COLOR TRANSFORMATIONS AND LOOKUP TABLES

Karl J. Heuberger, Zhou Mo Jing, Serdar Persiev*

I. Abstract

In electronic publishing, it becomes more and more important that the color rendering on monitors and on printing substrates is visually corresponding. This reduces time consuming repetitions and waste of material in the pre press stage and in printing. Different color spaces have to be used for the monitors and the output media. In this paper the RGB color space is used for presenting the colors on the monitors, the CIE-XYZ color space for the calculation and the CMY color space for presenting the colors on the printing substrates.

Once a picture from a monitor has to be presented on a printing substrate the color information of every pixel of the picture has to be transformed to the color space valid for the printing process. Three different methods for this data transformation between the two color spaces are explained in this paper.

The first method is calculating the transformation for every pixel by a mathematical model! e.g. "Neugebauer" which is accurate but time consuming.

The second method is calculating the transformation for every pixel by a mathematical analytical model with less accuracy but faster then the first method.

The third method is using measured color test charts as a lookup table. The transformation for every pixel is done by using the created lookup table. This method is fast for the transformation and its accurancy depends on the density of the lookup table.

Six mathematical models were tested in order to get informations about the accurancy of the color transfonnation between the RGB and the CMY color space. Three models work with the well known Neugebauer equations and the other three with first, second and fourth order matrix equations. Deviations of the resulted screen dot sizes calculated by the six models were compared by using two color test charts. In addition the results were also compared with deviations resulted by using the color test charts as low density lookup tables.

* EMPA/UGRA

2. Descriptions

2.1 Description of colors on monitors

The colors on monitors are produced by excitation of a red, a green and a blue phosphor. The colors are described in the RGB color space. Each color on the monitor is represented by the amounts of luminance of the three phosphors.

Since colors on monitors follow the laws of additive color-mixture, they can be described by linear equations. Based on this. color values from one monitor can be convened to another monitor by means of a linear transformation.

$$
\begin{pmatrix} R_1 \\ G_1 \\ B_1 \end{pmatrix} = (A_{ik}) \cdot \begin{pmatrix} R_2 \\ G_2 \\ B_2 \end{pmatrix}
$$
 [1]

 R_1, G_1, B_1 represent the phosphor values (basic vector) of the first monitor, R_2, G_2, B_2 those of the second monitor, and (A_n) is the 3 x 3 matrix of the basic transformation.

2.2 Description of colors on paper

In halftone printing the colors on paper are described by the dot areas of the primary inks. This is the color space CMY with the primary colors cyan. magenta and yellow. The color appearance on paper is a result of additive color-mixture (variation of the dot area) and of subtractive color mixture (superimposition of ink films). This produces a non-proportional sensation including a visually effective dot gain (the geometric dot area of the color on paper appears enlarged to the eye). As a result, the CMY color space is not linear. Generally, linear equations cannot be used in the CMY color space.

2.3 Description of colors in the CIE color space

The standard color space XYZ was defined to describe colors irrespective of a particular reproduction technology. It is based on the color sensitivity of the human eye. XYZ denote the virtual response functions of the CIE standard observer defined in 1931. These are in a linear relationsship with the real sensitivity functions of the human eye.

2.4 Transformation between various color spaces

The conversion of color coordinates from one color space to another is called color space transformation. The transformations are linear, if the involved color spaces are based on additive color mixture. For instance color values represented in the RGB color space can be transformed into the XYZ color space by a linear transformation. Its advantage is that the transfonnation can be mathematically in versed.

$$
\begin{pmatrix}\nR \\
G \\
B\n\end{pmatrix} = (A_{ik}) \cdot \begin{pmatrix}\nX \\
Y \\
Z\n\end{pmatrix}
$$
\n
$$
\begin{pmatrix}\nX \\
Y \\
Z\n\end{pmatrix} = (A_{ik})^{-1} \cdot \begin{pmatrix}\nR \\
G \\
B\n\end{pmatrix}
$$
\n[3]

If the transformation is described by a 3 x 3 matrix (A_n) , the values XYZ will be represented as values RGB in the RGB color space. The inverse matrix $(A_n)^{-1}$ transforms RGB values to XYZ.

Color transformations between additive and subtractive color-mixing systems are non-linear, especially if the CMY color space is involved. A linear conversion, as shown in equation 3, produces strong color deviations and is not suitable for general use. The Neugebauer equations (1937) are a useful approach for the transformation from the CIE color space XYZ to printing color space CMY. Here the color values XYZ are expressed by the eight combinations resulting between paper white and the three ink colors CMY. The formulas are shown in appendix I. While the values XYZ can be directely calculated from the dot area values cmy, the latter have to be calculated from XYZ by an iteration process. The investigation of the color transformation between RGB and CMY can be limited to the transformation from CMY to XYZand vice versa, because of the linearity between RGB and XYZ.

3. Methods

The first method describes three mathematical models based on the Neugebauer equations.

Model 1.1: Neugebauer equations

According to appendix I, the Neugebauer equations have been used without modifications. The dot area CMY is calculated from the color values *XYZ* by numerical iteration.

Model 1.2: Modified Neugebauer equations

The Neugebauer equations are modified by exponents, i.e. by an exponent e, f and g for each tristimulus value. The modified Neugebauer equations are shown in appendix 2. With three exponents 3 dot areas can be corrected.

In the present work the dot areas of the measuring point $(c,m,y) = (50,50,50)$ have been corrected. The dot gain from film to paper has been included with the choice of the exponents.

Model 1.3: "cellular" Neugebauer equations

For this case the CMY color space has been divided in 8 cells (see fig. I). These are determined by 27 corner points which are the combination of solid-tone and 50%-halftone colors of the 3 primary inks. For each measuring point the dot area is first calculated with model 1 what determines the cell. Then, the dot areas cmy can be calculated iteratively within the corresponding cell from the values *XYZ.*

(e,m,y) : Dot area of the primaries cyan, magenta. yellow

0 : Corner points of the eight cells in the CMY color space

The second method describes three mathematical models based on Matrix equations.

Model 2.1: First order matrix

Although the CMY color space is non-linear, the linear conversion is used for comparison and as easiest transformation. For reversible transformations between CMY and XYZ it is favorable, if the basic values in the XYZ space run parallel to the values in the CMY space. This applies to the color density and absorption values. **Jn** this investigation, fromXYZ virtual absorptions are defined. Hereby the paper white and black (solid-tone value from all three primaries) are included as comer points:

$$
A_x = 1 - \frac{X - X_s}{X_w - X_s}
$$
 [4]

 X is the tristimulus value belonging to the color to be transformed, X_n and X_n those of the paper white and black in the XYZ space. The same equations apply to the tristimulus values *Y* and Z. For the linear transformation, a 3 x 3 matrix is sufficient.

$$
\begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = (A_{ik}) \cdot \begin{pmatrix} c \\ m \\ y \end{pmatrix}
$$
 [5]

The base in the XYZ space are the virtual absorptions A_x , A_y , A_z as defined in formula [4]. while lhe dot areas c, m andy describe the CMY color space. The transformation matrix (A_{ik}) is defined by the absorption values $A_{k}A_{k}A_{j}$ belonging to the tristimulus values XYZ.

$$
(A_{ik}) = \begin{pmatrix} A_{xc100} & A_{xm100} & A_{xy100} \\ A_{yc100} & A_{ym100} & A_{yy100} \\ A_{zc100} & A_{zm100} & A_{zy100} \end{pmatrix}
$$
 [6]

The dot areas of the measuring point are calculated by the inverse matrix belonging to matrix (6] if the tristimulus values *XYZ* are given.

Model 2.2: Second order matrix

The advantages of the linear transformation (i.e. the backward transformation with an inverse matrix) and the higher accuracy of the quadratic transformation are combined in conversion [7].

 A_{α} , A_{α} and A_{β} are the virtual absorption coefficients in the XYZ color space, c is dot area, and A_{μ} are the coefficients to be determined.

The coefficients of the 9 x 9 matrix (A_{ik}) are determined by the 9 calibration points:

[7]

$$
\begin{pmatrix}\nc \\
m \\
y \\
c\n\end{pmatrix}\n= (A_{ik}) \cdot\n\begin{pmatrix}\nA_x \\
A_y \\
A_z \\
A_x A_y \\
A_x A_z \\
A_y A_z \\
A_y A_z \\
A_y A_z \\
A_z\n\end{pmatrix}
$$
\n[7]

100% c, 100% m, 100% y

50% c, 50% m, 50% y Ľ.

100% c+m, 100% c+y, 100% m+y

Model 2.3: Fourth order matrix

In Order to get better accurancy than the second order matrix formulas with fourth order are introduced. The fourth order polynoms are represented by the following matrix formula [8].

points: The coefficients of the 12 x 12 matrix (A_n) are determined by the 12 calibrations

The **third method** describes how the testforms are used as lookup tables.

Model 3.1: Lookup tables

For transformations with lookup tables the density of the lookup table is very important. Tn this work we used a lookup table with low density consisting of 27 look up points. For this case the look up points were the corner points of Fig. 1.

For every measured point the nearest four comer points were determined. After that a linear interpolation was made using the four comer points.

3.1 Comparison of the mathematical models

The comparison of the different mathematical models has been carried out with 2 test forms. They contain 20 x 20 mm respectively 15 x 20 mm large measuring points of the primary colors C,M,Y printed on coated paper in the off-press proofing system. All the measurements have been made with an Elrepho 2000 spectrophotometer under illurninant C/ 2° with an aperture of 12 mm. The tristimulus values *X*, *Y* and *Z* have been determined. In addition all monochrome measuring points have been measured with the Gretag densitom-

eter D 186. The dot area on paper has been determined from the specified halftone values on film and the measured dot gain.

In case of measuring points having superimposed colors the corresponding dot gains of the monochrome measuring points have been used.

3.2 Test process

The mathematical models are calibrated by the tristimulus color values XYZ of the. calibration points and the corresponding dot areas on paper (see fig. 2).

Fig. 2 Calibration of the mathematical models

Basically the models can be tested in two ways, i.e. by calculating the corresponding dot areas for the given *XYZ* values, or by calculating the *XYZ* values, if the dot areas are given. In the first case a color difference E and in the second case a dot area dif ference is obtained as measure for the accuracy of the model. In this work it has been decided to calculate the dot areas from XYZ (see fig. 3). ln this case the model is tested in a more ambitious way, because with the Neugebauer equations an iteration is necessary. In addition, the dot area is more meaningful, because in practice the models are also used for this purpose.

Fig. 3 Test of the mathematical models

4. Results

Table I gives a survey of the 6 mathematical models. For each model the required number of calibration points and the kind of transformation (directly or iteratively) is shown. Appendix 3 shows a list of the required calibration points for each model.

4.1 Accuracy

Table 2 and 3 show the accuracy of the mathematical models for the two test charts. In test chart 1 the mathematical model 1.3 ("cellular" Neugebauer equations) cannot be applied, because the test chart does not contain all the 27 calibration points.

The number of measured points were 33 in test chart I and 53 in test chart 2. The accuracy is defined as the difference between the calculated and the specified dot area. The dot area difference has been calculated for each single color (m, c, y) and for all three colors. For the latter a total dot area difference F has been defmed:

$$
F = (c \t2 + m \t2 + y \t2)^{1/2}
$$
 [9]

As a measure for the variation of the differences the standard deviation has been calculated. The total difference F for the dot area of all three primary colors corresponds to the mathematical defmition of the color difference E.

All mean values c, m, y and F have been calculated from the absolute dif ferences of the measuring points. The calculated dot areas of the single colors have been cut to obtain values between 0 and 100%. Negative values were taken as 0, values above 100 as 100.

Nr.	Model	Number of	Transformation			
	Name	calibration points	$XYZ \rightarrow CMY$	$XYZ \rightarrow CMY$		
1.1	Neugebauer equations	8	direct	iterative		
1.2	modified Neugebauer equations	9 direct		iterative		
1.3	"cellular" Neugebauer equations	27	direct	iterative		
2.1	first order matrix	5	direct	direct		
2.2	second order matrix	11	direct	direct		
2.3	fourth order	14	direct	direct		

Table I: List of the transformation models

Test chart 1	$c\%$		$m\%$		$V\%$		F%	
33 Measuring points	Mean value	Std. dev.	Mean value	Std. dev.	Mean value	Std. dev.	Mean value	Std. dev.
Model 1.1	1.3	1.8	1.9	3.6	2.9	4.5	5.0	5.0
Model 1.2	1.2	1.2	1.7	1.9	1.2	1.9	3.0	2.5
Model 1.3	-							$-$ *)
Model 2.1	3.3	5.8	8.7	18.1	13.9	20.3	22.4	24.9
Model 2.2	1.0	1.5	2.5	4.1	5.9	11.7	8.0	11.6
Model 2.3	0.7	0.9	0.7	0.8	5.4	8.2	6.5	8.1

Table 2: Deviation of the calculated areas for test chart I

*1 Test chart I has a too small number of measuring points

Test chart 2	$c\%$		$m\%$		$y\%$		F%	
53 Measuring points	Mean value	Std. dev.	Mean value	Std. dev.	Mean value	Std. dev.	Mean value	Std. dev.
Model 1.1	1.8	2.3	2.6	3.1	3.4	3.8	5.8	4.0
Model 1.2	2.2	2.2	2.7	2.8	2.3	2.5	5.1	3.2
Model 1.3	0.8	1.7	0.7	1.5	0.9	1.8	1.9	2.6
Model 2.1	5.8	7.1	14.9	17.7	21.0	20.3	29.7	24.2
Model 2.2	4.1	5.0	4.4	5.6	9.5	13.0	13.7	12.8
Model 2.3	1.9	2.4	4.2	4.2	6.2	7.0	8.6	7.5
Model 3.1							12.6	16.1

Table 3: Deviation of the calculated dot areas for test chart 2

4.2 Reproducibility

The mathematical model 1.1, 1.2, 1.3 and 2.1 are fully reversible, i.e. for each measuring point the transformation can be applied foreward and backward, and the result corresponds to the initial value.

For the quadratic transformation with a 9 x 9 matrix (model 5), however, a deviation occurs for all values except for the calibration points. The magnitude of the deviation is shown for

both test charts in table 4. The maximum deviations occur at measuring points being overlaps of the three primary colors.

Table 4: Irreversible deviations after one transformation cycle
 $(CMX \rightarrow YYZ \rightarrow CMY)$ \sim CMY)

4.3 Number of mathematical operations

The number of mathematical operations for the calibration of the models and for the transformation of one measuring point are listed separately.

Calibration of the mathematical models

The number of operations required to calibrate each model is shown in table 5. The calibration includes the necessary mathematical steps to transform XYZ to CMY and vice versa.

Mathematical steps for a transformation

The number of mathematical steps for a transformation is shown in table 6. For the transformation $XYZ \rightarrow CMY$ with the Neugebauer equations the number of iterations has to be included.

Table 5: Number of mathematical operations for the calibration of the transformation (XYZ \leftrightarrow CMY)

*1 estimated number of operations to optimize the exponents

5. Discussion of the results

Table 7 shows a summary of the characteristics of the 7 tested models.

5.1 Model 1.1

The difference between the calculated dot area and the measured dot area is comparable for both test charts in all colors. Test chart I has a slightly larger variation of the relults than test chart 2.

5.2 Model 1.2

If the Neugebauer equations are modified with exponents, an additional calibration point is included. As a result, the deviation in test chart I is considerably smaller, the one for test chart 2 slightly smaller. As can be expected, the deviations near the calibration point are smaller. but certain deviations near the comer points of the color space are larger.

	$XYZ \leftarrow CMY$ Number of operations:	$XYZ \rightarrow CMY$ Number of operations:	
	Operations for a transformation	Operations without iteration	Total with iteration $*)$
Model 1.1	105	402	281'400
Model 1.2	159	458	320'600
Model 1.3	105	840	588'000
Model 2.1	30	30	(30)
Model 2.2	66	66	(66)
Model 2.3	81	81	(81)

Table 6: Number of mathematical operations for a transformation

*' about 700 iterations are necessary to obtain an agreement of E 0.01

5.3 Model 1.3

The "cellular" Neugebauer equations could only be used on test chart 2 with 53 measuring points. To divide the color space in 8 cells, 27 measuring points have to be used as calibration points. As can be expected, the mathematical accuracy is considerably higher than with model 1 and 2. The "cellular" Neugebauer equations show among the 6 models the best results. but they require the largest number of mathematical operations.

5.4 Model 2.J

The linear conversion requires a much smaller number of mathematical operations, but it shows very large deviations. The linear transformation works very well, if measuring points consisting of only one primary color are used. In this case the maximal F is smaller than 3%, and the mean value is only 1.7% for test chart I. There are larger values in test chart 2, i.e. F is maximally 12% and as the mean value it is 4.9%. However , measuring points with 2 and 3 colors are only transformed with large deviations. The maximum deviation has been achieved in the combination fthe solid-tone colors. The measuring point (c,m,y) = $(100,100,100)$ is calculated with model 2.1 in test chart 1 to $(c,m,y) = (82.33,44)$ and in test chart 2 to $(c,m,y) = (75,44,44)$. The measuring point $(c,m,y) = (0,100,100)$ has the second largest deviation.

Table 7: Comparison of the mathematical models

5.5 Model 2.2

The quadratic conversion includes as calibration points the single color solid-tones and the combinations of two colors.

In test chart 1 the mean values for the deviations are slightly larger compared with the unmodified Neugebauer equations (table 2). The measuring point $(c,m,y) = (100,100,100)$ has, as in model 2.1. the maximum deviation.

ln test chart 2 (table 3) the mean values show deviations being twice as large as those with the regular Neugebauer equations (model 1.1). The measuring point (c,m,y) = $(100, 100, 100)$ and the measuring points with mainly magenta and yellow, such as (c, m, y) = (50, 100, 100), (0,30,30), (0,50,50) and (0,70,30) show the largest deviations.

Compared with the linear model, a considerable improvement of the accuracy can be seen in both test charts.

In test chart 1 the measuring points $(c,m,y) = (100,100,100)$ and $(c,m,y) = (0,25,25)$ show the largest irreversible deviations after a transformation cycle CMY \rightarrow XYZ \rightarrow CMY. The measuring points with three-color overlap show the largest irreversible deviations in test chart 2.

ä,

5.6 Model 2.3

The fourth order model gives a much better accurancy than the linear (model 2.1) and the second order matrix (model 2.2). In addition to the second order matrix three 50 % halftone overlaps were used as calibration points.

As in model 2.2 the measuring point $(c, m, y) = (100, 100, 100)$ shows the largest deviation and also the largest irreversible deviation.

5.7. Lookup tables

Results from transformations with the chosen low density lookup table are listed in table 3 (model 3.1). The deviations are slight bigger than the deviation with the quadratic matrix model. The deviations have to be compared with the density of the chosen lookup table. The density of the lookup table is (50% smalest distance between neighbour points).

6. Conclusions

The "cellular" Neugebauer equations show the best results. Good results which should be obtained in practice, ought to be smaller than 2%, expressed as total mean difference of the dot area. Moreover, the standard deviation should be smaller than 2%. In this case 90% of all transformed values show a smaller deviation than \pm 4%. These mean values are obtained by the "cellular" Neugebauer equations. But the variation is too large.

The large amount of mathematical steps required for the models 1.1, 1.2 and 1.3 is caused by the iteration process necessary for the Neugebauer equations. Matrix transformations only require one mathematical effort, namely the calibration.

Matrix transformations are therefore suitable for fast transformation processes "on the fly", but the accuracy has to be improved.

For good results with look up tables the smalles distance between neighbour points should be 6% when using linear interpolation and 2% without interpolation. That means that between 5800 and 132'600 points have to be measured or calculated.

Appendix I

Neugebauer equations

Primary colors n

i.e. :

Appendix 2

Modified Neugebauer equations

Primary colors **n**

i.e.:

Appendix 3

i.

List of the calibration points

882

 $\frac{1}{2}$

 $\sim 10^{-10}$