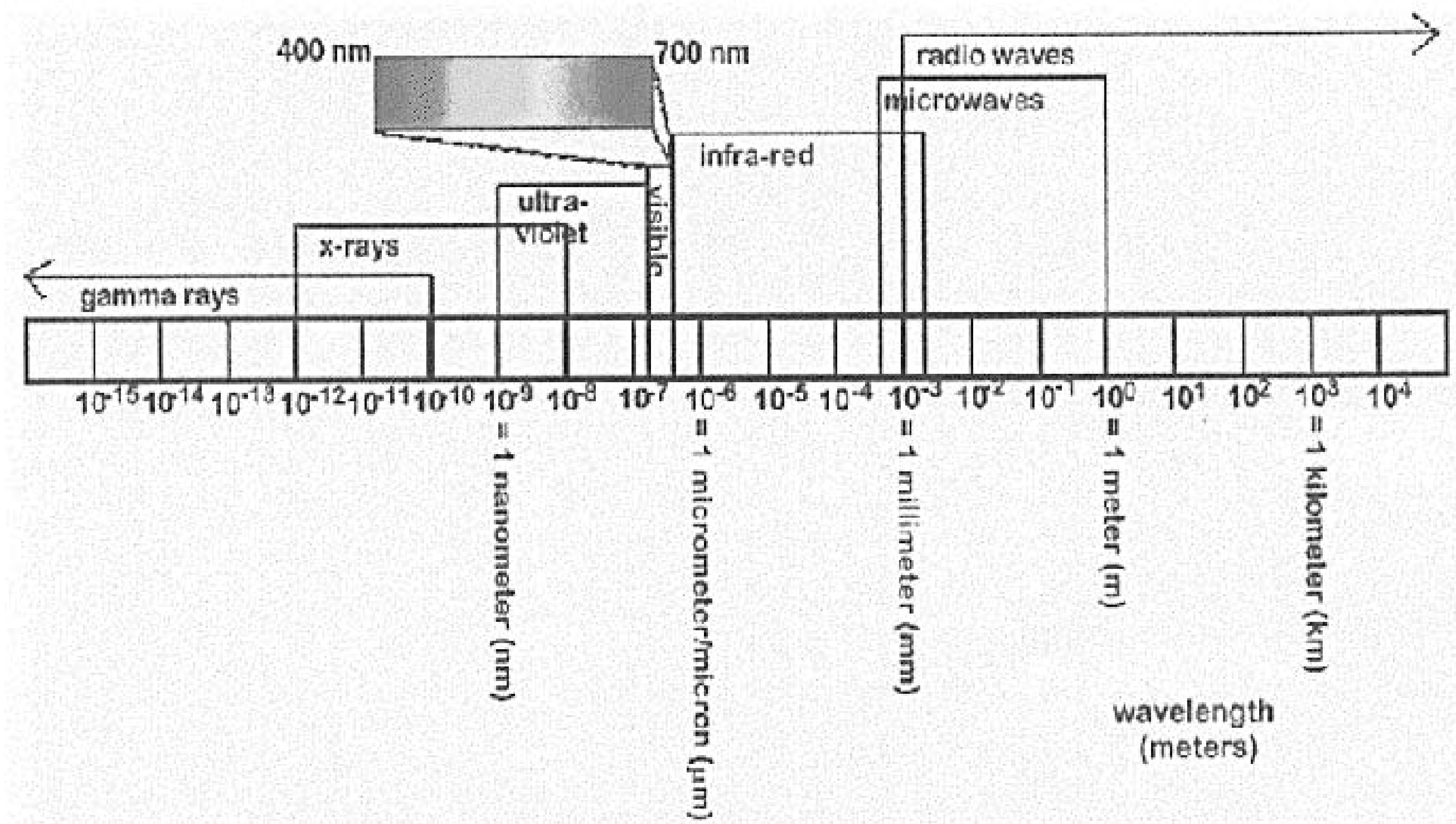
LECTURE 1 - Part 1
BASICS of ANALOG and DIGITAL VIDEO

- Light, color perception
- Human visual system
- Analog Video
- Digital Video
- Digital Video Standards

Light



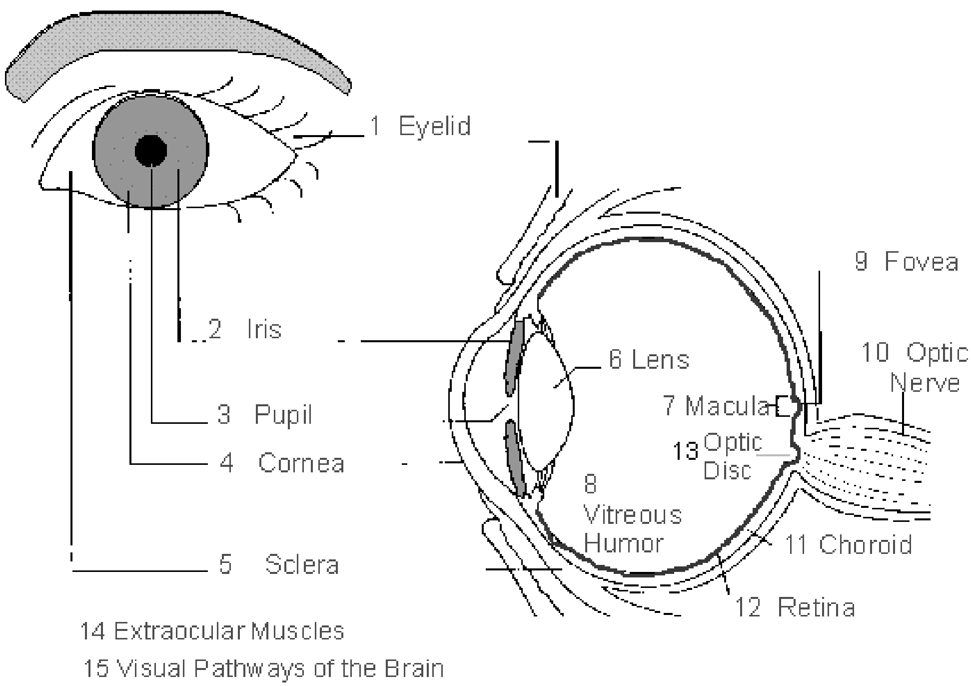
400 nm
700 nm
radio waves
microwaves
infra-red
ultra-violet
visible
x-rays
gamma rays
10^{-15} 10^{-14} 10^{-13} 10^{-12} 10^{-11} 10^{-10} 10^{-9} 10^{-8} 10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4}
= 1 nanometer (nm)
= 1 micrometer/micron (µm)
= 1 millimeter (mm)
= 1 meter (m)
= 1 kilometer (km)
wavelength
(meters)

Illuminating and reflecting sources

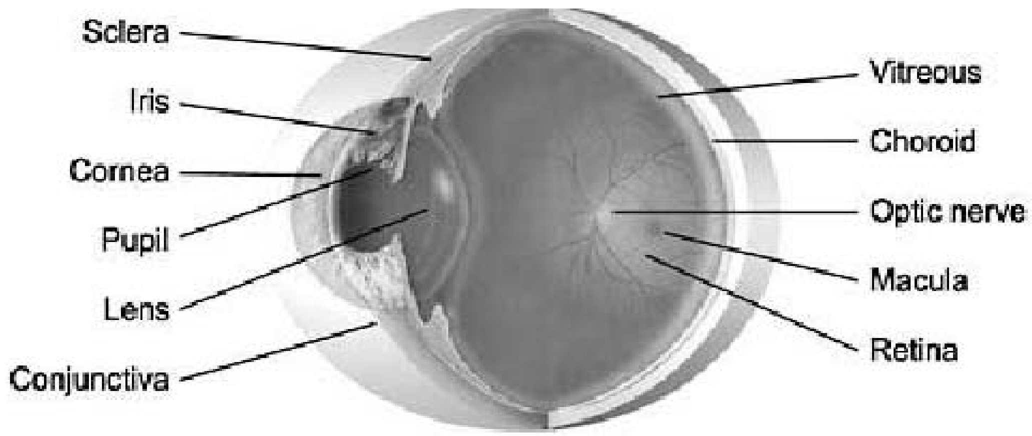
- Illuminating sources:
  - emit light (e.g. the sun, light bulb, TV monitors)
  - perceived color depends on the emitted freq.
  - follows additive rule (http://javaboutique.internet.com/ColorFinder/)
- Reflecting sources:
  - reflect an incoming light (e.g. the color dye, matte surface, cloth)
  - perceived color depends on reflected freq (=emitted freq-absorbed freq.)
  - follows subtractive rule

Vision: The eye

The eye can be viewed as a dynamic, biological camera: it has a lens, a focal length, and an equivalent of film.



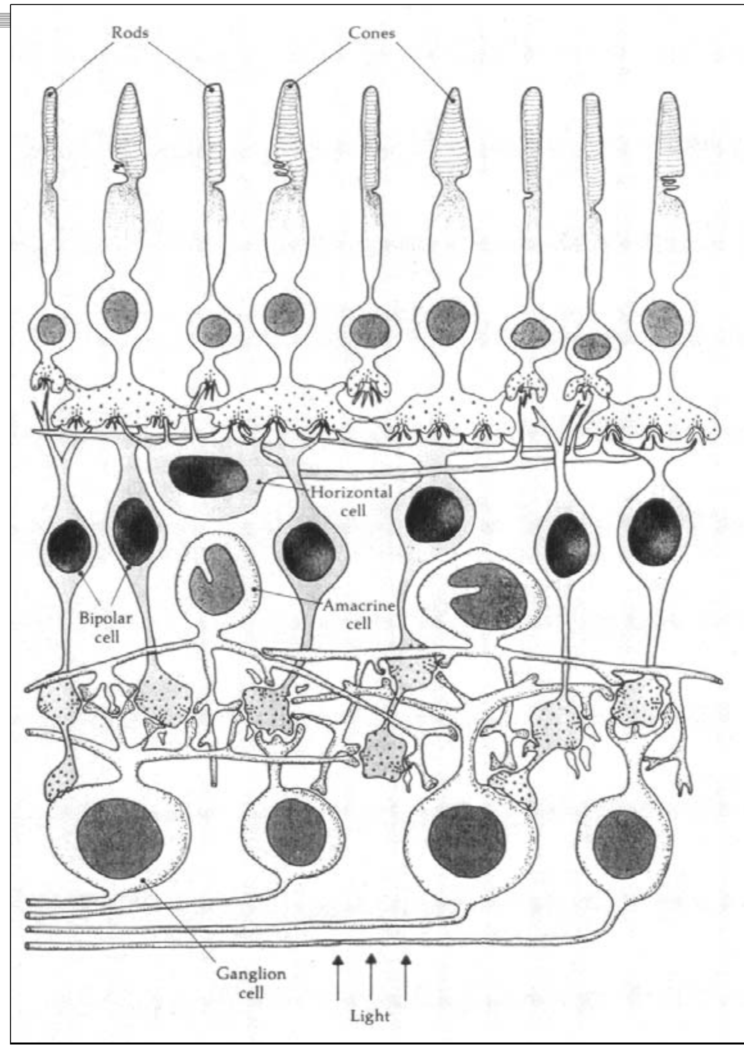
Eye vs camera



Camera components	Eye components
Lens	Lens, cornea
Shutter	Iris, pupil
Film	Retina
Cable to transfer images	Optic nerve send the info to the brain

- The working principal of eye is very similar to that of a pinhole camera
  - The iris, acts the same way as a diaphragm by controlling the size of the pupil and adjusts the quantity of light that enters into the eye.
  - Behind the iris we find the crystalline lens which exactly acts like a camera's lens.
  - The retina functions as the eye's "film". It is covered with cells sensitive to light. These cells turn the light into electrochemical impulses that are sent to the brain. There are two types of cells, rods and cones.

The Anatomy of the Human Visual System

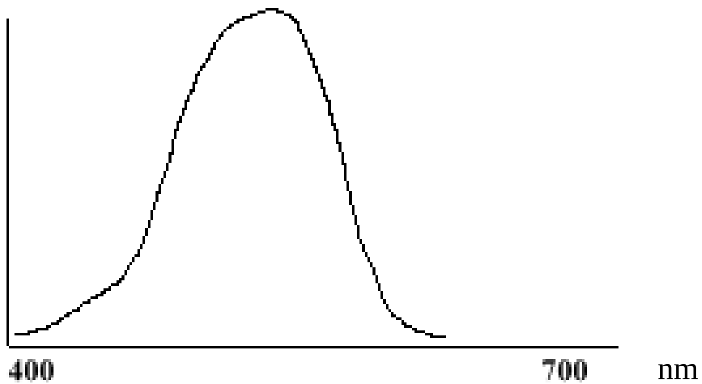


Retina

- The image on the retina is upside down and switched left/right
- Three major layers
  - Photoreceptors, Bipolar cell layer, Ganglion cell layer
- Photoreceptors
  - Rods and cones
  - All photoreceptors contain photosensitive pigments
    - All rods have a single photo pigment
    - There are 3 types of cones, each with different photo pigment
    - Light that hits the rod/cone bleaches the photo pigment
  - There are more rods than cones

Rods

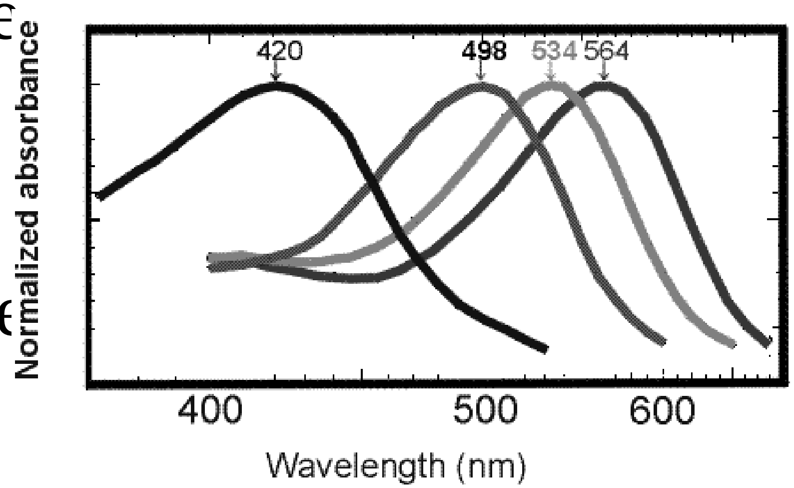
- Sensitive to most visible frequencies (brightness) (maximum sensitivity at 500nm)
- About 120 million in eye
- Most located outside of fovea, or center of retina.
- Used in low light (theate night) environments, result in achromatic (b&w) vision. – percieve brightness only



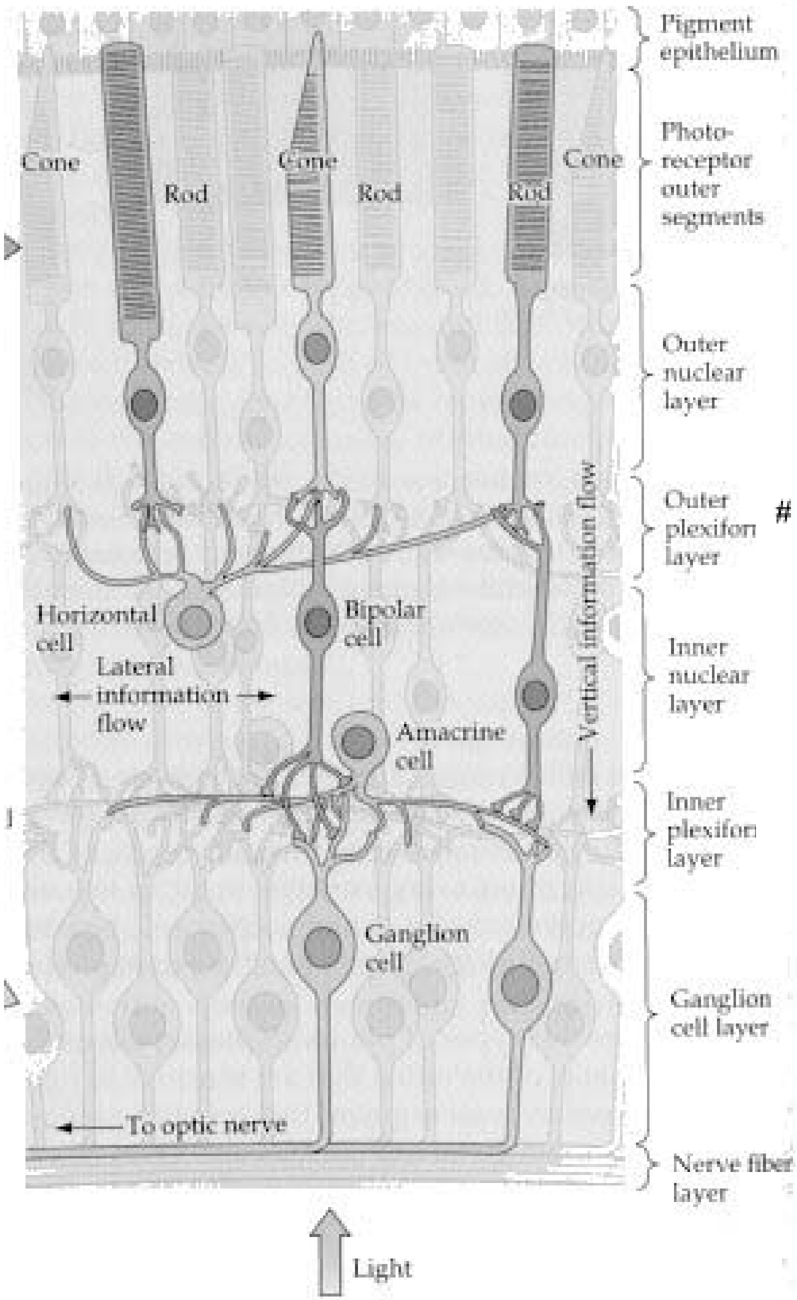
rod absorption function

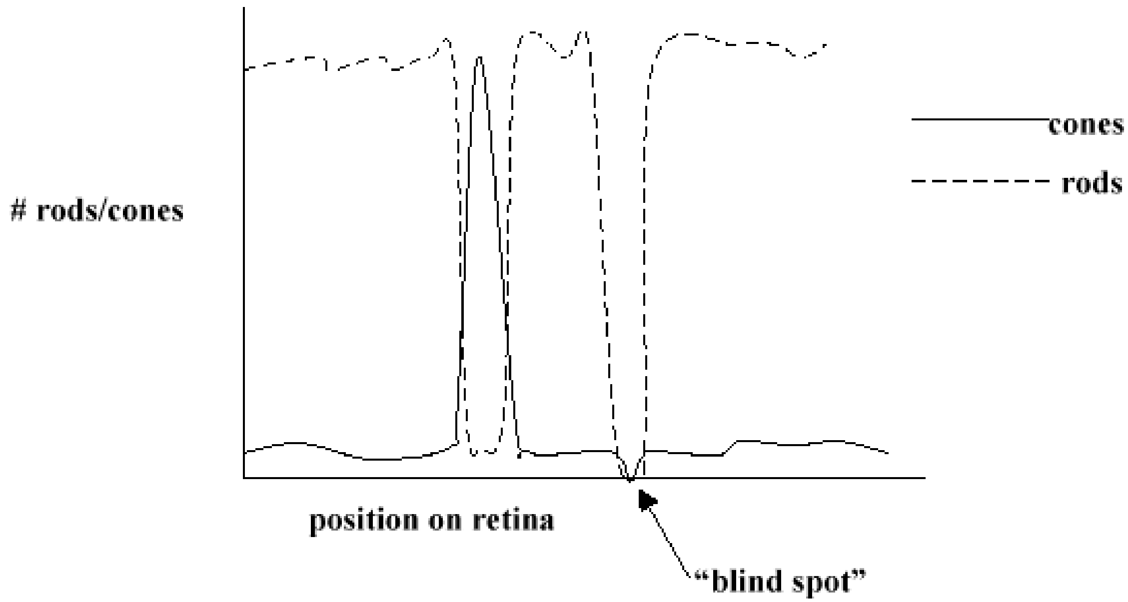
Cones

- R cones are sensitive to long wavelengths, G to middle and B to short.
- About 8 million in eye.
- Highly concentrated in fove with B cones more evenly distributed than the others.
- Used for high detail color vision.



After Bowmaker & Dartnall, 1980

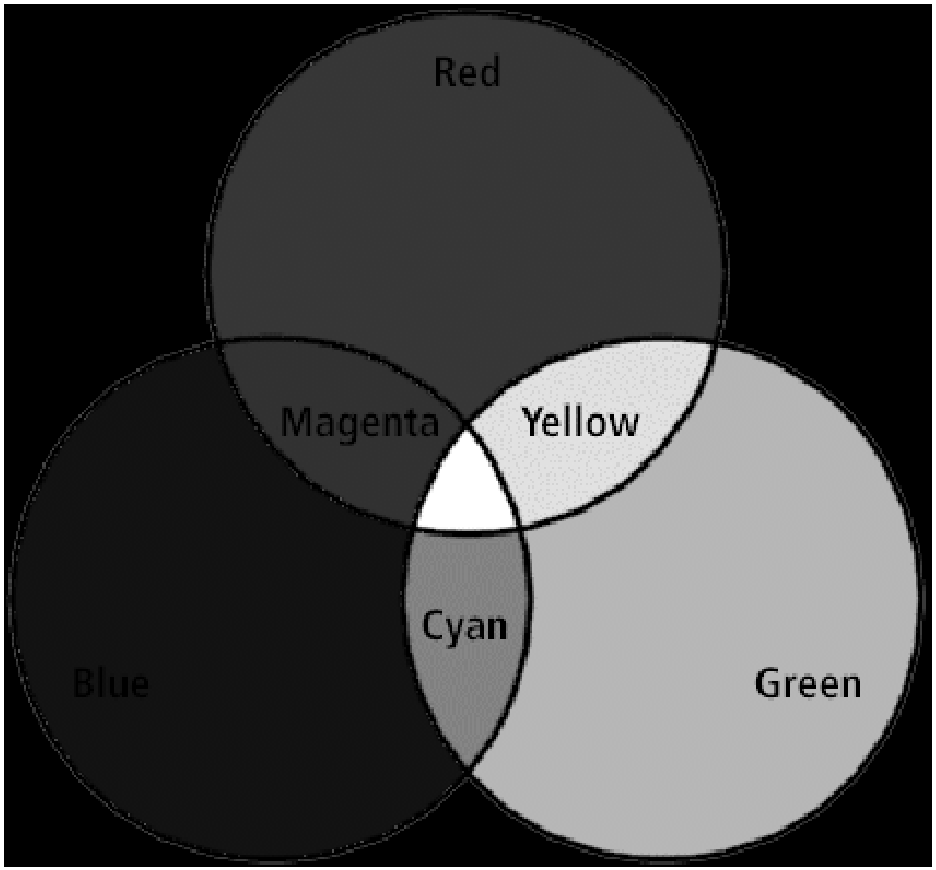


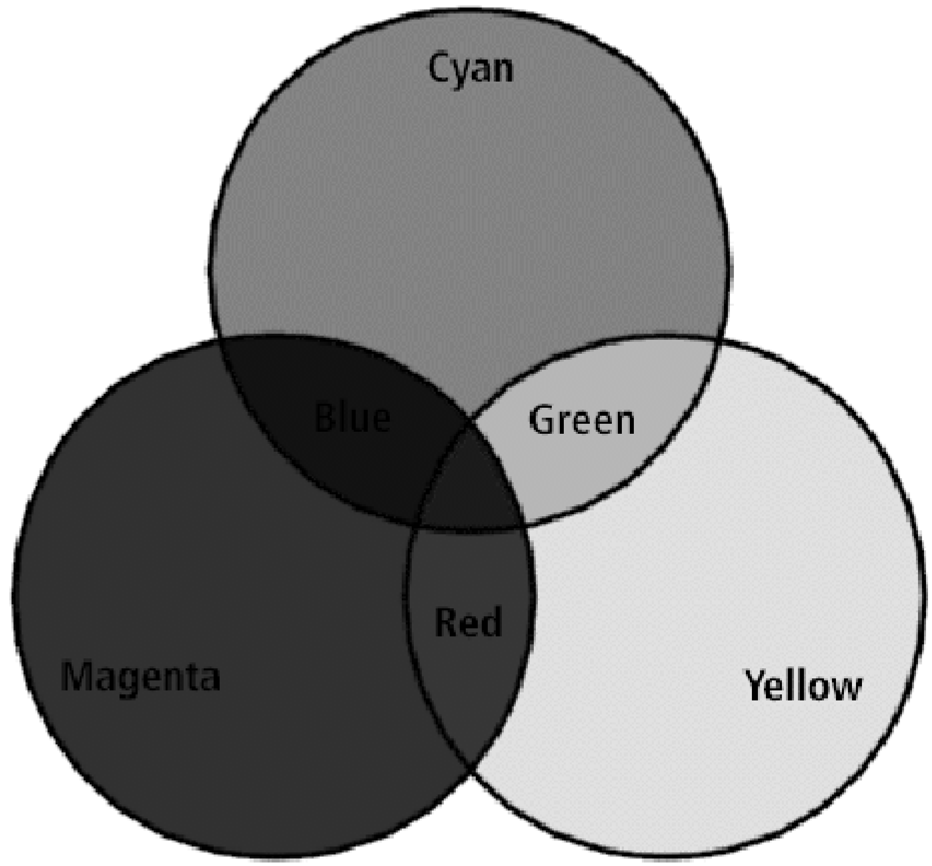


ROD/CONE DISTRIBUTION

Trichromatic Theory of Color Vision

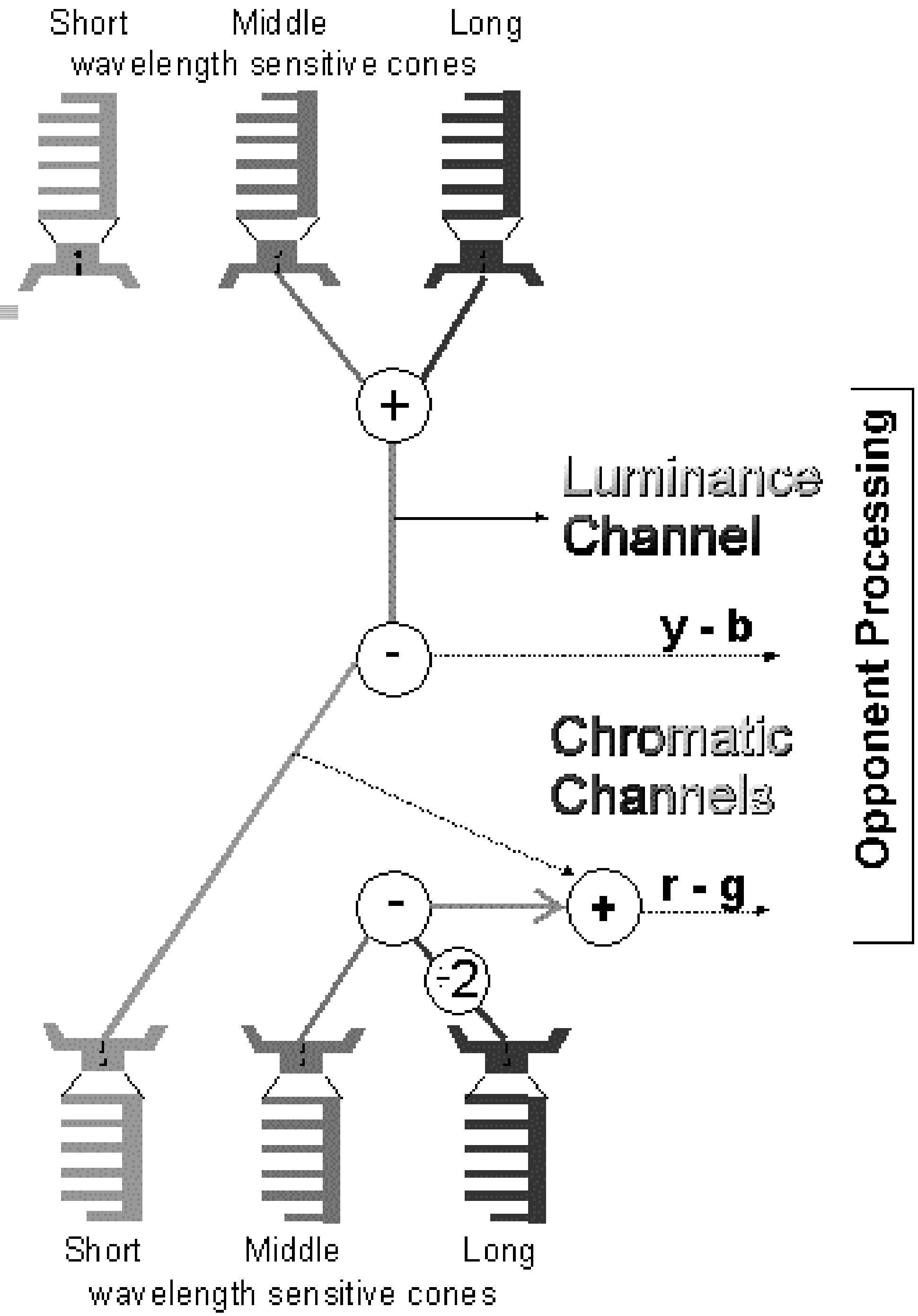
- It is possible to match all of the colors in the visible spectrum by appropriate mixing of three primary colors.
- Primary colors for illuminating sources:
  - Red, Green, Blue (RGB)
- Primary colors for reflecting sources:
  - Cyan, Magenta, Yellow (CMY)


Red
Magenta
Yellow
Cyan
Blue
Green


Cyan
Blue
Green
Red
Magenta
Yellow

Color Opponency

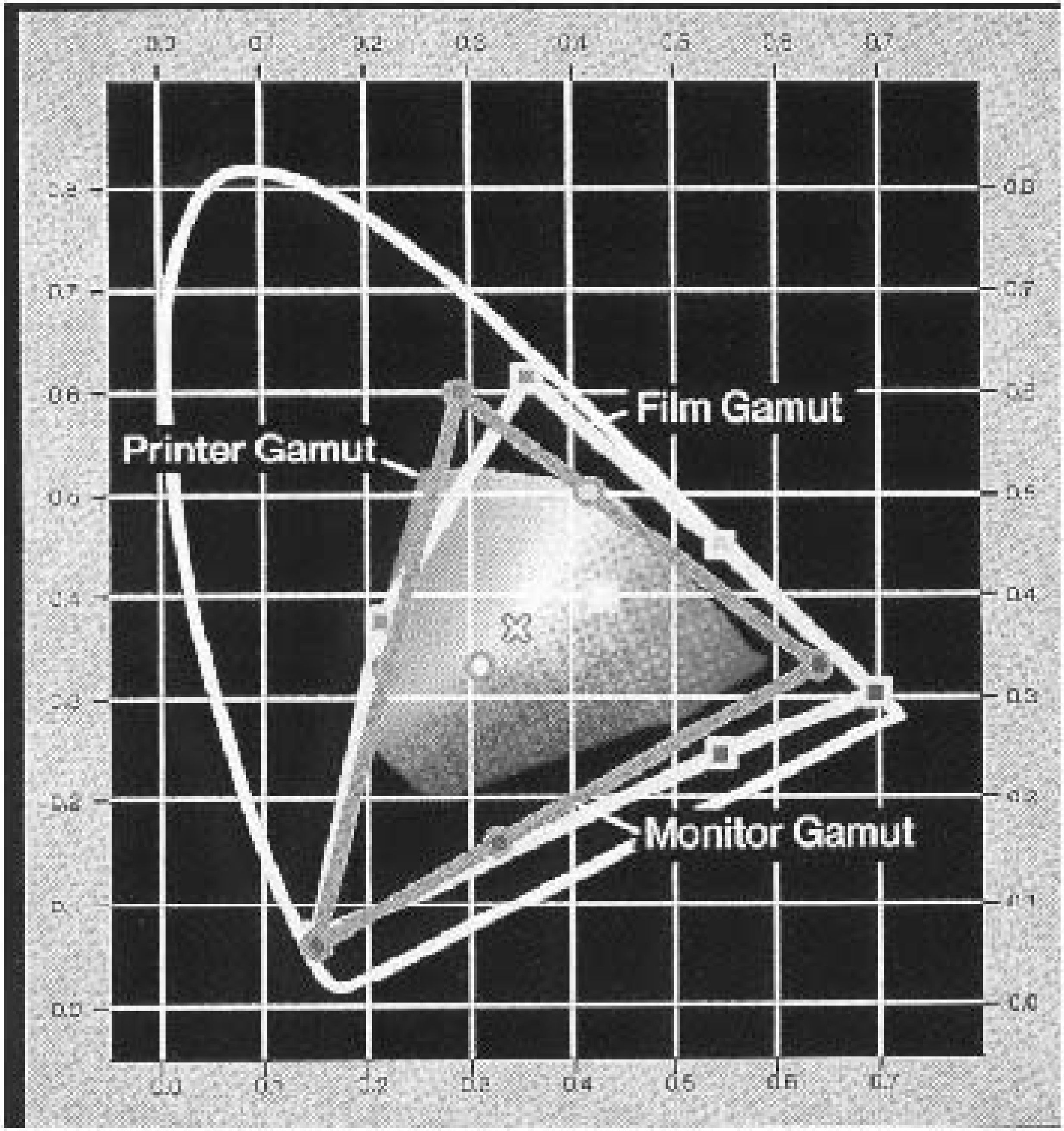
- After color is turned into three signals by the R, G and B cones, it is turned into three parallel signals:
- Achromatic: R + G
- Blue-yellow: R + G - B
- Red-green: B + R - G
- The blue/yellow and red/green pairs are called complementary colors. Mixing the proper shades of them in the proper amounts produces white light.



Short
Middle
Long
wavelength sensitive cones
+
Luminance
Channel
-
y - b
Chromatic
Channels
-
+
r - g
÷2
Opponent Processing
Short
Middle
Long
wavelength sensitive cones

CIE Chromaticity diagram

- Used to describe color as accurately as possible
- Make use of the fact that colors can be described by combinations of three basic colors, called primary colors.
- Color perceptions are measured by giving subjects various combinations of 3 standard CIE primary colors and measuring their perception. These perceptions are plotted against an x-y diagram.



Film Gamut
Printer Gamut
Monitor Gamut

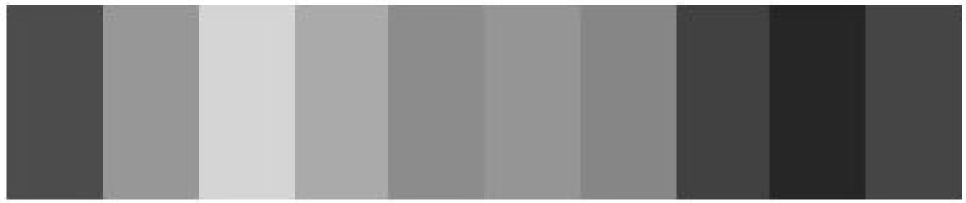
Color representation models

- Specify the tristimulus values associated with the three primary colors
  - RGB
  - CMY
- Specify the luminance and chrominance
  - XYZ
  - HSI (Hue, saturation, intensity)
  - YIQ (used in NTSC color TV), YCbCr (digital color TV)
- Amplitude specification:
  - 8 bits for each color component, or 24 bits total for each pel
  - Total of 16 million colors
  - A true RGB color display of size 1Kx1K requires a display buffer memory size of 3 MB

Psychological Models: HSV

- Not easy to describe a color in R,G,B.
- H,S,V is more convenient since it models how human being visualizes color.
  - Hue: Specific individual pure color (described by peak wavelength).
  - Value: How light or dark a color is related to gray scale (overall strength of light, intensity of light reflected from or transmitted by a color image.)
  - Saturation: How pure the color is (ratio of dominant wavelength to others)(strength of color).

Hue Changes

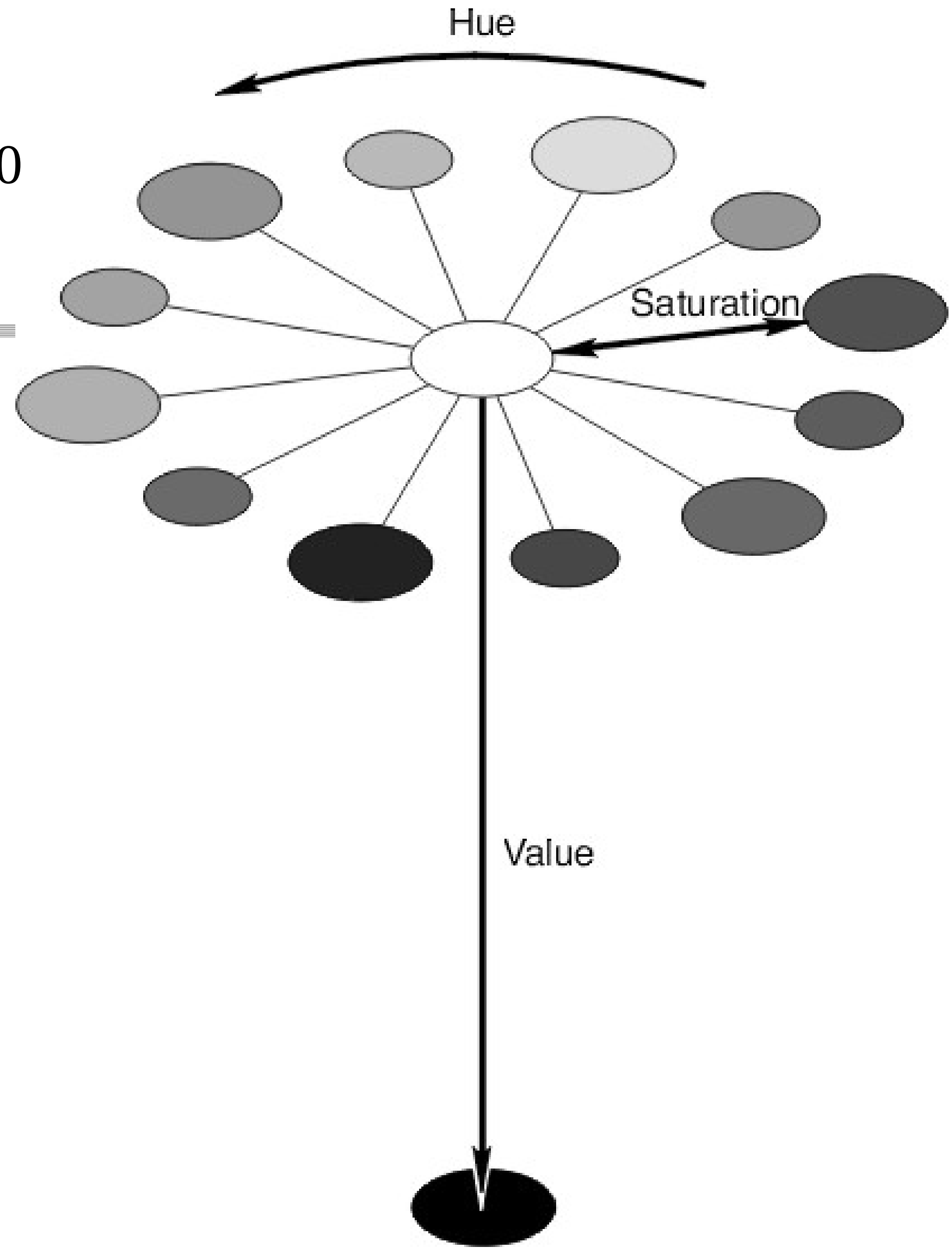


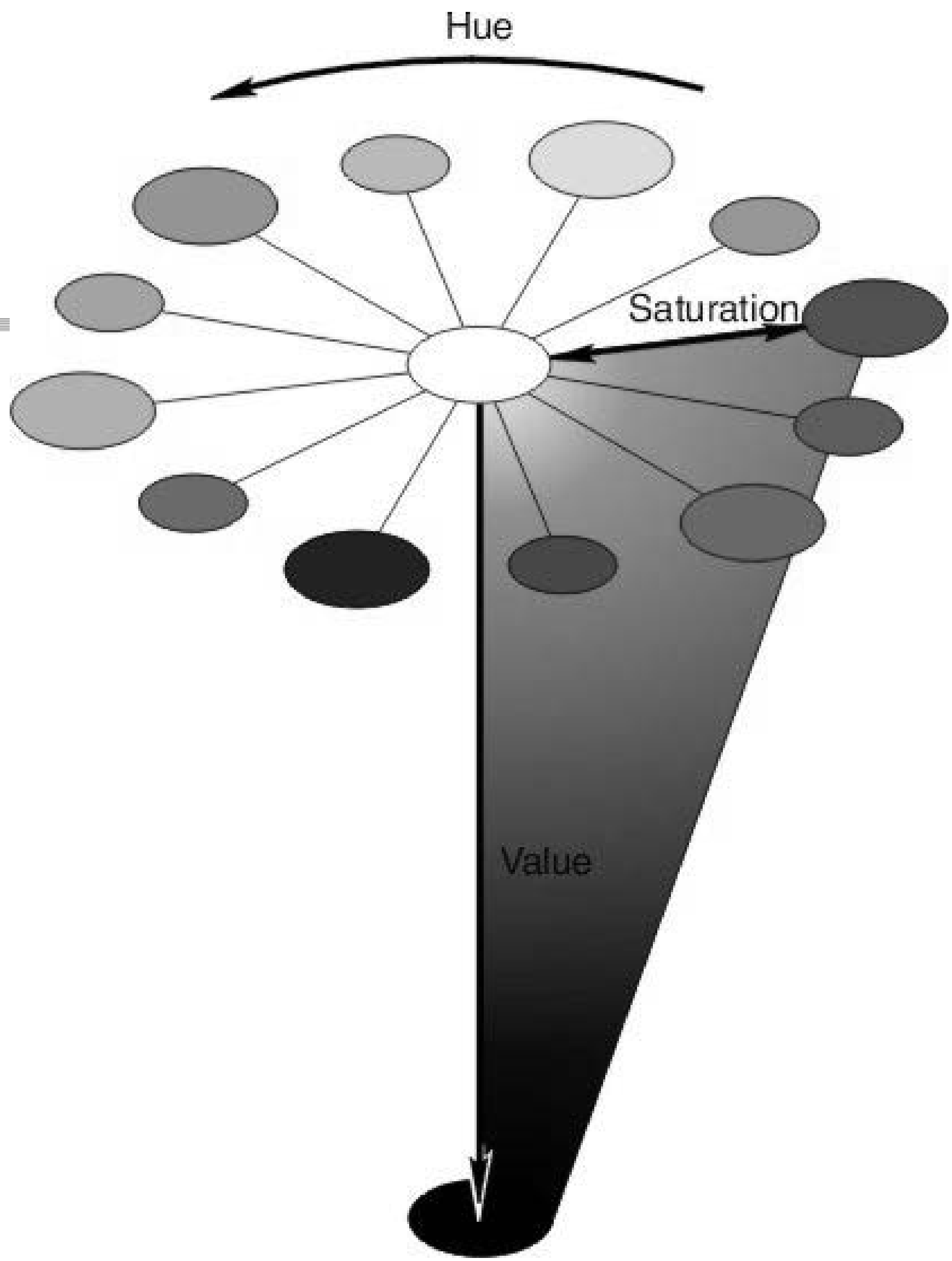
Saturation Changes

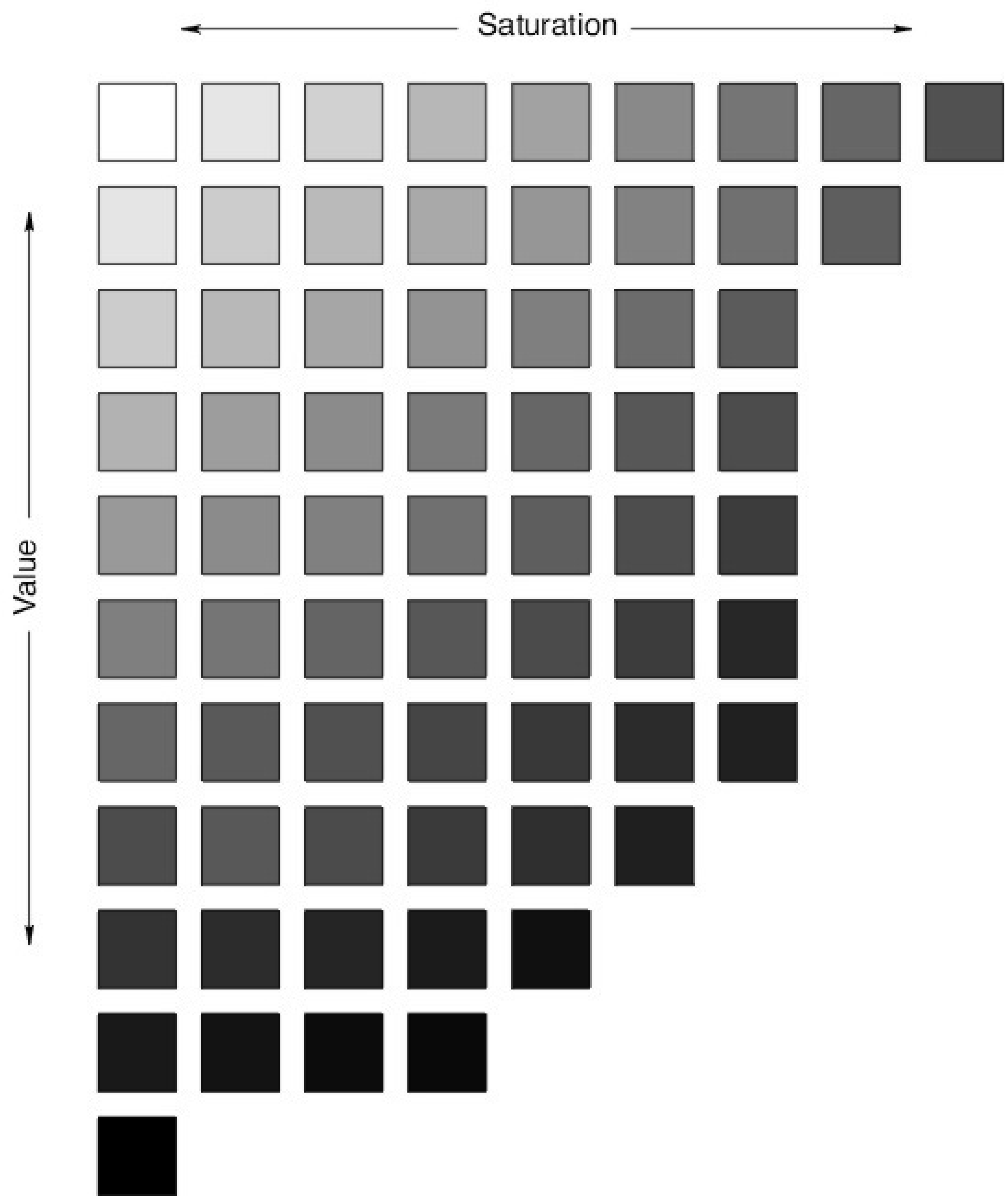
Brightness Changes

Hue varies between 0 and 360 degrees, saturation and value between 0 and 1

The eye can see about 128 different hues, and about 130 different saturations. The number of values varies between 16 (blue) and 23 (yellow)



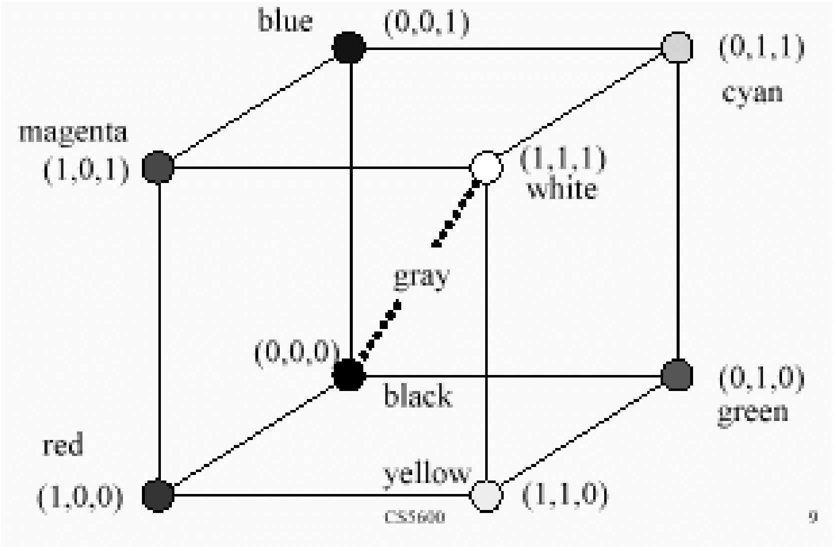

Hue
Saturation
Value



Demo

Engineering Models: RGB

- The RGB stands for red, green and blue
- All parameters vary between 0 and 1
- Hue is defined by the one or two largest parameters
- Saturation can be controlled by varying the collective minumum value of R, G and B
- Luminance can be controlled by varying magnitudes while keeping ratios constant



Engineering Models: CMY

- The CMY stands for cyan (blue + green), magenta (red + blue), and yellow (red + green)
- All parameters vary between 0 and 1
- [CMY] = [WWW] - [RGB]
- Used for reflective media, like printing
- Since reflective media absorb light, this model does also. It is subtractive. Thus red is M + Y: M subtracts green, Y subtracts blue

Engineering Models: CMYK and YIQ

- CMYK is a variant of CMY that enhances darkness by setting K = min(C,M,Y), and printing black with intensity K.
- YIQ is the model used for broadcast television. This model can be encoded by black and white or color TV sets. Y contains the luminance values and is the only component used in a black and white TV.

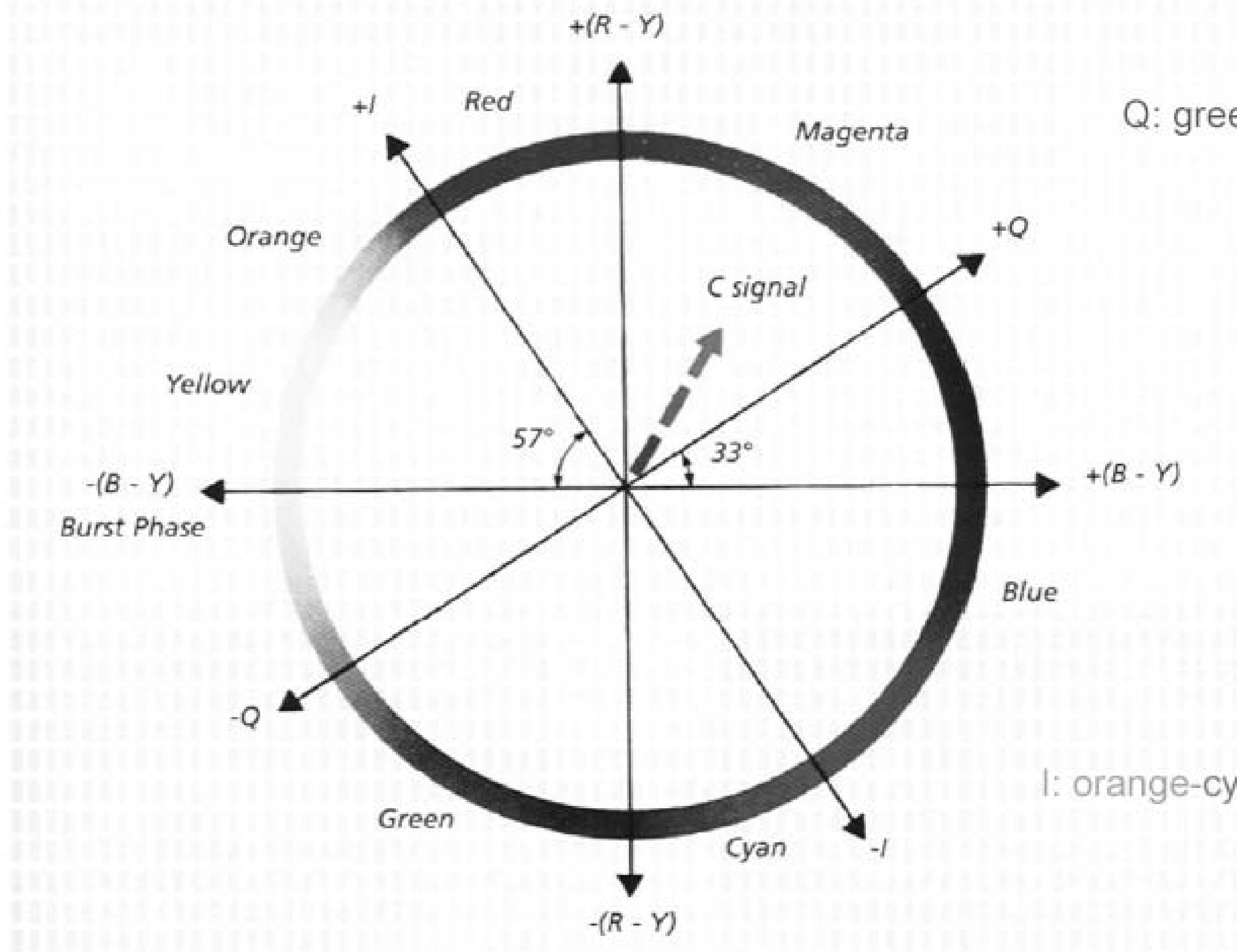
Why not RGB?

- R,G,B components are correlated
  - Transmitting R,G,B components separately is redundant
- RGB->YC1C2 transformation
  - Decorrelating: Y,C1,C2 are uncorrelated
  - C1 and C2 require lower bandwidth
  - Y (luminance) component can be received by B/W TV sets
- YIQ in NTSC
- YUV in PAL
- YCbCr in digital domain

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- U α (B-Y)
- V α (R-Y)
- I, Q rotated versions of U, V
- I: orange-to-cyan
- Q: green-to-purple (human eye is less sensitive)
- Phase=Arctan(Q/I) = hue
- Magnitude=sqrt (I^2+Q^2) = saturation



Q: green-purple

I: orange-cyan

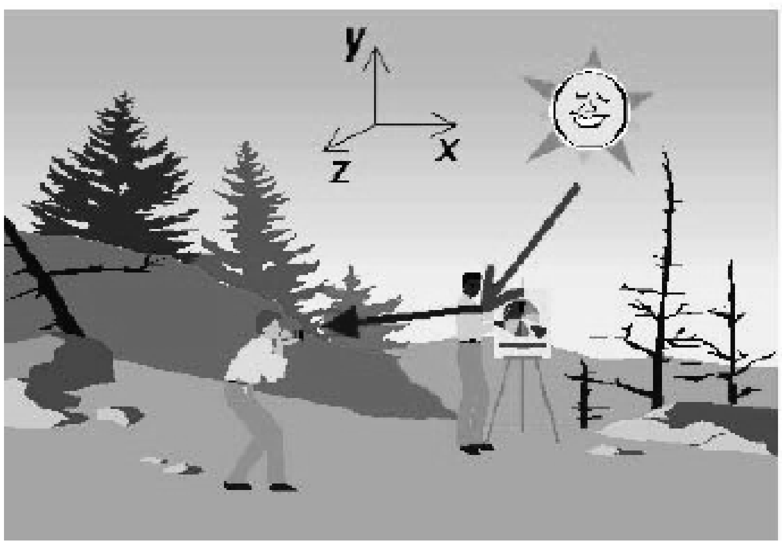
Persistance of vision

- Human eye retains image for a fraction of a second after it views it.
- Motion pictures : 24frames/sec
- PAL : 25 frames/sec
- NTSC : 30 frames/sec

- To increase frame rate: Rotating shutters
- Scan

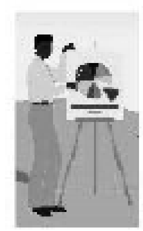
Video capture

- For natural images we need a light source *(λ: wavelength of the source)* :
  - $E(x, y, z, \lambda)$: incident light on a point *(x, y, z world coordinates of the point)*
- Each point in the scene has a reflectivity function.
  - $r(x, y, z, \lambda)$: reflectivity function
- Light reflects from a point and the reflected light is captured by an imaging device.
  - $c(x, y, z, \lambda) = E(x, y, z, \lambda) \times r(x, y, z, \lambda)$: reflected light.



$\rightarrow$ $E(x, y, z, \lambda)$

$\rightarrow$ $c(x, y, z, \lambda) = E(x, y, z, \lambda) \cdot r(x, y, z$

Camera$(c(x, y, z, \lambda))$ **=** 

- Reflected light to camera
  - Camera absorption function

$$\overline{\psi}(\mathbf{X},t) = \int C(\mathbf{X},t,\lambda)a_c(\lambda)d\lambda$$

  - Projection from 3-D to 2-D

$$\mathbf{X} \underset{P}{\rightarrow} \mathbf{x}$$

$$\psi(P(\mathbf{X}),t) = \overline{\psi}(\mathbf{X},t) \quad \text{or} \quad \psi(\mathbf{x},t) = \overline{\psi}(P^{-1}(\mathbf{x}),t)$$

    - The projection operator is non-linear
      - Perspective projection
      - Othographic projection

ANALOG VIDEO

1-D analog video signal f(t) contains both timing and intensity information

It is obtained by sampling a 3-D signal

$s(x_1, x_2, t)$ in x_2 and t dimensions.

Cameras and displays

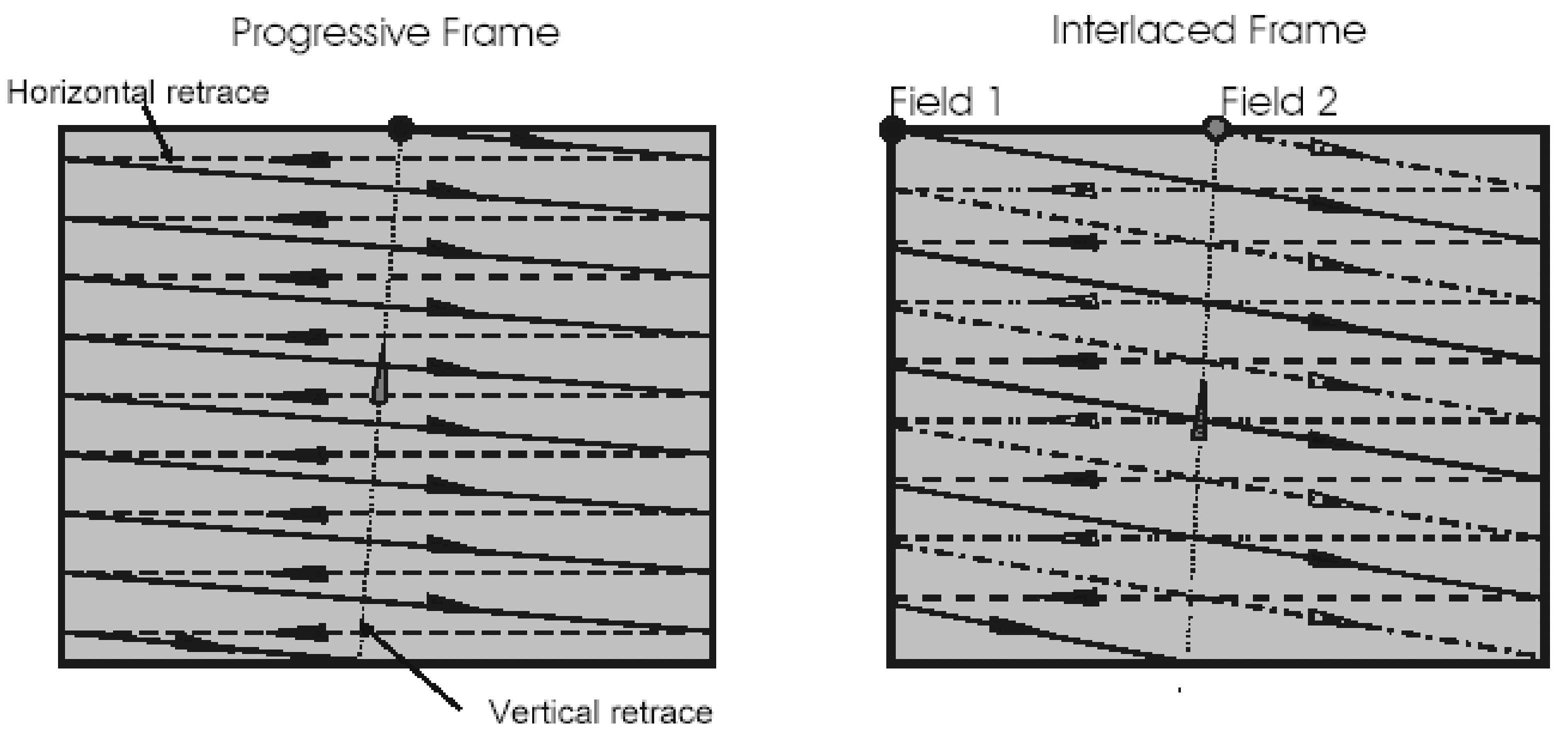
- Tube-based
- CCD
- CRT
- LCD – TFT
- Plazma

Composite vs component video

- Component analog video
  - RGB
  - YCrCb (YIQ or YUV)
- Composite Video
- S-Video

Analog video raster

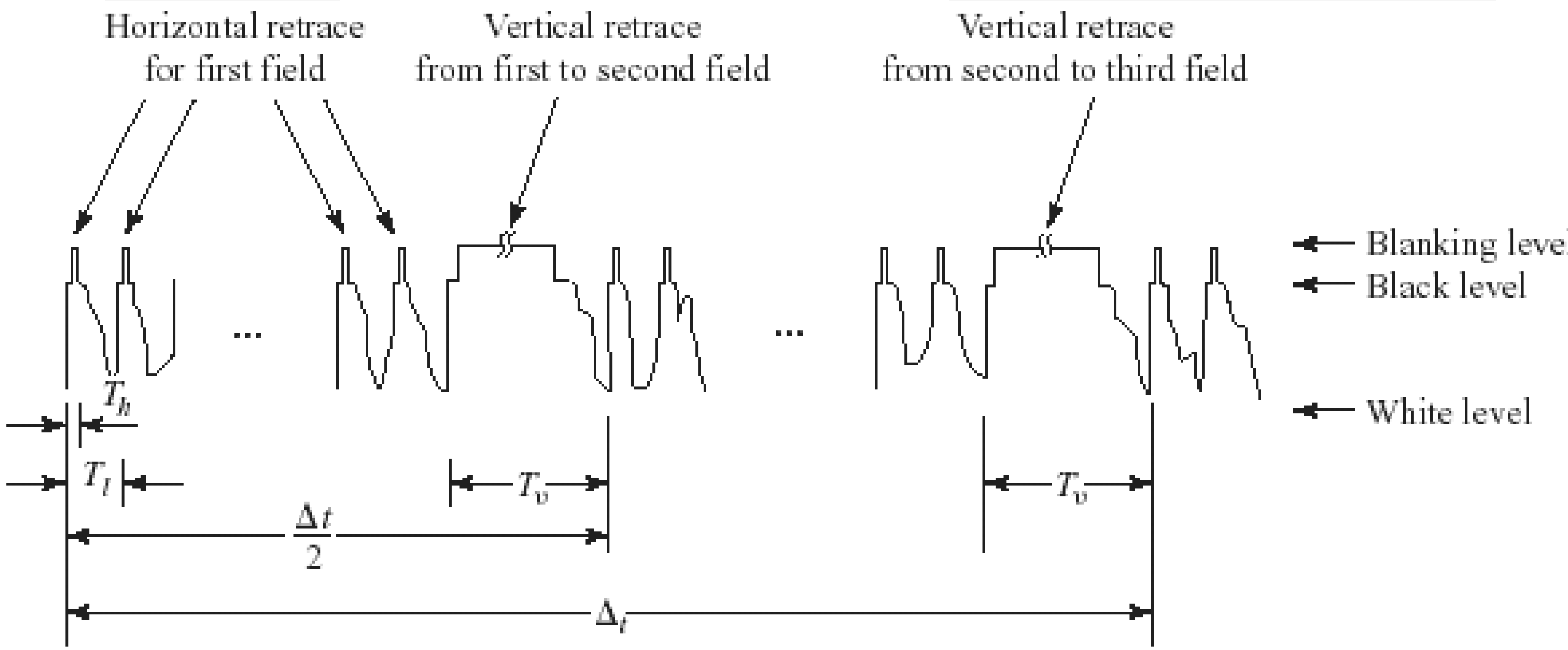
- Analog video is stored in the raster format
  - Sampling in time: consecutive sets of frames
  - To render motion properly, >=30 frame/s is needed
  - Sampling in vertical direction: a frame is represented by a set of scan lines
  - Number of lines depends on maximum vertical frequency and viewing distance, 525 lines in the NTSC system
- Video-raster = 1-D signal consisting of scan lines from successive frames


Progressive Frame
Horizontal retrace
Vertical retrace
Interlaced Frame
Field 1
Field 2

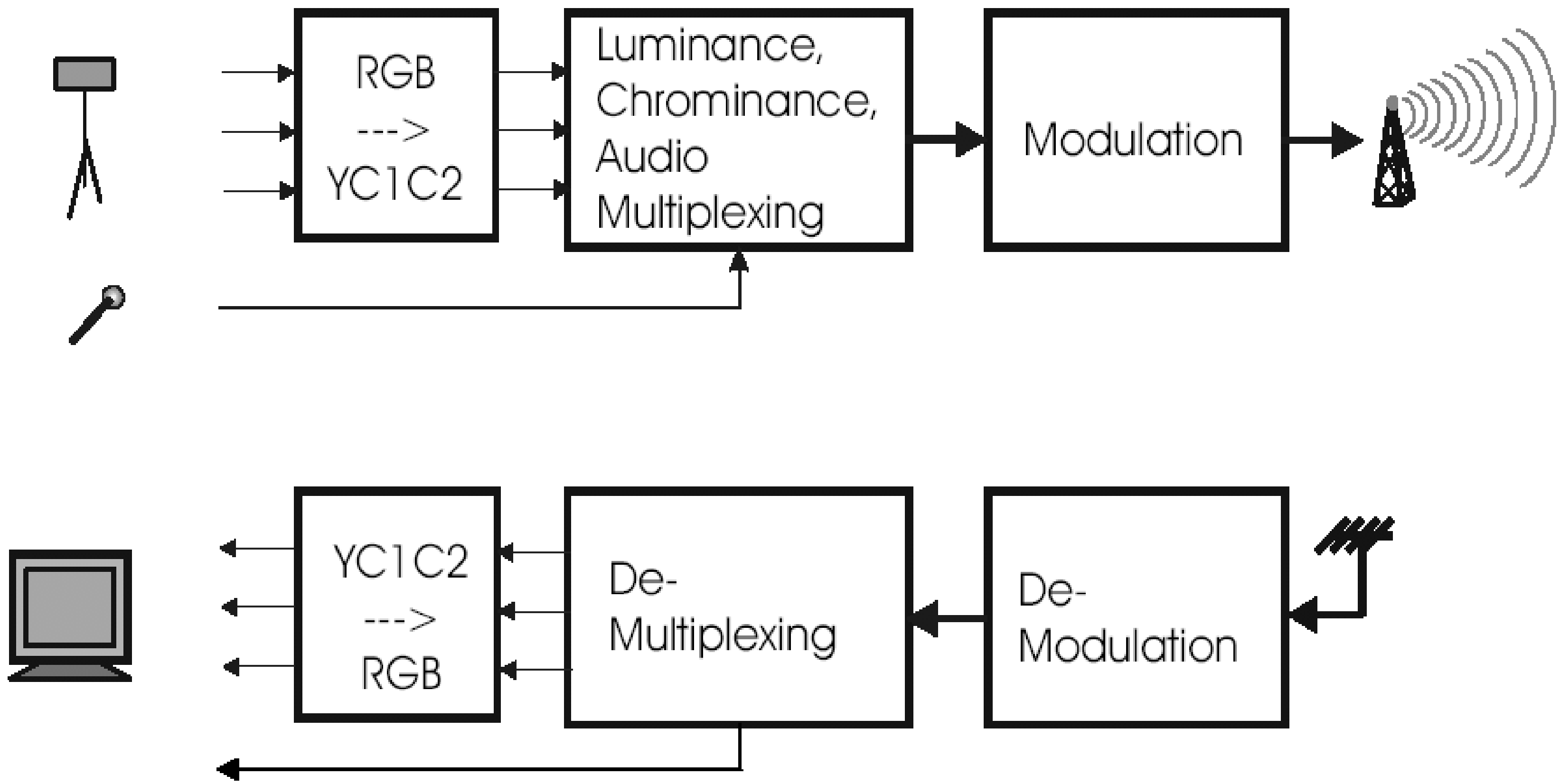
Characterization of video raster

- Frame rate (fps, Hz): $f_{s,t}$
- Line number (lines/frame): $f_{s,y}$
- Line rate (lines/sec): $f_l =$
- Frame interval: $\Delta_t =$
- Vertical sampling interval: $\Delta_y =$
- Line interval: $T_l =$
- Actual scanning time : $T'_l = T_l - T_h$
- Active lines: $f'_{s,y} = (\Delta_t - Tv) / T_l$

Basic Black and White TV



Color TV Broadcasting



Spatial and Temporal resolution

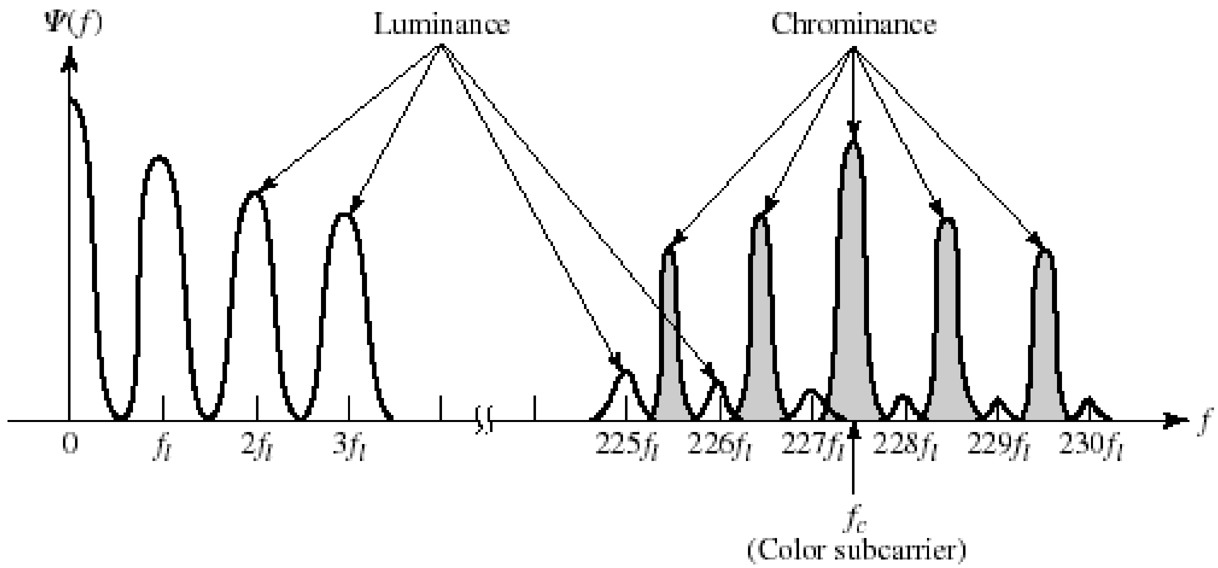
Parameters	NTSC	PAL	SECAM
Field rate	59.94	50	50
Line number/frame	525	625	625
Line rate (line/s)	15,750	15,625	15,625
Image aspect ratio	4:3	4:3	4:3
Color coordinate	YIQ	YUV	YDbDr
Luminance bandwidth (MHz)	4.2	5.0, 5.5	6.0
Chrominance bandwidth (MHz)	1.5 (I), 0.5 (Q)	1.3 (U, V)	1.0 (U, V)
Color subcarrier (MHz)	3.58	4.43	4.25 (Db), 4.41 (Dr)
Color modulation	QAM	QAM	FM
Audio subcarrier (MHz)	4.5	5.5, 6.0	6.5
Composite signal bandwidth (MHz)	6.0	8.0, 8.5	8.0

CV

- NTSC (National Television Standards Committee)

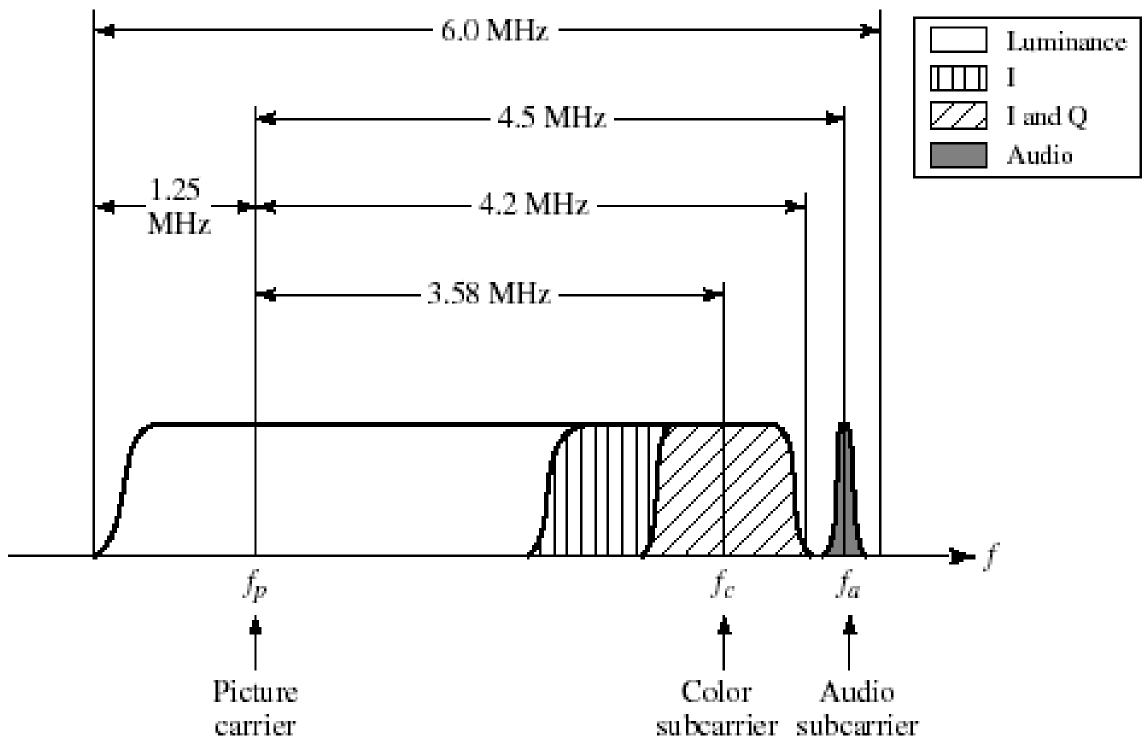
525 lines/frame, 60 fields/sec, 4:3 aspect ratio

- Horizontal sweep frequency: 15.75 KHz
- Th : 10μsec, Tv = 1333 μsec (21 scan lines/field)
- Number of active lines: 483/frame
- Max freq that can be rendered by a system
  - Fvmax = K $f_{s,y}$ (cycles/picture height)
  - Fmax = IAR. K. $f_{s,y}$ / ($2T_l$) Hz
- BW : 4.2MHz - Total: 6 MHz



Ψ(f)
Luminance
Chrominance
0
f_l
$2f_l$
$3f_l$
$225f_l$
$226f_l$
$227f_l$
$228f_l$
$229f_l$
$230f_l$
f
f_c
(Color subcarrier)

(a)



6.0 MHz
4.5 MHz
1.25 MHz
4.2 MHz
3.58 MHz
Luminance
I
I and Q
Audio
f
f_p
f_c
f_a
Picture carrier
Color subcarrier
Audio subcarrier

- PAL - SECAM

  Phase Alternating Lines

  Sequential Color and Memory

BW : 8 MHz

625 lines/frame

50 fields/sec

NTSC/525 Advantages

- Higher Frame Rate
- Less inherent picture noise

NTSC/525 Disadvantages

- Lower Number of Scan Lines
- Smaller Luminance Signal Bandwidth.
- Susceptablity to Hue Fluctuation
- Lower Gamma Ratio
- Undesirable Automatic Features

PAL/625 Advantages

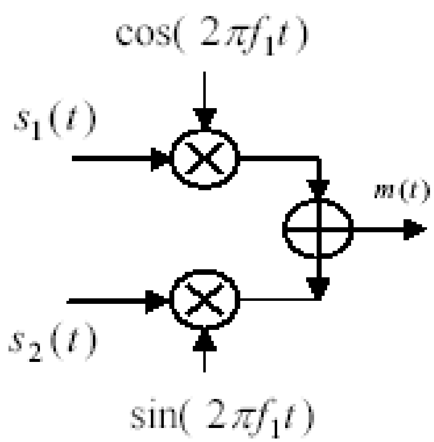
- Greater Number of Scan Lines
- Wider Luminance Signal Bandwidth
- Stable Hues
- Higher Gamma Ratio

PAL/625 Disadvantages

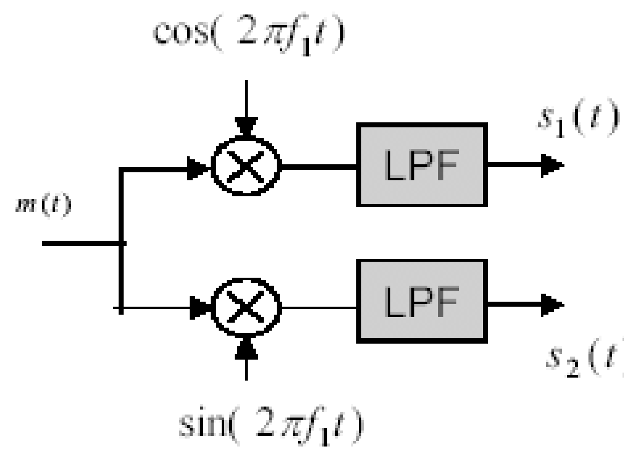
- More Flicker
- Lower Signal to Noise Ratio
- Loss of Colour Editing Accuracy
- Variable Colour Saturation

QAM

A method to modulate two signals onto the same carrier frequency, but with 90^{o} phase shift

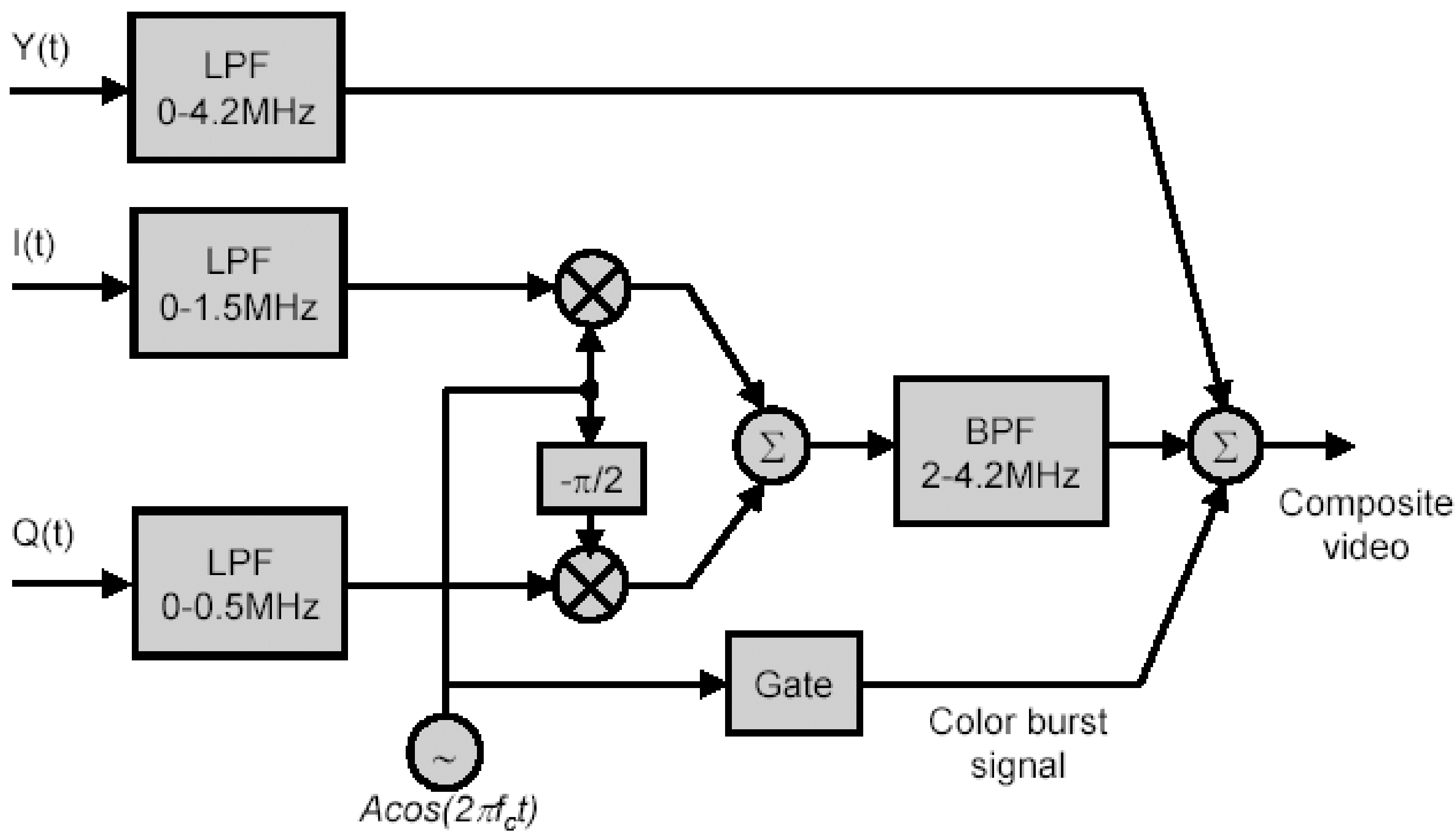


QAM modulator

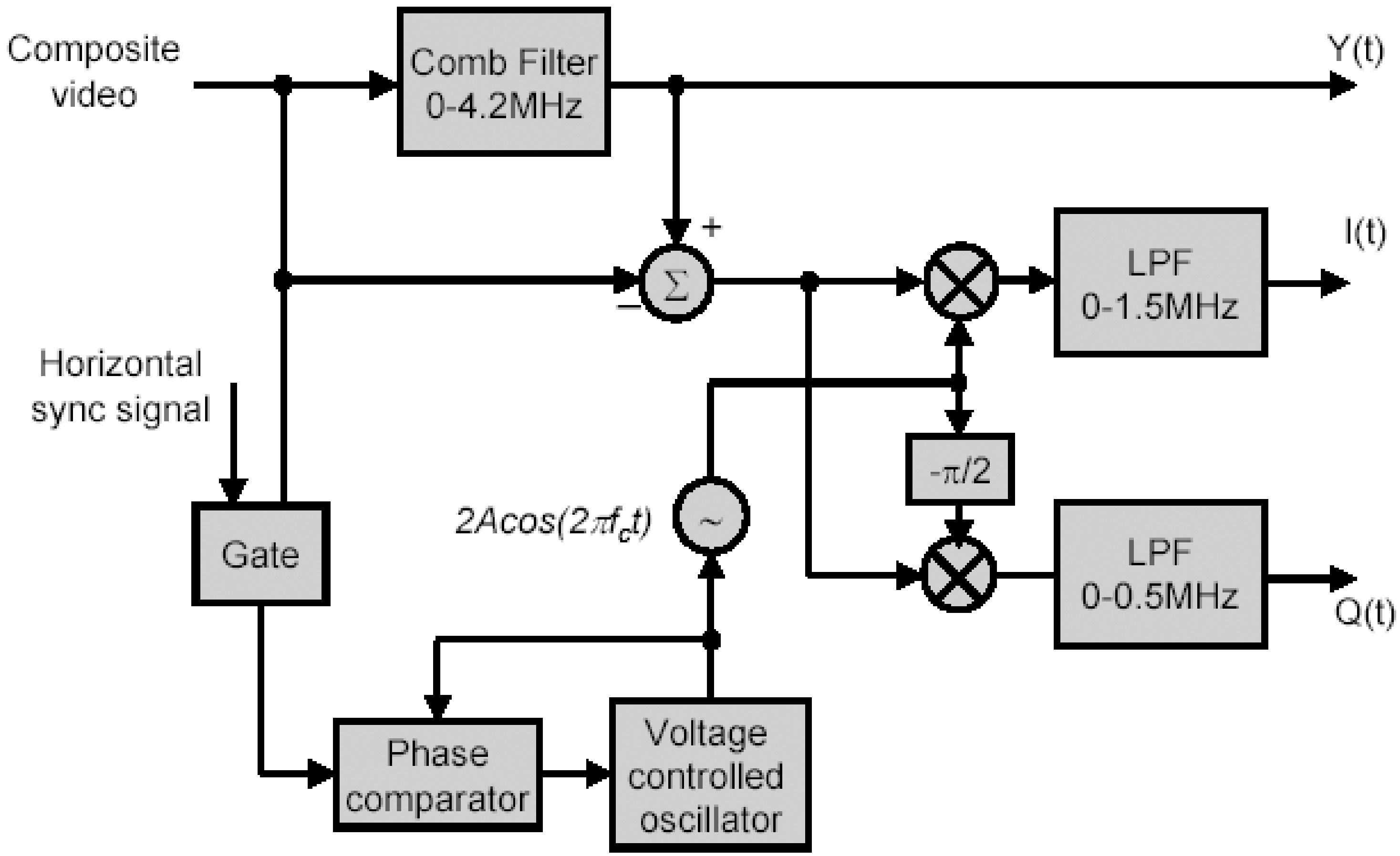


QAM demodulator

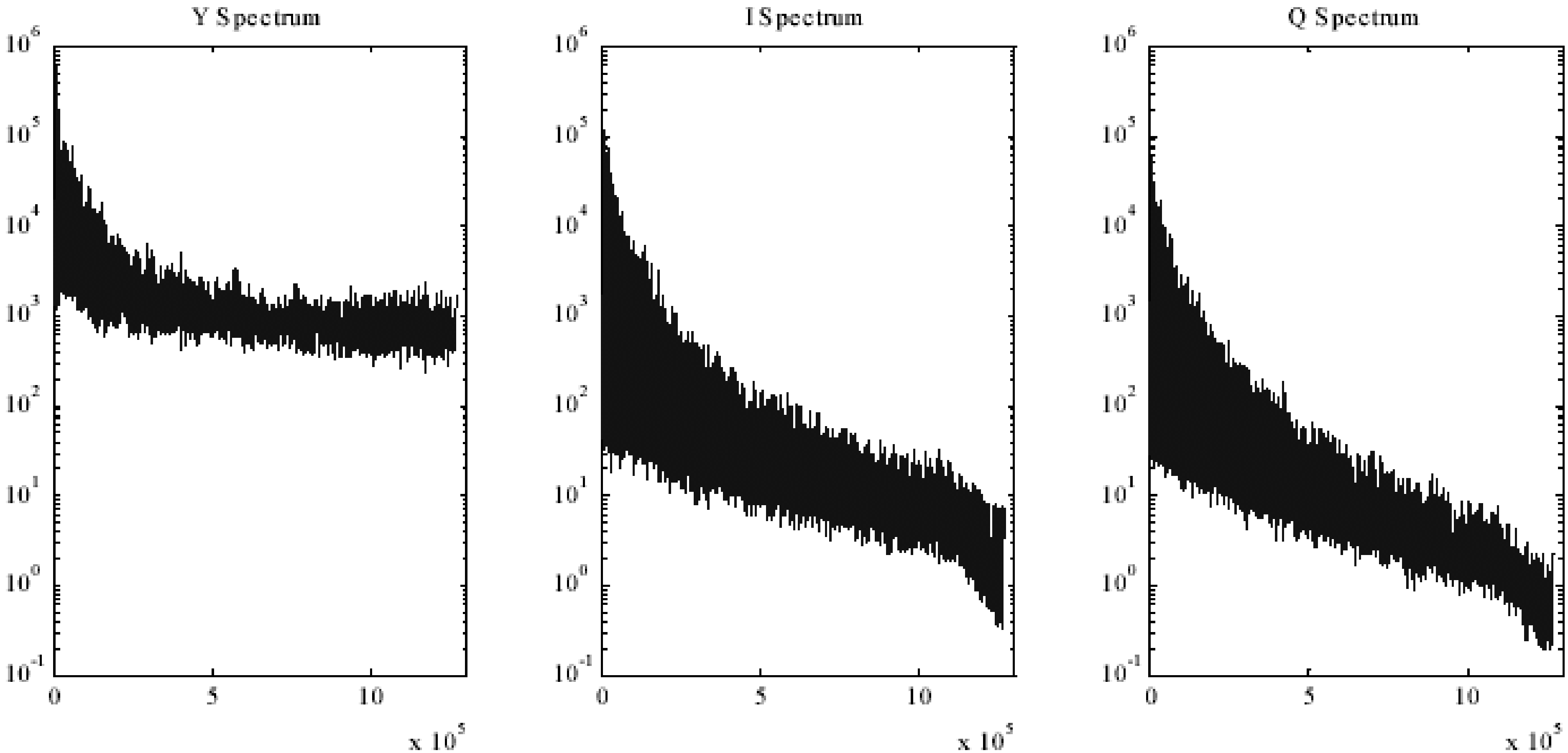
Multiplexing



Demultiplexing



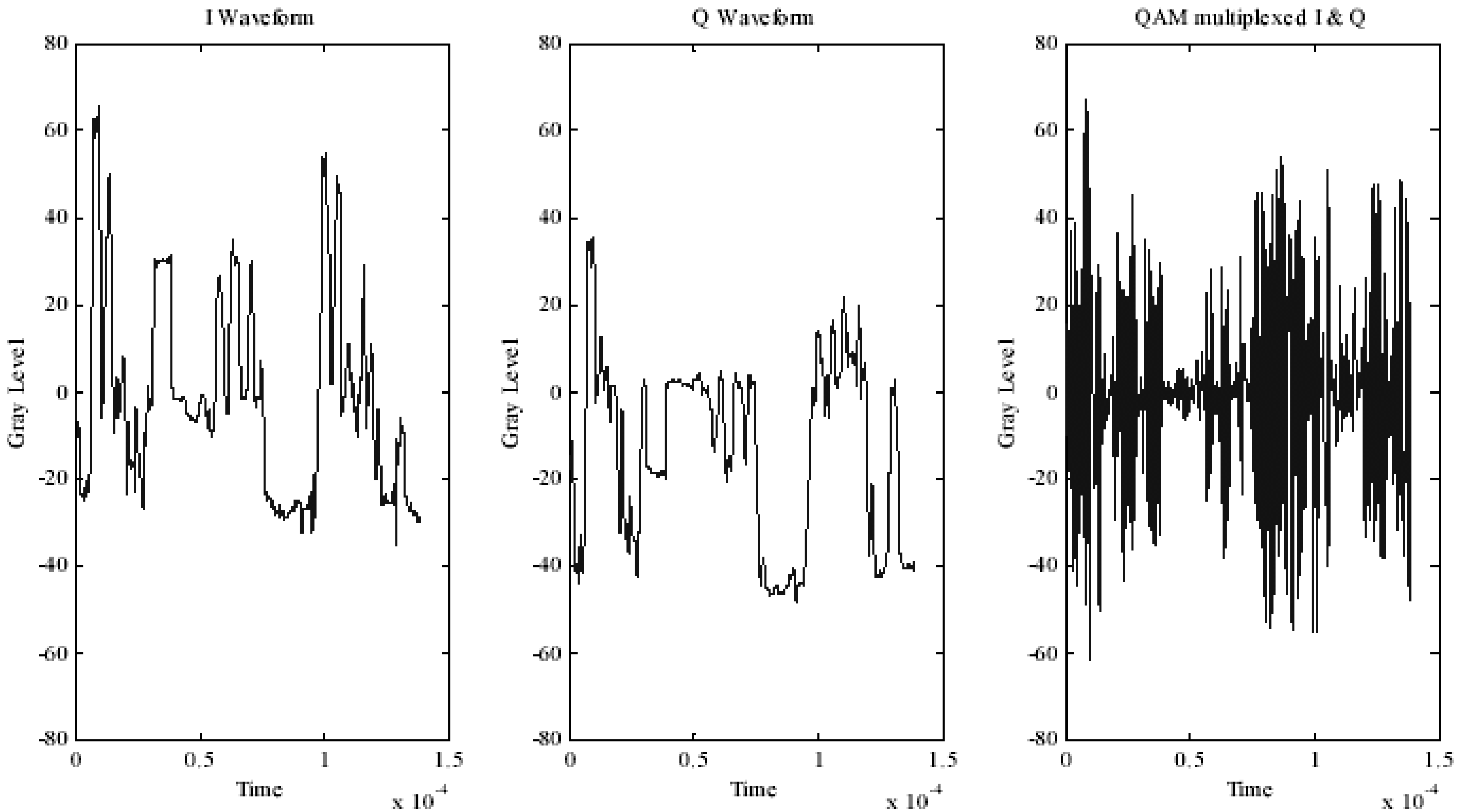
Spectrum of Y, I, Q



Spectrum of Y, I, and Q components, computed from first two progressive frames of “mobilcal”, 352x240/frame

Maximum possible frequency is 352x240x30/2=1.26 MHz.

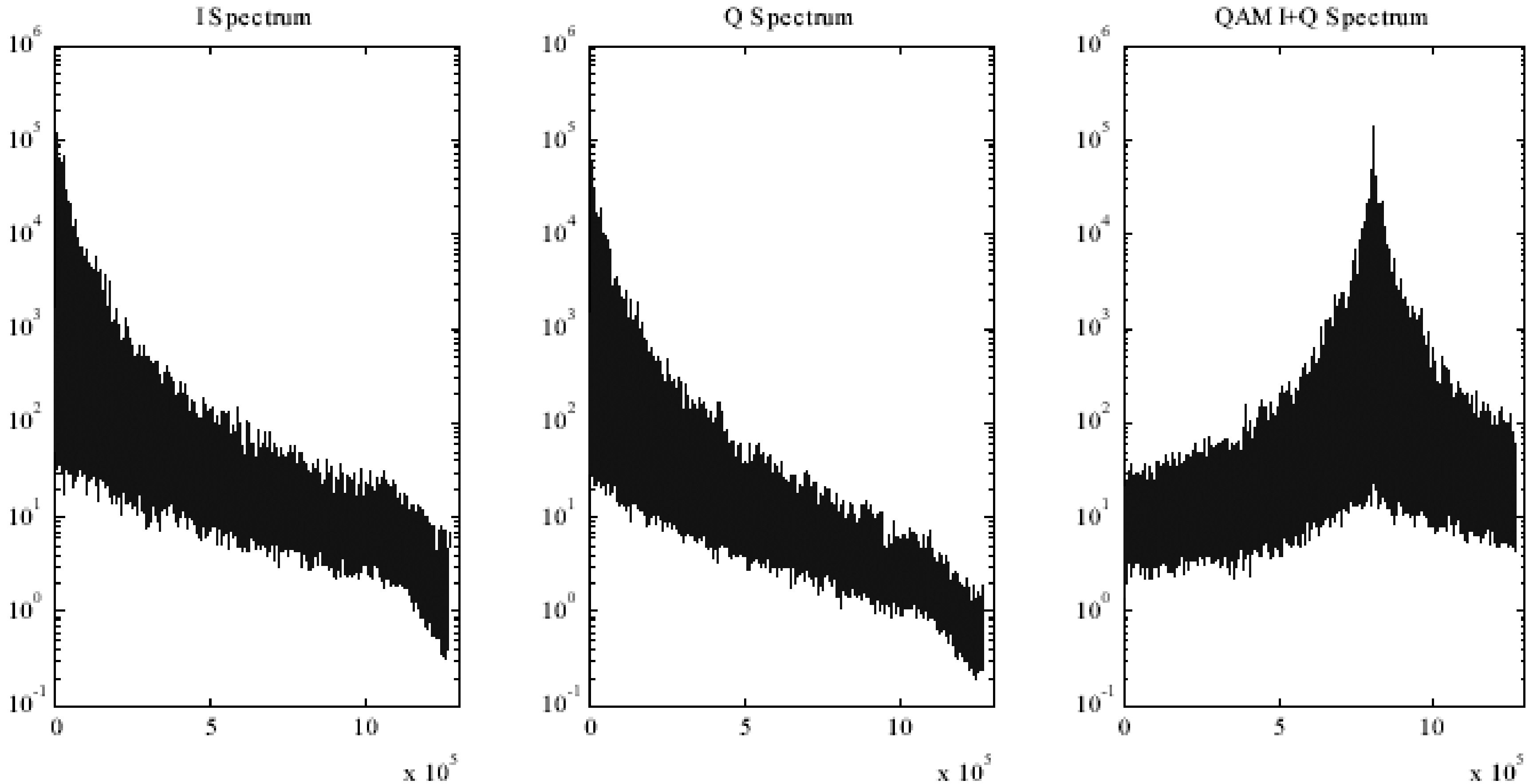
Notice bandwidths of Y, I, Q components are 0.8,0.2,0.15 MHz, respectively, if we consider 10^3 as the cut-off magnitude.



Line rate f_l=30*240; Luminance f_{max}=30*240*352/2*0.7=.89 MHz, The color subcarrier f_c=225*f_l/2=0.81MHz.

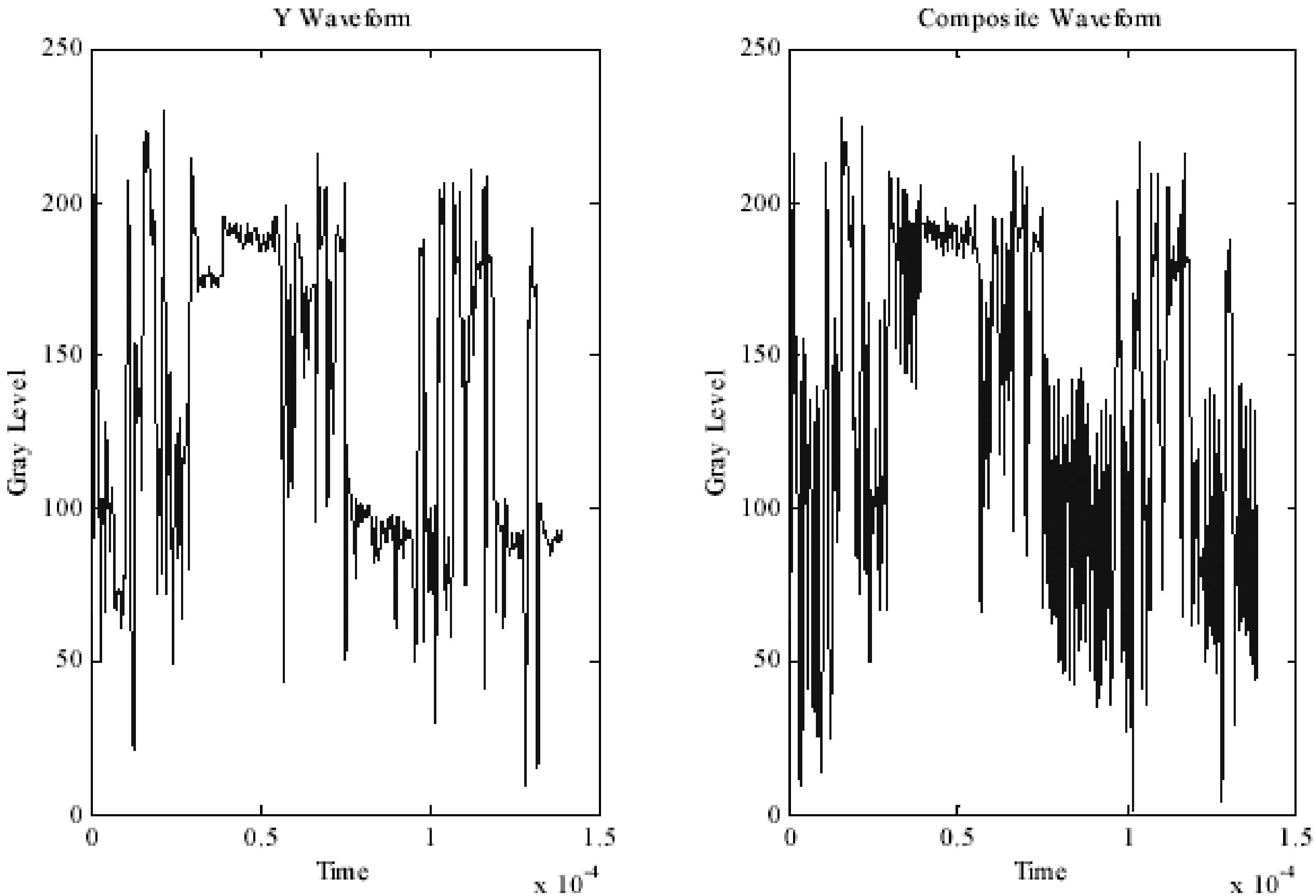
$M(t)=I(t)*cos(2\pi f_c t)+Q(t)*sin(2\pi f_c t)$

QAM of I and Q: Spectrum

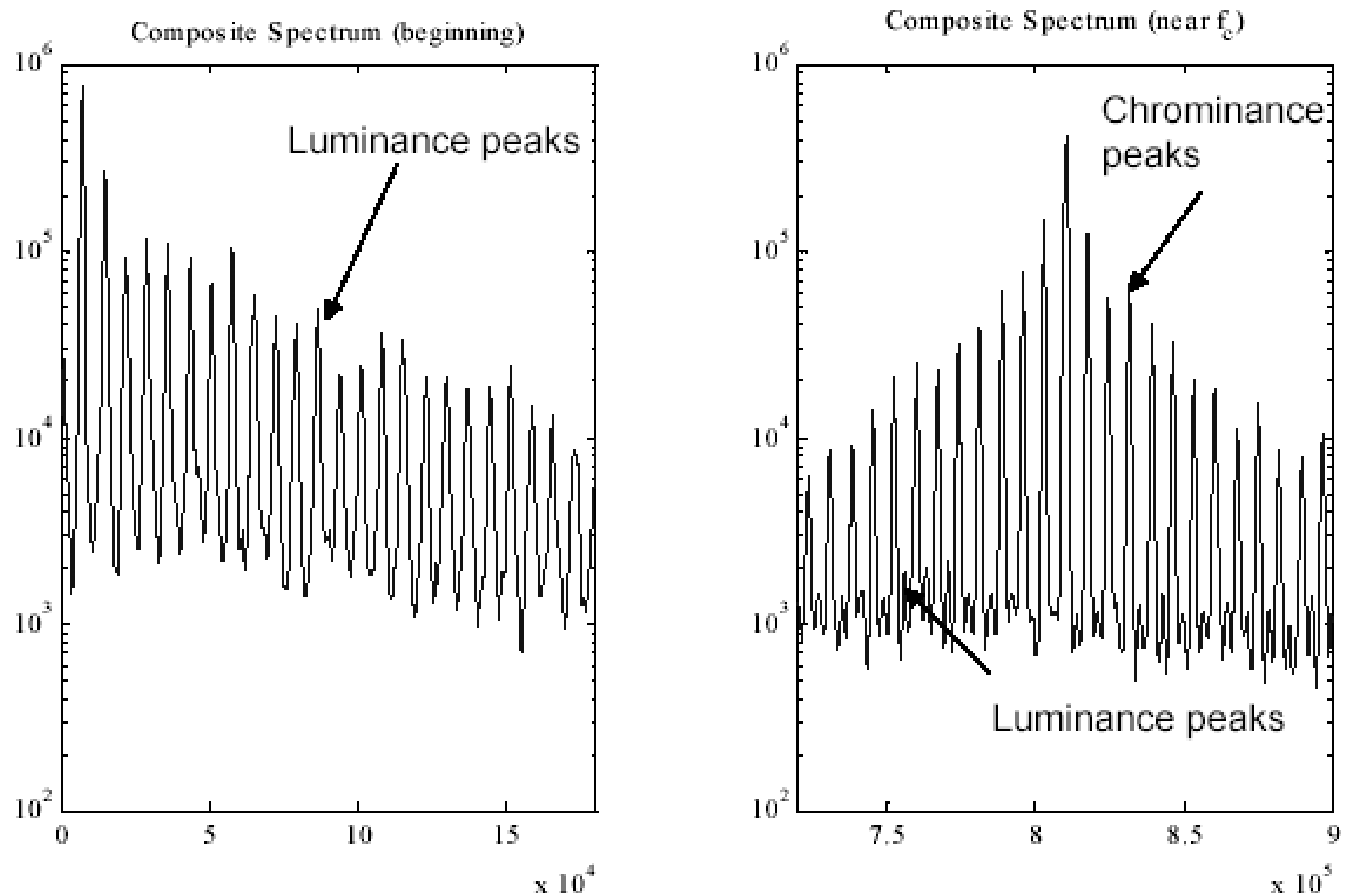


Spectrum of I, Q, and QAM multiplexed I+Q, fc=225*fl/2=0.81 MHz

Composite Video: Waveform

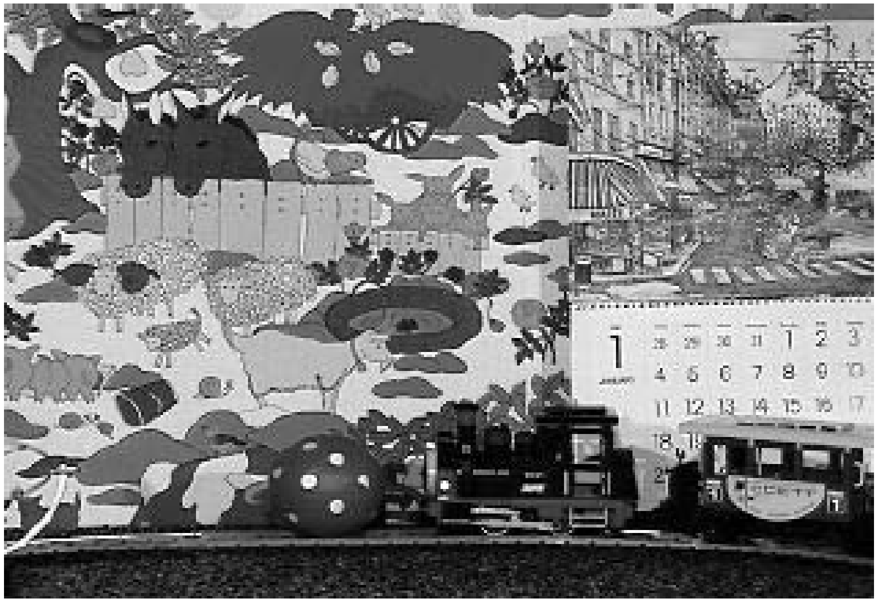


Waveform of the Y signal *Y(t)* and the composite signal *V(t)=Y(t)+M(t)*. 1 line

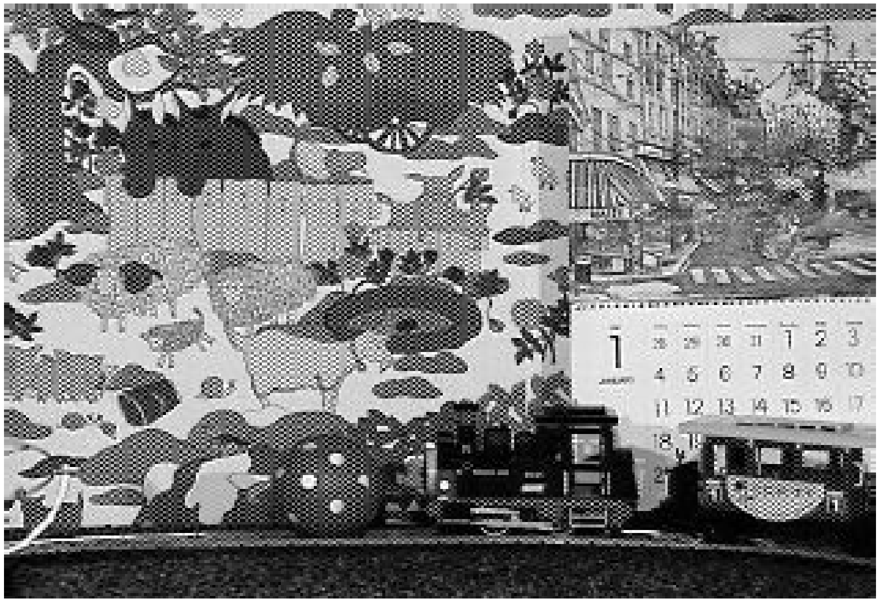


Notice the harmonic peaks of Y and M interleaves near fc

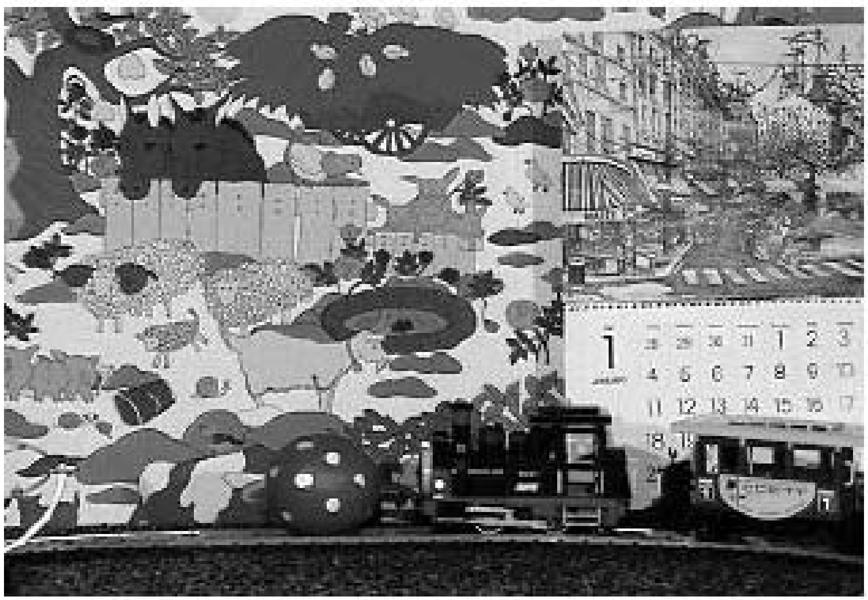
Original Y



Composite Signal as Y



Original Y



Recovered Y

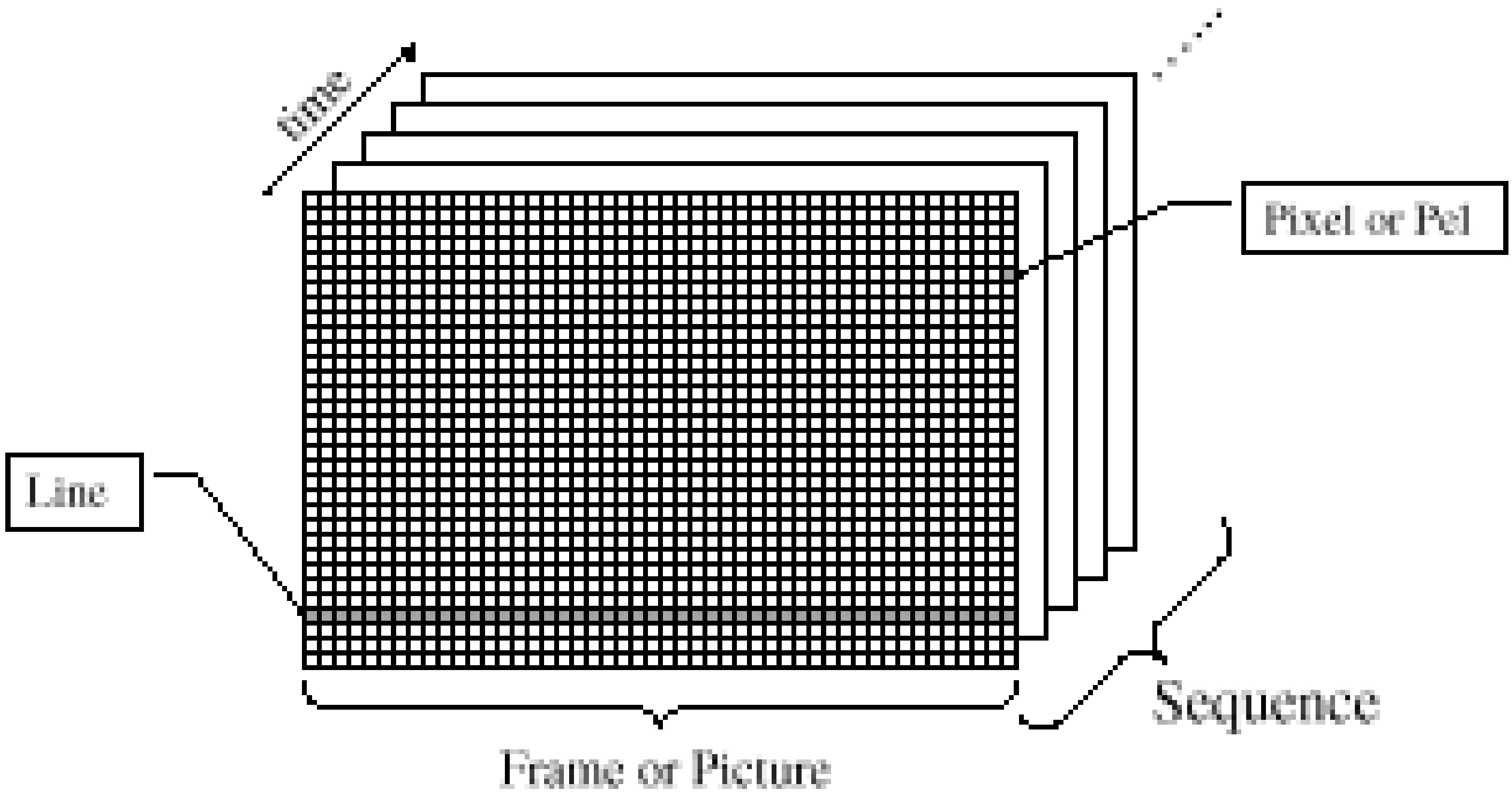
On the right is what a B/W receiver will see if a lowpass filter with cutoff frequency at about 0.75 MHz is applied to the baseband video signal. This is also the recovered Y component by a color receiver if the same filter is used to separate Y and QAM signal.

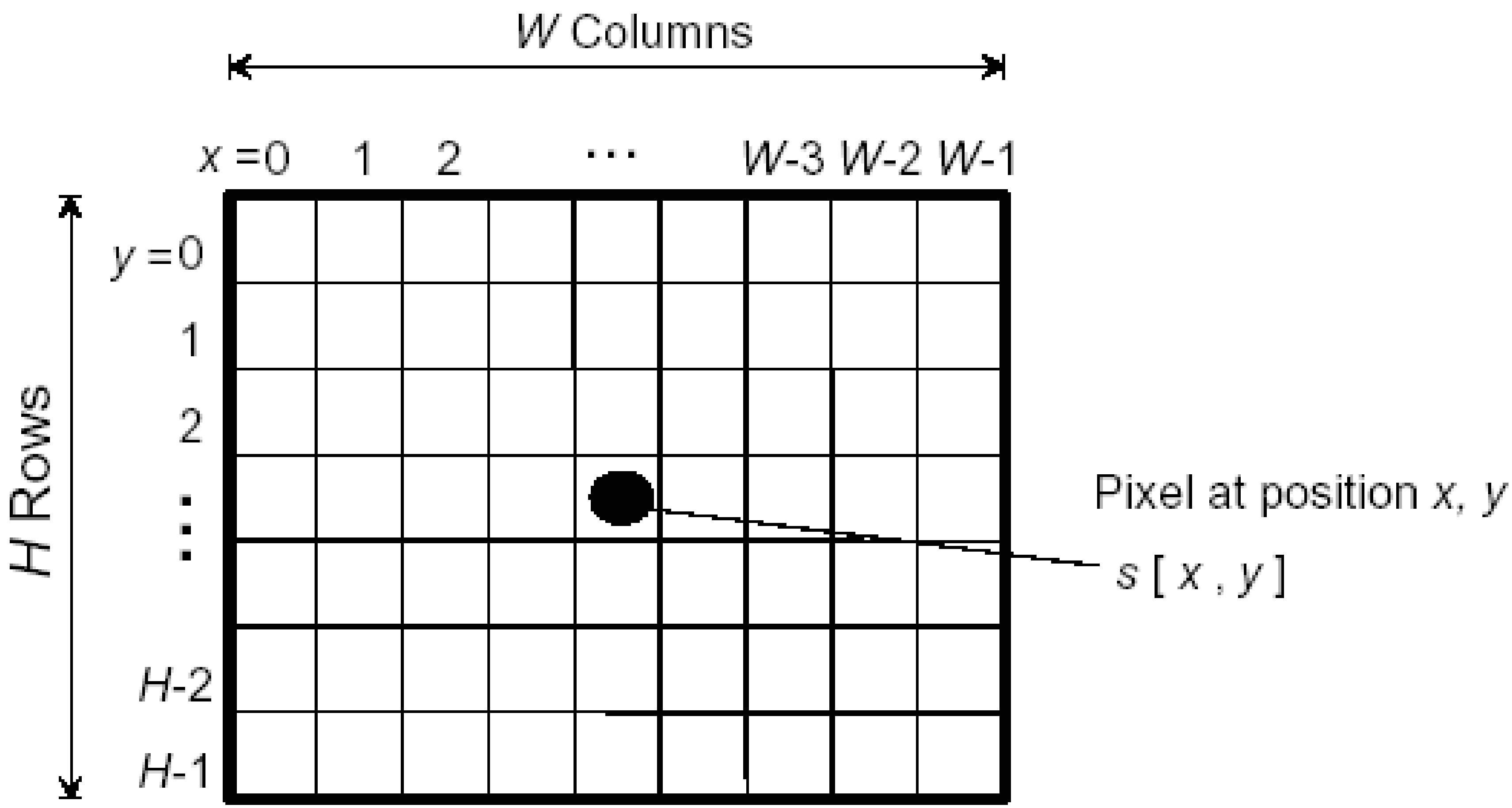
$$Y'(t)=\mathrm{conv}(V(t),LPF(t))$$

DIGITAL VIDEO

- Digital data communication (computer networks, e-mail)
- Digital audio (CD players, digital telephony)
- Digital video

Images and video



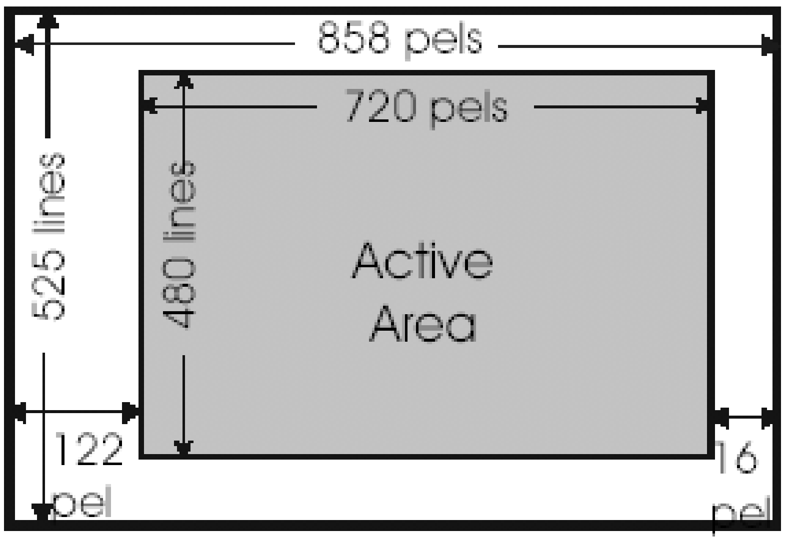


W Columns
x =0
1
2
...
W-3
W-2
W-1
H Rows
y =0
1
2
H-2
H-1
Pixel at position x, y
s [x , y]

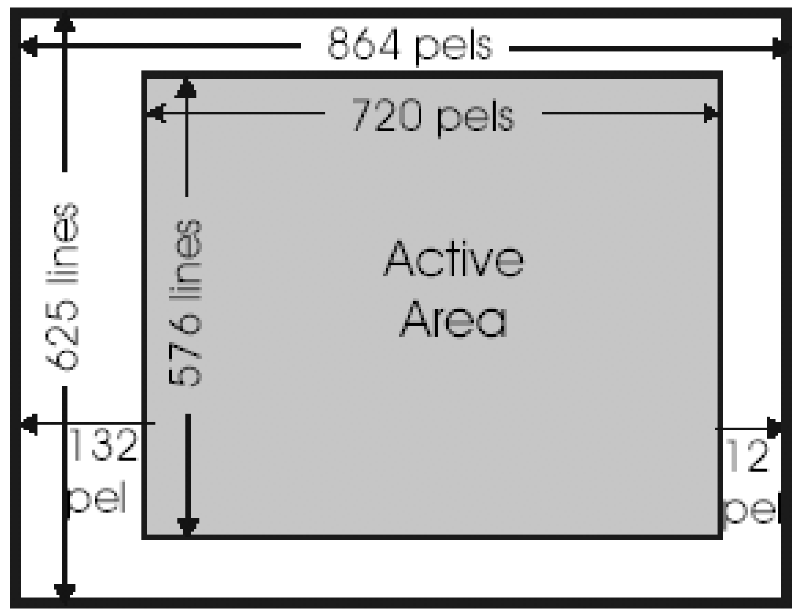
Digital Image Formats

	QCIF	*CIF*	*ITU-R 601*	*ITU-R 709*
Pixel / row (*Y*)	176	352	720	1920
Number of rows (*Y*)	144	288	576 (480)	1080
Pixel / row (*U,V*)	88	176	360	960
Number of rows (*U,V*)	72	144	576 (480)	1080
Aspect ratio	4:3	4:3	4:3	16:9
Pictures per second [*Hz*]	5-15	10-30	25 (30)	25 (30)
Bits per picture [*kbyte*] bei 8Bit-PCM	38,02	152,1	829,4 (691,2)	4.424 (3.686)
Bit-rate for image sequence [*Mb/s*]	0,84 - 3,8	10,1 - 30,4	165,9	884,7

601 video format



525/60: 60 field/s

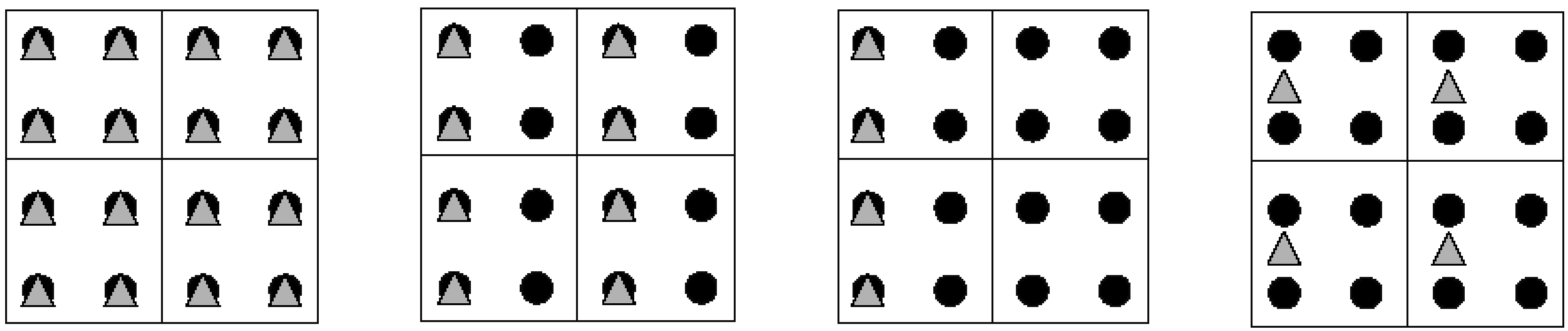


625/50: 50 field/s

Color conversion

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

Chrominance subsampling



4:4:4
For every 2×2 Y pixels
4 Cb & 4 Cr pixels
(no subsampling)

4:2:2
For every 2×2 Y pixels
2 Cb & 2 Cr pixels
(subsampling by 2:1 horizontally only)

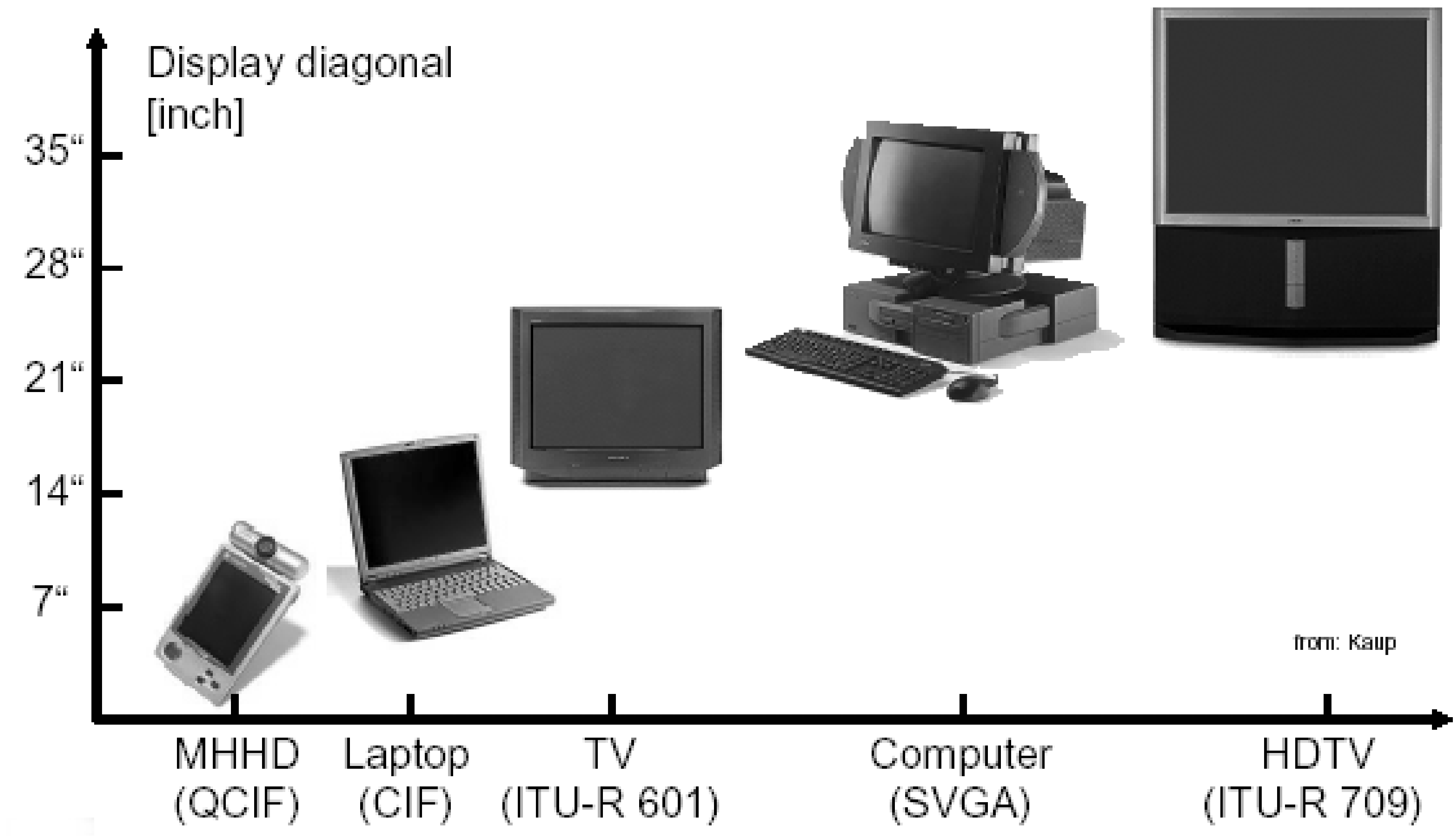
4:1:1
For every 4×1 Y pixels
1 Cb & 1 Cr pixel
(subsampling by 4:1 horizontally only)

4:2:0
For every 2×2 Y pixels
1 Cb & 1 Cr pixel
(subsampling by 2:1 both horizontally and vertically)

Applications

- All Digital HDTV (20Mbits/sec)
- Multi-media (1.5 Mbits/s)
- Videoconferencing (384 kbits/s)
- Videophone (8-16Kbits/s)
- Medical imaging
- Education, military,trafic systems,...

Digital image formats and applications



Bottleneck

- HDTV - 1440x1050 lum
  720x525 chrom
  30 frames/s x 8 bits/pel/channel

  545 Mbps

Bitrate requirements

- Conventional phone - 0.3 - 56kbits/s
- ISDN - 64-144 kbits/s
- T1 - 1.5 Mbits/s
- Ethernet - 10 Mbits/s

Compression is needed

- H.261
- MPEG1
- MPEG2
- MPEG4
- AVI
- Quicktime

Advantages

- Open architecture video systems
- Interactivity
- Variable-rate transmission on demand
- Easy sw conversion from one standard to another
- Integration
- Editing capabilities
- Robustness to channel noise and ease of encryption

HOMEWORK 1

- Use any video sequence you can fing from internet (raw video) forma composite video signal and then perform filtering to seperate luminance and chrominance signals.
- Show the reconstructed and original B&W and color frames