C I & S S L a b

UNIVERSITÀ PARTHENOPE
DiST
Dipartimento di Scienze e Tecnologie

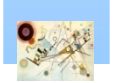
Intelligent Signal Processing

Basics of Image and Video

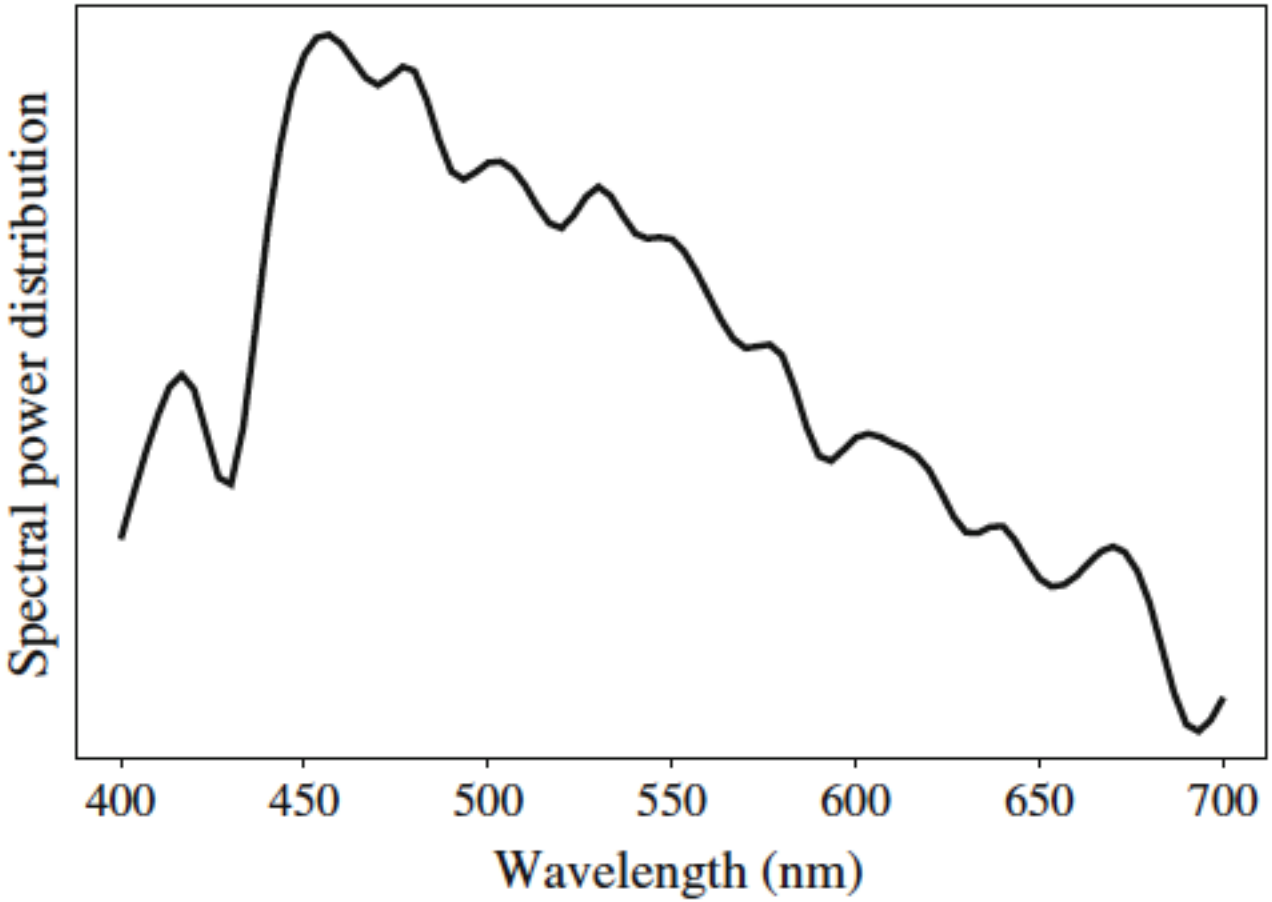
Angelo Ciaramella

Light and spectra

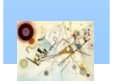
- Light is an electromagnetic wave
  - humans cannot detect all light, but just contributions that fall in the "visible wavelengths"
  - Visible light is an electromagnetic wave in the range 400–700 nm
    - nanometer is 10^{-9} m
- The eye works like a camera
- The retina consists of an array of rods and three kinds of cones
  - The rods come into play when light levels are low and produce a image in shades of gray
  - The three kinds of cones are most sensitive to red (R), green (G), and blue (B) light



Spectral Power Distribution



Spectral power distribution of daylight ($E(\lambda)$ where λ is the wavelength)



Light and spectra

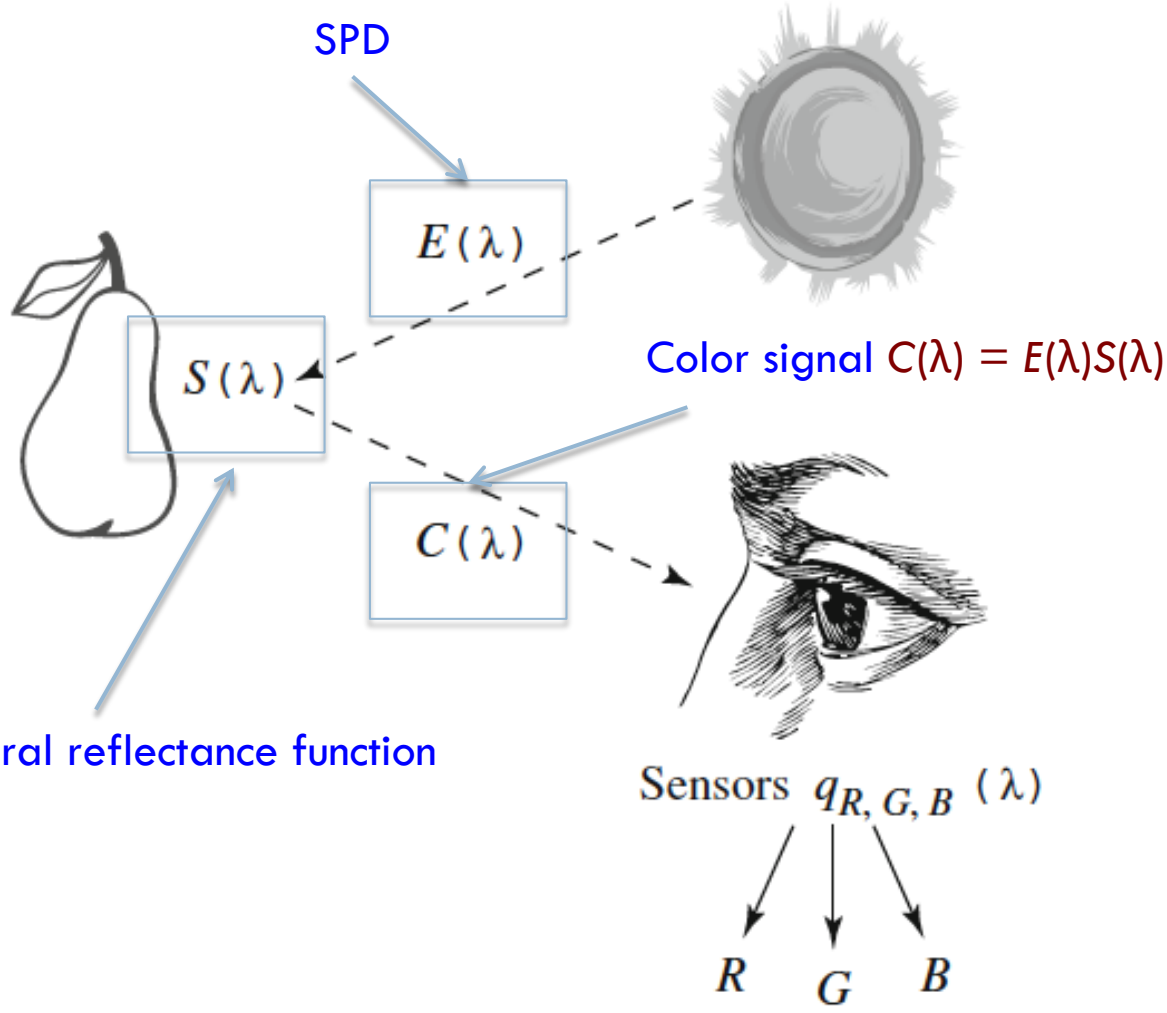


Image formation model

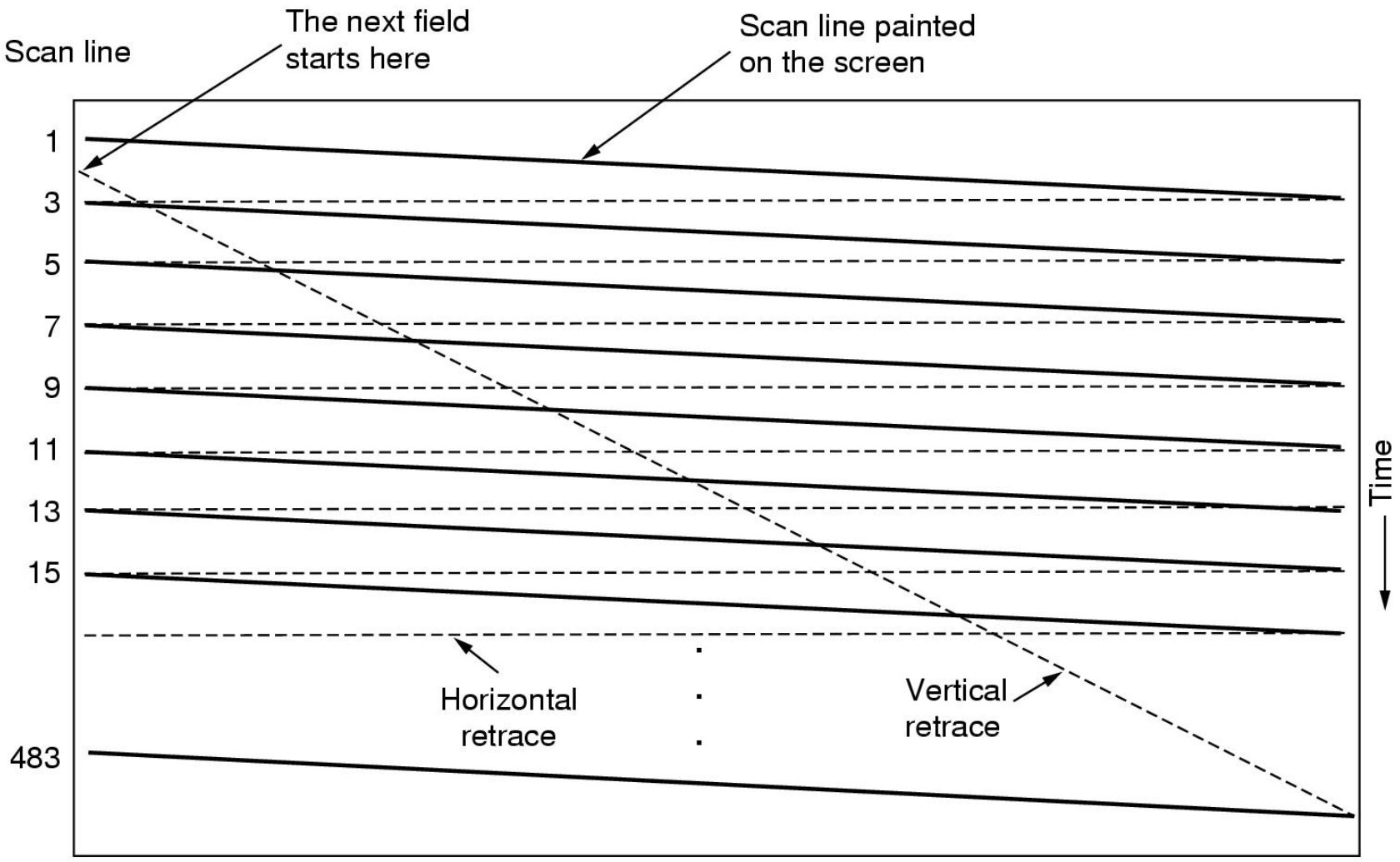
Video encoding

- Up until last decade, most TV programs were sent and received as an analog signal
- The human eye has the property that
  - when an image is flashed on the retina it is retained for some number of milliseconds before decaying
- If a sequence of images is flashed at 50 or more images/sec, the eye does not notice that it is looking at discrete images
  - All video- and film-based motion picture systems exploit this principle to produce moving pictures

Black-and-white television

- It represents two-dimensional images as a one-dimensional voltage as a function of time
  - The camera scans an electron beam rapidly across the image and slowly down it, recording the light intensity as it goes
  - At the end of the scan, called a frame, the beam retraces
- This intensity as a function of time could be broadcast, and receivers repeat the scanning process to reconstruct the image

Black-and-white television



The scanning pattern used for NTSC video and television

Scanning parameters

System	n. lines	n.frames/sec	Dimension ratio
NTSC	525 (485)	30	4:3
PAL	625 (576)	25	4:3
SECAM	625 (576)	25	4:3

Scanning parameters. The top few and bottom few lines are not displayed

Interlacing

- 25 frames/sec is enough to capture smooth motion
  - many people, especially older ones, will perceive the image to *flicker*

- Interlacing
  - First all the odd scan lines are displayed then the even ones
  - Each of these half frames is called a field

Color video

- Same scanning pattern as monochrome
  - three beams moving in unison
  - one beam is used for each of the three additive primary colors: red (R), green (G), and blue (B)
  - any color can be constructed from a linear superposition of red, green, and blue with the appropriate intensities

Digital video

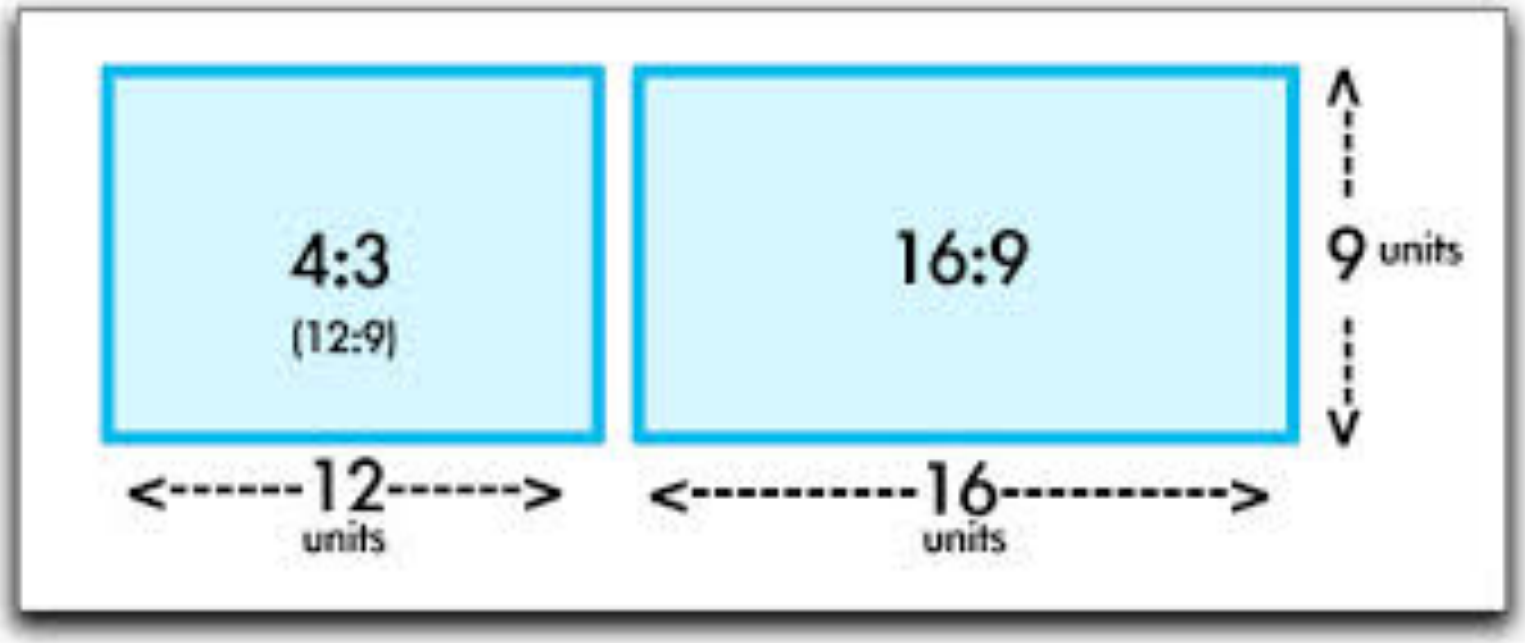
- sequence of frames, each consisting of a rectangular grid of picture elements, or pixels
- 8 bits per pixel are used for each of the RGB colors, giving $2^{24} \approx 16$ million colors
- To produce smooth motion
  - at least 25 frames/sec
  - interlacing is not needed
    - rescan the screen from images stored in video RAM at 75 times per second or more

Digital video

- *smoothness* of motion is determined by the number of different images per second
- *flicker* is determined by the number of times the screen is painted per second
- The significance of these two parameters becomes clear when we consider the bandwidth required for transmitting digital video over a network

Digital video

- Many computer monitors use the 4:3 aspect ratio (i.e., the ratio of picture width to height)



Aspect ratio

Digital video

- Common configurations are for example
  - 640 x 480 (VGA)
  - 800 x 600 (SVGA)
  - 1024 x 768 (XGA)
  - 1600 x 1200 (UXGA)

- interlacing is needed for analogic broadcast television
  - analog signals cannot be stored in RAM without first converting them to digital form

NTSC video

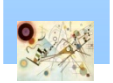
- NTSC TV standard is mostly used in North America and Japan
  - 4:3 aspect ratio and 525 scan lines per frame at 30 fps
  - it follows the interlaced scanning system
    - each frame is divided into two fields with 262.5 lines/field
  - blanking information is placed into 20 lines reserved for control information at the beginning of each field
    - the number of active video lines per frame is only 485
  - NTSC uses the YIQ color model
    - 4.2MHz band of Y is overlapped and interleaved with the 1.6MHz to I and 0.6MHz to Q

PAL video

- PAL (Phase Alternating Line) is a TV standard originally invented by German scientists
  - 625 scan lines per frame, at 25 fps (or 40 ms/frame), with a 4:3 aspect ratio and interlaced fields
  - It uses the YUV color model
    - 8MHz channel
    - allocating a bandwidth of 5.5MHz to Y and 1.8MHz each to U and V

SECAM video

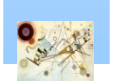
- Electronique Couleur Avec Memoire
  - invented by the French, is the third major broadcast TV standard
  - 625 scan lines per frame, at 25 fps, with a 4:3 aspect ratio and interlaced fields
  - SECAM and PAL are similar
    - In SECAM U and V signals are modulated using separate color subcarriers at 4.25MHz and 4.41MHz, respectively



Digital video

- The advantages of digital representation for video are many
  - Storing video on digital devices or in memory, ready to be processed (noise removal, cut and paste, and so on) and integrated into various multimedia applications
  - Direct access, which makes nonlinear video editing simple
  - Repeated recording without degradation of image quality
  - Ease of encryption and better tolerance to channel noise

- Modern digital video generally uses component video
  - RGB signals are first converted into a certain type of color opponent space
  - The usual color space is YCbCr

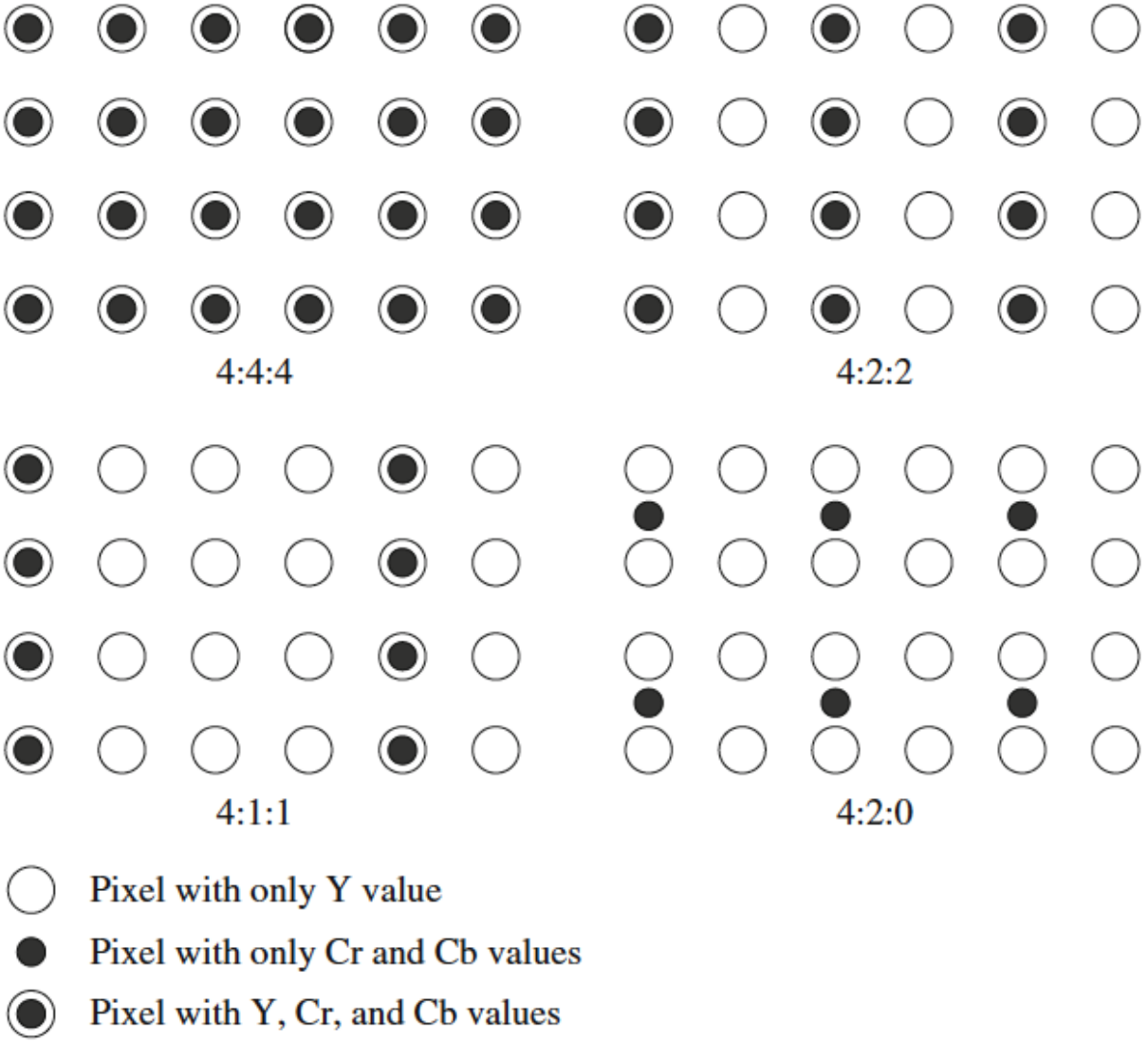


Chroma Subsampling

- Humans see color with much less spatial resolution than black and white
  - it makes sense to decimate the chrominance signal

- chroma subsampling scheme
  - 4:4:4 indicates that no chroma subsampling is used
  - 4:2:2 indicates horizontal subsampling of the Cb and Cr signals by a factor of 2
  - 4:1:1 subsamples horizontally by a factor of 4
  - 4:2:0 is commonly used in JPEG and MPEG
    - subsamples in both the horizontal and vertical dimensions by a factor of 2

Chroma Subsampling



Chroma subsampling

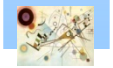
ITU-R digital video specifications

Common Intermediate Format

Quarter CIF

	Rec. 601 525/60 NTSC	Rec. 601 625/50 PAL/SECAM	CIF	QCIF
Luminance resolution	720 × 480	720 × 576	352 × 288	176 × 144
Chrominance resolution	360 × 480	360 × 576	176 × 144	88 × 72
Color subsampling	4:2:2	4:2:2	4:2:0	4:2:0
Aspect ratio	4:3	4:3	4:3	4:3
Fields/sec	60	50	30	30
Interlaced	Yes	Yes	No	No

ITU-R digital video specifications



Ultra and High-Definition TV

- High-Definition TV (HDTV)
  - increase the visual field especially its width
  - much wider aspect ratio of 16:9 instead of 4:3
  - its move toward progressive (non-interlaced) scan
- Ultra High Definition TV (UHDTV)
  - provide superior picture quality but it will require a much higher bandwidth and/or bitrate
  - the aspect ratio is 16:9