecting each tristimulus vector \((X, Y, Z)\) onto the plane connecting points \((1, 0, 0)\), \((0, 1, 0)\), and \((0, 0, 1)\).
- Fig. 4.11 shows the locus of points for monochromatic light.
(a) The color matching curves each add up to the same value — the area under each curve is the same for each of $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$.

(b) For an $E(\lambda) = 1$ for all $\lambda$, — an “equi-energy white light” — chromaticity values are $(1/3, 1/3)$. Fig. 4.11 displays a typical actual white point in the middle of the diagram.

(c) Since $x, y \leq 1$ and $x + y \leq 1$, all possible chromaticity values lie below the dashed diagonal line in Fig. 4.11.
Video Color Transforms

(a) Largely derive from older analog methods of coding color for TV. Luminance is separated from color information.
(b) For example, a matrix transform method called YIQ is used to transmit TV signals in North America and Japan.
(c) This coding also makes its way into VHS video tape coding in these countries since video tape technologies also use YIQ.
(d) In Europe, video tape uses the PAL or SECAM codings, which are based on TV that uses a matrix transform called YUV.
(e) Finally, digital video mostly uses a matrix transform called YCbCr that is closely related to YUV.
YUV Color Model

(a) YUV codes a luminance signal (for gamma-corrected signals) equal to $Y'$ in Eq. (4.20). the "luma".

(b) **Chrominance** refers to the difference between a color and a reference white at the same luminance. → use color differences $U, V$:

$$U = B' - Y', \quad V = R' - Y'$$  \hspace{1cm} (4.27)

From Eq. (4.20),

$$
\begin{bmatrix}
Y' \\
U \\
V
\end{bmatrix} =
\begin{bmatrix}
0.299 & 0.587 & 0.144 \\
-0.299 & -0.587 & 0.886 \\
0.701 & -0.587 & -0.114
\end{bmatrix}
\begin{bmatrix}
R' \\
G' \\
B'
\end{bmatrix}
$$  \hspace{1cm} (4.28)

(c) For gray, $R' = G' = B'$, the luminance $Y'$ equals to that gray, since $0.299 + 0.587 + 0.114 = 1.0$. And for a gray ("black and white") image, the chrominance $(U, V)$ is zero.
(d) In the actual implementation $U$ and $V$ are rescaled to have a more convenient maximum and minimum.

(e) For dealing with composite video, it turns out to be convenient to contain $U, V$ within the range $-1/3$ to $+4/3$. So $U$ and $V$ are rescaled:

\[
U = 0.492111 \cdot (B' - Y')
\]
\[
V = 0.877283 \cdot (R' - Y')
\]

The chrominance signal = the composite signal $C$:

\[
C = U \cdot \cos(\omega t) + V \cdot \sin(\omega t)
\]  

(f) Zero is not the minimum value for $U, V$. $U$ is approximately from blue ($U > 0$) to yellow ($U < 0$) in the RGB cube; $V$ is approximately from red ($V > 0$) to cyan ($V < 0$).

(g) Fig. 4.18 shows the decomposition of a color image into its $Y'$, $U$, $V$ components. Since both $U$ and $V$ go negative, in fact the images displayed are shifted and rescaled.
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°:

\[
I = 0.492111(R' - Y') \cos 33° - 0.877283(B' - Y') \sin 33°
\]
\[
I = 0.492111(R' - Y') \sin 33° + 0.877283(B' - Y') \cos 33° \quad (4.31)
\]

(c) This leads to the following matrix transform:

\[
\begin{bmatrix}
Y' \\
I \\
Q
\end{bmatrix} =
\begin{bmatrix}
0.299 & 0.587 & 0.144 \\
0.595879 & -0.274133 & -0.321746 \\
0.211205 & -0.523083 & 0.311878
\end{bmatrix}
\begin{bmatrix}
R' \\
G' \\
B'
\end{bmatrix} \quad (4.32)
\]

(d) Fig. 4.19 shows the decomposition of the same color image as above, into YIQ components.
Fig. 4.19: $I$ and $Q$ components of color image.
YCbCr Color Model

- The Rec. 601 standard for digital video uses another color space, $YC_bC_r$, often simply written YCbCr — closely related to the YUV transform.

  (a) YUV is changed by scaling such that $C_b$ is $U$, but with a coefficient of 0.5 multiplying $B'$. In some software systems, $C_b$ and $C_r$ are also shifted such that values are between 0 and 1.

  (b) This makes the equations as follows:

  $$C_b = ((B' - Y')/1.772) + 0.5$$
  $$C_r = ((R' - Y')/1.402) + 0.5$$

  (4.33)

  (c) Written out:

  $$\begin{bmatrix}
  Y' \\
  C_b \\
  C_r
  \end{bmatrix} =
  \begin{bmatrix}
  0.299 & 0.587 & 0.144 \\
  -0.168736 & -0.331264 & 0.5 \\
  0.5 & -0.418688 & -0.081312
  \end{bmatrix}
  \begin{bmatrix}
  R' \\
  G' \\
  B'
  \end{bmatrix} +
  \begin{bmatrix}
  0 \\
  0.5 \\
  0.5
  \end{bmatrix}$$

  (4.34)
(d) In practice, however, Recommendation 601 specifies 8-bit coding, with a maximum $Y'$ value of only 219, and a minimum of $+16$. Cb and Cr have a range of $\pm112$ and offset of $+128$. If $R'$, $G'$, $B'$ are floats in $[0..+1]$, then we obtain $Y'$, $C_b$, $C_r$ in $[0..255]$ via the transform:

$$
\begin{bmatrix}
Y' \\
C_b \\
C_r
\end{bmatrix}
= \begin{bmatrix}
65.481 & 128.553 & 24.966 \\
-37.797 & -74.203 & 112 \\
112 & -93.786 & -18.214
\end{bmatrix}
\begin{bmatrix}
R' \\
G' \\
B'
\end{bmatrix}
+ \begin{bmatrix}
16 \\
128 \\
128
\end{bmatrix}
$$  \hspace{1cm} (4.35)

(f) The YCbCr transform is used in JPEG image compression and MPEG video compression.
**Video Signals**

- **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.

  (a) Most computer systems use Component Video, with separate signals for R, G, and B signals.

  (b) For any color separation scheme, Component Video gives the best color reproduction since there is no “crosstalk“ between the three channels.

  (c) This is not the case for S-Video or Composite Video, discussed next. Component video, however, 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 and then the two color components can be further recovered.

  d) 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.
S-Video

• **S-Video**: as a compromise, (Separated video, or Super-video, e.g., in S-VHS) uses two wires, one for luminance and another for a composite chrominance signal.

• As a result, there is less crosstalk between the color information and the crucial gray-scale information.
S-Video

• The reason for placing luminance into its own part of the signal is that black-and-white information is most crucial for visual perception.
  – In fact, humans are able to differentiate spatial resolution in gray-scale images with a much higher accuracy 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. 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 traced. This results in “odd” and “even” fields - two fields make up one frame.
  b) 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.
c) Figure 5.1 shows the scheme used. 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.

d) The jump from Q to R, etc. in Figure 5.1 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.
• 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.

• 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.
• 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.
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 525x29.97 approx 15,734 lines/sec, so that each line is swept out in 1/(15,734 x10³) sec 63:6 μsec.

  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.
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).

d) 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 specified active lines. With interlaced scan, this could be as low as 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, as listed in Table 5.1.

<table>
<thead>
<tr>
<th>Format</th>
<th>Samples per line</th>
</tr>
</thead>
<tbody>
<tr>
<td>VHS</td>
<td>240</td>
</tr>
<tr>
<td>S-VHS</td>
<td>400-425</td>
</tr>
<tr>
<td>Betamax</td>
<td>500</td>
</tr>
<tr>
<td>Standard 8 mm</td>
<td>300</td>
</tr>
<tr>
<td>Hi-8 mm</td>
<td>425</td>
</tr>
</tbody>
</table>
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 \( f_{sc} \approx 4.43 \text{ MHz} \).
  (b) 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".
  (c) This facilitates the use of a (line rate) comb filter at the receiver - the signals in consecutive lines are averaged 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** stands for *Systeme Electronique Couleur Avec Memoire*, 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

<table>
<thead>
<tr>
<th>TV System</th>
<th>Frame Rate (fps)</th>
<th># of Scan Lines</th>
<th>Total Channel Width (MHz)</th>
<th>Bandwidth Allocation (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTSC</td>
<td>29.97</td>
<td>525</td>
<td>6.0</td>
<td>Y: 4.2, I or U: 1.6, Q or V: 0.6</td>
</tr>
<tr>
<td>PAL</td>
<td>25</td>
<td>625</td>
<td>8.0</td>
<td>Y: 5.5, I or U: 1.8, Q or V: 1.8</td>
</tr>
<tr>
<td>SECAM</td>
<td>25</td>
<td>625</td>
<td>8.0</td>
<td>Y: 6.0, I or U: 2.0, Q or V: 2.0</td>
</tr>
</tbody>
</table>
Digital Video

(a) 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;

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

(c) Repeated recording does not degrade image quality;

(d) 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.

• Different (but not necessarily informative!) names have arisen to label the different schemes used.

• 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. 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).

(c) The scheme “4:1:1" subsamples horizontally by a factor of 4.

(d) 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.
Fig. 5.6: 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.

- Table 5.3 shows some of the digital video specifications, all with an aspect ratio of 4:3. The CCIR 601 standard uses an interlaced scan, so each eld has only half as much vertical resolution (e.g., 240 lines in NTSC)
• CIF stands for Common Intermediate Format specified by the CCITT.

  (a) The idea of CIF is to specify a format for lower bit-rate.
  (b) CIF is about the same as VHS quality. It uses a progressive (non-interlaced) scan.
  (c) 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.
  (d) 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: Digital video specifications

<table>
<thead>
<tr>
<th></th>
<th>CCIR 601 525/60 NTSC</th>
<th>CCIR 601 625/50 PAL/SECAM</th>
<th>CIF</th>
<th>QCIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminance resolution</td>
<td>720 × 480</td>
<td>720 × 576</td>
<td>352 × 288</td>
<td>176 × 144</td>
</tr>
<tr>
<td>Chrominance resolution</td>
<td>360 × 480</td>
<td>360 × 576</td>
<td>176 × 144</td>
<td>88 × 72</td>
</tr>
<tr>
<td>Color Subsampling</td>
<td>4:2:2</td>
<td>4:2:2</td>
<td>4:2:0</td>
<td>4:2:0</td>
</tr>
<tr>
<td>Aspect Ratio</td>
<td>4:3</td>
<td>4:3</td>
<td>4:3</td>
<td>4:3</td>
</tr>
<tr>
<td>Fields/sec</td>
<td>60</td>
<td>50</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Interlaced</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
HDTV (High Definition TV)

• The main thrust of HDTV 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) 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.

(d) It is also anticipated that high quality HDTV signals will be transmitted using more than one channel even after compression.
A brief history of HDTV evolution

(a) In 1987, the FCC decided that HDTV standards must be compatible with the existing NTSC standard and be connected to the existing VHF (Very High Frequency) and UHF (Ultra High Frequency) bands.

(b) In 1990, the FCC announced a very different initiative, i.e., its preference for a full-resolution HDTV, and it was decided that HDTV would be simultaneously broadcast with the existing NTSC TV and eventually replace it.

(c) Witnessing a boom of proposals for digital HDTV, the FCC made a key decision to go all-digital in 1993. A "grand alliance" was formed that included four main proposals, by General Instruments, MIT, Zenith, and AT&T, and by Thomson, Philips, Sarno and others.

(d) This eventually led to the formation of the ATSC (Advanced Television Systems Committee) responsible for the standard for TV broadcasting of HDTV.

(e) In 1995 the U.S. FCC Advisory Committee on Advanced Television Service recommended that the ATSC Digital Television Standard be adopted.
The standard supports video scanning formats shown in Table 5.4. In the table, "I" mean interlaced scan and "P" means progressive (non-interlaced) scan.

Table 5.4: Advanced Digital TV formats supported by ATSC

<table>
<thead>
<tr>
<th># of Active Pixels per line</th>
<th># of Active Lines</th>
<th>Aspect Ratio</th>
<th>Picture Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,920</td>
<td>1,080</td>
<td>16:9</td>
<td>60I 30P 24P</td>
</tr>
<tr>
<td>1,280</td>
<td>720</td>
<td>16:9</td>
<td>60P 30P 24P</td>
</tr>
<tr>
<td>704</td>
<td>480</td>
<td>16:9 &amp; 4:3</td>
<td>60I 60P 30P 24P</td>
</tr>
<tr>
<td>640</td>
<td>480</td>
<td>4:3</td>
<td>60I 60P 30P 24P</td>
</tr>
</tbody>
</table>
• For video, MPEG-2 is chosen as the compression standard.
• For audio, AC-3 is the standard. It supports the so-called 5.1 channel Dolby surround sound, i.e., five surround channels plus a subwoofer channel.
• The salient difference between conventional TV and HDTV:
  (a) HDTV has a much wider aspect ratio of 16:9 instead of 4:3.
  (b) HDTV moves toward progressive (non-interlaced) scan. The rationale is that interlacing introduces serrated edges to moving objects and flickers along horizontal edges.
The FCC has planned to replace all analog broadcast services with digital TV broadcasting by the year 2006. 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 Table 5.4.
- **HDTV (High Definition TV)**: 720 active lines or higher.
Digital TV in Australia
(Ref: http://www.digitaltv.com.au/)

• 26 May 1993
  – First meeting of the Australian Broadcasting Authority Digital Terrestrial Television Specialist Group.

• 1995

• 30 January 1997

• July 1997
  – Australian Broadcasting Authority releases its response to the Specialist Group report.
• **Late 1997**
  - The Department of Communications and the Arts and the Federation of Australian Commercial Television Stations in conjunction with the ABA and the ABC conducted tests of potential digital television technology for use in Australia. These were the first trials of both DVB and ASTC systems in the one country.

• **18 June 1998**
  - Digital Terrestrial Television Broadcasting (DTTB) Selection Panel unanimously agrees to recommend that the European DVB-T system be used in Australia.

• **3 July 1998**

• **15 July 1998**
  - Television Broadcasting Services (Digital Conversion) Bill 1998 and the Datacasting Charges (Imposition) Bill 1998 passed by the Australian House of Representatives in their new form as amended by the Senate.
• **30 July 1998**
  – FACTS Specialist Group-DTTB announces further technical element recommendations. Namely: the preferred video format for HDTV production purposes is 1920/1080/50Hz interlaced with a total line count of 1125 lines; Dolby AC-3 should be used for audio encoding; and the Service Information data standard will be based on the DVB-SI standard.

• **1 January 2001**
  – Deadline for commercial and national Free-to-Air broadcasters to commence DTTB broadcasts in metropolitan areas.

• **1 January 2004**
  – Deadline for commercial and national Free-to-Air broadcasters to commence DTTB broadcasts in regional areas.

• **31 December 2006**
  – Prohibition of new Free-to-Air broadcasters expires.
Why Digital?

- To overcome problems experienced with Analog Television
- To provide enhanced television services
DVB-T

- DVB-T is the digital broadcast standard that has been chosen for use in Australia. It comes from the "Digital Video Broadcasting" family of standards.
- This family covers things like Satellite and Cable broadcasting, but in our case we're interested in terrestrial broadcasting, the "over the air" broadcasting most people will be familiar with.
- DVB-T's main feature is that it uses COFDM modulation.
COFDM

- COFDM stands for **Coded Orthogonal Frequency Division Multiplex**.
- COFDM uses many closely spaced carriers to transmit a signal. This gives it two important properties. Firstly, it's fairly resistant to ghosting-style interference (technically called "multi-path" interference... which as far as technical expressions go, is pretty non-technical).
- Secondly, if there is another nearby transmitter transmitting the same COFDM signal on the same frequency, this will not be much different to "multi-path" interference.
- We can transmit the same signal on the same frequency from different sites and get away with it. The interference from the two signals will have minimal effect.
SFN - Single Frequency Networks

• What's so good about SFN?
• In the Australian context it generally means all "repeaters" and "translators" within the one viewing area can operate on the same frequency. This makes for much more efficient use of spectrum and will free up a lot of the stuff for future use.
HDTV is not Digital Television

- A lot of people get the impression that HDTV and Digital Television are the same thing. This is not the case. Digital Television offers many things and High Definition is one of them.
- High Definition brings a widescreen format to the home television as well as a sharper picture.
- The picture is sharper because with HDTV, there is much more information sent to describe many more individual pixels or dots which make up this picture.
Advantages of Digital TV

• Widescreen
• Reduced/No Interference
• Clearer Pictures and Sound
• Multiple Channels/Views
• Easy access to captions
• Datacasting/Interactive TV