Copyright 2000, IS&T

Comparison between Different Color Transformations for JPEG 2000

Sang-Eun Han, Bo Tao†, Ted Cooper†, and Ingeborg Tastl†*
** Stanford University, Stanford, California USA*
† Sony US Research, San Jose, California USA

Abstract

The future JPEG 2000 international standard supports the compression of multi-component images. However, the current implementation of JPEG 2000, so-called verification model (VM), has not been fully tested for color images. Furthermore, the color space issue has not been completely resolved within the JPEG 2000 standard. This work compares different color transformations for JPEG 2000 image compression, including RGB, YCbCr, CIELAB, CIELUV and the Karhunen-Loeve transformation (KLT) by using both objective measure and subjective testing. The results indicate that the KLT excels in both objective and subjective tests, while RGB gives the worst performance as expected. The results also show that YCbCr is a good practical choice.

1. Introduction

The JPEG 2000 committee[9] is developing a new compression standard for still images based on wavelet technology. It has received much enthusiasm, both because of good image quality at low bitrates[7] and the useful functionalities it provides. However, there has been relatively less research effort going on with regard to color image compression, due to the decision by JPEG 2000 committee early on to separate color from the rest of the compression system. Even within the image compression community as a whole, more works have been performed on gray-scale images, while human color vision properties have yet to been well married to other compression techniques.

This work takes on the issue of color transformation for image compression by using the state-of-the-art image compression system offered by JPEG 2000. On one hand, from the rate-distortion theory, it is well known that color transformation is necessary to decorrelate a multi-channel signal in order to achieve good rate-distortion results. The optimal transformation is the Karhunen-Loeve transformation (KLT). On the other hand, from human vision studies, it is known that color decorrelation is performed by our human vision system by doing so-called opponent color extraction. A number of color transformations and spaces have been devised based on this study, such as CIELAB and CIELUV. It is of great interests to find out whether the use of these human vision-based color spaces will give good subjective compression results.

In the rest of this paper, we first briefly describe in Section model the JPEG 2000 verification model (VM) on which the work is conducted, especially its support of multicomponent image compression. Color spaces and color transformations are discussed in Section space. We present experimental results in Section results which lead to the conclusion of this paper in Conclusion.

2. Color Mixing in JPEG 2000 Verification Model

JPEG 2000 would support the compression of colored images with any combination of component in the future. For non-reversible systems, where the Wavelet transforms are implemented with convolution kernels as in traditional subband decomposition, the sufficient condition is that the number of input and output image components are identical. For reversible systems, which have integer wavelets with fixed precision arithmetic, it is necessary that the bit-depth of each output image component must be identical to the bit-depth of the corresponding input image component.

The current implementation of the component mixing is, however, applicable only for identical dimensioned image components. Through the command switch, the color space for the input image has been changed from RGB to luminance-chrominance by YCbCr transform with non-reversible decompositions and Reversible Color Transform (RCT) transform with reversible decompositions. In this work, only non-reversible decompositions are used since all added experimental color transforms are neither integer-valued nor reversible.

3. Color Spaces and Transformations

A color space is the way we can specify, create and visualize color.[4,6] There are several color spaces which calls our attention in terms of human's perception as well as compression.

RGB is the most popular color space, which stands for Red/Green/Blue. It is device-dependent and normally used on monitors. YCbCr is the most commonly used color space for compression, for example, JPEG compression uses in these coordinates as a default. Y is the luminance component and Cb and Cr are the chrominance components. Since the human visual system is more sensitive to changes in luminance than to changes in chrominance, images are

 Copyright 2000, IS&T

compressed more effectively, via subsampling or interpolation, with this space.

Given the primary RGB inputs (R, G and B in [0,255] or [0,1]),

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

and given a YCbCr input (Y in [0,1] and Cb, Cr in [-0.5, 0.5]),

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.0 & 1.4021 \\ 1 & -0.3441 & -0.7142 \\ 1 & 1.7718 & 0.0 \end{bmatrix} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix}$$

XYZ was introduced in 1931 by the CIE-Commission Internationale de l'Eclairage-to represent all visible colors using only positive three values of X, Y and Z, where Y is identical to luminance, and X and Z give coloring information.[3] These values are based on measurements of the color-matching abilities of the average human eyes. It is completely device independent, while it is very complicated to implement the space.

RGB values in a particular set of primaries can be transformed to and from CIE XYZ via a 3x3 matrix. However, before the transform could proceed, the raw RGB data has to be inverse gamma corrected in order to get the actual RGB values that were converted to CIEXYZ.[2] The 8-bit RGB values are converted to non-linear sR'G'B' values, by the inverse gamma correction, which represents the appearance of the image as displayed on the reference display in the reference viewing condition. Then, this matrix transforms the image from sR'G'B' to XYZ,

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

and the inverse transformation matrix is as follows:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 3.2406 & -1.5372 & -0.4986 \\ -0.9689 & 1.8758 & 0.0415 \\ 0.0557 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

RGB values can be transformed from sR'G'B' through gamma correction. Note, this matrix has negative coefficients. Some XYZ color may be transformed to RGB values that are negative or greater than one, while the range for valid R, G, B values is [0,1]. This means that not all visible colors can be produced using the RGB system.

CIE L*a*b* is a color space introduced by the CIE in 1976, in which L* is the luminance component, a* and b* are respectively red/blue and yellow/blue chrominancies. Equal distances in this space represent approximately equal color difference.

CIE 1976 L*a*b* is based directly on CIE XYZ and is an attempt to linearize the perceptibility of color differences. The non-linear relations for L*, a*, and b* are intended to mimic the logarithmic response of the eyes. Coloring information is referred to the color of the white point of the system, subscript n.

$$L^* = 116\, f(Y/Y_n) - 16 \tag{1}$$

$$a^* = 500[f(X/X_n) - f(Y/Y_n)] \tag{2}$$

$$b^* = 200[f(Y/Y_n) - f(Z/Z_n)] \tag{3}$$

where

$$f(w) = \begin{cases} w^{1/3} & \text{for } w > 0.008856 \\ 7.787w + 16/116 & \text{otherwise} \end{cases}$$

Here X_n, Y_n, and Z_n denote the tristimulus values for reference white. For the CIE standard illuminant D65, X_n = 95.05, Y_n = 100 and Z_n = 108.91.[1] The reverse transformation (for $Y/Y_n > 0.008856$) is

$$Y = \begin{cases} Y_n L^*/903.3 & \text{for } L^* < 8.0 \\ Y_n[(L^*+16)/116^3 & \text{otherwise} \end{cases}$$
$$X = X_n\left[a^*/500 + (Y/Y_n)^{1/3}\right]^3$$
$$Z = Z_n\left[(Y/Y_n)^{1/3} - b^*/200\right]^3$$

CIE L*u*v* is also defined by the CIE in 1976. Likewise, L* defines the luminancy, u* and v* define chrominancy. This is another attempt to linearize the perceptibility of color differences.

CIE 1976 L*u*v* is based directly on CIE XYZ like CIE L*a*b*. The non-linear relations for L*, u* and v* are given below:

$$L^* = 116 f(Y/Yn) - 16 \tag{7}$$

$$u^* = 13L^*(u' - u_n') \tag{8}$$

$$v^* = 13L^*(v' - v_n') \tag{9}$$

$$\text{where } \begin{cases} u' = 4X/(X+15Y+3Z) \\ v' = 9Y/(X+15Y+3Z) \end{cases} \tag{10}$$

The transformation from CIELUV to XYZ is performed as following:

$$u' = u^*/(13L^*) + u_n \tag{11}$$

$$v' = v^*/(13L^*) + v_n \tag{12}$$

$$Y = \begin{cases} Y_n L^*/903.3 & \text{for } L^* > 8.0 \\ Y_n[L^*+16)/116^3 & \text{otherwise} \end{cases} \tag{13}$$

$$X = 9Yu^*/v^* \tag{14}$$

$$Z = 4X/3u^* - X/3 - 5Y \tag{15}$$

Copyright 2000, IS&T

The Karhunen-Loeve transform is not a color space conversion, but a linear transform where the basis functions are taken from the statistics of the signal, and can thus be adaptive.[5,7] It is optimal in the sense of energy compaction; i.e., it places as much energy as possible in as few coefficients as possible. The KLT , also called principal component analysis, is equivalent to a singular value decomposition. The KLT matrix is defined as

$$\text{where} \quad \begin{aligned} T &= U' \\ U &= \left[u_1 \;\; u_2 \;\cdots\; u_3\right], \end{aligned} \tag{16}$$

and the columns of U are the eigenvectors of the autocorrelation matrix of the input vector X. The output $Y = TX$ will have uncorrelated components. The KLT matrix is image-dependent. The adaptiveness is not used in the coder, and the basis functions are calculated off-line.

Experimental Results

We use the baseline version of JPEG 2000 VM 4.2 to compress and decompress the images. To be consistent, we use the same settings for all approaches except changing the color space from RGB to YCbCr, CIELAB, CIELUV and KLT. Four test images are chosen as shown in Fig. 1. These images are selected so that a wide range of colors and textures would be included in the experiment. All the computations were performed in floating point in order to minimize round-off error.

Figure 1. Original images used for subjective test—they are Miniflow, Dad, Hotel and Boat respectively.

 Copyright 2000, IS&T

RMS error is used to evaluate the quality of the decompressed image, which makes sense from a signal processing perspective. These quantities provide objective measure of the image quality. First of all, we can see the trend of distortion between the original image, Miniflow, and the decompressed image, with varying the bitrate. The rate-distortion curve in Fig. 2 shows that the KLT is about the same quality in RMS error at 0.25 bpp (RMS error 20) compared to the RGB at 0.5 bpp resulting in 2:1 improvement in compression bit rate. The YCbCr and CIELAB achieve very close bitrate keeping the same quality of images. We can save more bits when compression gets closer to lossless coding.

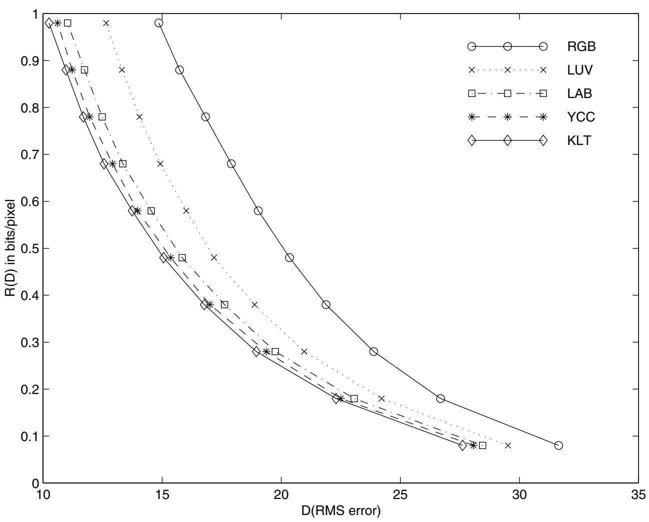

Figure 2. Rate-Distortion curve for the test image "Miniflow"

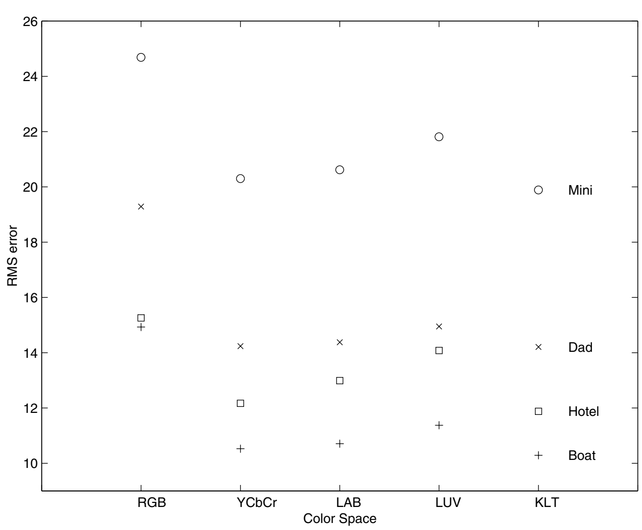

Figure 3. RMS error for 0.25 bpp

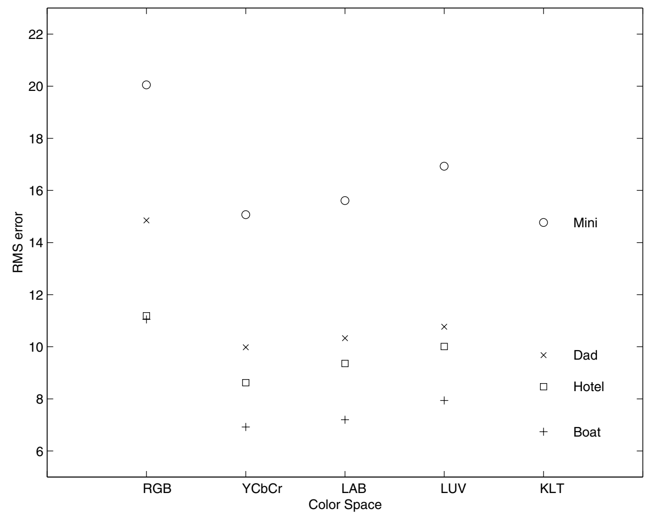

Figure 4. RMS error for 0.5 bpp

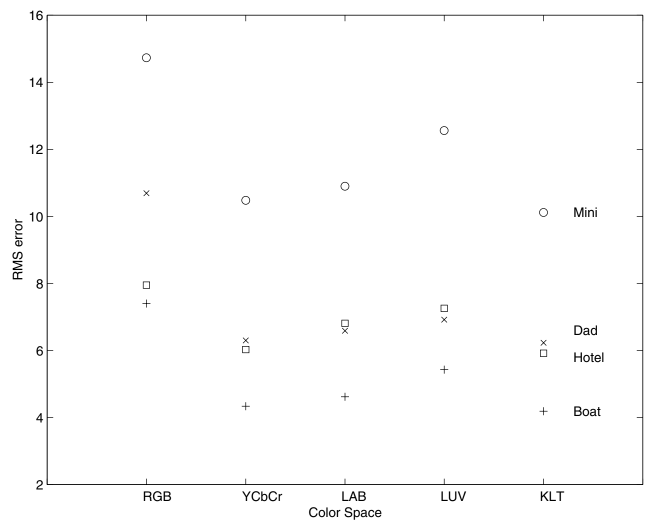

Figure 5. RMS error for 1 bpp

Since the wavelet-based compression is known for the good quality at low bitrates, the four experimental images are all compressed to 3 target bitrates of 0.25, 0.5 and 1 bpp. In all cases, for all images and at all bitrates, KLT performs the best and RGB color space is the worst. Fig. 3, Fig. 4, and Fig. 5 also show that the YCbCr and CIELAB have the similar performance for four test images.

RMS error can give a view of image quality, however, it does not make a lot of sense as a measure of quality in real world situations since most image consumption is by human eyes. Therefore, subjective tests are following.

We evaluate the visual quality based on prints of the SONY UP-D70A SCSI printer at a resolution of 300dpi. A set of 6 printed images are given at 1 bpp, which is close to the bitrate given by digital cameras in the range of high-compression in JPEG, per a test image; the original and each image produced by using one of the color

Copyright 2000, IS&T

transformations described in Section 3. The observers were asked to order the 6 images in terms of ascending visual quality. In this way, we evaluate a ranking for each image. Similar to the objective measure, in general, the KLT is the best and YCbCr is the next. CIELAB is very close to YCbCr, but still worse than that. CIELUV is worse than CIELAB and RGB is the worst color space for compression. The bigger observer group needs in the future.

Conclusion

Both objective measurement and the limited subjective analysis conducted in this work show that the KLT is the best color transformation for JPEG 2000 compression. The YCbCr is the second best, and CIELAB and CIELUV are following the next. The RGB is the worst color space for JPEG 2000 compression.

The JPEG 2000 color sub-group[8] already proposed that the multi-component decorrelators by the KLT are added into VM4 and the results showed a significant increase in compression efficiency, while keeping the same level of quality for highly correlated hyper-spectral images with 225 bands. The compression efficiency increase is not as good, but this work shows that the addition of decorrelator with KLT also gives a gain to the color image compression. An increase in computation and memory needs to be considered for the future work.

Acknowledgments

The authors thank Professor Robert Gray at Stanford University for his helpful comments on subjective tests. Sang-Eun Han also thanks Sony for providing a summer internship.

References

1. V. Bhaskaran and K. Konstantinides, *Image and Video Compression Standards-Algorithms and Architectures*, 2nd ed., Kluwer Academic Publishers, Norwell, MA, Appendix A.
2. IEC 61966-2-1: Multimedia Systems and Equipment-Colour Measurement and Management-Part 2-1:C Management-Default RGB Color Space-sRGB.
3. M.D. Fairchild, *Color Appearnce Models*, Addison Wesley, Reading, MA, Nov. 1997.
4. H. R. Kang, *Color Technology for Electronic Imaging Devices*, SPIE Press, Bellingham, WA, 1997.
5. A. Gersho and R. M. Gray, *Vector Quantization*, Kluwer Academic Publishers, Norwell, MA, 1992.
6. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison Wesley, 1992.
7. M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, Englewood Cliffs, NJ, 1995.
8. M. Gormish, Color Sub-Group Results for Experiments Defined in Seoul, ISO/IEC JTC1/SC29/WG1 N1297, Jul., 1999.
9. ISO/IEC CD 15444-1, Information Technology: JPEG 2000 Image Coding System, 1999.

Biography

Sang-Eun Han (sehan@leland.stanford.edu) received her B.S. degree in electrical engineering from Seoul National University, Korea, in February 1997 and her M.S. degree in electrical engineering from Stanford University in June 1999. She is currently a Ph. D. student in electrical engineering of Stanford University and is also a member of the Signal Compression and Classification Group. Her research interests include image processing and multimedia communication.