FUNDAMENTALS OF MULTIMEDIA TECHNOLOGIES

Lecture slides

BME Dept. of Networked Systems and Services

2018.

Synopsis

- Psychophysical fundamentals of the human visual system
- Color spaces, video components and quantization of video signal
- ITU-601 (SD), and ITU-709 (HD) raster formats, sampling frequencies, UHDTV recommendations
- Basics of signal compression: differential quantization, linear prediction, transform coding
- JPEG (DCT based transform coding)
- Video compression: motion estimation and motion compensated prediction, block matching algorithms, MPEG-1 and MPEG-2,
- H-264/MPEG-4 AVC: differences from MPEG-2

What is our aim, and why?

- Components of a video format
  - size (resolution, raster)
  - frame rate
  - representation of a color pixel (color space, luma/chroma components)
- Two main goals:
  - finding a video format, that ensures indistinguishable video quality from the real, original scene
  - derive a source encoder methodology resulting in unnoticeable errors for **Human Visual System**
- Suitable representation/encoding method, adapted to **human vision** → important to know basics of human vision

Visible spectrum

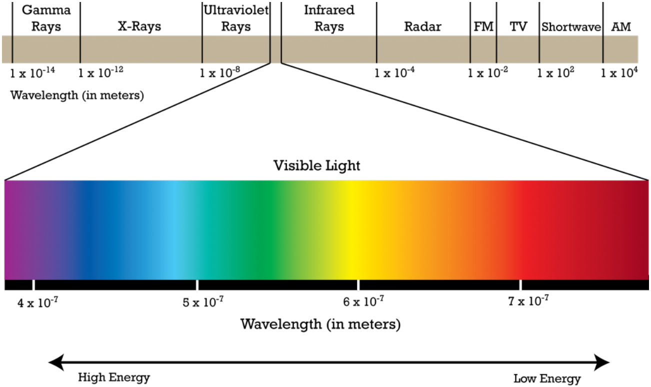
Visible light and colors

- The Human Visual System (HVS) is sensitive to a narrow frequency band of the electromagnetic spectrum:
  - frequencies/wavelengths between ultraviolet (UV) and infrared (IR): between ca. 400 and 700 nm of wavelength
- Different wavelengths: different color experiences for HVS
- Basic perceived colors:
  - **blue, cyan, green, yellow, orange, red, purple.**
  - White (and different shades of grey) lights are the mixtures of the basic colors.

Visible spectrum

Visible lights and color

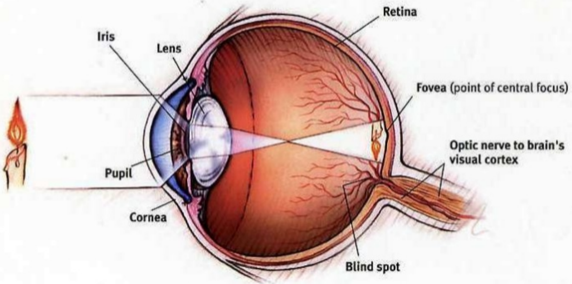
- Light with shortest visible wavelength: blue (violet),
- Light with longest visible wavelength: red,



Structure of the human eye

- HVS includes eyes, optic nerves, parts of brain converting light excitation to stimulus
- Here only parts of the eye, related light perception discussed
- Light rays, passing through the pupil are collected and focused (refracted) by the lens, projected to the retina, creating a real, reduced, inverted image

Producing an Image



The converging lens produces an inverted image which the brain interprets as being upright

Structure of the human eye

Blind spot experiment

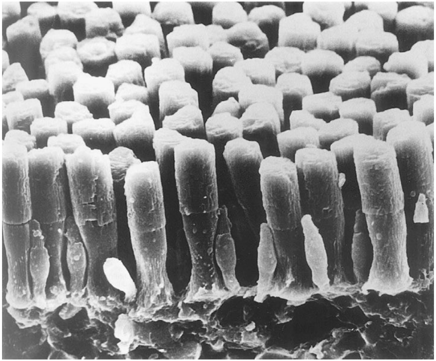
- Draw an image similar to this one, with the dot and cross being about 6 inches (15 cm) apart

- close your right eye and focus on the cross with the left eye
- hold the image about 20 inches (50 cm) away from your face and move it slowly towards you
- the dot should disappear

Role of rods and cones

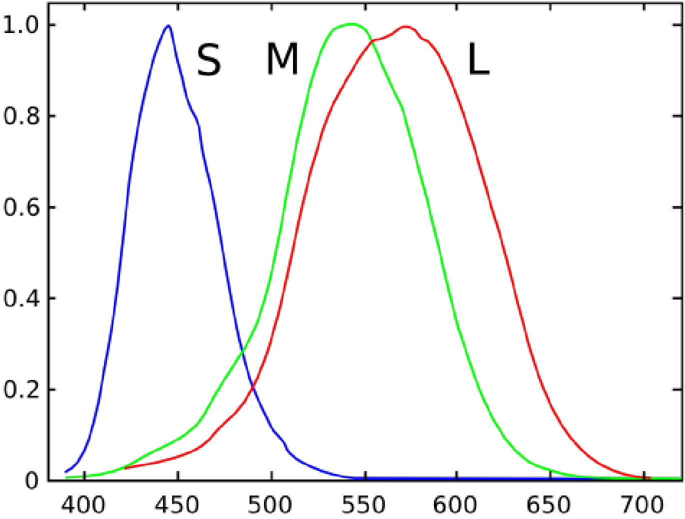
Photopic and scotopic vision

- Photopic: at daylight: perceived by cones - three types of cones (L,M,S) ensure color vision
- Scotopic: "night vision": received by rods - scotopic luminosity function’s maximum translated to blue colors



Role of rods and cones II

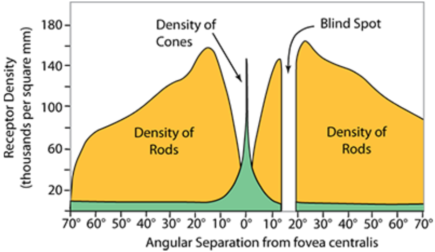
Sensitivity function of L,M,S cones



Role of rods and cones

Location of rod and cone photoreceptor cells

- Cones: in the center of the retina (at the fovea: place of color vision) at high density
- Rods: distributed towards the edge of retina, responsible for peripheral vision. More sensitive, responds faster than cones → flicker fusion threshold is higher at peripheral vision



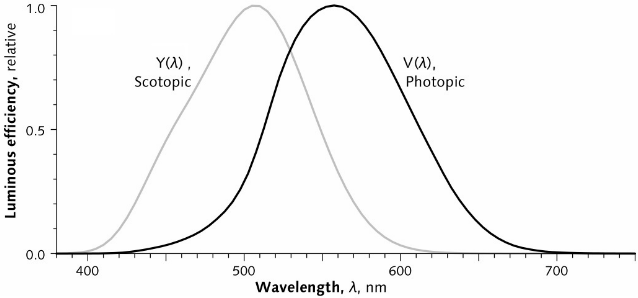
Variance of visual perception among people

Aim: to give a unified model describing the average human vision

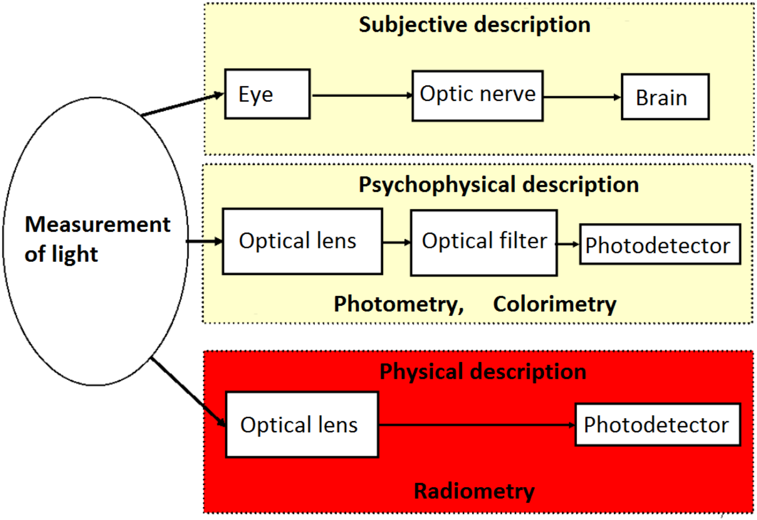
- structure of the visual system: identical for each person
- operation principle of the HVS: identical for each person
- perceived vision: various among people
- therefore: each person perceives the world in slightly different colors
- basic attribute of HVS in the aspect of image/video capture, transmission and reproduction: **luminosity function**

Luminosity function

- The luminosity function $V(\lambda)$ or luminous efficiency function describes the average spectral sensitivity of human visual perception of **brightness**
- Perceived brightness for different wavelengths *excited with constant intensity*
- Averaged, measured for numerous observers (*CIE standard observer*)
- Maximum of the curve at $\sim 555\mathrm{nm}$ (green) —$V(\lambda)$ is normalized to this value



Measurement of light



Subjective description

Eye → Optic nerve → Brain

Measurement of light

Psychophysical description

Optical lens → Optical filter → Photodetector

Photometry, Colorimetry

Physical description

Optical lens → Photodetector

Radiometry

Measurement of light

Subjective attributes

- Measurement "device": Human Vision System
- Purely subjective quantities

Psychophysical quantities

- Measurement device: optoelectric sensors
- Optical filter (or the transfer function between the source and the sensor) corresponds to the luminosity function $V(\lambda)$

Physical quantities

- Measurement device: optoelectric sensors
- Purey radiometric quantities, independent of the properties of human vision

Physical quantities

- Radiometry
  - Measurement device: optoelectric sensors
  - Radiometric techniques in optics characterize the distribution of the radiation's power in space
  - Optical radiation: between $\lambda = 1\ \mathrm{nm}$ X-ray and $\lambda = 1\ \mathrm{mm}$ radio waves
- (Some) radiometric quantities
  - Radiant flux (flux) $[W]$
  - Radiant intensity $[Wsr^{-1}]$ - radiant flux emitted, reflected, transmitted or received, per unit solid angle (directional)
  - Irradiance $[Wm^{-2}]$ - radiant flux received by a surface per unit area.
  - Radiance $[Wsr^{-1}m^{-2}]$ - Radiant flux emitted, reflected, transmitted or received by a surface, per unit solid angle per unit projected area (directional)

Physical quantities

Radiometric measurements

- Measurement of radiometric quantities: radiometer/roentgenometer
- Measurement of radiometric quantities in narrow wavelength intervals: spectroradiometer

Photometric measurements

- Measurements of photometric (perceived by HVS) quantities: photometer
- Measurement of photometric quantities in narrow wavelength intervals: spectrophotometry

Psychophysical quantities

Photometry

- measurement of light in the visible wavelength domain $(400 - 700 \mathrm{~nm})$ in terms of its perceived brightness to the human eye
- optoelectric transfer function (or optical filter): luminosity function $V(\lambda)$

- photometric quantities: luminous flux, luminous intensity, illuminance, **luminance**

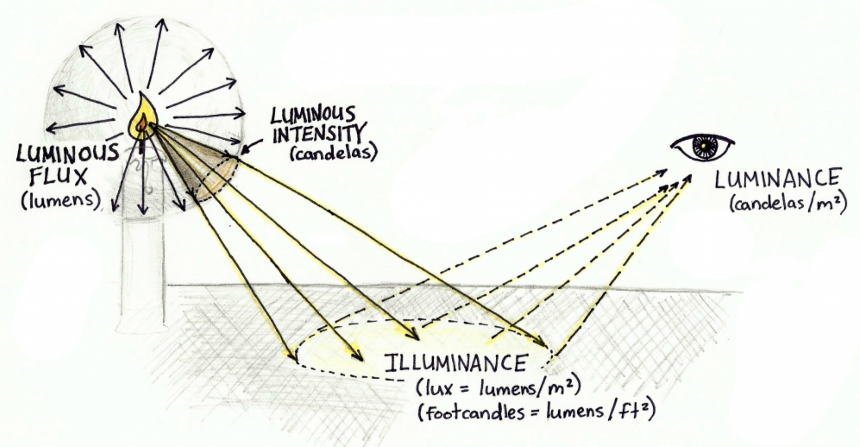
Psychophysical quantities

Relation of radiometric and photometric quantities

Radiometric quant.	unit	Photometric quant.	unit
Radiant flux (Φ_e)	W	Luminous flux (Φ_v)	lm (lumen)
Radiant intensity (I_e)	Wsr^{-1}	Luminous intensity (I_v)	cd (candela)
Irradiance (E_e)	Wm^{-2}	Illuminance (E_v)	lx (lux)
Radiance (L_e)	$\mathrm{Wsr}^{-1}\mathrm{m}^{-2}$	**Luminance** (L_v)	cdm^{-2} (nit)

in the aspect of video the most important quantity: **luminance.**

Psychophysical quantities



Subjective quantities

Brightness

Attribute of a visual perception according to which an area appears to exhibit more or less light. (Adjectives: bright and dim.) Brightness refers to an absolute term (strongly related to luminance)

Lightness

Brightness of a surface compared to a white surface in the same scene (strongly related to relative luminance) Lightness: relative term

Subjective quantities

Hue

Degree to which a stimulus can be described as similar to or different from stimuli that are described as red, green, blue, and yellow

Colorfulness

Attribute of a visual perception according to which the perceived color of an area appears to be more or less chromatic

Chroma

colorfulness of an area judged as a proportion of the brightness of a similarly illuminated area that appears white or highly transmitting

Saturation

colorfulness of an area judged in proportion to its brightness

Subjective quantities

$$Chroma = \frac{Colorfulness}{Brightness(White)}$$

$$Saturation = \frac{Colorfulness}{Brightness}$$

$$Lightness = \frac{Brightness}{Brightness(White)}$$

$$Saturation = \frac{Chroma}{Lightness}$$

$$Saturation = \frac{Colorfulness}{Brightness(White)} \cdot \frac{Brightness(White)}{Brightness}$$

Subjective quantities

Flicker and flicker fusion rate

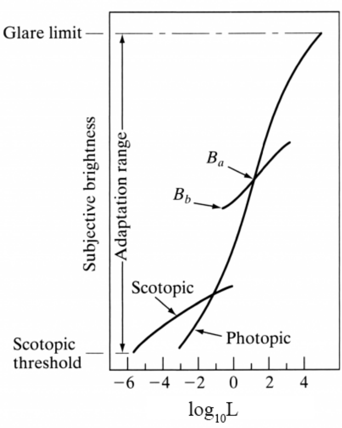
- Flicker: usually unpleasant impression due to the rapid change of the perceived brightness or hue, faster than several Hz, but below the flicker fusion rate
- Flicker detection depends on
  - the frequency of the modulation
  - the amplitude or depth of the modulation
  - the average illumination intensity
  - the wavelength (or wavelength range) of the illumination
  - the position on the retina at which the stimulation occurs (due to the different distribution of photoreceptor types at different positions)
  - the degree of light or dark adaptation
  - physiological factors such as age and fatigue
- Flicker fusion threshold (or flicker fusion rate): the frequency at which an intermittent light stimulus appears to be completely steady to the average human observer

Brightness adaptation

Brightness dynamic range of HVS

- Full luminance range of human vision: 8-10 orders of magnitude (full dark to direct sunlight)
- Within full range: adaptation
  - eye's iris changing pupil diameter (fast, within $\sim$ 1 decade)
  - photochemical adaptation of photoreceptors (slow)
- At a particular state of adaptation, vision can discern different luminances across about a 100:1 range. When viewing a real scene, adaptation changes depending upon where in the scene your gaze is directed increasing this to 1000:1.
- Loosely speaking, luminance levels less than 1% of peak white appear just “black”:

Brightness adaptation and subjective brightness



- B_a: Brightness adaptation level
- B_b: in adapted state the smallest noticeable luminance (ratio of about 100:1!)
- scotopic vision not color sensitive, more sensitive to light and mainly peripheral (rods)
- photopic vision color sensitive, less sensitive to light and mainly central (cones)

Contrast sensitivity

Ratio of the just noticeable luminances between two adjacent patches

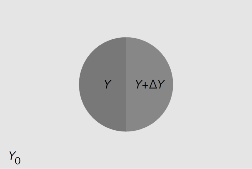
depends on

- eye's stability, e.g.:
  - eye is fixed
  - under normal motion
  - moves between two points forth and back
- temporal changes of contrast
- spatial changes of contrast
- different for brightness and hue
- **average luminance, at high adaptation level contrast sensitivity is higher**

Contrast sensitivity

Contrast sensitivity as the function of surround luminance

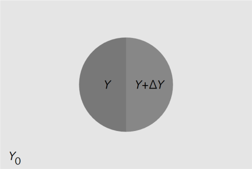
- Most of the observer's field of vision is filled by a surround luminance level, Y_0, fixing the observer's state of adaptation.
- In central area of the field of vision are placed two adjacent patches having slightly different luminance levels, Y and $Y + \Delta Y$.



Contrast sensitivity

Contrast sensitivity as the function of surround luminance

- The experimenter presents stimuli having a wide range of Y/Y_0 values. At each test luminance, the experimenter presents to the observer a range of luminance increments with respect to the test stimulus, that is, a range of $\Delta Y/Y$ values. Result: contrast sensitivity $\Delta Y/Y$ as function of Y_0: smallest noticeable relative luminance (contrast sensitivity threshold)



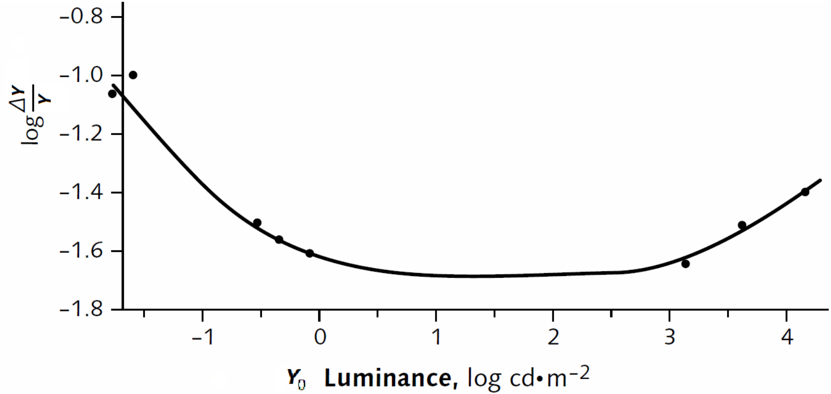
Contrast sensitivity

Contrast sensitivity as the function of surround luminance

- within 2 decades range of Y_0 the contrast sensitivity is 1 %
- in this range: two surfaces are not distinguishable if the ratio of their luminance is smaller than 1.01
- constant luminance ratio: in this range (at average surround luminances) the perceived brightness is appr. logarithmic
- at smaller or higher surround luminances: contrast sensitivity degrades

Contrast sensitivity

Contrast sensitivity as the function of surround luminance



Lightness

Description of subjective lightness in the function of luminance

- based on the foregoing: perceived lightness as the function of relative luminance is approximately logarithmic (with logarithmic *x*-axis it's graph is linear)
- more accurate description: a power function with an appropriately chosen exponent

The CIE Lightness definition

Lightness (L^*) according to CIE (1976)

$$L^* = \begin{cases} \left(\frac{29}{3}\right)^3 Y/Y_n, & Y/Y_n \leq \left(\frac{6}{29}\right)^3 \\ 116\,(Y/Y_n)^{1/3} - 16, & Y/Y_n > \left(\frac{6}{29}\right)^3 \end{cases}$$

- L^* is the power function of relative luminance (Y/Y_n),
- around the black level (below $L^* = 8$, $Y/Y_n = 1\%$) the function is extrapolated with a linear segment to cross $L^*(Y/Y_n = 0) = 0$.
- values of L^* are between 0 and 100
- Y_n: luminance of a freely chosen white point

The CIE Lightness definition

Lightness (L^*) according to CIE (1976)

$$L^* = \begin{cases} \left(\frac{29}{3}\right)^3 Y/Y_n, & Y/Y_n \leq \left(\frac{6}{29}\right)^3 \\ 116\,(Y/Y_n)^{1/3} - 16, & Y/Y_n > \left(\frac{6}{29}\right)^3 \end{cases}$$

- $\frac{Y}{Y_n}$: **relative luminance**: In image reproduction – including photography, cinema, video, and print – we rarely, if ever, reproduce the absolute luminance of the original scene. Instead, we reproduce luminance approximately proportional to scene luminance, up to the maximum luminance available in the reproduction medium.
- **The overall curve is approximated better with a 0.4-power function**

Contrast sensitivity (again)

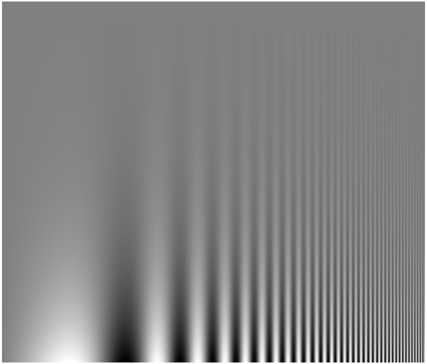
Ratio of the just noticeable luminances between two adjacent patches

depends on

- eye's stability, e.g.:
  - eye is fixed
  - under normal motion
  - moves between two points forth and back
- temporal changes of contrast
- **spatial changes of contrast**
- **different for brightness and hue**
- average luminance, at high adaptation level contrast sensitivity is higher

Contrast sensitivity

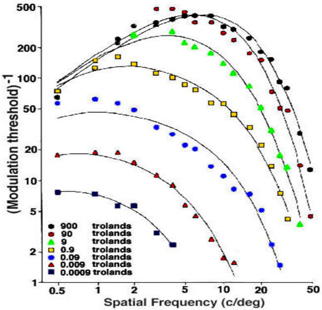
Contrast sensitivity function (CSF) as function of spatial frequency



Luminance varies over space with given spatial frequency between L_{min} and L_{max}. Contrast is then given as $C = \frac{L_{max} - L_{min}}{\bar{L}}$, with $\bar{L} = \frac{L_{max} + L_{min}}{2}$ being the average luminance. In the figure $\bar{L}$ is constant, contrast increases from top to the bottom, spatial frequency increases from left to right.

Contrast sensitivity

Contrast sensitivity function (CSF) as function of spatial frequency

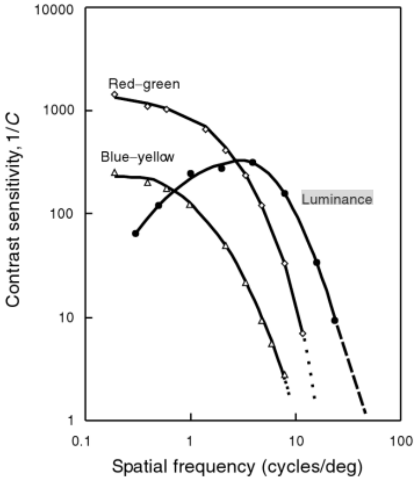


- troland: unit of conventional retinal illuminance.
- The modulation threshold is the smallest contrast value where modulation is perceived.
- Sensitivity: modulation threshold^{-1}

CSF varies as the function of average surround luminance and the spatial frequency. All curves vanish above 60 c/deg: vision isn’t capable of perceiving spatial frequencies above that

Contrast sensitivity

CSF as function of spatial frequency for luminance and hue

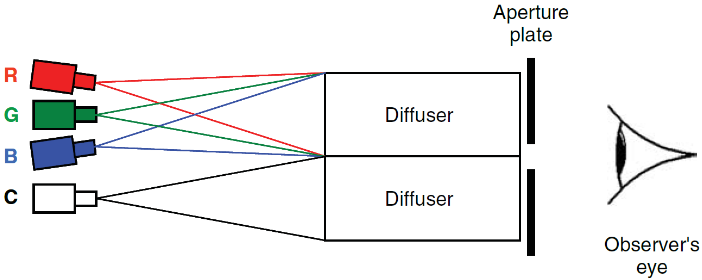


The highest spatial frequency limit of CSF is lower for colors (hue) than for luminance: resolution of human visual system is lower for colors than for luminance

CIE RGB color space

CIE RGB color matching experiment (1931)

The experiments were conducted by using a circular split screen 2 degrees in diameter, (angular size of the human fovea). On one side of the field a test color was projected and on the other side, an observer-adjustable color was projected. The adjustable color: mixture of three primary colors, with fixed chromaticity, adjustable brightness.



CIE RGB color space

CIE RGB color matching experiment (1931)

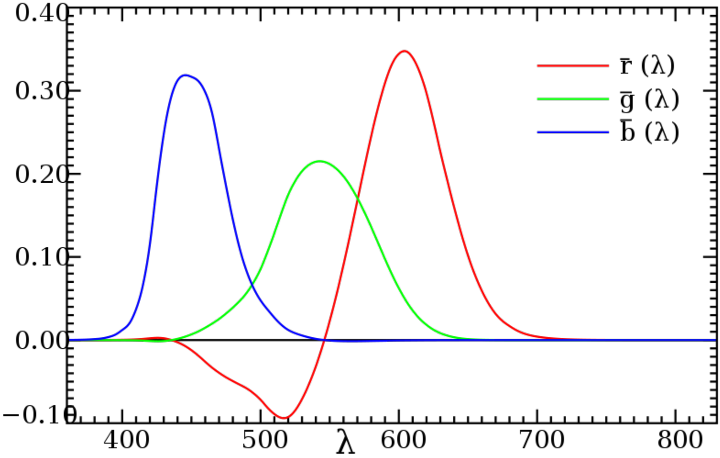
- 2° standard observer: indicator of color measurement (subjective evaluation)
- The observer would alter the brightness/intensity of each of the three primary beams until a match to the test color was observed (the color of circular split would look homogeneous)
- The intensity/brightness of the primary sources when matched: objective description of a given color
- Wavelengths of the primaries: $\lambda_{\mathbf{R}} = \mathbf{700nm}, \lambda_{\mathbf{G}} = \mathbf{546, 1nm}, \lambda_{\mathbf{B}} = \mathbf{435, 8nm}$
- Realization: filtered from mercury vapor discharge (mercury-vapor lamp) with liquid crystal filter

CIE RGB color space

CIE RGB color matching experiment (1931)

- problem: colors below 546.1 nm could not be matched using this technique
- the adjustable mixture looked more red even at $R = 0$ than the reference spectral color
- solution: a variable amount of the red primary was added to the test color. In those cases R was considered to be negative
- furthermore equal-energy white, or "illuminant E" could be realized with the ratio of RGB primary intensities: $L_R : L_G : L_B = 1 : 4,59 : 0,06$

The CIE RGB color-matching functions



- curves were normalized by $L_R : L_G : L_B = 1 : 4,59 : 0,06$

Additive color model

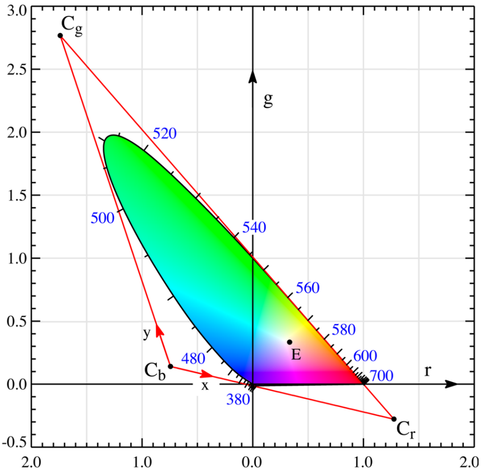
Grassman’s law:

- Human color vision/color sensation can be described in a 3 dimensional linear space
- Given the $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ spectral color-matching functions
- The RGB tristimulus values for a color with the spectral power distribution $\phi(\lambda)$ are given as
  - $R = \int \phi(\lambda)\bar{r}(\lambda)\, d\lambda$
  - $G = \int \phi(\lambda)\bar{g}(\lambda)\, d\lambda$
  - $B = \int \phi(\lambda)\bar{b}(\lambda)\, d\lambda$
- Furthermore if $\phi_1(\lambda) \rightarrow R_1, G_1, B_1$ and $\phi_2(\lambda) \rightarrow R_2, G_2, B_2$, then the sum of the colors is $\phi_1(\lambda) + \phi_2(\lambda) \rightarrow R_1 + R_2, G_1 + G_2, B_1 + B_2$

CIE rg chromaticity space

CIE RGB chromaticity diagram

- $r = \frac{R}{R+G+B}$
- $g = \frac{G}{R+G+B}$
- $b = \frac{B}{R+G+B}$
- Transform: projection of the 3D RGB space to the plane, described by $R+G+B=1$



The CIE XYZ color space

Requirements for constructing the CIE XYZ color space

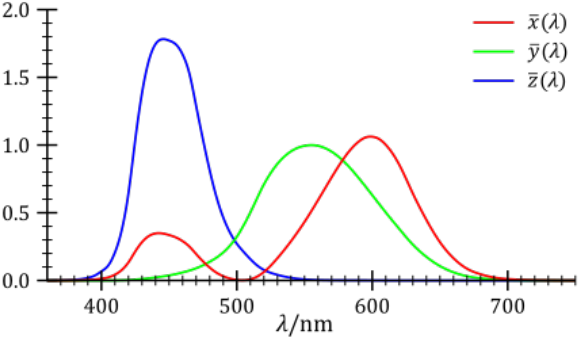
- no negative values present
- the color matching function $\bar{z}(\lambda)$ should be zero above 650 nm
- $\bar{y}(\lambda)$ should be exactly equal to the photopic luminous efficiency function $V(\lambda)$, hence Y would describe luminance
- constant-energy white point should be at $X = Y = Z$

Calculating CIE XYZ coordinates from CIE RGB

- by means of a linear transform

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \left[\underline{\underline{K}}\right] \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{CIE} = \begin{bmatrix} 2{,}77 & 1{,}75 & 1{,}13 \\ 1 & 4{,}59 & 0{,}06 \\ 0 & 0{,}06 & 5{,}6 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{CIE}$$

The CIE XYZ color matching functions



Calculation of XYZ coordinates:

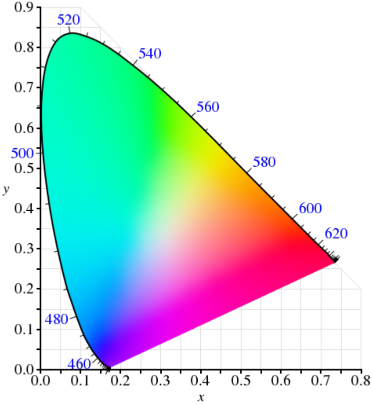
- $X = \int \phi(\lambda)\bar{x}(\lambda)\, d\lambda$
- $Y = \int \phi(\lambda)\bar{y}(\lambda)\, d\lambda$
- $Z = \int \phi(\lambda)\bar{z}(\lambda)\, d\lambda$

The CIE xy coordinate system

Derivation of CIE xy chromaticity coordinates

$x = \frac{X}{X+Y+Z}$ és $y = \frac{Y}{X+Y+Z}$

The CIE xy chromaticity diagram

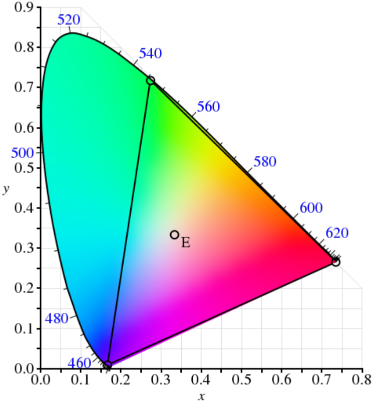


The CIE xy coordinate system

Derivation of CIE xy chromaticity coordinates

$x = \frac{X}{X+Y+Z}$ és $y = \frac{Y}{X+Y+Z}$

Position of illuminant E ($x = \frac{1}{3}, y = \frac{1}{3}$) and CIE RGB primaries



Visible colors in the CIE xy diagram

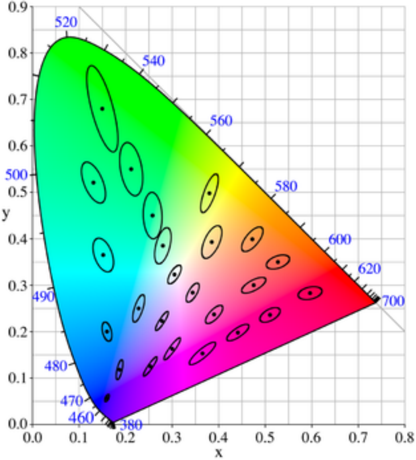
- Area of visible colors: bounded by spectral colors and the line of purples
- Achromatic region: colors, perceived as white in the central part of the diagram
- spectral colors: colours on the horseshoe-shaped curve on the outside of the diagram
- line of purples: the bottom straight line: no single wavelength produces purple color

MacAdam ellipse, CIE uv, CIE LAB

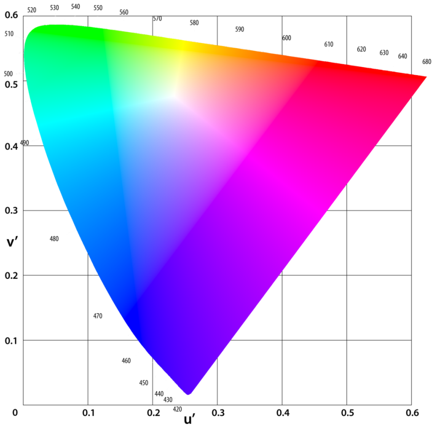
- The CIE XYZ coordinate system and xy diagram is *non-uniform*: Colors in the chromaticity diagram are not perceptually uniform: perceptual difference between color points is not proportional to coordinate-distances
- The minimal alteration in hue, perceived by the HVS: Just Noticable Difference: JND.
- Colors inside MacAdam ellipses are not distinguishable by average human eye
- In 1960 CIE introduced the Uniform Color Space, modified several times later on. Current form: CIE 1976 UCS
- Color spaces adapted more to human color vision: e.g. CIE LAB (based on opponent color models), and CIE Color Appearance Models.

MacAdam ellipse, CIE uv, CIE LAB

MacAdam ellipses



CIE Luv (UCS) chromaticity diagram



Primary colors of TV technology

- metamerism: perceived matching of the colors with different (nonmatching) spectral power distributions
- consequence: natural colors can be reproduced (in a perceived sense) as the mixture of properly chosen primaries (additive reproduction)
- this is the basic principle of color reproduction techniques (TV-tech., printing, photography, etc.)
- choosing 4 primaries: the obtained quadrilateral covers almost all the visible colors in the chromacity driagram
- Uncovered areas: mainly green colors, for which resolution of HVS is low $\rightarrow$ 3 primaries are enough for proper color reproduction

Primary colors of TV technology

- obvious choice: CIE RGB primaries for TV
- practical problem: no phosphors—used in cathode ray tubes—in this colors were available
- compromise: ensure both practicability and maximize covered area of colors

TV/video RGB primaries

Definition of RGB color space

An additive RGB system is specified by the chromaticities of its primaries and its white point. The extent – or gamut – of the colors that can be mixed from a given set of RGB primaries is given in the [x, y] chromaticity diagram by a triangle whose vertices are the chromaticities of the primaries.

RGB colos spaces

Név	White	x_R	y_R	x_G	y_G	x_B	y_B
CIE (1931) RGB	E	0.7347	0.2653	0.2738	0.7174	0.1666	0.0089
ITU-R BT.709 / sRGB	D65	0.64	0.33	0.30	0.60	0.15	0.06
NTSC (1953) / FCC 1953	C	0.67	0.33	0.21	0.71	0.14	0.08
NTSC (1987) / (SMPTE C / SMPTE 170M)	D65	0.63	0.34	0.31	0.595	0.155	0.07
PAL/SECAM (1970) / (EBU / ITU 601)	D65	0.64	0.33	0.29	0.60	0.15	0.06
Apple RGB	D65	0.625	0.34	0.28	0.595	0.155	0.07
ROMM RGB	D50	0.7347	0.2653	0.1596	0.8404	0.0366	0.0001

TV/video RGB color spaces I.

RGB primaries

With the development of CRT TV and phosphore technology the RGB primaries (first standard: NTSC/FCC 1953) underwent major changes. Currently: SMPTE C (USA), EBU (Europe) (ITU-601, SDTV) and ITU-709 (HDTV) recommendations are the most widespread.

Luminance (Y) from the RGB primaries

Relative luminance can be formed as a properly weighted sum of RGB tristimulus components

$$\begin{aligned} Y_{601} &= 0.299R \quad &+0.587G \quad &+0.114B \\ Y_{709} &= 0.212R \quad &+0.715G \quad &+0.072B \end{aligned}$$

comment: NTSC/FCC coefficients were adopted for the ITU-601 Rec., and the EBU rec. (with new RGB primaries and white point), so in the followings these coefficients are used

The color difference components

Luminance (Y) from the RGB primaries

Relative luminance can be formed as a properly weighted sum of RGB tristimulus components

$$Y = 0.3R + 0.6G + 0.11B$$

Rearranging the equation

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$

The color difference components

color-difference signals

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$

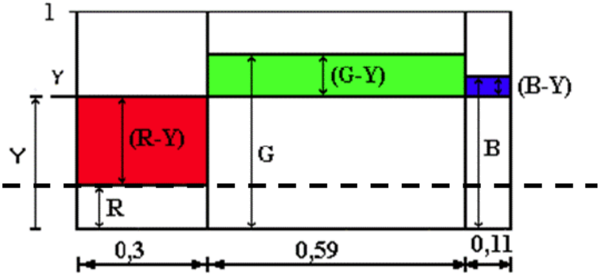
- The (R-Y), (G-Y) and (B-Y) the color difference components of video, with no luminance info
- Resolution of HVS is lower for color information → feasible to handle luminance and color info separated
- the color space of video:
  - one component: luminance
  - two components: two color difference components

The color difference components

Graphic interpretation

The two equations above can be illustrated in an area chart for a given RGB triplet. The area below the line of Y equals to the area above it.

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$



The color difference components

Some properties

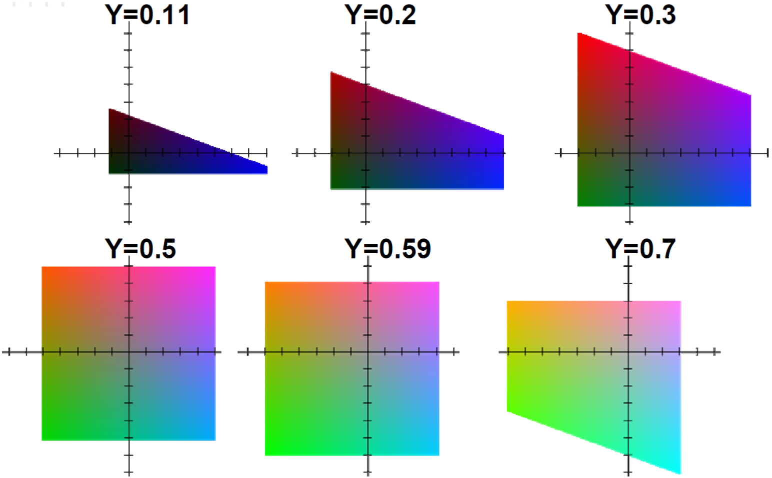
- signed quantities, on the area chart above the Y line: positive, below: negative
- if two of them equals zero → all of them zero
- in this latter case $R = G = B$: white point
- if we are not in the white point: at least two color differences are non-zero

The color difference components

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$

- Dynamic range of color differences:
  - $-0,7 \leq (R - Y) \leq +0,7$
  - $-0,89 \leq (B - Y) \leq +0,89$
  - $-0,41 \leq (G - Y) \leq +0,41$
- the dynamic range of green color difference is the smallest
- always the larger dynamic range is transmitted (for high SNR)
- therefore: color information is transmitted with $R - Y$ and $B - Y$
- frequently used depiction of colors: in 2D coordinate system with the basis $R - Y$ and $B - Y$ (e.g. vectorscope)

Example for B-Y,R-Y diagram



TV-Luminance, -hue, and -saturation

How can we define the above subjective quantities?

- straightforward choice: TV-luminance: Y component with the dynamic range of 0-1 (0:black, 1: white)
- RGB coordinates: the same dynamic range (see defining equation)

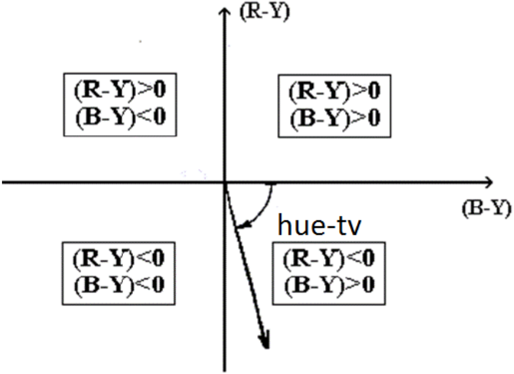
TV-Luminance, -hue, and -saturation

- $B - Y, R - Y$ representation of colors: at a fixed Y coordinate
- thus: for any $0 \leq Y \leq 1$ value a unique $B - Y, R - Y$ diagram can be drawn
- obviously, at a fixed Y not all the colors are present with 100% saturation, e.g. saturated blue: $[R, G, B] = [0, 0, 1]$ $Y = 0.11$, i.e. at a Y value of e.g. 0.5 only mixture of blue and white is present
- origin of diagram: white (grey), where $R - Y = B - Y = 0$ (and $G - Y = 0$)
- on half-lines, launched from the origin: mixtures of a saturated color and white
- hence on these half-lines: colors with the same hue but different saturation
- therefore: TV-hue definition: polar angle of the point $B - Y, R - Y$

TV-Luminance, -hue, and -saturation

- Hue$_{\mathrm{TV}}$:

$$\arctan \frac{(R-Y)}{(B-Y)}$$



TV-Luminance, -hue, and -saturation

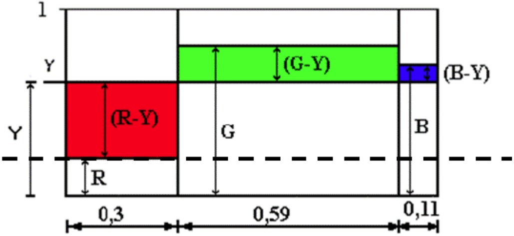
Quasi-spectral colors

- quasi-spectral colors: counterpart of spectral colors along with the line of purples in TV tech.
- common property: do not contain all three primaries: (only one or two R,G,B component describes them)
- hence, quasi-spectral colors are located on the edge or apex of the gamut (triangle of reproducible colors in a colorspace)

Saturation$_{\mathrm{TV}}$

- preconcept: Saturation$_{\mathrm{TV}}$ should be zero in the white point and 1 for a quasi-spectral color, i.e. at the edge of gamut
- one or two RGB components are zero: color point under discussion is on the edge of gamut
- no zero RGB component: saturation cannot be 1, because the point is somewhere inside the gamut
- i.e. saturation: describes how far is the color point from the edge of gamut, influenced mainly by smallest RGB component

Saturation$_{\mathrm{TV}}$

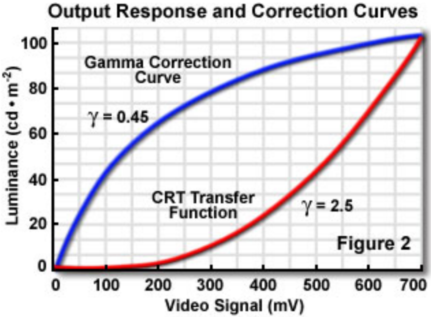


- Interpretation on area chart:
  - find the smallest RGB component (here: R)
  - Draw a horizontal line at its level, so that we divide the color under investigation into two parts:
  - we get R amount of white (R=G=B) and a quasi-spectral color, not containing R
  - the luminance of white: R, luminance of the quasi-spectral color: Y-R, which is $|R - Y|$
  - Saturation$_{\mathrm{TV}}$ is defined as the luminance of the quasi-spectral color and the luminance of white part:Saturation$_{\mathrm{TV}} = \frac{|\min(R,G,B) - Y|}{Y}$

Gamma correction

Gamma distortion of CRT

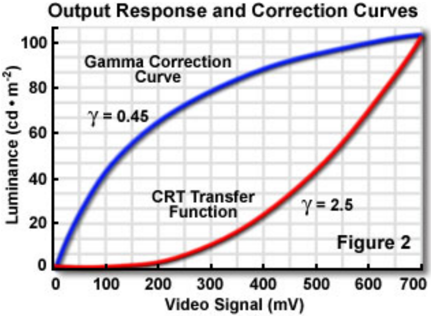
- Luminance of cathode ray tube screens was non-linear function of driving voltage (triode) characteristics: $Y = k\ U^{\gamma}$. This non-linear characteristic has/had to be compensated for
- principle of correction: before transmission the signal is pre-distorted with an inverse power function
- this gamma corrected signal will flow through the whole transmission path



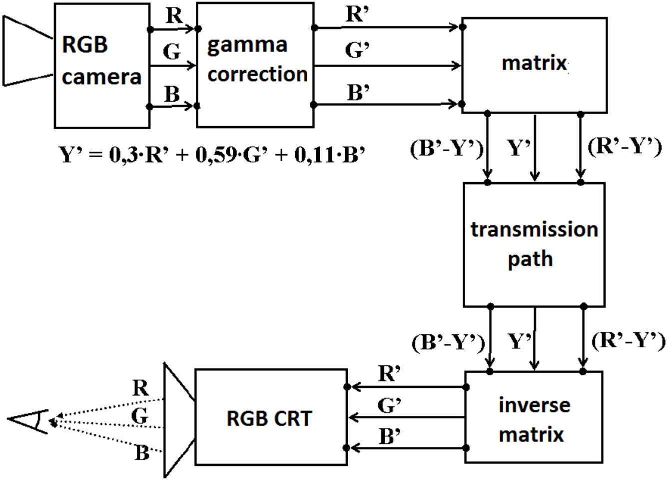
Gamma correction

Gamma distortion of CRT

- in practice: first process the gamma-corrected RGB tristimulus signals (R'=$R^{\frac{1}{\gamma}}$, G'=$G^{\frac{1}{\gamma}}$ and B'=$B^{\frac{1}{\gamma}}$)
- then: from these gamma compensated components the "gamma corrected " luminance and color differences are calculated (matrixed) (Y', R'-Y', B'-Y' respectively, where ' denotes gamma correction)

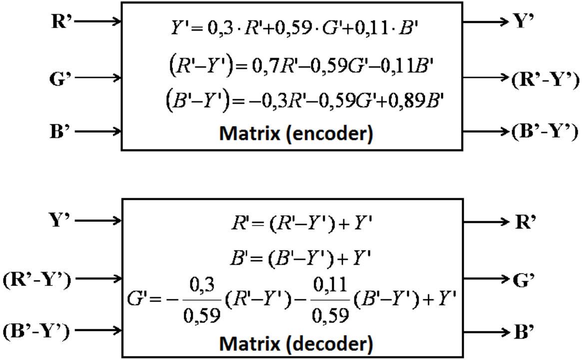


Implementation of Gamma-correction



$Y' = 0{,}3 \cdot R' + 0{,}59 \cdot G' + 0{,}11 \cdot B'$

Matrix and inverse matrix



R’

G’

B’

$$Y' = 0{,}3 \cdot R' + 0{,}59 \cdot G' + 0{,}11 \cdot B'$$

$$(R' - Y') = 0{,}7R' - 0{,}59G' - 0{,}11B'$$

$$(B' - Y') = -0{,}3R' - 0{,}59G' + 0{,}89B'$$

Matrix (encoder)

Y’

(R’-Y’)

(B’-Y’)

Y’

(R’-Y’)

(B’-Y’)

$$R' = (R' - Y') + Y'$$

$$B' = (B' - Y') + Y'$$

$$G' = -\frac{0{,}3}{0{,}59}(R' - Y') - \frac{0{,}11}{0{,}59}(B' - Y') + Y'$$

Matrix (decoder)

R’

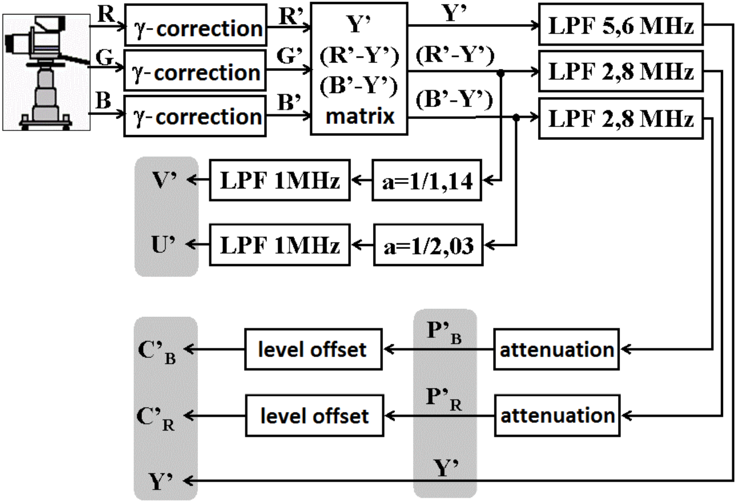
G’

B’

Video components

- Y’ component: weighted sum (linear combination) of non-linearly distorted RGB tristimulus signals
- terminology: Y’: **luma**, R’-Y’, B’-Y’: **chroma signals**
- in component video systems: 3 components are stored, transmitted and handled independently
- HVS resolution is lower for colors: the luma (Y’) component is captured at full bandwidth and the chroma components are bandlimited (to the half of Y’-s bandwidth, or even lower)

SD video-components



Video component formats

- Y’U’V’: composite format in PAL and NTSC (Y’I’Q’) systems. Signals U’ and V’ are derived from (B’-Y’) and (R’-Y’) with rescaling and correction, so that the modulated chroma signal, added to the luma would fit the dynamic range of the interface. Today it is more or less obsolete, used only for analog video broadcast
- Y’ Pb’ Pr’: analog component system, the chroma signals are rescaled/corrected and bandlimited to the half of the luma signal’s bandwidth with analog filters. The actual interface varies among manufacturers
- **Y’ Cb’ Cr’**: digital component system, the chroma signals are rescaled/corrected so it would fit to the dynamic range of A/D conversion, then sub-sampled digitally according to the choma-subsampling structure (e.g. 4:2:2, 4:2:0, see later)

Video signal dynamic ranges I.

Component	Definition	Dynamic range
Y’ =	0,299·R’ + 0,587·G’ + 0,114·B’	0 … 1
V’ =	(R’-Y’)/1,14 = 0,877·(R’ - Y’)	-0,62 … +0,62
U’ =	(B’-Y’)/2,03 = 0,493·(B’ - Y’)	-0,44 … +0,44

Dynamic range of Y' 0-1 in voltage : 0 - 700 mV

Video signal dynamic ranges II.

Component	Definition	Dynamic range
Y'	$0{,}299 \cdot R' + 0{,}587 \cdot G' + 0{,}114 \cdot B'$	$0 – 1$
P'_R	$(0{,}5/0{,}7) \cdot (R' - Y') = 0{,}713 \cdot (R' - Y')$	$-0{,}5 \ \ldots \ +0{,}5$
P'_B	$(0{,}5/0{,}89) \cdot (B' - Y') = 0{,}564 \cdot (B' - Y')$	$-0{,}5 \ \ldots \ +0{,}5$
C'_R	$P'_R + 0{,}5 = 0{,}713 \cdot (R' - Y') + 0{,}5$	$0 – 1$
C'_B	$P'_B + 0{,}5 = 0{,}564 \cdot (B' - Y') + 0{,}5$	$0 – 1$

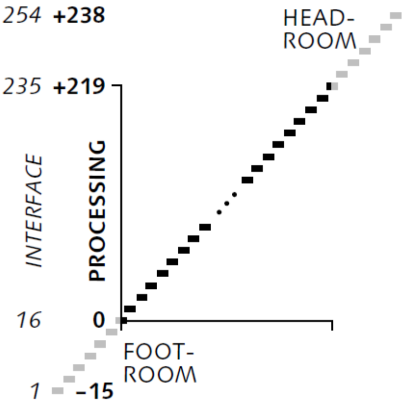
Dynamic range of Y' in voltage:	0 - 700 mV
Dynamic range of P'B, P'R in voltage:	+- 350 mV
Dynamic range of C'B, C'R in voltage:	0 - 700 mV

YCbCr digital dynamic ranges

- in case of representing YCbCr on 8 bits, it would be straightforward to assign the range 0..255 to the analog range 0...1 of the Y’ luma signal.
- this is only used in several JPEG/JFIF standards
- studio video standards provide footroom below reference black, and headroom above reference white because of possible overshoots, that are liable to result from processing by digital and analog filters
- Black level for Y’: 16 (offset), white level: 16+219=235
- Zero level of Cb/Cr: 128 (offset), max level 128+ 112= 240, min. level: 128-112=16

YCbCr digital dynamic ranges

Figure 2.4 **Footroom and headroom** are provided in digital video standards to accommodate filter undershoot and overshoot. For processing, black is assigned to code 0; in an 8-bit system, *R′*, *G′*, *B′*, or luma (*Y′*) range 0 through 219. At an 8-bit interface according to Rec. 601, an offset of +16 is added (indicated in italics). Interface codes 0 and 255 are reserved for synchronization; those codes are prohibited in video data.



No. of bits/sample and quantizer characteristics I.

Perceptual aspects (again)

- in about two decades range the contrast sensitivity of HVS 1%:
- in this range: vision cannot distinguish two luminance levels if the ratio between them is less than about 1.01
- thus: contrast sensitivity of HVS is nearly logarithmic
- in practice: contrast sensitivity is approximated by power functions (e.g. CIE L^*)

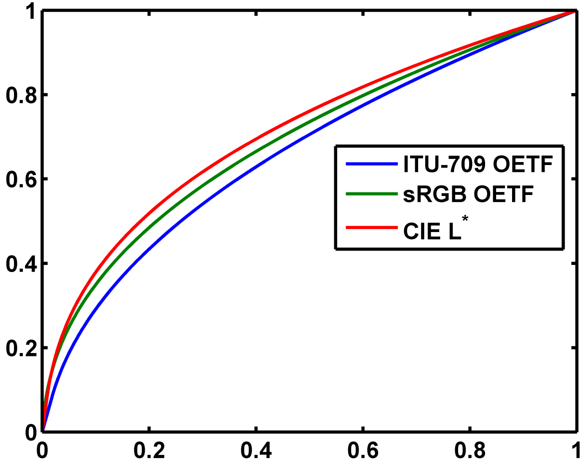
No. of bits/sample and quantizer characteristics I.

CIE Lightness (L^*) definition

$$L^* = \begin{cases} \left(\frac{29}{3}\right)^3 Y/Y_n, & Y/Y_n \leq \left(\frac{6}{29}\right)^3 \\ 116\,(Y/Y_n)^{1/3} - 16, & Y/Y_n > \left(\frac{6}{29}\right)^3 \end{cases}$$

The entire curve (power function of $1/3$ and the linear) corresponds to approximately a power function of $Y^{0.4}$

No. of bits/sample and quantizer characteristics II.



No. of bits/sample and quantizer characteristics III.

- consider pixel values proportional to luminance, where code zero represents black, and the maximum code value of 255 represents white
- the boundary between a region of code 100 samples and a region of code 101 samples is likely to be visible (1%)
- codes below 100: the difference in luminance between adjacent codes becomes increasingly perceptible (e.g: 21/20= 105 %): visible jumps in luminance produce artifacts known as contouring or banding.
- codes above 100 suffer no banding artifacts, however, as code value increases toward white, the codes have decreasing perceptual utility (e.g. 201/200=0.01 %): not efficient

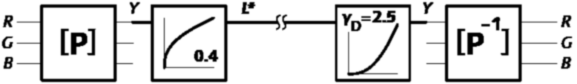
No. of bits/sample and quantizer characteristics III.

- codes below 100: the difference in luminance between adjacent codes becomes increasingly perceptible (e.g: 21/20= 105 %): visible jumps in luminance produce artifacts known as contouring or banding.
- codes above 100 suffer no banding artifacts, however, as code value increases toward white, the codes have decreasing perceptual utility (e.g. 201/200=0.01 %): not efficient
- this is the **code-100 problem**
- solution: quantization characteristic should take luminance-lightness transfer of HVS into consideration (e.g. CIE L^*)

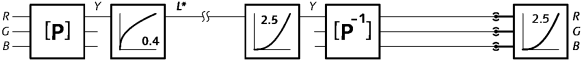
No. of bits/sample and quantizer characteristics IV.

perceptual quantization

- Theoretical solution of perceptual based quantization



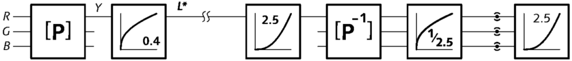
- Inclusion of γ-distortion of CRT displays not solved:



Gama distortion: not Y component is displayed

CRT GAMMA

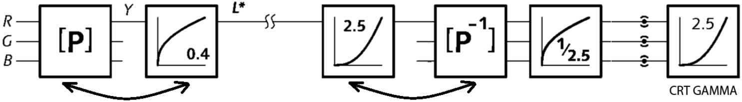
impractical and expensive



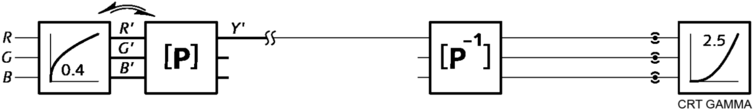
CRT GAMMA

No. of bits/sample and quantizer characteristics V.

perceptual quantization



- lucky coincidence: luminance-lightness characteristics of HVS = inverse of CRT γ-distortion (power function of 0.4)
- modified, final system block diagram:



No. of bits/sample and quantizer characteristics V.

perceptual quantization

- role of gamma correction nowadays: not the compensation of CRT characteristics, but ensuring quantization, adapted to HVS characteristics, in case of digital representation
- for modern displays (LCD-TFT, etc): driving voltage-luminance transfer not identical to that of CRTs → gamma correction is neutralized with an artificial power function of 2.5 (e.g. based on a look up table)

No. of bits/sample and quantizer characteristics VI.

perceptual quantization

- Effeciency of quantization adapted to HVS (to perception of brightness) is better
- Linear quantization to 8 bits: not feasible (banding at dark shades)
- Aim: quantize dynamic range of 100:1, ensuring resolution of 1 %
- Linear quantization: 100x100 = 10.000 code words are required → 14 bit linear quantization
- Logarithmic quantization: ratio of adjacent codes = 1.01 → $1.01^x = 100$, $x = 463$, coded with 9 bits
- Gamma-distortion: power function of 0.4 → +1 bit → 10 bits are required

No. of bits/sample and quantizer characteristics VI.

perceptual quantization

- Logarithmic quantization: ratio of adjacent codes = 1.01 → $1.01^x = 100$, $x = 463$, coded with 9 bits
- Gamma-distortion: power function of 0.4 → +1 bit → 10 bits are required
- Original studio formats used non-linear quantization at 8 bits
- Nowadays: 10 bits non-linear quantization is general (sometimes even more, see UHDTV)

No. of bits/sample and quantizer characteristics VII.

Standardized transfer characteristics of Rec. ITU-709

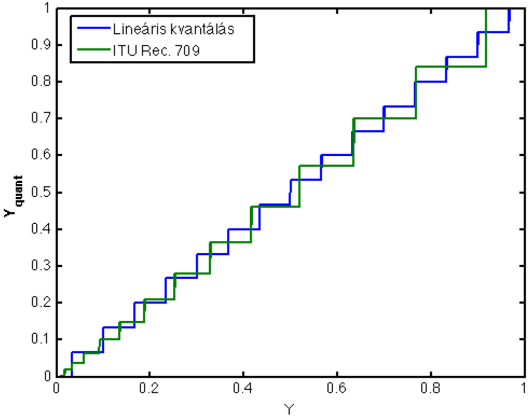
- terminology: Optoelectronic Transfer Function (OETF)

$$V' = \begin{cases} 4.500\,V & V < 0.018 \\ 1.099\,V^{0.45} - 0.099 & V \geq 0.018 \end{cases}$$

- with V denoting R,G,B components
- comment 1: entire curve: power function of 0.5 (differs from 0.4 for the compensation of Hunt effect)
- comment 2: linear part: noise suppression (otherwise curve slope would be ∞ around origin)

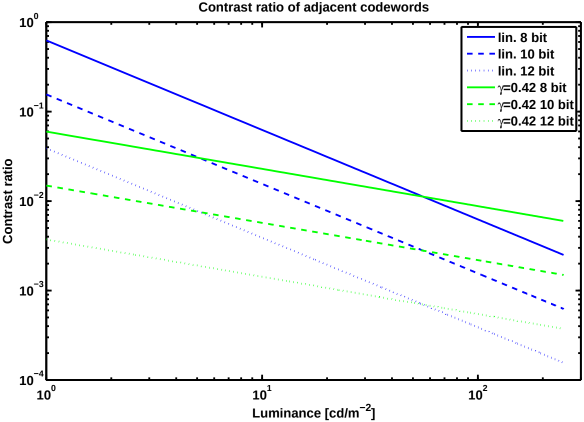
No. of bits/sample and quantizer characteristics VII.

Example for ITU-709 quantization (for didactic purpose only: on 4 bits)



Contrast ratio of adjacent codes

At linear OETF and power function of 1/2.4=0.42 OETF



Chroma subsampling

Principle

Analog component systems (YUV, YPbPr): bandwidth limitation of chroma components → digital system: subsampling (decimating) the chroma components

Subsampling schemes

- $J : a : b$: alpha
  - J: horizontal sampling reference (width of the conceptual region). Usually, 4.
  - a: number of chrominance samples (Cr, Cb) in the first row of J pixels.
  - b: number of changes of chrominance samples (Cr, Cb) between first and second row of J pixels.
  - Alpha: horizontal factor (relative to first digit). May be omitted if alpha component is not present, and is equal to J when present.

Chroma subsampling

Principle

Analog component systems (YUV, YPbPr): bandwidth limitation of chroma components → digital system: subsampling (decimating) the chroma components

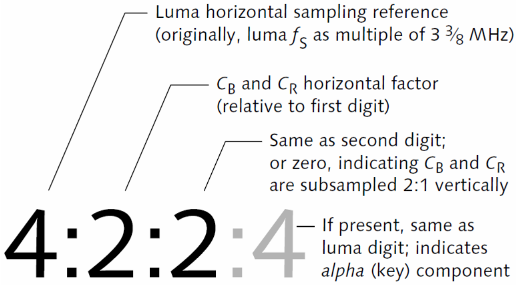
Subsampling schemes

- e.g. 4:2:0 or 4:2:2
  - *J*: horizontal sampling reference (width of the conceptual region). Usually, 4.
  - Subsampling ratio of chroma samples in the horizontal direction (4:2= 2:1 subsampling horizontally)
  - 0, or identical to previous one. 0 means vertical subsampling at the rate of 2:1, identical to the previous one: no vertical subsampling

Chroma subsampling

Subsampling schemes

Figure 10.3 **Chroma subsampling notation** indicates, in the first digit, the luma horizontal sampling reference. The second digit specifies the horizontal subsampling of C_B and C_R with respect to luma. The third digit originally specified the horizontal subsampling of C_R. The notation developed without anticipating vertical subsampling; a third digit of zero now denotes 2:1 vertical subsampling of both C_B and C_R.



4:2:2:4

Luma horizontal sampling reference (originally, luma f_S as multiple of 3 $\frac{3}{8}$ MHz)

C_B and C_R horizontal factor (relative to first digit)

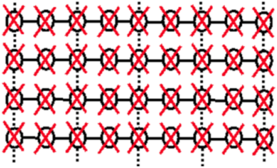
Same as second digit; or zero, indicating C_B and C_R are subsampled 2:1 vertically

If present, same as luma digit; indicates *alpha* (key) component

4:4:4 subsampling scheme

X : Cr', Cb' samples (chroma)

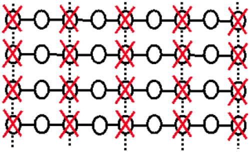
O : Y' samples (luma)



- Sampling structure is identical in each row
- Unit: 1 pixel
- 8, 10 bit/sample/component
- in case of 8 bits: 3 bytes (24 bits)
- in case of 10 bits: 3 *10 = 30 bits

4:2:2 ITU-601

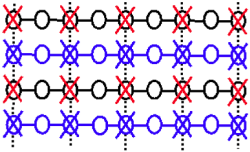
4:2:2 progressive



✕ : C'_R, C'_B chroma samples

○: luma samples

4:2:2 interlaced



✕ : chroma samples, field 1.

○ : luma samples, field 1.

✕ : chroma samples, field 2.

○ : luma samples, field 2.

8, 10 bits/sample/component

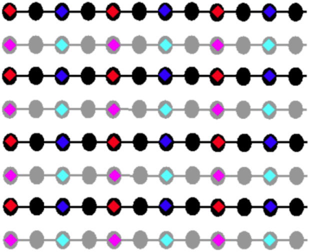
8 bits: 2+1+1 bytes: 32 bits/2pixels

10 bits: (2+1+1)*10 bits = 40 bits/2pixels

4:4:4 to 4:2:2 subsampling: compression factor of 3:2

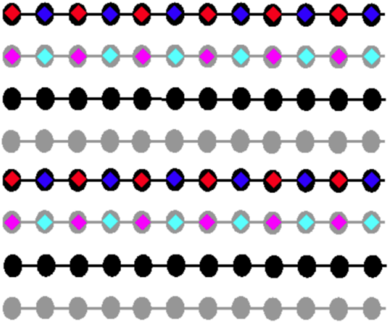
4:1:1 DV-USA

- each row has the same structure
- unit: 4 pixels in a row
- chorma components is subsampled horizontally by factor of 4
- remaining chroma samples coincide vertically with a Y' sample
- C'_R and C'_B samples are shifted
- 8 bits: 4+1+1 = 6 bytes = 48 bits / 4 px
- 10 bits = (4+1+1)*10 = 60 bits / 4 px
- compression factor from 4:2:2 = 4 : 3



◆ : C'_R samples field 1.
◆ : C'_B samples field 1.
● : Y' samples field 1.
◇ : C'_R samples field 2.
◇ : C'_B samples field 2.
○ : Y' samples field 2.

4:2:0 DV-EU (interlaced)



◆ : C'_R samples field 1.

◆ : C'_B samples field 1.

● : Y' samples field 1.

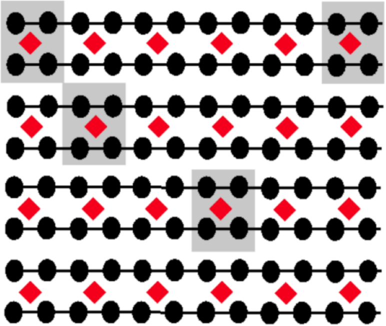
◆ : C'_R samples field 2.

◆ : C'_B samples field 2.

● : Y' samples field 2.

4:2:0 MPEG-1 (progressive) and JPEG

- unit: 2x2 block of pixels
- chroma component subsampled horizontally and vertically by a factor of 2
- chroma samples: positioned vertically between to lines, horizontally between 2 luma samples
- 8 bits: (4+1+1) byte = 48 bits/4px
- 10 bits: (4+1+1)*10 = 60 bits/4px
- Compression factor from 4:2:2 = 4 : 3

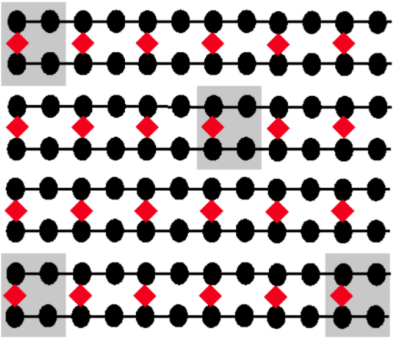


◆ : C'_R, C'_B samples

● : Y' samples

▨ : unit

4:2:0 MPEG-2



◆ : C’$_R$, C’$_B$ samples

● : Y' samples

▨ : unit

- unit: 2x2 block of pixels
- chroma component subsampled horizontally and verticall (by 2)
- chroma samples positioned vertically between two lines (filter length is odd)
- chroma samples coincide horizontally with Y' samples (horizontal filter length is even)
- 8 bits: (4+1+1)byte = 48 bits/4px
- 10bits: (4+1+1)*10 = 60 bits/4px
- Compression factor from 4:2:2 = 4:3

TV raster formats

- Field of view of HVS: 9 – 10° (main FOV, without pheripheral vision)
- Aspect ratio of FOV: circa 4:3, (larger for the horizontal dimension)
- Average spatial resolution of HVS is limited (see earlier)

TV raster formats

- aim for choosing resolution: pixel structure should not be noticeable (smaller than HVS resolution)
- angle between the light rays from adjacent pixels should be smaller than the angular resolution of HVS (about $1/60°$)
- color picture: if pixel resolution is high enough, additive color mixing is done by the HVS (not has to be done within the display)
- in order to have the angular resolution smaller than $1/60°$: display has to be watched from a minimal viewing distance, defined by the distance of pixels/rows
- this is the starting point for the development of each TV raster format standard (SDTV, HDTV, UHDTV)

Frame rate and scanning method

- in order to avoid flickering the rate of changing the picture content should be higher than flicker fusion threshold ($50 - 60$ Hz)
- under fusion threshold: flickering is perceived
- phi phenomenon/beta movement: is the optical illusion of perceiving a series of still images, when viewed sequentially as continuous motion
- in order to ensure illusion of continuous motion $10 - 20$ displayed phase of motion per second is enough
- thus: displaying the same picture 2-3 times: no flickering perceived and motion looks continuous
- In cinemas: dual-, or triple blade shutters: 24fps $\rightarrow$ 48 and 72fps

Frame rate and scanning method

- in order to avoid flickering the rate of changing the picture content should be higher than flicker fusion threshold $(50 - 60\ \mathrm{Hz})$
- in order to ensure illusion of continuous motion $10 - 20$ displayed phase of motion per second is enough
- thus: displaying the same picture 2-3 times: no flickering perceived and motion looks continuous
- In TV technology:
  - transmitting the same frame multiple times: low bandwidth efficiency
  - storing one entire frame: was not realizable at the time of analog technology
  - solution: interlaced scanning

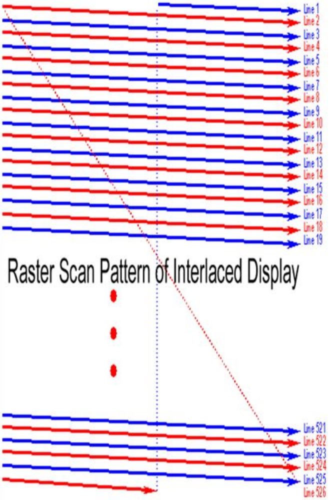
Interlaced scanning

- Basic idea of interlaced video:
- Frame is partitioned into two fields
  - odd lines of frame: odd field
  - even lines of frame: even field
- The entire frame changes with a frame rate of 25-30 Hz
- Fields are transmitted at a two times higher rate (at 50-60 Hz): first the odd, then the even fields → flickering is avoided

Interlaced scanning

Raster Displays

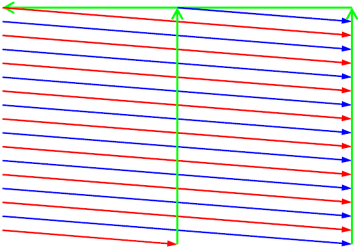
- Electron beam traces over screen in **raster scan order**
- Each left –to right trace is called a **scan line**
- Each spot on the screen is a **pixel**
- When the beam is turned off to sweep back, that is **a retrace, or a blanking interval**
- B/W TVs are basically oscilloscopes (with a hardwired scan pattern)
- Entire screen painted 30 times/sec
- Screen is traversed 60 times/sec
- Even/Odd lines on alternate scans (called fields)
- Smooth motion on dynamic scenes
- High Resolution on static scenes
- Optimize bandwidth



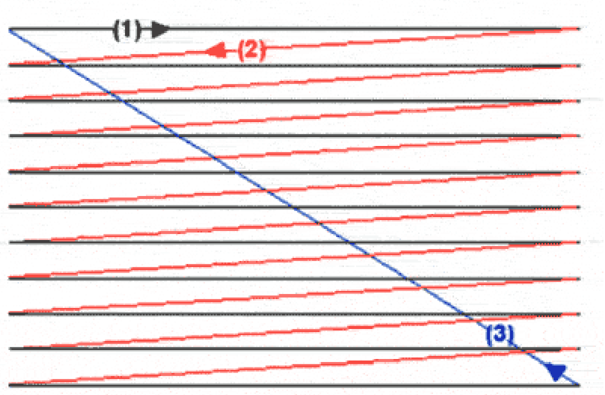
Raster Scan Pattern of Interlaced Display

Interlaced scanning with odd no. of lines

- only one type of horizontal and vertical retrace
- Whole frame consists of odd no. of lines, therefore
  - last line ends at the center of frame
  - at that position: vertical retrace (in the **blanking time/interval**)
  - next first line: also a half line



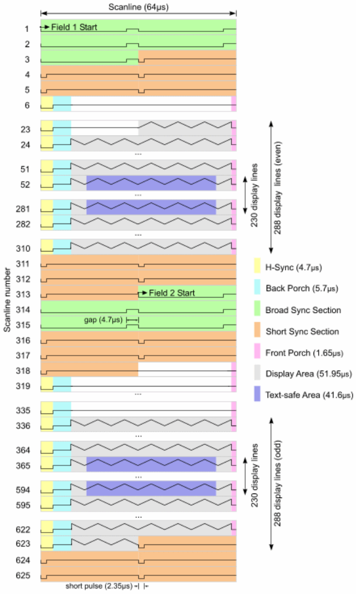
Progressive scanning



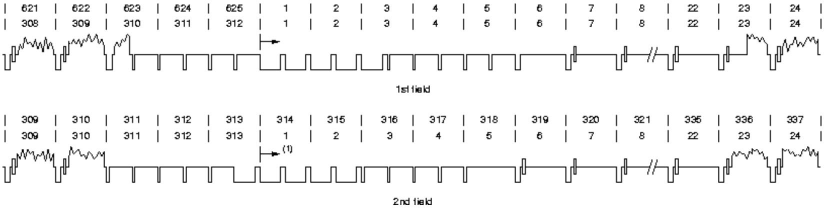
SD formats

- originally field rate was chosen based on the standard mains/line voltage frequency ("from the network socket") (30 Hz for America, 50 Hz for Europe)
- Effects of voltage ripple on the displayed video: if supply frequency differs from field/frame rate: additive moving noise picture
- if two frequencies equal: still noise picture (non-moving): this type of noise is less disturbing
- **USA, Japan**: 525 lines per frame (480 active), 60 Hz field, 30 Hz frame rate, 525x30=15750 Hz line frequency. notation: 480i
- **Europe, Asia**: 625 lines per frame (576 active lines carrying picture content), 50 Hz field rate, 25 Hz frame rate, 625x25=15625 Hz line frequency, format notation: 576i
- Number of active columns: 720 (704). **Non-square pixels!**

Structure of analog SD TV signal



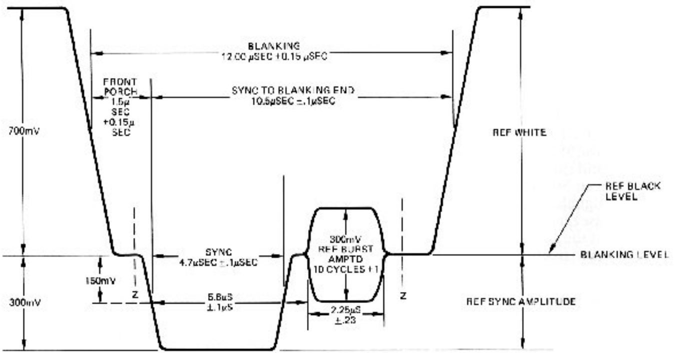
Structure of analog SD TV signal



- Synchronisation pulses ensure that the video image is locked on a video monitor (or VCR etc) vertically and horizontally without any jitter or rolling.
- If vertical sync is lost, picture may move/tear vertically.
- If horizontal sync is lost, picture may move/tear horizontally.
- Synchronisation pulse ensure that all the equipment used in a CCTV set up like VCRs, Monitors, Multiplexers and cameras are locked together

Structure of analog SD TV signal

Horizontal synchronization:



- Ref. burst: to aid separation of colour signals from composite video signal , a colour burst signal is added to the back porch of Horizontal synchronization signal

Sampling frequency of SD video I.

Requirement

- common sampling frequency is required for European (PAL) and USA (NTSC) systems
- orthogonal sampling structure: whole number of sampling period (pixel) in a TV-line
- $f_s = n\ f_h^{EU}$ és $f_s = k\ f_h^{USA}$, where n, k are integers, f_h: line frequency
- $f_h^{EU} = 625 \times \frac{50}{2}\ \mathrm{Hz} = 15625\ \mathrm{Hz} = 5^6\ \mathrm{Hz}$
- $f_h^{USA} = 525 \times \frac{60}{2} \times \frac{1000}{1001}\ \mathrm{Hz} = 15734, 2\dots\ \mathrm{Hz}$
- $1001 = 7 \times 11 \times 13$

Sampling frequency of SD video II.

- Least common multiple: $144 f_h^{USA} = 143 f_h^{EU} = 2,25\mathrm{MHz}$
- Nyquist criterion: sampling frequency should be larger, than twice the bandwidth of video signal (6 MHz)
- thus: we are looking for the smallest multiple of 2.25 MHz, greater than 12 MHz: *13.5 MHz!!!*
- sampling frequency of luma signal: 13.5 MHz,
- by taking the properties of HVS into consideration chroma is bandlimited to half of the luma bandwidth $\rightarrow$ chroma sampling frequency: 6.75 MHz

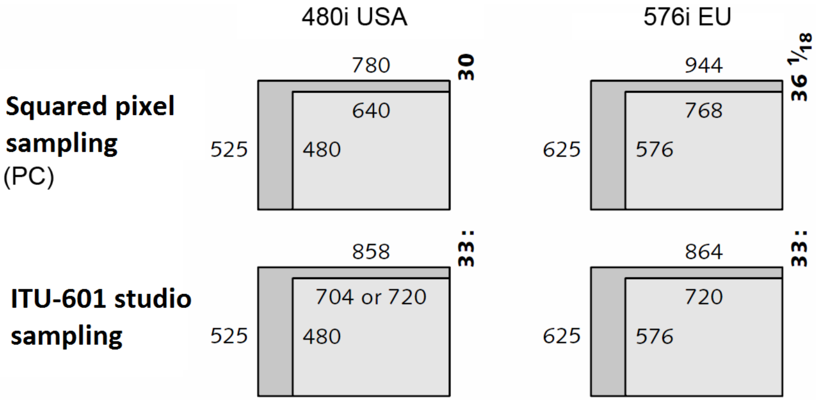
Sampling frequency of SD video II.

$$\begin{tikzcd}
f_{mv}^{Y} = 13{,}5MHz \arrow[r, Rightarrow, "/2"] \arrow[d, Rightarrow] & f_{mv}^{U,V} = 6{,}75MHz \arrow[d, Rightarrow] \\
f_{B}^{Y} = 5{,}6MHz & f_{B}^{U,V} = 2{,}8MHz
\end{tikzcd}$$

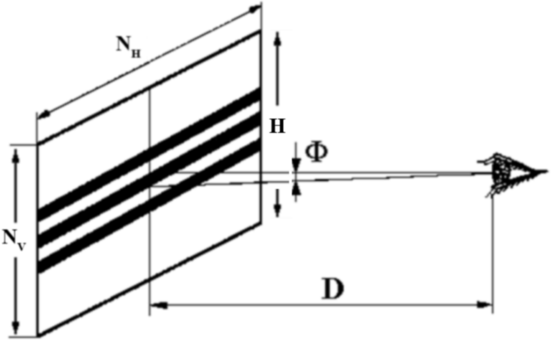
$$n_{H}^{USA} = \frac{f_{mv}^{Y}}{f_{H}^{USA}} = 858 \qquad n_{H}^{EU} = \frac{f_{mv}^{Y}}{f_{H}^{EU}} = 864$$

- Active interval/period of a TV-line choosen to be $53,3\ \mu\mathrm{sec} \rightarrow$ 720 samples/pixels per line
- due to transients in signal history real active content: 704 samples (rise time/fall time: 8-8 samples)
- This is common for EU and USA systems: therefore in case of format conversion no need of horizontal conversion

Studio vs. PC raster format



Optimum viewing distance I.



- height and width of picture in pixels: N_V and N_H, height in meters: H
- angle between rays from adjacent pixels at the viewer's eye: Φ
- viewing distance: D
- aim: $\tan(\frac{1}{60}) = \frac{H/N_V}{D}$

Optimum viewing distance II.

If $\tan(\frac{1}{60}) = \frac{H/N_V}{D}$, than

- in EU SD system (576i)
  $D_{576} = \frac{1}{N_V \tan(\frac{1}{60})} H = \frac{1}{576 \times 2.9 \times 10^{-4}} H \approx 6H$
- 720p HD system (720 lines)
  $D_{720} = \frac{1}{N_V \tan(\frac{1}{60})} H = \frac{1}{720 \times 2.9 \times 10^{-4}} H \approx 5H$
- 1080i/p HD system (180 lines)
  $D_{1080} = \frac{1}{N_V \tan(\frac{1}{60})} H = \frac{1}{1080 \times 2.9 \times 10^{-4}} H \approx 3H$

Horizontal field of view (Θ)

$\tan(\frac{\Theta}{2}) = \frac{\frac{A}{B}/2}{D}$, and from this
$\Theta = 2\arctan(\frac{A}{2BD})$, where $\frac{A}{B}$ is the image aspect ratio (SD: 4:3, HD: 16:9)

Optimum viewing distance III.

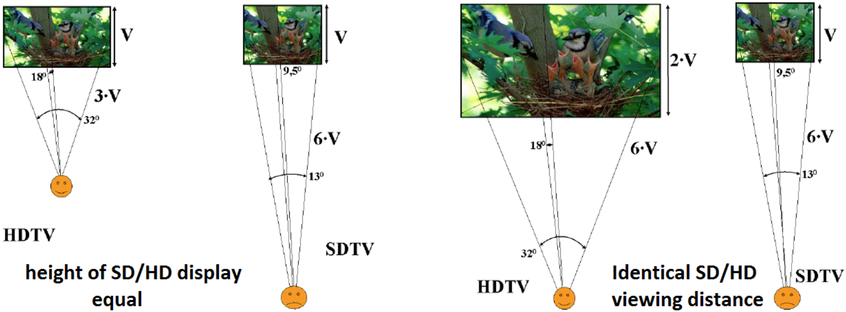
	USA SD	EU SD	HDTV
No. of lines	480	576	1080
Viewing distance	7 times height	6 times height	3 times height
Viewing distance	4.25 x diagonal	3.6 x diagonal	1.5 x diagonal
Horizontal FOV ≈	11°	13°	32°
Vertical FOV	8°	9.5°	18°

HDTV

Motivation and principles of HDTV format

- the primary motivation behind creating HDTV format increasing the "perceived reality" of the reproduced scene, mainly by filling an increased field of view with video content compared to SDTV
- thus the goal is not to squeeze six times the number of pixels into the same visual angle! Instead, the angular subtense of a single pixel should be maintained, and the entire image may now occupy a much larger area of the viewer's visual field. HDTV allows a greatly increased picture angle.

HDTV / SDTV comparison I.



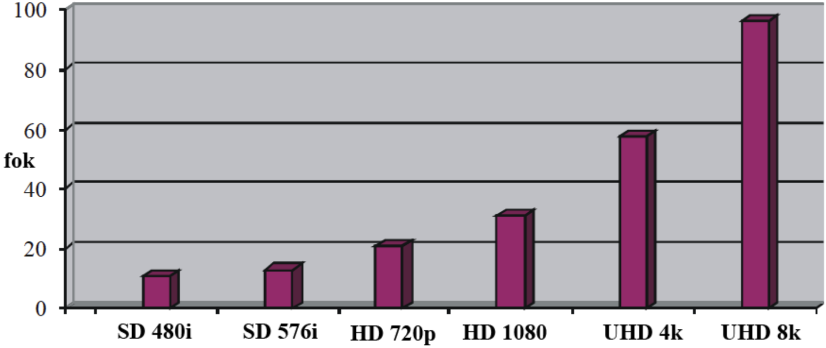
height of SD/HD display equal

Identical SD/HD viewing distance

thus the aim is not to squeeze six times the number of pixels into the same visual angle! Instead, the angular subtense of a single pixel should be maintained, and the entire image may now occupy a much larger area of the viewer’s visual field (**filling even the pherpheral vision**)

Horizontal Field of View

Horizontal FOV in case of optimum viewing distance



HDTV: Required display size

Suppose a fixed viewing distance: 3m (typical distance in living rooms (measured))

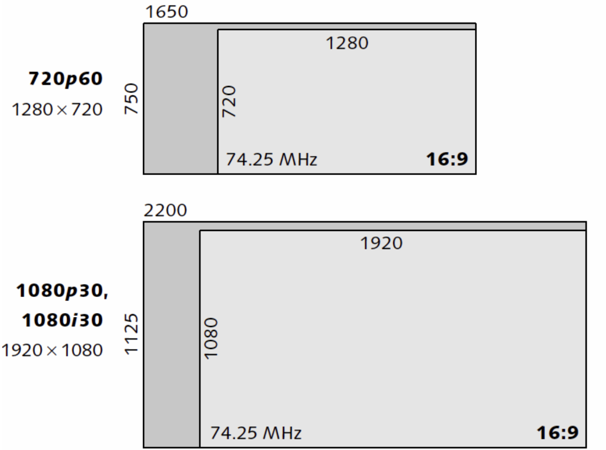
Image height: $H = DN_V \tan \frac{1}{60}$

Image width: $W = H \times \frac{A}{B}$,

where $\frac{A}{B}$ is the aspect ratio (SD: 4:3, HD: 16:9)

- for EU SD 576i system: H=50cm, W=67cm
- for 720 line HD system: H= 62cm, W= 110cm
- for 1080 line HD system: H=94 cm, W=168cm

HDTV formats



The HDTV RGB format

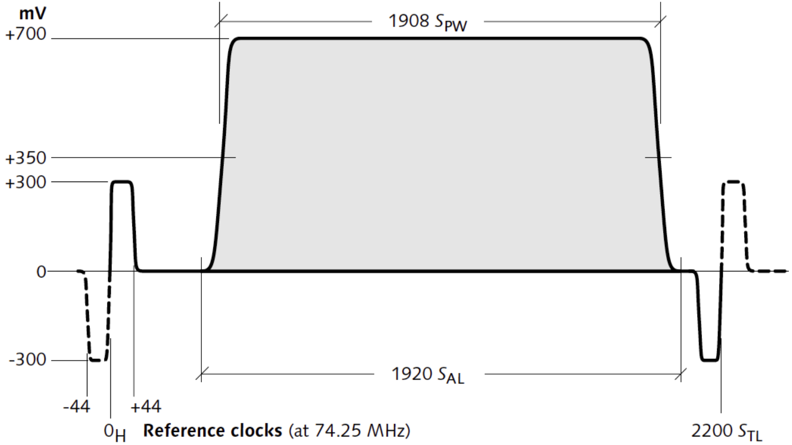
The RGB color space and gamma characteristics are standardized in ITU Rec.709

- Standardized luma coefficients:
  $Y_{709} = 0.212R + 0.715G + 0.072B$
- Overall opto-electronic transfer characteristics:

$$V' = \begin{cases} 4.500\,V & V < 0.018 \\ 1.099\,V^{0.45} - 0.099 & V \geq 0.018 \end{cases}$$

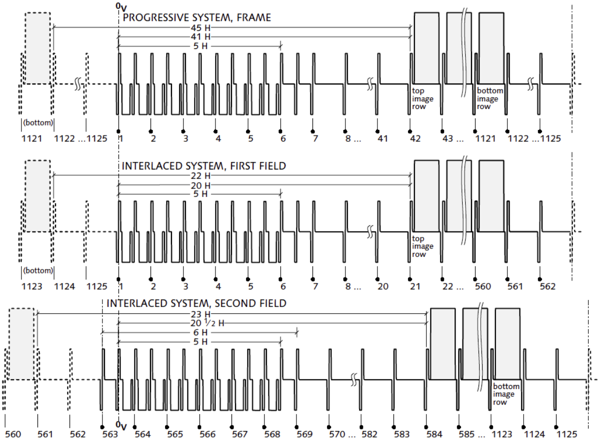
Structure of one line in analog HD 1080i signal

1080/30p, and 1080/60i example

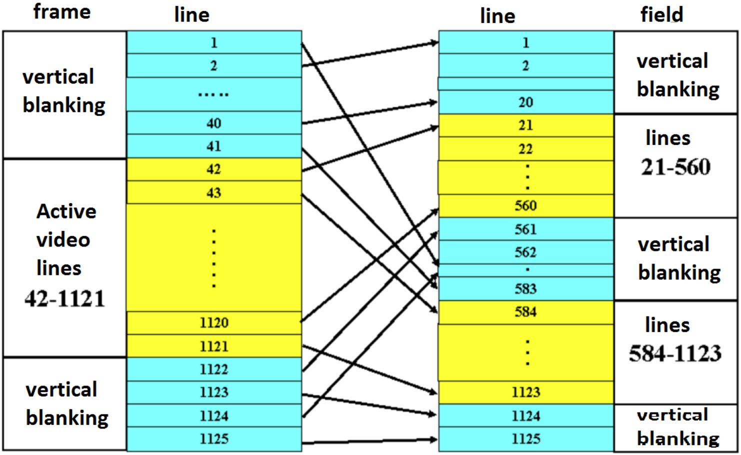


Structure of analog HD 1080 frame/field

Progressive and interlaced case:



Connection between HD frame and field lines



HDTV YCbCr code ranges

Identical with ITU-601

$$R'_{digi} = 219 \cdot D \cdot R' + 16 \cdot D$$

$$G'_{digi} = 219 \cdot D \cdot G' + 16 \cdot D$$

$$B'_{digi} = 219 \cdot D \cdot B' + 16 \cdot D$$

$$Y'_{digi} = 219 \cdot D \cdot Y' + 16 \cdot D \qquad D = 2^{n-8}$$

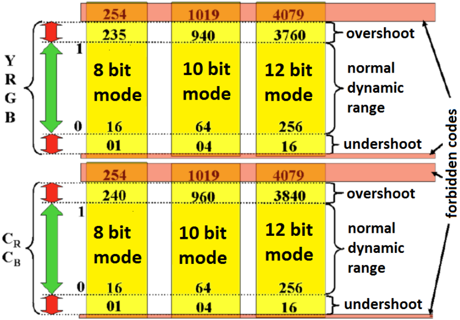
$$C'_{R\,digi} = 224 \cdot D \cdot C'_{R} + 128 \cdot D$$

$$C'_{B\,digi} = 224 \cdot D \cdot C'_{B} + 128 \cdot D$$

R',G',B',Y' dynamic range : 0 - 1

C'_R, C'_B dynamic range here : -0,5 – +0,5

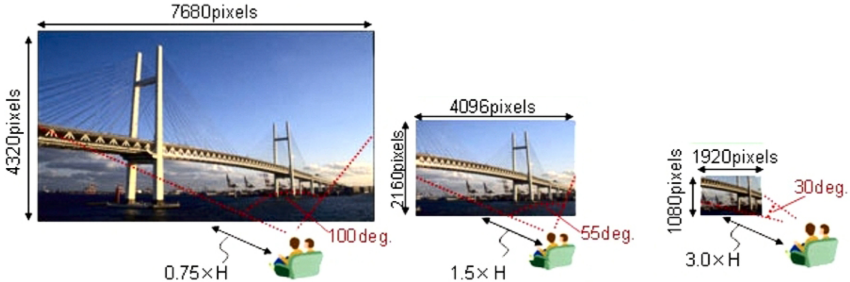
n : bit / sample : 8, 10, 12 bit



UHDTV - motivation

- further enhanced visual experience primarily by having a wide field of view both horizontally and vertically with appropriate screen sizes relevant to usage at home and in public places.
  - even larger field of view
  - enhanced motion reproduction (higher frame rates)
  - extended color space/gamut
- based on the foregoing: spatial resolution requirement: $1/60^\circ$
- formats:
  - UHD-1 (4K): 3840 x 2160 - FOV: 58°, optimal display diameter from 3m viewing dist.: 1.88 m x 3.35 m
  - UHD-2 (8k): 7680 x 4320 - FOV: 96°, optimal display diameter from 3m: 3.76 m x 6.70 m (!!!)

UHDTV field of view



UHDTV - realization

ITU Rec. 2020

- new RGB primaries: pure monochromatic RGB primaries (theoretically)
  - R (CIE x=0.708,y=0.292), 630 nm
  - G (CIE x=0.170,y=0.797), 532 nm
  - B (CIE x=0.131,y=0.046), 467 nm
- new luma coefficients: $Y' = 0.2627\ R' + 0.6780\ G' + 0.0593\ B'$
- Opto-electronic Transfer Characteristics: identical with Rec. 709 (HDTV)
- YCbCr transform and digital representation identical with Rec. 709
- 10 bit/sample and 12 bit/sample formats
- 4:2:0, 4:2:2, és 4:4:4 (RGB) formats
- 60 Hz to 120 Hz (!) only progressive frame rates

UHDTV color space