Color space, digital coding, and sampling schemes for video signals

There are various ways to represent video information. This note describes some aspects of different color spaces, conversion between them, and normalized digital coding.

RGB at video camera output

The principal signal components of color camera or scanners, or other imaging pickup devices are Red, Green and Blue, RGB. These are also the principal components for video signal reproduction (i.e. picture display) at the monitor, as the CRT phosphors are comprised of these colors. But there is a non-linear relation between the camera signal pickup function (light input) and the CRT signal display function (light output). The transfer function is approximately exponential, and commonly referred to as "gamma" curve. Gamma is mainly a light reproduction function of the CRT.

$$R_{display} = R_{camera}^{\gamma}$$
$$G_{display} = G_{camera}^{\gamma}$$
$$B_{display} = B_{camera}^{\gamma}$$

During the development of the video transmission standards it was decided to compensate for this gamma-curve at the source side (camera, studio), and not to burden the television receiver with this effort and cost. The NTSC standard defines a gamma of 2.2, the PAL and SECAM standards defines a gamma of 2.8. Normally this gamma-correction is performed directly in the camera.

$$R_{transmit} = R_{pickup}^{1/\gamma}$$
$$G_{transmit} = G_{pickup}^{1/\gamma}$$
$$B_{transmit} = B_{pickup}^{1/\gamma}$$

The gamma-pre-corrected RGB signals at the camera output are stretched in the darker range and compressed in the lighter signal range. This has, as a side effect, a positive effect on noise influence on the transmission channel. The human eye is more sensitive to noise in dark areas, where the gamma behavior of the CRT reduces visibility.

Computer graphics generation is defined normally in "linear" RGB color space. The computer monitor of today has often a smaller gamma factor than used by the television standard definition, but there is no standard value. Sometimes it is compensated in the monitor itself, or by means of the look-up tables of the graphics RAMDAC, or not at all. The human eye is not very sensitive against gamma mismatch.

If video (camera) RGB gets merged with computer RGB, it is preferably be done in the same RGB space, including the assumed gamma. The anti-gamma compensation, as implemented in the Philips scaling ICs , compensates for a gamma-pre-correction of 1.4 only. The remaining gamma factor is assumed to be still performed by the computer monitor. A greater value of gamma-correction-compensation would lose more digital codes in the available 8-bit number range, and produce larger quantization steps in bright areas, which is not acceptable.

RGB can assume only positive values, and generate a cube like color space. The RGB components are commonly normalized to unity (e.g. 1 Volt peak-peak as analog signal). If any of the components is 0, it means there is no color of this component, if it is 1, there is full (100%) saturation of this color. All components equal zero represents the color 'black', all components equal 1 represents bright 'white'. The RGB cube is an additive color space.

Matrix to YUV (YCbCr)

In order to allow a compatible migration from black&white television to color television the YUV color space was utilized. Y stands for the luminance (lightness) information, and is compatible to black&white (and gray) signal. U and V are the so-called color difference signals B-Y an R-Y, and carry the additional color information (additive color space). The YUV representation of video information is also oriented on the human perception of visual information, whereby RGB representation is more based on the technical reproduction of color information. The human eye senses luminance and color with different receptors. There are less color receptors, and they have significant less spatial resolution. The YUV color space representation can take advantage of that fact, by spending less bandwidth for color difference information than for luminance information (see sampling schemes, later in this note).

Luminance Y can be positive only, the color difference signals U and V can be positive or negative. Commonly YUV is also normalized to unity (peak-to-peak = 1). The following matrix equation transforms gamma-pre-corrected and normalized RGB into normalized YUV (see also CCIR recommendation 601).

$$\begin{aligned} Y &= 0.299 * R + 0.587 * G + 0.114 * B \\ U = C_b = (B - Y) &= -0.169 * R - 0.331 * G + 0.500 * B \\ V = C_r = (R - Y) &= 0.500 * R - 0.419 * G - 0.081 * B \end{aligned}$$

(**NOTE:** For analog signal processing often un-normalized signals are used, which results in different number in the matrix equations, but does not change the cross relationship between RGB and YUV.)

U and V form a square color plane. But for colors of natural pictures and due to some restrictions in the video standards NTSC and PAL, this square color plane is reduced to a color circle plane. The vectors of natural colors don’t point into the extreme corners of the square UV plane. The size of that circle is further restricted, if luminance values are close to minimum or maximum. There can’t be any color in black or white e.g.. (Artificial YUV signals, e.g., test signals can use those extreme combinations). The YUV color space is best represented by a round column, with the dimension of luminance Y as axle in its center, and this round YUV color space column is shaped to a point at the bottom and at the top.

CCIR rec. 601 describes also how to represent these YUV signals by digital codes. It is recommended not to use the entire available number range for nominal signal values, but leaving some margin, room for digital signal processing, e.g. for over and under shoots. In an 8 bit system, luminance Y black is coded with 16 decimal (= 10 hexadecimal), 100% white is coded with 235 decimal (= EB hexadecimal). The color difference signals Cb and Cr are coded in offset binary, which ‘offsets’ the ‘no color’ point into the middle of the number range to code 128 (80 hex). 100% color saturation uses the codes from 16 (10 hex) to 240 (F0 hex). 75% color saturation uses only codes from 44 (2C hex) to 212 (D4 hex) (see also data sheet SAA7151B, Fig.13, for example).

The codes 00 hex and FF hex should not be used for video signal coding. These two codes are reserved for synchronization purposes (see CCIR rec 656).

(**Note regarding nomenclature:** The terms “YUV” and “YCbCr” are referring to the same color space and cross relationship to RGB. The expressions “B-Y” and “R-Y” are normally used for non-normalized color difference signals. It is not part of any standard specification, but some literature is using the term “YUV” to indicate analog signal representation, and the term “YCbCr” for its digital representation. Most data sheets and documents in this book are using both terms interchangeable for digital signal representation of normalized signals.)

The CCIR recommendation 601 (re-printed elsewhere in this book) gives an example of a digital RGB to YUV conversion. It is assuming digital sampled RGB, defined in codes like luminance signal Y, i.e., between 16 for black and 235 for full saturation. The given equation assumes a matrix realization by means of 8x8bit multipliers, which is only approximating the correct relationship. This equation system should not be used as reference to construct the inverse matrix from YUV to RGB. Today’s technology allows matrix implementation by means of look-up tables, avoiding the limiting multiplier resolution and truncation problem.

The accurate digital RGB to digital YCrCb conversion is described by the following matrix:

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ \left(\frac{0.701}{(2*0.701)} * \frac{224}{219}\right) & \left(-\frac{0.587}{(2*0.701)} * \frac{224}{219}\right) & \left(\frac{0.114}{(2*0.701)} * \frac{224}{219}\right) \\ \left(-\frac{0.299}{(2*0.886)} * \frac{224}{219}\right) & \left(-\frac{0.587}{(2*0.886)} * \frac{224}{219}\right) & \left(\frac{0.886}{(2*0.886)} * \frac{224}{219}\right) \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The digital RGB ranges from 16 to 235, i.e. over 219 possible values. The digital CrCb goes from 16 to 240, uses 224 possible values. This causes a re-normalization factors.

The inverse matrix from digital YCrCb to digital RGB (16 to 235) calculates to :

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1.371 & 0 \\ 1 & -0.698 & -0.336 \\ 1 & 0 & 1.732 \end{bmatrix} * \begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix}$$

YIQ, and other YUV related color spaces

YIQ color space is similar to YUV color space except that it has the I and Q color axes rotated 33 degrees with the respect to the U and V axes of the YUV definition.. “I” means “in phase”, and “Q” means “quadrature phase”. This color space was adopted by early NTSC systems to take full advantage of the human eye color response with respect to color bandwidth capability.

$$I = V * \cos(33^\circ) - U * \sin(33^\circ)$$
$$Q = V * \sin(33^\circ) + U * \cos(33^\circ)$$

The Philips digital decoder have fully adjustable “hue” control. The demodulation angle can be programmed to any value, and can achieve an I-Q demodulation, i.e., generating I and Q outputs instead of U and V.

Some other color space approaches (like HSI, or HSV, or HSL etc.) describe the UV plane in polar coordinates by means of a vector, its length(S = saturation) and its angle(H = hue). The luminance (Intensity, Value, Lightness) corresponds to the Y of YUV space. This color space representations are related to the quadrature encoding of U and V onto a color subcarrier, in the transmission standards NTSC and PAL.

CMYK for color printer

CMYK color space is a subtractive color space used for color printing. CMYK stands for Cyan, Magenta, Yellow and Black. It describes, which color component is removed from white, to generate a certain wanted/printed color. In theory, only the CMY portion is required, however, in actual printing ink applications, black ink is added to enhance the contrast ratio and purity of the black portion of the image. K is defined as min(CMY), that is, K is equal the lowest value of C, M, or Y.

The relation of CMY to RGB is given vectorally as :

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

4:4:4 sampling (RGB, YCbCr)

Figure 1 illustrates the sampling positions for 4:4:4 sampling, which is mainly used for RGB, but can also be used for YUV or YCbCr. At each pixel a sample is taken for R, G, and B, or Y, U, and V etc.

All three components have the same spatial resolution (bandwidth). If 8 bits per component is used, a 24 bit system is required.

4:2:2 YCbCr sampling

Figure 2 represents a more effective sampling format, in which Y samples are measured at each pixel position, and Cb and Cr samples only at every second pixel position. By that the color information has horizontally a resolution, that is half of that of luminance. The human eye does not perceive chrominance with the same clarity as luminance,. therefore this type of data reduction causes very little visual loss of content. The 4:2:2 sampling scheme reduces the data bandwidth need by a third.

Cb and CR samples are co-sited with every second Y samples, but starting with the first Y sample of each line. If 8 bits per component is used, a 16 bit system is required.

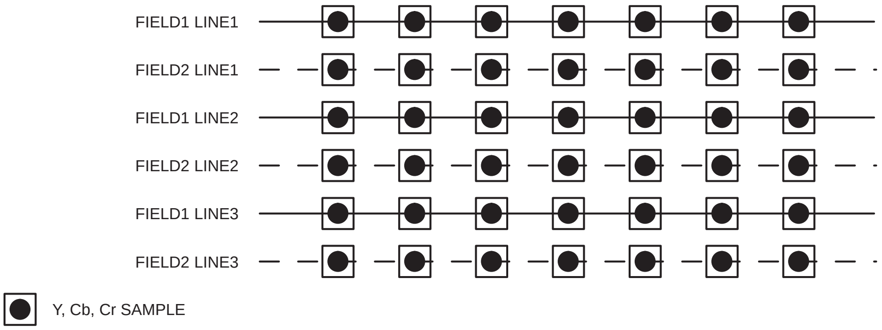

Figure 1. 4:4:4 Sampling Scheme

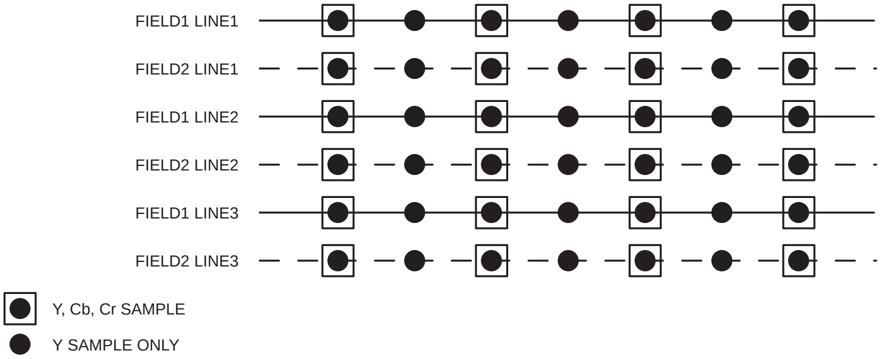

Figure 2. 4:2:2 Sampling Scheme

4:1:1 YCbCr (orthogonal) sampling

Figure 3 is an example of 4:1:1 sampling, often used in consumer type video products. The achievable color bandwidth in this case is only one third that of luminance. But in broadcasted video (NTSC, PAL, or SECAM), or in tape-recorded video, there is normally not more chroma bandwidth supported/available.

The CbCr samples are taken co-sited with every fourth luminance pixel, but starting with the first luminance sample of each line. An 8 bit per component system is capable of fitting into a 12 bit wide frame buffer. 7 bit per component and 6 bit per component systems are also used in combination with 4:1:1 sampling, which reduces the needed frame buffer capacity even more (e.g., for PIP function on television sets).

4:2:0 YCbCr (spatial) sampling

This sampling scheme is used generally for MPEG and H-261 compression standards, and is also called “coded picture sampling”. Figure 4 shows the two dimensional 2:1 sub-sampling of color pixels relative to luminance pixels. The CbCr samples are not co-sited with a luminance sample, but representing the color information for a quartet of four Y pixels, ordered in a square. The CbCr values are normally derived (calculated) from a 4:4:4 or 4:2:2 sampling scheme by both horizontal and vertical filtering and interpolation. Usually the CbCr values are transported only every second scan line with pairs of Y samples, the other line carries only Y samples (4:2:0). The overall data bandwidth of 4:2:0 sampling is identical to 4:1:1 sampling.

In the example in Figure 4 a non-interlaced video source is represented, as those compression standards know only ‘pictures’ and use whole frames, or just one field.

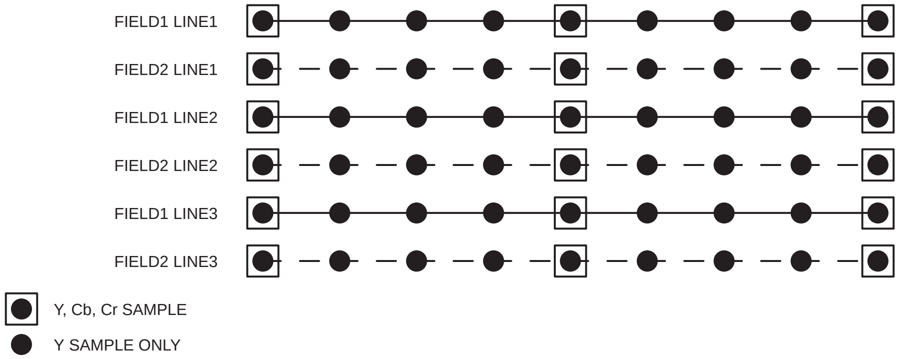

Figure 3. 4:1:1 Sampling Scheme

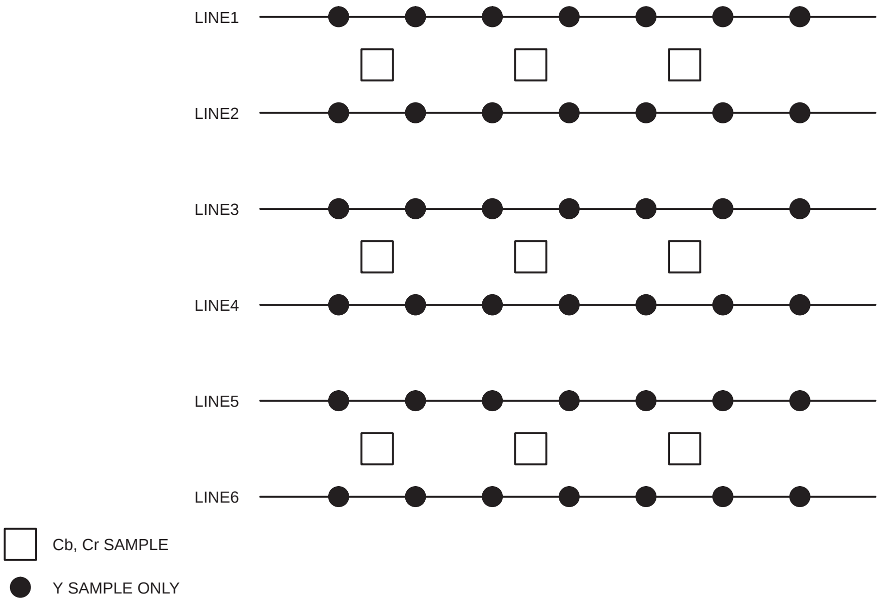

Figure 4. 4:2:0 Sampling Scheme