CIE XYZ COLOUR SPACE USED FOR IMAGE COMPRESSION

Ilkka Yläkoski *
Håkan Rönngren **

Abstract

In non-reversible coding of colour images, a suitable colour space conversion will improve the compression ratio. A new method is proposed for compressing cmyk images, using the Neugebauer equations to convert from the cmyk space to the CIE XYZ space. The images are then converted from XYZ to the YUV 4:2:2 colour space proposed in the JPEG/ADCT standard. The inverse mapping from XYZ back to cmyk is done using an iterative method. With these colour space conversions, the compression ratio may be roughly twice as high as in the case without any colour space conversions at all, with the same overall quality. The use of the XYZ space to compress RGB images, due to the difference between the phosphorus specifications of the CCIR digital TV and the filter specifications of RGB scanners is also discussed.

Introduction

Compression of continuous tone colour images has typically been used in transmission of digital TV signals. The methods developed for compressing moving RGB images are useful also for compressing four colour cmyk images. There are, however, a few differences. One of them is the colour space.

Compression can also be done without colour space conversions. On the other hand, it is known that neither RGB, nor cmyk resemble the way human eye perceives colour differences. Experiments have also shown that these

* Research engineer, Technical Research Centre of Finland, Graphic Arts Laboratory

** Research engineer, Swedish Graphic Arts Research Laboratory

colour spaces are intercorrelated, meaning that every colour component of the RGB space has information from the other two (three in cmyk) components [Pratt71]. Furthermore, in cmyk images, this correlation is even more significant because, according to colour theory, only three primaries are needed to describe all colours.

Colour space conversions also introduce errors caused by numerical rounding operations, but they affect only the least significant bit of the pixels. This disadvantage is practically invisible compared to the effects of the whole compression scheme.

The criteria of the optimal colour space for image compression may be defined as follows [Frei&Baxter77]:

1. Visual uniformity, an approximately linear relationship with the perception of colour differences (such that quantization errors disperse uniformly in the perceptible space)
2. Concentration of signal energy. Most of the information that is common to all colour components should be concentrated into a single component.
3. Uncorrelated components, meaning that the colour components do not have common information. Redundant information is removed and only true colour information is coded.
4. Black and white compatibility. This is needed because the algorithm should be used also for b/w images.
5. Ease of implementation

Different colour spaces have been placed into a diagram (figure 1) in which the horizontal axis describes the visual uniformity and the vertical axis describes the orthogonality of colour spaces. Uniformity increases to the right and internal correlation increases upwards. The distance from the upper left corner describes approximately the complexity of the conversion.

Criterion 4 can be fulfilled by many colour spaces. For example, the colour spaces used in transmission of video signals have this feature. The video colour spaces like YIQ/NTSC (National Television System Comittee) and YUV/PAL (Phase Alternation Line) have been used widely. The optimum solution to criterion 3 is the Karhunen-Loeve transformation while CIE-Lab satisfies criterion 1 well. CIE-Lab has been proposed as a suitable colour

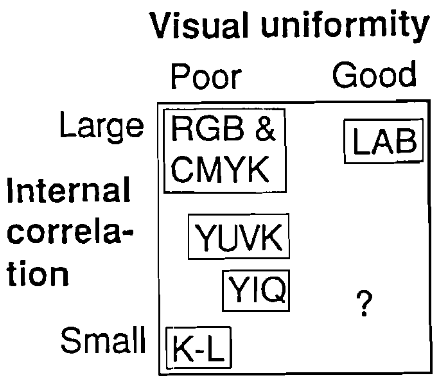


Figure 1. Correlation versus colour uniformity. In this figure, YIQ stands for NTSC, PAL, CCIR-601 etc.

space — because of the uniformity, the quantization errors would be less visible [Schreiber&Buckley81].

There has also been attempts to use a model of the human visual system in the design of an efficent colour image coding system [Frei&Baxter77]. Basically, the optimal colour space should be perceptively a uniform and an orthogonal system. Very likely there is no such colour space. The reason for this is that the way the human eye perceives colours is non-linear and the colour space is non-orthogonal, too. Moreover, the colour space should be simple to implement. Unfortunately, the models that describe the colour perception of the human eye are quite complex.

JPEG/ADCT does not specifically define the colour space but it suggests the conversion matrix defined in CCIR standard 601. The transformation is as follows:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Since CCIR 601 is only a part of the whole digital video transmission system, the RGB values are normally gamma and bias corrected (e.g. NTSC or PAL primaries). This approach has been used in Kodak's PhotoYCC colour encoding scheme to encode Photo CD image data. PhotoYCC is based both on CCIR 601 and CCIR 709 standards.The above mentioned schemes can not be used as such for cmyk data unless the cmyk space is mapped to the XYZ space.

Colour space used in compression of cmyk images

Except for the Karhunen-Loeve transformation, the former mentioned methods are basically feasible only for specific RGB monitor primaries. The Karhunen-Loeve transformation has been suggested [Gilge 1987] to be used as a colour space conversion in compression of cmyk images. It could always be used without any knowledge about the spectral features of the image data. The Karhunen-Loeve transformation is defined as eigenvectors of the covariance matrix of the data. It produces a matrix that rotates the original axis system so that the base vectors are set parallel to the maximum variances of the original data (see figure 2). Because the new space is not a normal colour space, the subsampling of the less significant components can not be used.

The Karhunen-Loeve transformation is based on the covarinace of the image data, so it is dependent on its statistical features and the transformation matrix must be calculated for each image separetly. Even though an average transformation matrix is proposed, it is not a suitable solution, because the statistics between different images may vary greatly. Gilge has also proposed that only the three largest components be coded. The assumption that k is approximately a linear combination of c, m and y is not generally valid, so the linear projection to a three dimensional space after Karhunen-Loeve transformation may cause a severe colour impairment. [Yläkoski91]

The cmy values could always be transformed to e.g. YIQ values even though scanner filters are not similiar to NTSC camera filters. It can be shown that some data reduction is achieved by the YIQ colour space conversion. Even so, the black component must be compressed separately. [Yläkoski91]

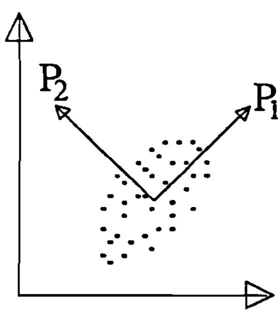


Figure 2. Rotation of two-dimensional data in order to minimize the correlation

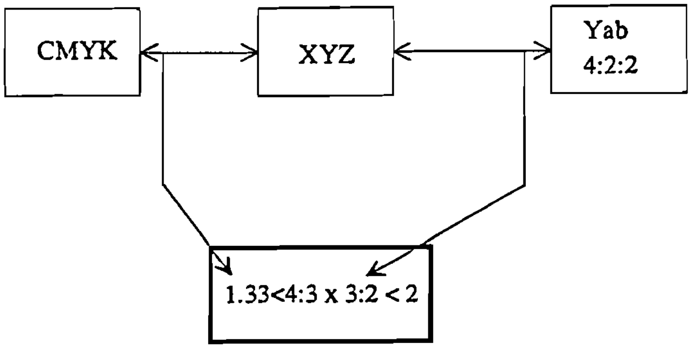


Figure 3. Data reduction by XYZ space conversion

If cmyk data is mapped to a video colour space through the XYZ space, a significant data reduction is achieved according to figure 3. The first mapping from cmyk to XYZ drops the "extra" fourth component and the further 2:1 subsampling of chrominance gives further data reduction with a factor 1.5. Thus, the overall improvement in compression is two. Since this improvement is independent on the compression algorithm used, it will theoretically double the compression ratio maintaining the same overall quality.

The Neugebauer Equations used for XYZ – cmyk mapping

Background

The Neugebauer equations presuppose that screen dots are distributed over a substrate without correlation between different screens. This turns out to be almost true for normally angled screens (although it is totally false for flamenco, for instance). For actual c, m and y values, the CIE X value is calculated in the following way:

$$\begin{aligned} X = \; & (1-c)(1-m)(1-y)X_n + c(1-m)(1-y)X_c + \\ & + (1-c)m(1-y)X_m + (1-c)(1-m)yX_y + \\ & + cm(1-y)X_{cm} + (1-c)myX_{my} + \\ & + c(1-m)yX_{cy} + cmyX_{cmy} \end{aligned}$$

Y and Z are calculated accordingly. The eight terms in each equation represent reflected X, Y or Z values from the different combinations of inks, multiplied by the share of a local area that is covered by the corresponding ink combination. X_n does not correspond to an ink combination, but to unprinted paper. Since XYZ is additive, this model works quite well.

Metameric effects

Obviously, there are several possible combinations of cmyk for a given XYZ. They correspond to different spectral reflectances, but due to metameric effects, they are all visually equivalent.

Since there are only three equations, it is possible to determine the values of three independent variables. Therefore, only c, m and y are used in the equations. When solving the equations, a suitable value for black should be determined in another way; presumably before the equations are actually solved. The desired XYZ values may then be adjusted according to the amount of black that is already chosen.

Solving the equations

As a result of the metameric effect described above, there are a great number of different methods for determining cmyk values, and nobody could motivate why one is better than another. The following model has been used for testing within this project. It consists of two parts. In the first part, the amount of black is determined in a way that a high value is possible only in neutral tones. In the second part, the best cmy values are found using any optimizing method.

Determine normalized XYZ values for estimation of the "neutrality":

1. X* := X / Xn
   Y* := Y / Yn
   Z* := Z / Zn

Define a measure for the neutrality and suppress low values:

2. neutrality := min(X*, Y*, Z*) / max(X*, Y*, Z*)
   (if min = max = 0, then let neutrality = 1)

3. alpha := neutrality * neutrality * neutrality

Determine the amount of black to be used:

4. K := alpha * (1 – max(X*, Y*, Z*))

Adjust the desired XYZ values, according to the already chosen amount of black:

5. $X' := (X - kX_k) / (1 - k)$
   $Y' := (Y - kY_k) / (1 - k)$
   $Z' := (Z - kZ_k) / (1 - k)$

Use any optimization method to find the optimal cmy for the modified XYZ:

6. (We have used a gradient method [Bradley *et al* 77])

Comment to step 1 – 4: As mentioned above, a given colour does not have a specific amount of black, but rather an interval. Ideally, the lower limit of that interval should be zero, since c, m and y should ideally be able to reach any colour by themselves. In practice, black is needed in very dark colours (to make them dark enough) and in neutral colours (to assure that they stay neutral). The formula chosen here uses a lot of black in neutral colours (that is, $X / X_n = Y / Y_n = Z / Z_n$) and suppresses black efficiently (making use of the power three in the definition of alpha) for non-neutral colours. Combined with an optimization method to find the best cmy solution, this method has given good results.

Comment to step 5: Let a given combination of c, m and y be described by a combination of X, Y and Z. Consider what will happen when k is printed upon the actual area. The new XYZ values X', Y' and Z' will be approximately

$$X' = (1 - k) X + kX_k$$

$$Y' = (1 - k) Y + kY_k$$

$$Z' = (1 - k) Z + kZ_k$$

These equations look like the Neugebauer equations. In step 5, we use this expression, rearranged in order to express the necessary XYZ values for the cmy combination, given the X'Y'Z' values for the actual cmyk combination that will be printed. X_k, Y_k and Z_k are, quite arbitrarily, set to 0.01 (which would mean that the density for black is 2).

Comment to step 6: Since an optimization method is used to find cmy values, this step could be used to test the model used to choose an amount of black. For every XYZ combination that has been calculated from a combination of cmyk, it must be possible to find a k value such that there is a cmy combination that satisfies the given XYZ value with a zero error. If this turns out to be impossible for some pixels in an image, the model for K is not good. The reason why the described model is accepted without further theoretical studies is that no such problems have shown up.

Taking care of dot gain

There is a way to take care of dot gain already in the Neugebauer equations, using Yule-Nielsen exponents. Then, the equations are called "modified Neugebauer equations". Unfortunately, this slows down the calculations seriously. Therefore, the dot gain is regarded later, using one-dimensional transfer functions. Separating the dot gain compensation from the XYZ to cmyk transformation also gives freedom to choose output system after that the cmyk file is ready. In this case the test images used in the visual test are Signature colour proofs, so the transfer functions are chosen to match the dot gain for Signature.

Compression and quantization

The compression algorithm used in this survey was JPEG/ADCT [ISO88] even though the results are not dependent on it. The reason is that the colour space conversion is always done before compression, so any algorithm, that compresses colour images one colour component at a time, will do.

The compression principles of the baseline JPEG/ADCT are as follows:

The input image is divided into a stream of 8×8 pixel blocks, and each one is transformed with the 8×8 DCT to produce 64 output coefficients. Each coefficient is independently quantized (divided) by a uniform quantizer with a step size determined by the corresponding visibility threshold from a 64-element quantization matrix. This means that every DCT coefficient is divided by the equivalent threshold value of the quantization matrix.The result is rounded to the nearest integer value, which must be in the range from –127 to +127.

The zero frequency or the DC coefficient is encoded differentially (DPCM) from the DC term of the previous block. Of the 63 other coefficients only the non-zero values are coded using Huffman codes to reduce statistical redundancy. The quantization matrix is the only parameter that can be modified — or needs to be modified — to change the compression result. The same quantization matrix is used for every DCT coefficient block regardless of their position or spatial activity.

The adaptivity in the ADCT algorithm is based on the fact that there are less coefficients to be encoded for low activity blocks and more for high activity blocks. This means that the amount of coded information varies from block to block even though the quantization matrix is the same for every block. This quality factor, which can be chosen in many JPEG implementations, is basically a scaling factor with which one only makes the threshold values of the quantization matrix larger or smaller in order improve the final quality of the compressed image. After scaling, the quantization matrix is used as described above.

One of the major advantages of video colour spaces like YIQ/NTSC and YUV/PAL/SECAM is that the chrominance components can be subsampled and quantized coarser than the luminance. This approach has been used in both analog[Wezel87] and digital TV [CCIR601-1] [Limb *et al* 77]. On the other hand, the situation is a bit different in the case of a printed still image. It has been shown that the compression ratio improved by subsampling the luminance, but a coarser quantization of the chrominance did not make a noticeable improvement in the compression ratios [Yläkoski91]. In these experiments the same quantizer is used for both luminance and chrominance. The quantization matrix suggested in the ISO standard draft [ISO88] is used in the survey.

Experiments

The test image is a 512×512 pixel version of an image showing three ladies with musical instruments. It is scanned in a Crosfield scanner, output in a Scitex system and proofed with Signature. Due to the low resolution of the image, the output format is 5×5 cm. With a larger format, the pixels and the compression blocks would be unrealisticly big.

The image is compressed in four different spaces: cmyk, Karhunen-Loeve transformed cmyk, YUV (CCIR-601) and a colour space which is here named Yab. The Yab space is defined so that the Y is the same Y component as in the XYZ space, $a = X - Y$ and $b = Z - Y$. This definition of Yab is based on the analogy of RGB and video colour spaces like YIQ in which chrominance components are defined as colour difference signals. Since XYZ and RGB have certain similarities, the a and b components could be defined as e.g. I and Q in NTSC.

It can be seen in the following tables that Yab and CCIR spaces have almost equal data reduction properties. The internal correlation of different colour spaces can be estimated by computing the covariance matrix C_x after each colour space conversion. The formula is as follows:

$$C_X = E\{[(X - E(X)] * [(X - E(X)]^T\}$$

where X is the colour vector after a certain colour conversion (the dimension is N), E is the expected value and C_X is the covariance matrix.

For non-correlated vectors X the matrix has only N non-zero elements along the diagonal. The following results were obtained for cmyk data that was first converted to the XYZ space.

XYZ:	+2648.8	+2614.2	+1967.7
	+2614.2	+2602.9	+2007.0
	+1967.7	+2007.0	+1769.2

(energy per component: +0.377 +0.371 +0.252)

Yab:	+23.278	+11.324	–50.619
	+11.324	+2602.9	–595.88
	–50.619	–595.88	+358.08

(energy per component: +0.008 +0.872 +0.120)

CCIR:	+2600.0	+298.12	–539.79
	+298.12	+113.86	–116.64
	–539.79	–116.64	+167.21

(energy per component: +0.902 +0.040 +0.058)

The final compression ratios were 10, 15, 20, 25 and 30 for cmyk and Karhunen-Loeve and 20, 30, 40, 50 and 60 for CCIR and Yab. For the two last mentioned the 2:1 subsampling of chrominance was applied. A simple filter is applied prior to subsampling and interpolation for suppressing moiré.

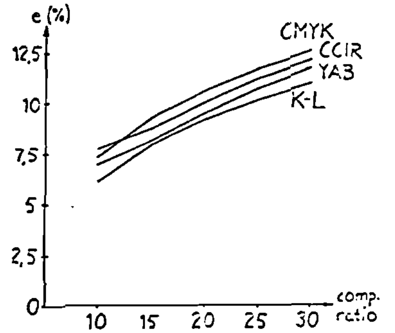
Results

For evaluation, measurements have been made of the relative mean-square error between the original XYZ tristimulus planes and the XYZ tristimulus planes of the coded images (figure 4). The formula used is as follows:

$$e_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\Delta X_i^2 + \Delta Y_i^2 + \Delta Z_i^2\right)}$$

Both of the chrominance-subsampled colour spaces (CCIR and Yab) give smaller quantization errors than compression of raw cmyk data does. In these cases, the total compression ratio has actually been twice as high as indicated in the diagram, because of the data reduction of the colour space conversion. The Karhunen-Loeve transformation gives a superior result compared to the raw cmyk, with the same compression ratio.

As a complement, a visual test has been made with a test panel consisting of twelve persons, mostly image researchers (figure 5). They were asked to group the images into three classes: 0 for no difference compared to an original image, 1 for a just noticeable difference and 2 for an obvious difference.



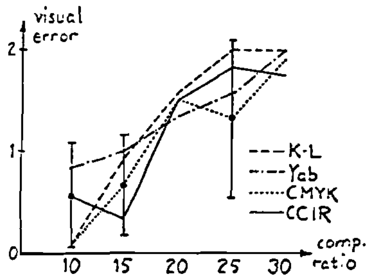


Figure 4. RMS errors in the CIE XYZ colour space.

Figure 5. Visual test results, defined as mean values in a three level scale.

Conclusions and discussion

As indicated by the standard deviations marked in figure 5, this test is certainly not a strong argument for a certain ranking between colour space transformations for this purpose — at each compression ratio, the difference between the quality points is generally less than the standard deviation. The only thing that could be seen in the diagram is that the CCIR and Yab colour spaces, which actually have about twice as high compression ratios as seen in the diagram, still have about the same visual quality as the less compressed cmyk data.

Thus, the mapping from the cmyk space into a subsampled video colour space like YUV 4:2:2 significantly improves the compression ratio. Whether the theoretical improvement of 100% is feasible or not is not sure. The reason is that the 2:1 subsampling of the chrominance does not increase ratio by 50%, because the transformation from RGB to YUV already concentrates most of the signal energy to luminance, so that the subsampling does not give a full increase in the compression ratio.

On the other hand, the colour space used need not be according to CCIR-601 — colour spaces like YIQ will do as well. An interesting result is that the CIE-XYZ space could be useful for compression without further transformation. The only transformation that is needed is subtraction of Y from the other two components. The similarities of RGB and XYZ spaces support this result. The results suggest that CIE XYZ could be a suitable colour space for compressing cmyk images with non-reversible algorithms.

The use of XYZ space might be useful also for the sake of the device dependency, even though the XYZ space does not offer a complete solution. The possibility to choose an appropriate cmyk separation for a given screening technology makes the colour appearance with different screens more similar.

There are some questions that remain to be answered. Were the test material and the test methods reliable enough? How uniform is the Yab space and is the subsampling of a and b justified? If the original cmyk values could not be restored, will it make the whole idea inconvenient? On the other hand, the method developed makes it possible to define black values before hand, so the algorithm may contain any kind of UCR (Under Colour Removal) and GCR (Grey Component Replacement). Unfortunately, the backward calculation of Neugebauer equations is time consuming, so a real application needs special purpose hardware.

As far as RGB images are concerned, the same colour portability applies. Theoretically, the quality of compressed RGB images would improve too, because the use of e.g. the CCIR-601 transformation is suboptimal for RGB images with primaries unequal to the CCIR monitor primaries.

References/literature

Bradley, S. P., Hax, A. C., Magnanti, T. L., "Applied Mathematical Programming", Addison-Wesley 1977

CCIR Recommendation 601-1, Encoding Parameters for Digital television For Studios

Gilge M. "Adaptive Transform Coding of Four-Color Printed Images", Proceedings: ICASSP 87, IEEE 1987 Acoustics, Speech and Signal Processing, vol 3 pp 1374–1377

Frei W., Baxter B. "Rate-Distortion Coding Simulation for Color Images", IEEE TRANS. ON COMMUN., vol 25, no 11, 1977 pp 1385–1392

ISO/STC1/SC2/WG8 N800, May 1988, Initial Draft for Adaptive Discrete Cosine Transform Technique for Still Picture Data Compression Standard

Laihanen, P., Rönngren, H., "Colour Space Conversions in Graphic Arts Industry", Subreport 2/3 of the project Digital Image Processing Adapted to Graphic Production, Helsinki University of Technology, Swedish Graphic Arts Research Laboratory, September 1988

Limb J. O., Rubinstein C. B., Thompson J. E. "Digital Coding of Color Video Signals -A Review", IEEE TRANS. ON COMMUN., vol 25, no 11, 1977 pp 1349–1384

N.N., "Kodak Releases Data on PhotoYCC and Photo CD", Dunn Report, August 1991, vol IX no 8, pp 16–20

Pratt W.,K. "Spatial Transform Coding of Color Images", IEEE TRANS. ON COMMUN., vol 19, no 6, 1971 pp 980–992

Schreiber W., F., Buckley R., R., "A Two-Channel Picture Coding System: II - Adaptive Companding and Color Coding", IEEE TRANS. ON COMMUN., vol 29, no 12, 1981 pp 1849–1858

van Wezel, R., "Video Handbook", London: Heinemann, 1987, ch 1

Yläkoski, I., "Compression of Monochrome and Colour Images", Subreport 1b/3 of the project Digital Image Processing Adapted to Graphic Production, Graphic Arts Laboratory, Technical Research Centre of Finland, July 1991