Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG
(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)
JVT PExt Ad Hoc Group Meeting: 22-24 July 2003, Trondheim

Document: JVT-I014r3
Filename: JVT-I014r3.doc

Title:	**YCoCg-R: A Color Space with RGB Reversibility and Low Dynamic Range**
Status:	Input Document to JVT
Purpose:	Proposal

Author(s) or Contact(s): Henrique Malvar and Gary Sullivan

One Microsoft Way
Redmond, WA 98052 USA

Tel: +1 425 882 8080
Email: malvar@microsoft.com
garys@microsoft.com

Source: Microsoft Corp.

0.0 Introduction

At the Geneva meeting we presented the YCoCg color space [1], including its simple transformation equations relative to RGB and its improved coding gain relative to both RGB and YCbCr. We also discussed the reversibility of RGB to YCoCg conversion process in the case that two additional bits of precision are used for YCoCg relative to the precision used for source RGB data. In this proposal, we additionally present a variation of this color space, which we call YCoCg-R.

YCoCg-R has properties very similar to those of YCoCg, but in addition it enables transformation reversibility with a smaller dynamic range requirement than does YCoCg. In particular, YCoCg-R produces a Y component with the same dynamic range as the input RGB data, and requires only a one-bit expansion of dynamic range support for the Co and Cg components.

Note: Compared to our previous “r1” version, this version presents a corrected coding gain table (Table 1), a correction of a typo equation error at the end of section 2.0, and provides a non-normalized autocovariance matrix rather than a normalized one, because the coding gain should be computed from the non-normalized autocovariance matrix. Although this changes the computed coding gains somewhat, it does not affect the relative performance of the different color transforms.

1.0 The YCoCg Color Space

As we mentioned in [1], encoding of RGB data directly in RGB space does not lead to the best compression results, because it doesn’t take advantage of the statistical properties of the source video content.

In the same way that the use of transforms improves PSNR in macroblock coding, a transformation applied to RGB can also improve overall PSNR. If we map the original RGB channels into one luma and two chroma channels, the chroma components will require fewer bits, so the overall compression will be better (more than 3 dB improvement in PSNR in terms of theoretical coding gain, i.e., for simple non-predicted intra coding) than encoding directly in the RGB space.

Color spaces such as YCrCb provide good decorrelation, but even better results can be obtained with the YCoCg (luminance + offset orange + offset green).

We recall from [1] that the direct YCoCg color space transform is defined by:

$$\begin{bmatrix} Y \\ Co \\ Cg \end{bmatrix} = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & -1/2 \\ -1/4 & 1/2 & -1/4 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

and the inverse YCoCg color space transform is:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} Y \\ Co \\ Cg \end{bmatrix}$$

Thus, the encoder needs only additions and shifts to convert to YCoCg, and the decoder needs just 4 additions per pixel to convert back to RGB, using:

$$G = Y + Cg;\;\; tmp = Y - Cg;\;\; R = tmp + Co;\;\; B = tmp - Co.$$

2.0 New Reversible Transform: YCoCg-R

For encoding with the smallest quantization step sizes, or for lossless encoding, it would be useful to employ a reversible transform, that is, one that allows the mapping of a triplet of integer RGB numbers into a triplet of integer YCoCg numbers in such a way that the input RGB integers are exactly recoverable.

We could generate a reversible RGB ↔ YCoCg mapping by multiplying all elements of the direct transform matrix above by four. At the inverse transform, we would divide the final results by four. However, that would generate color components that would need two more bits of precision to represent, and there would also be a loss of coding gain relative to using an equivalent color transform without an increase in bit depth (because the increase in the dynamic range of Y also causes an increase in its entropy).

We can convert the YCoCg transform above to an integer-reversible mapping by applying S-transform concepts [2], [3] (also known as lifting) to factors of the matrices above. For example, given two integer numbers x and y, their integer difference d and truncated average m can be computed by the formulas

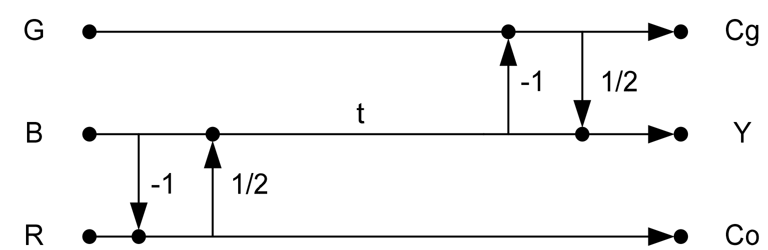
$$\begin{matrix} d = x - y \\ m = y + (d >> 1) \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} y = m - (d >> 1) \\ x = y + d \end{matrix}$$

where >> is the arithmetic shift operator, which approximates a division by 2. Using this concept, we can convert the RGB ↔ YCoCg mappings to a reversible form by first scaling the Co and Cg components by a factor of two, so the S-transform above can be directly applied.

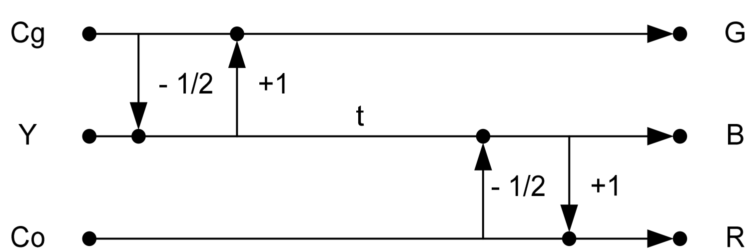
The final result is:

$$\begin{array}{c} Co = R - B \\ t = B + (Co >> 1) \\ Cg = G - t \\ Y = t + (Cg >> 1) \end{array} \quad \Leftrightarrow \quad \begin{array}{c} t = Y - (Cg >> 1) \\ G = Cg + t \\ B = t - (Co >> 1) \\ R = B + Co \end{array}$$

The corresponding flowgraph is shown in the Figure below:



FORWARD TRANSFORM



REVERSE TRANSFORM

Figure 1: Flowgraph of the reversible form of the RGB ↔ YCoCg color mapping. Its main advantage is that Y has the same dynamic range as the RGB channels, and Co and Cg require only one additional bit, without loss of coding gain performance.

We can see that the reversible mapping above provides an approximation to the construction

$$\begin{aligned} Y &= G/2 + (R+B)/4 \\ Cg &= G - (R+B)/2 \\ Co &= R - B \end{aligned}$$

which is identical to the original definition, but with Co and Cg scaled up by a factor of two.

2.1 Dynamic Range

It is easy to see from Figure 1 that if each of the RGB channels are integer values with an *N*-bit range, then the luma channel Y requires *N* bits, and the chroma channels require $N+1$ bits. That's a minimal dynamic range expansion.

3.0 Coding Gain

Besides being the simplest possible color space transform to implement, in terms of decoder operations, as we saw above, the YCoCg space is justified by its better compression performance.

Assuming that the original data was in RGB domain, let us consider the PSNR for the reconstructed RGB values. As it is well known in coding theory, we can get an increase is PSNR by transforming each RGB triplet to an appropriate domain, and coding in that domain. The *coding gain* is the high-rate increase in PSNR by encoding in that domain. It is well known that the optimal transform, in terms of maximizing the coding gain, is the Karhunen-Loève transform (KLT), whose basis vectors are the eigenvectors of the autocovariance matrix for the RGB components.

Let us consider a reasonably diverse data set of images, e.g. the widely used "Kodak set" of 24 images of size 512x768, captured with a high-quality 3-CCD camera (no Bayer mosaic artifacts). The RGB autocovariance matrix for that set is

$$\begin{bmatrix} 0.9943 & 0.9130 & 0.7727 \\ 0.9130 & 1.0571 & 0.9183 \\ 0.7727 & 0.9183 & 0.9486 \end{bmatrix}$$

Given that autocovariance matrix, it is easy to compute the coding gain. In Table 1 we compare the coding gain of YCoCg to that of the several YCbCr transforms specified in H.264/AVC Annex E, and we see that the YCoCg leads to a 0.7 dB improvement in coding gain when compared to those transforms. We have also included the coding gain for the reversible color transform used in JPEG2000.

Color Transform	Coding gain, dB
Optimal (KLT)	4.97
ITU-R BT.709 [1]	3.79
FCC [4]	3.91
ITU-R BT.470-2 [5]	3.92
SMPTE 170M [6]	3.92
SMPTE 240M [7]	3.83
JPEG2000 RCT	4.31
Proposed YCoCg or YCoCg-R	4.54

Table 1: Theoretical coding gain for several color space transforms, for the 24 RGB images in the Kodak test set. KLT is the statistically-optimal Karhunen-Loève transform. The numbers in square brackets are the values of the parameter **matrix_coefficients**, in Table E.5 of H.264/AVC.

We see from Table 1 that the proposed YCoCg-R color space transformation has three important features:

- Exactly reversible in integer arithmetic.
- Minimal increase in dynamic range; no increase for Y and only 1 bit increase for Co and Cg.
- Higher coding gain than other color spaces, including RCT.

References

[1] H. S. Malvar and G. J. Sullivan, "Transform, Scaling & Color Space Impact of Professional Extensions," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6 Document JVT-H031, Geneva, May 2003.

[2] P. Lux, "A novel set of closed orthogonal functions for picture coding," *Arch. Elek. Ubertragung*, vol. 31, pp. 267–274, 1977.

[3] P. Piscaglia and B. Macq, "Multiresolution lossless compression scheme," *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996, vol. 1, pp. 69–72.

[4] The Kodak image set is available at ftp://ftp.ipl.rpi.edu/stills/kodak/color.

[5] D. S. Taubman and M. W. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards, and Practice*. Boston, MA: Kluwer, 2000.