WHY DO COLOR TRANSFORMS WORK?

John Seymour
Quad/Tech, International
Sussex, WI, USA

ABSTRACT

Numerous papers have been written regarding techniques to translate color measurements from an RGB device (such as a scanner or camera) into some standard color space. The papers seem to ignore the mathematical "truth"... that the translation is impossible. Do these color transforms work, or under what conditions do they work, and what are the limitations? Why is it that they work (if they do)?

In this paper, light emitting diodes (LEDs) are viewed with a color video camera. It is seen that, for the spectra of LEDs, transforming from the camera color space to XYZ tristimulus space leads to very large errors. The problem stems from the fact that the RGB filter responses are not a linear combination of XYZ responses.

Next, it is shown that the transformation of CMYK halftones does not pose such difficulties. Here, it is found that a simple linear transform is relatively accurate, and some options to improve this accuracy are investigated.

Several explanations are offered as to why transforms of CMYK are more accurate than transforms of LEDs. To determine which of the explanations is the most likely, linear transforms are applied to a variety of collections of colors.

Keywords: Color transform, colorimetry, device independent color

1. EXAMPLE OF AN "IMPOSSIBLE TRANSFORM"

A handful of LEDs and a color video camera are all that is required to demonstrate that color transforms are "impossible". LEDs were selected with peak wavelengths of 555 nm (green), 568 nm (yellow-green), 585 nm (yellow), 610 nm (amber), 635 nm (orange-red) and 660 nm (red). The video camera used was a commercially available three CCD camera. The output of the camera was fed into a frame grabber for analysis. The camera was color balanced so that white paper under fluorescent office lighting registered equal intensity in all three channels.

Each of the six LEDs was placed in front of the camera, and the resulting images were analyzed to determine the average relative intensities in the red, green and blue channels of the camera. Table 1 shows the results. The intensities were scaled for unit red channel response.

Two idiosyncrasies are evident from Table 1. First, it is clear that, despite the obvious difference in visual appearance, the camera is incapable of distinguishing among the amber, orange-red and red LEDs. Second, despite the fact that the the green and the yellow-green LEDs are relatively close in color appearance, there is a huge difference in the camera response between the two (5.700 vs. 1.602).

One potential explanation for these idiosyncrasies is that they are artifacts of the nonlinearity of the color processing in the human visual system. It is known that colors which are close are not necessarily “close” when their reflectances are compared. This leads to two hypotheses: (1)

<table>
<thead>
<tr>
<th>LED Color</th>
<th>Wavelength</th>
<th>Green Response</th>
<th>Blue Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>555 nm</td>
<td>5.700</td>
<td>0.041</td>
</tr>
<tr>
<td>Yellow-Green</td>
<td>568 nm</td>
<td>1.602</td>
<td>0.017</td>
</tr>
<tr>
<td>Yellow</td>
<td>585 nm</td>
<td>0.115</td>
<td>0.007</td>
</tr>
<tr>
<td>Amber</td>
<td>610 nm</td>
<td>0.015</td>
<td>0.007</td>
</tr>
<tr>
<td>Orange-Red</td>
<td>635 nm</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>Red</td>
<td>660 nm</td>
<td>0.007</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Table 1 - Relative response of video camera to various LEDs

---

*a* Email - jseymour@qtiworld.com
That amber, orange-red and red are close in terms of their reflectances, and that the human visual system exaggerates the differences among them, and (2) That green and yellow-green are relatively far from each other in terms of their reflectances, and the human visual system minimizes the difference between them.

To test these hypotheses, the closest matches to the color of each of the LEDs were selected from a Pantone Matching System booklet. If the idiosyncrasies are caused by the nonlinearity of the human visual system, then the same idiosyncrasies would be found in comparing the camera's response to the Pantone patches. The visual matches were performed under fluorescent office lighting, and the same lighting was used to illuminate the patches for the camera. In general, a lower voltage on the LEDs was needed in order to make a match.

Table 2 shows the results of analyzing the Pantone patches with the video camera. It is seen in this table that the transition from green to red is considerably smoother. The conclusion is that the idiosyncrasies seen in the LED are not an artifact of the nonlinearity of the human visual system, but are an artifact of the video camera. The video camera does not see color the same way we do. In essence, the video camera is colorblind when it comes to LEDs.

The spectra of LEDs fall into the class of spectra which are impossible to transform from the video camera to the human visual system. They just don’t follow the rules!

### Table 2 - Relative response of video camera to PMS equivalents

<table>
<thead>
<tr>
<th>LED Color</th>
<th>Equivalent Pantone</th>
<th>Green Response</th>
<th>Blue Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>358</td>
<td>1.546</td>
<td>0.157</td>
</tr>
<tr>
<td>Yellow-Green</td>
<td>375</td>
<td>1.341</td>
<td>0.630</td>
</tr>
<tr>
<td>Yellow</td>
<td>380</td>
<td>1.094</td>
<td>0.201</td>
</tr>
<tr>
<td>Amber</td>
<td>1235</td>
<td>0.522</td>
<td>0.101</td>
</tr>
<tr>
<td>Orange-Red</td>
<td>164</td>
<td>0.276</td>
<td>0.162</td>
</tr>
<tr>
<td>Red</td>
<td>192</td>
<td>0.148</td>
<td>0.221</td>
</tr>
</tbody>
</table>

2. **THE PURIST’S STANDPOINT - WHY ARE THESE TRANSFORMS IMPOSSIBLE?**

To verify the conclusion from the previous section, a crude method was used to determine the spectral characteristics of the video camera. We used a linear variable filter, which is a narrow-band interference filter with the center of passband changing linearly across the filter. When this filter is backlit, a rainbow pattern is seen. The filter was backlit with an incandescent light of known color temperature. When the camera is focused on this filter, the average intensity vs. position on the filter is indicative of spectral response of the camera. A correction was made for color temperature. Calibration of position in image to wavelength was performed by introducing additional interference filters of known wavelength.

Figure 1 shows the camera spectral response determined in this way. It is readily apparent from this graph why the amber, orange-red and red LEDs were indistinguishable. It can be seen that only the red channel has any appreciable response above 600 nm. The three LEDs, at 610 nm, 635 nm, and 660 nm, varied in intensity due to red channel sensitivity (and LED efficiency), but the hue and saturation, which are reflected in the relationships between the channels, remained the same.

The large difference between the green and yellow-green LEDs can also be explained by Figure 1. In the region from 555 nm to 568 nm, the green channel is quickly falling off, and the red channel is quickly rising. This accounts for the rather abrupt change in the green-to-red ratio.

Why is it that these spectral characteristics were selected? Isn’t it possible to make a video camera with truer color?

First we must ask a seemingly simple question. What would we like the spectral response of the video camera to be? We would...
like for the color seen on the monitor to look like whatever the camera is pointing at. Since the phosphors for a monitor have been defined, this criteria is enough to decide the spectral response for video cameras.

Figure 2 shows the required spectral responses for a camera which feeds a display with SMPTE phosphors [1, 2, 3, 4, 24]. The disturbing part of these plots is that large parts of these three curves are negative. For example, the red response between 450 and about 540 nm is negative. This makes the camera a bit difficult to build! Such a camera had been envisioned as early as 1951 [3]; however, commercial camera designers generally favor a simpler design which is less colorimetrically precise. Broadcast quality cameras usually use the more complicated scheme [4].

In Figures 3 through 5, the SMPTE color matching functions from Figure 2, are individually compared with the spectral responses of the video camera. It is plain that the filters approximate the positive parts of the curves while completely ignoring the negative parts.

In the example of the amber, orange-red and red LEDs, we saw that a color transform is not possible, not only because the video camera’s spectral response is somewhat less than ideal, but also because the camera threw some information away. This information is crucial for discriminating among the three LEDs. The video camera has missed the mark (in part) because of a difficult requirement: it must have negative response at some wavelengths. Is it possible to define a set of physically realizable (that is, all positive) spectral responses for a video camera which do not throw any color information away?

The answer is “yes”. Figures 6 and 7 are two examples of such spectral responses. Figure 6 is the CIE tristimulus curves. A camera with these spectral responses (which could be built) would directly output XYZ values, so the transformation to XYZ would be trivial. (To display the colors properly on a video monitor would, however, require a 3X3 matrix transform.) Figure 7 is a linear transformation of the XYZ tristimulus curves, constructed with the following formula:
A video camera with the $fgh$ response in Figure 7 could provide XYZ numbers simply by multiplying the $fgh$ numbers by the inverse of the matrix in equation 1. The spectral responses depicted in Figures 6 and 7 meet the Luther-Ives condition in that they are linear combinations of the tristimulus functions [2, 5, 6].

The conclusion of the purist: The spectral responses of a video camera or scanner must exactly meet the Luther-Ives condition in order for a color transformation to exist. If the Luther-Ives condition is not met, large errors are possible because information has been lost.

3. THE PRAGMATIST’S STANDPOINT - A BRIEF REVIEW OF THE LITERATURE

The literature contains many references to color transforms. A variety of techniques are used including: polynomial transformation, interpolation from a look-up table, neural networks, and a model-based method. The following is a representative sampling of papers:

3.1. Polynomial transform

Wandel and Farrell [7] converted the RGB data from two scanners into XYZ. They use a 3X3 matrix transform and measure average $\Delta E$ of 3-6. By taking advantage of some knowledge about the typical errors, they have reduced this to $\Delta E$ of 2.

Engeldrum [8] transformed RGB (using hypothetical scanner spectral responses) into XYZ values with $\Delta E$ of 1.2 to 2.7.

Chu [9] converted RGB into XYZ, using a 3X3 matrix transform. The calibration set was of 216 CMY patches, and the test set was of 66. A $\Delta E$ of 3.9 is reported.

Balasubramanian and Maltz [10] use a “locally linear transform” to convert from CMYK into CIELAB and report a $\Delta E$ of 2.6.

Södergård et al [11] converted RGB from two video cameras and a scanner into XYZ. They calibrated with 236 test patches, and tested with the same patches. Their $\Delta Es$ are reported between 1.8 to 9.5.

Mumzhiu and Bunting [12] reported using RGB from a CCD camera to compute XYZ values. They transform with a second order polynomial, but no results were published.

3.2. Interpolation from a look-up table

Agar and Allebach [13] used interpolation on non-uniform grid to convert from CMY into CIELAB. They report a $\Delta E$ of 1 for 1000 grid pts and 0.4 for 15,000.

Kasson et al [5] and Kasson et al [6] investigated a large variety of interpolation schemes for look-up tables. They reported $\Delta Es$ of 1 to 2.

3.3. Neural nets

Abe and Marcu [14] and Marcu and Iwata [15] used neural networks to convert CMYK into monitor RGB values. They reported $\Delta E$ of 15 to 20 if the black was unconstrained, and a $\Delta E$ of 3 if the black level is fixed with a Gray Component Replacement (GCR) technique.

Tominaga [16] used neural networks to convert from CIELAB to CMYK with a mean $\Delta E$ of 2.5. In a separate paper [17], he used a neural network with 183 parameters to make a conversion from CMY to XYZ with a mean $\Delta E$ of 2.6. His training set included 216 samples, and the test set contained 125 samples.

Kang and Anderson [18] used a neural network to convert from RGB (from a color scanner) to XYZ. Many results are provided, but the most useful data is their “generalization” data, where the net is trained with one set of data and tested with another. In the tests where they trained on 34 CMYK patches and tested on 202, they report a $\Delta E$ of 8 to 12 (depending on parameters they select), using 3X6 matrix transform, a $\Delta E$ of 3.3 and 3X14 matrix transform, a $\Delta E$ of 4.5.

Arai et al [19] converted LAB to CMY dot area with a neural network. The net was trained with 125 samples, and tested with 360. They report a $\Delta E$ of 2.9.
Chu and Feng [20] used the same data as in [9] to convert from RGB to CIELAB. The training set was 216 patches, and the test set 66. A ΔE of 1.2 was reported.

3.4. Model-based

Berns and Shyu [21] developed a transform based on physical models to convert RGB scanner data from transparencies and from photographs into CIELAB. Mean ΔEs of 0.4 to 1.0 are reported.

Conclusion of the pragmatist: The spectral responses are close enough for transforms to work. They might not all be linear transforms, but they work.

4. EXPLAINING THE CONTRADICTION

One possible explanation is that all the devices meet the Luther-Ives condition. Other authors have stated that this is often the case for commercial video cameras and scanners [5, 6, 8, 22]. My experience has been that this is not the case for video cameras. If the Luther-Ives condition was met, then a linear transformation would be adequate, and the authors would not have resorted to more complicated schemes.

Other possible explanations:

1) The transforms don’t really work. The transforms might only work when applied to the data set which was used for calibration. This explanation does not appear to be very likely, considering the mass of papers claiming that the transforms do work!

2) The transforms work only for a limited set of pigments. The combinations here do not include all possible spectra. For the limited set of spectra, such as those spectra which can be produced with combinations of CMYK, a one to one correspondence exists, so a transform can be done.

3) The transforms work because the reflectance spectra of most solid objects are fairly smooth. Because of this, the dimension of spectral space that the transforms must work in is rather small.

5. EXPERIMENT #1 - DO THE TRANSFORMS REALLY WORK?

In this first experiment, the first explanation was tested by using different sets of data for calibration and testing. Spectra were collected from four different sets of test targets, all of which were printed by web offset press, but on different presses. XYZ values and an estimate of the RGB response for a hypothetical camera were computed. The first set of data was used to calibrate a conversion from RGB space into XYZ space. This conversion was then used to convert the other sets from RGB to XYZ. The results were analyzed.

The four data sets are as follows:

(1) A set of 27 patches of cyan, magenta and yellow ink. All possible combinations of 0%, 50% and 100% halftones were in this set.

(2) A set of 414 patches which sample CMYK space.

(3) A set of halftone scales. Each scale is comprised of 16 levels, equally spaced from 0% to 100% halftone. Each of these scales was printed at seven different inking levels so that the entire set covers a reasonable range in ink film thickness and the complete range in halftone for each of the four inks. There are 448 patches in this set (16 halftone levels X 7 inking levels X 4 inks).

(4) A collection of 106 neutral gray patches of assorted intensity, with gray being made up of CMY, CMYK and K. This set was chosen specifically to test whether the additional pigment (black) poses a problem, since with four inks there are multiple ways to produce the same shade of black.

The camera spectra (see Figure 1) were used to estimate the camera response to the samples under D50 lighting. The computational approach was chosen over direct measurement with a camera because it was less work, and this allowed separating experimental errors from the effect of spectral differences.

In some sense, the four methods described are equivalent. Given enough degrees of freedom (as in the degree of polynomial, storage locations in a look-up table, elements in a neural net, or complicated enough functions), the methods can have the same level of performance. Since the thrust of this paper is a question common to all methods (whether they can work, and why), only matrix conversions were chosen, primarily because the tools for matrix arithmetic were readily available.

The 27 RGB and 27 XYZ values from the first data set were used to generate the following least-squares conversion matrices:
Table 1 summarizes the average CIELAB ΔE errors when equations 2 and 3 were applied to the four data sets. Note that the optimization was performed in reflectance space, so the transforms were not optimized for CIELAB. Optimizing in CIELAB space would reduce the ΔE [21], but the computations become considerably more difficult.

![Equation 2](equation2.png)

![Equation 3](equation3.png)

<table>
<thead>
<tr>
<th></th>
<th>Mean ΔE, 3X3</th>
<th>Mean ΔE, 9X3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration set</td>
<td>6.3</td>
<td>1.4</td>
</tr>
<tr>
<td>CMYK Sampler</td>
<td>4.4</td>
<td>1.7</td>
</tr>
<tr>
<td>Tone scales</td>
<td>4.8</td>
<td>1.9</td>
</tr>
<tr>
<td>Assorted grays</td>
<td>3.4</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Table 1 - Results from the first experiment

One of the first observations from Table 1 is that (paradoxically) the 3X3 conversion matrix calibrated for data set 1 works better for the other three data sets than it does for data set 1! This is merely a reflection of the fact that the calibration set, which covers the gamut of CMY space, is inherently more difficult than the other sets. It can be generally observed that the calibration data set in the 3X3 case translates quite well to the other three data sets.

In the 3X9 case, the calibration set did not perform quite as well. In particular, data sets 2, 3 and 4 all performed worse than data set 1. That said, it must be noted that the 3X9 matrix did significantly improve performance in all cases.

The fact that data set 4 had a ΔE of comparable size to the other sets indicates that having the additional pigment is not a major problem for conversion. It is interesting that the gray set was the most easily matched with a 3X3 transform, but it was the most difficult set for the 3X9 transform. This suggests that the reason that the 3X3 transform works, and the reason that the extra parameters improve the match for the 9X3 transform are due to different mechanisms.

The conclusion of this experiment is that these simple transform methods do work. One can use a small calibration set to determine a conversion matrix which can be applied to other data. The 3X9 transform is an improvement over the 3X3, but the amount of improvement depends on the data set.

6. **EXPERIMENT #2 - HOW WIDELY CAN THE TRANSFORM BE ADAPTED?**

This experiment was performed to test explanation number 2. How well does the previous calibration work on other pigments? To test this, I collected spectra from four other sets of samples, including the following:

(5) The 24 color patches from a MacBeth color checker.

(6) A selection of 20 saturated colors in a Munsell color tree.

(7) A selection of 24 patches from the Pantone Matching System, including the primaries.

(8) The 96 crayons from a Crayola “Big Box of Crayons”.

Table 2 summarizes the results from the second experiment. The errors for the 3X3 conversion are roughly on the same order or slightly larger than the web offset data sets in the first experiment. Thus, the calibration for a 3X3 matrix conversion can be applied to a wider collection of color sources.
The errors for the 3X9 case are a different matter. It will be noted that, in three of the four data sets, the 3X9 matrix actually performed worse than the 3X3 matrix did and performed only marginally better for the remaining data set.

Why did the 3X3 transform work so well? The tentative conclusion is that the 3X3 matrix conversion is a “universal”. The fancier 3X9 transform, on the other hand, does not translate well to other color sources. Its extra “magic” depends upon the spectral characteristics of the samples, and not on intrinsic relations between the video camera and the tristimulus curves. To quote Kang and Anderson, “...the color space conversion does follow some analytical expressions, such as the Neugebauer equations.” The coefficients in the matrix in equation 3 are essentially the coefficients for a three dimensional Taylor series expansion of this analytical expression.

7. EXPERIMENT #3 - RETURNING TO THE LEDS

This paper began by demonstrating that measuring the color of LEDs is a tough test for a video camera. This was followed by a quantitative investigation of how accurate a video camera can be at measuring color. We need to return to the LEDs and quantify the colorimetric accuracy of the video camera measuring LEDs.

A set of spectra which represent a “ideal set of LEDs” was constructed. Ideally, each LED would give off light with a wavelength range of about 20 nm. Ideally, LED would have peak wavelengths centered at every 10 nm from 390 nm up to 720 nm. These spectra were constructed as a data file. The peaks of the spectra were scaled to a peak height of 1.0 so that the spectra would all be physically realizable (albeit uncommon) reflectance spectra.

The first thing discovered when the software was run was that some of the XYZ values delivered from the transform were negative! This is an artifact of the transform. The matrix which performs the transform (equation 2) has negative values in it. It can be seen that Y can be negative if B is greater than zero, and R and G are either very small or zero. This is the case for many of the LEDs in the range 400 nm to 500 nm.

Since CIE has not defined $L^*a^*b^*$ when one of the tristimulus values is negative, a check was added to the software which sets any negative XYZ values to zero. With this in place, it was found that the mean $\Delta E$ was 40.29, with a maximum value of 161, with nine above 50! The correct $L^*a^*b^*$ value at this maximum was {11.94, 35.21, -90.47}, whereas the transformed value was {-16., 184.27, -145.53}. Presumably the calculated value is a super-purple outside of the physically realizable gamut of color space.

The conclusion is that the 3X3 transform is not universal. It will not work for all possible spectra. It fails miserably on this concocted set of LEDs. Why did the 3X3 transform work so well for the diverse sets of colors that were chosen, and so poorly in the admittedly contrived case of the “ideal set of LEDs”?

8. EXPERIMENT #4 - WHAT IS THE DIMENSION OF REFLECTANCE SPECTRAL SPACE?

To recap the conclusions so far, it has been seen that color transforms are theoretically “impossible”. From a practical standpoint, however, they work. A simple 3X3 matrix transform gives fair results in all cases except the LEDs. In special cases (where the pigments are limited to CMYK, for example) a more complicated 3X9 transform can improve the results.

What has not been established is whether the fair results of the 3X3 transform is a result of the camera nearly meeting the Luther-Ives condition, or if the limited set of pigments chosen are just not a very exhaustive test of color transforms. The results of experiment #3 suggest that the set of available spectra is a key issue.

Following the lead of Wandell [7, 23], The dimension of the spectral spaces in experiments #1 and #2 were determined using singular value decomposition. This is a technique which can be used to determine a set of basis spectra for a larger collection of spectra. These basis vectors are spectra which can be combined linearly to approximate all the spectra in the set to some specified tolerance.

The 995 CMYK spectra (from the first experiment) can be approximated (with 1% tolerance) with only 5 basis spectra. One might think at first that there should be only four basis spectra, one for each of the inks used. This would be the case if the spectra of the inks combined linearly. Unfortunately, the spectra combine in much more complicated ways.

The collection of 165 spectra from the second experiment can be similarly approximated only when 12 basis spectra are used, but a modest approximation can be made with the first three. This suggests that it may not be possible to increase the accuracy of these color transforms beyond the accuracy of the 3X3 transform. The basis being limited to 12 spectra is related to the fact that the 3X3 transform performed reasonably well on these samples. The basis is small.
because the reflectance spectra of these samples are in general fairly smooth. The causes of this smoothness are discussed by Rossotti [25].

For comparison, the same singular value decomposition was performed on a set of ideal LEDs. The dimension of this space was determined to be 31.

9. SUMMARY

Subjects for color transforms can be grouped into three categories: well behaved, marginally well behaved, and ill behaved. Well behaved subjects include CMYK samples. It can be expected that small calibration (or training) sets will generalize well to other collections of CMYK samples, and higher order methods will improve performance. Marginally well behaved subjects include collections where a wider choice of pigments is used. The low order methods will generalize from CMYK training sets to these sets, but higher order methods will probably need to be individually calibrated. Ill behaved subjects, such as LEDs, are by far the most challenging. Unless the camera or scanner meets the Luther-Ives condition, color transforms of ill behaved subjects are hopeless.

10. REFERENCES

2. Wil Plouffe, private communication.
9. C. Chu, Preliminary studies on color space conversion from RGB to L*a*b* using least square fit, Quad Tech Internal Report, May 10, 1996
11. C. Södergård, M. Kuusisto, Y. Xiaohan, and K. Sandström, On-line control of the color print quality guided by the digital page description, IARIGAI’s 22nd International Conference in Munich, Sept. 1993
20. C. Chu and X. Feng, Preliminary studies on color space conversion from RGB to L*a*b* using nonlinear neural networks, Quad Tech Internal Report, April 30, 1996

8—IS&T’s 49th Annual Conference