Robust Processing of Color Target Measurements for Device Characterization

Raja Bala, Gaurav Sharma and Dennis Venable
Xerox Innovation Group, Webster, New York

Abstract

Device characterization typically involves generating one or more targets each comprising a number of color patches and making colorimetric measurements of these patches. In order to derive the characterization function, one has to obtain the correspondence between the set of colorimetric measurements and a set of device signals. In the case of output device characterization, the device signals are the CMY(K) values used to create the target; for input device characterization the device signals are the captured RGB values. Establishing the correspondence requires knowledge of the spatial target layout, the order in which the measurement instrument measures the patches, and the order in which the input or output device captures or renders respectively the target patches. Since the measurement process requires human intervention, errors can conceivably be introduced at this step. A common example is misorientation of the target on the measurement stage. This results in an incorrect correspondence between device and colorimetric signals, thus producing an erroneous characterization.

This paper presents some practical approaches to detect and correct for measurement errors due to misorientation. One approach reserves a number of patches in the target to encode the target layout and orientation. The measurements from this so-called "patch code" are then decoded to retrieve, and if necessary, correct the orientation. Another approach anticipates common misorientations of the target, and uses a pattern matching approach to select the correct orientation. This is done either automatically with a prior device characterization, or with visual validation by the user. Both techniques are successful in correcting for target misorientation with minimal burden to the user.

1. Introduction

A primary goal of any color management system is to achieve consistent color among multiple devices. This requires that the color response of each device is characterized. Device characterization typically involves generating one or more targets each comprising a number of color patches and making colorimetric measurements of these patches using a colorimeter, radiometer or spectrophotometer.[1] In order to derive the characterization function, one has to obtain the correspondence between the set of colorimetric measurements (e.g. CIEXYZ or CIELAB) and a set of device signals. In the case of output device characterization, the device signals are the CMY(K) values used to create the target; for input device characterization the device signals are the captured RGB values. To correctly establish the correspondence between colorimetric and device signals, one must have knowledge of the spatial target layout, the order in which the measurement instrument measures the patches, and the order in which the input or output device captures or renders respectively the target patches. Since the measurement process requires human intervention, errors can conceivably be introduced at this step. A common example is misorientation of the target on the measurement stage. Another example of incorrect correspondence can occur when the order in which the measurement instrument measures the patches does not match the order in which the patches were rendered or captured by the device. Incorrect correspondence between device and colorimetric signals can produce a grossly erroneous characterization.

This paper presents two practical approaches to detect and correct for incorrect correspondence between colorimetric and device signals. One approach reserves a number of patches in the target to encode the target layout and orientation. The measurements from this so-called "patch code" are then decoded to retrieve, and if necessary, correct the orientation. Another approach anticipates common misorientations of the target, and selects the correct orientation either automatically with a prior device characterization, or with visual validation by the user. Although most of the discussion focuses on printer characterization, the techniques invariable extend in a straightforward fashion to input device characterization.

2. Patch Codes

It would be desirable to encode onto the printed target useful information such as orientation, layout, etc. There are many techniques of encoding information onto a printed page. Ideally it is desirable to encode information in a fashion that can be naturally decoded by the scanning device used in the given application. A common example is the barcode usually read by handheld laser scanners. For color characterization applications, the scanning device is usually a colorimeter or spectrophotometer that is designed to measure the color of uniform patches. (For illustration, we will refer in this paper to a spectrophotometer mounted on a stage, such as the Gretag Spectrolino™.) The logical scheme is therefore to use certain dedicated color patches on the target to encode the information. We refer to this set of patches as a patch code.[2]

The design of the patch code requires that the sequence of patches is readily identified and distinguished from other sequences of patches on the target. This must hold true for a range of devices one can expect to encounter in a given application. One would expect that the basic 8 colors formed from the 0% and 100% combinations of C, M, Y are readily distinguishable amongst a wide variety of printers. In the case of CMYK devices, the black colorant can be used in place of the CMY overprint in order to avoid potential problems from excessive colorant area coverage. Figure 1 is an a*-b* plot showing these 8 colors for 6 different output devices. (Note: the figures in the online version of this paper are in color, the figures in the printed version in the proceedings are in monochrome.) These include 4 different models of laser printers, a thermal inkjet, and solid inkjet printer. The figure shows that hue angle alone can be used in most cases to easily distinguish among the 6 colored overprints, while chroma and lightness can be used to detect white and black. Distinguishing these patches from other patches in the target would require a detection scheme that examines all three attributes of lightness, chroma and hue.

In the scheme described above, each patch in the patch code sequence can take on one of 8 colors, thereby representing 3 bits of information. A single row of 20 patch codes can encode up to 60 bits of data; four such rows can encode 240 bits of information. One can also consider using a higher bit depth to encode more information in each patch. For example, including 50% area coverage of each of C, M, Y would allow 27 possible colors, or 4.75 bits per patch. However detection of each patch with respect to other patches in the code, as well as patches in the remainder of the target, becomes more challenging.

There are two ways to improve the patch detection accuracy:

i. incorporate explicit knowledge about a specific type of device being characterized. This will in effect reduce the inter-cluster variance shown in Fig. 1, hence allowing more color combinations to be detected at a given level of accuracy.

ii. Use easily distinguishable patches as "start codes" to encode the device type. An example would be to use pure C, M, Y. Since there are 3! = 6 possible sequences of these 3 colors, 6 device types can be encoded into the start codes. (In the case of a CMYK printer, 4! = 24 device types can be encoded.) For a given device type, the expected measurements of the other colors in the patch code (e.g. R, G, B) can be looked up from a database. Alternatively, or additionally, the measurements of C, M, Y can be used to obtain a refined estimate of the other colors in the patch code hence improving detection accuracy. Thus the decoding scheme is continually improved based on previous successful decoding operations.

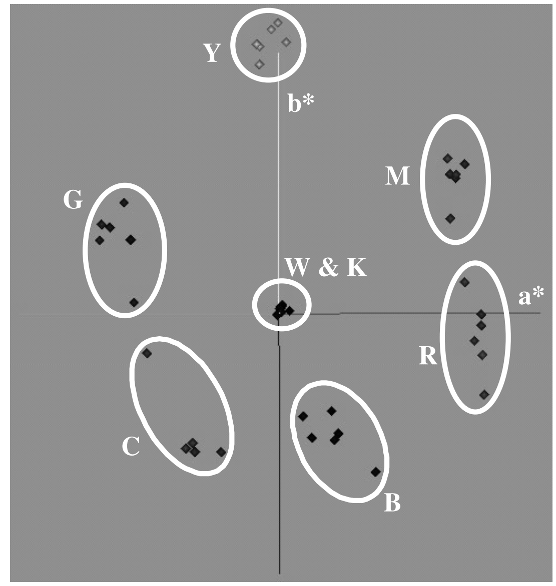

Figure 1. a vs. b* plot of solid C, M, Y, R, G, B, K and white for 6 output devices. Media-relative colorimetry was used in the CIELAB measurements.*

We must define a protocol for the use of patch codes. Figure 2 represents a possible encoding of binary values into patch codes.

Figure 2. Exemplary patch code encodings.

As suggested earlier, the first few patch codes represent start codes, given by pure C, M, Y. If the measurement software cannot resolve these first few codes, the software will acknowledge a possible problem. Scanning for start codes provides several advantages in addition to the aforementioned improvement in decoding accuracy:

- Proper functioning of all the separations of the printer can be quickly checked. Significant errors in printing can be detected simply by checking for the start codes.
- The software can quickly detect the absence of patch code information. If there are no start codes in the expected location, the software can either fail gracefully, or search for patch codes in another corner of the page. This is to account for the possibility that

the user may inadvertently measure patches in an incorrect manner. If for example the Spectrolino is used for measurements, it is possible to orient the page incorrectly on the stage. The measurement software would check several locations that correspond to rotations of the page, or mis-positioning of the page. Figure 3 shows search locations for possible rotational errors of the target printout. If the start codes are located at any one of the given orientations, the measurement software can easily modify the measurement processing algorithms to read the patch codes as well as the remainder of the target in the new orientation.

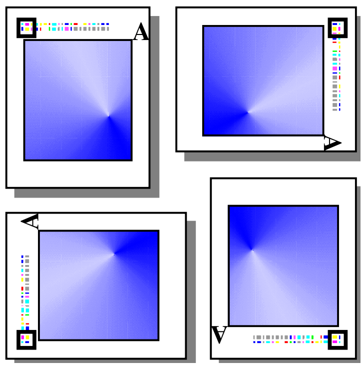

Figure 3. Alternative target positions corresponding to rotations or other mis-orientations. The black rectangle represents the search position.

A sample format for patch codes is described next for a CMY device. With reference to Fig. 4, the patch code can be divided into several fields. The first field comprises three start code patches (i.e. pure C, M, Y). The next field contains two patches indicating the number of rows (in octal) in the patch code. The example shows a Cyan-Yellow pair of codes, indicating a total number of rows as 2. The following two patch codes represent the number of columns, or the width of the patch codes. The example shows a Yellow-Red code indicating an octal 24, or 20 patches across. The next field contains two codes identifying a patch code version number, currently set to 1. The version would be incremented appropriately as addition fields are added to the currently defined patch code. From this point on, patch codes represent actual data where the specific format is dependent on the patch code version number. An exemplary set of fields is shown in Table 1. Unused codes are filled with values of 7, i.e., "mid-gray".

The two-row 20-column patch code shown in Fig. 4 can encode 93 bits of job information, with an additional 27 bits used for start codes, row/height, and version. The optimal number of rows and columns used in the patch code depends on several factors, including:

- the amount of information that the user wishes to encode in the patch code;
- the bit depth of each patch in the code;
- the size of the printable portion of the page;
- the smallest patch size that can be reliably scanned by the measurement device;
- the size of the remainder of the target.

Table 1: An Example of the Fields Encoded in a Patch Code for a CMY Device.

Field	Number of Bits	**Information encoded**
Target ID	30	Unique ID used as a key to access the state of a given calibration.
Date	30	Time stamp
Page Number	6	Applies in the case of multi-page targets
Target Type	21	Can include number of rows, columns in the target. Can also identify special targets, e.g. Kodak Q60, IT8, etc.
Qualifer	6	Other useful information such as halftone, substrate etc.

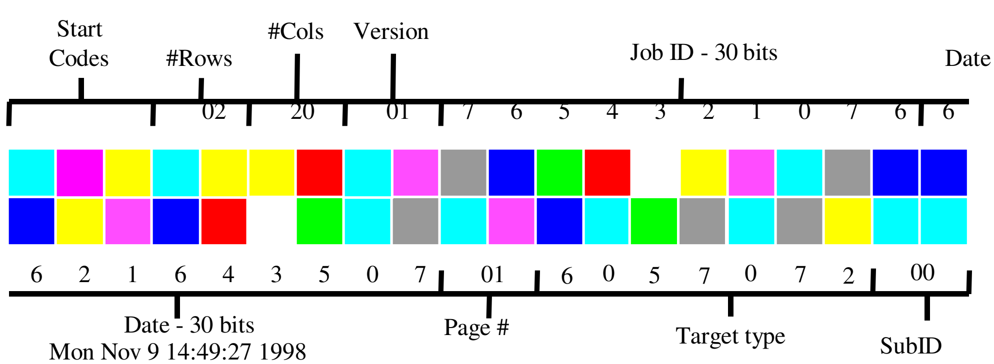

Figure 4. Sample patch code format

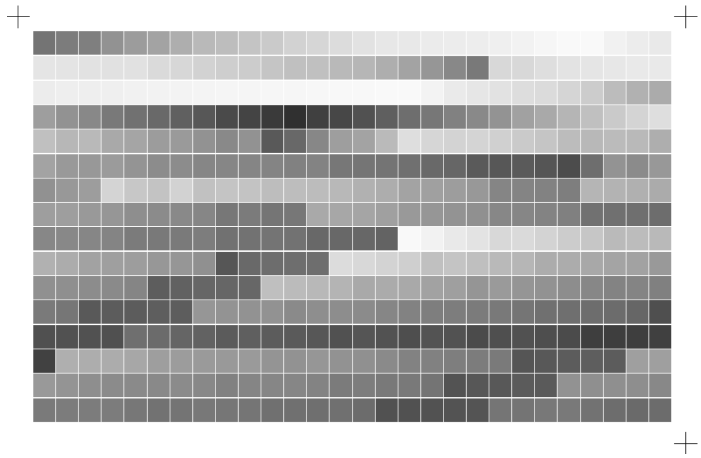

Figure 5. Sample printer characterization target

In order to make maximal use of the printed page, the patches in the patch code should be considered as part of the characterization data, and need not be repeated elsewhere on the target.

This system requires that the patch code interpreter interacts with the driver that operates the measurement device. The particular implementation of this interaction will depend on several factors, such as access to the driver software, the operating system, etc.

3. Pattern Matching

The second approach anticipates all common orientations of the target resulting from different target placements on the measurement stage and scanning order by the measurement device. The measurements are then reordered according to each orientation, and matched with a reference pattern of target colors. The orientation with the closest match to the reference is chosen as the correct one.[3] The pattern matching can be done either visually by the user or fully automatically. Both approaches will be described shortly. In either case, the need for the user to re-measure the target is eliminated.

For illustrating the idea, in the following description it will be assumed that the characterization target is a single page. The method can however be readily generalized to multiple target pages by applying the same technique to each printed page and over pages (possibly by partitioning raw measurements into pages using the known number of patches per page from the target generation step and the reasonable assumption that measurements corresponding to a page are contiguous). A sample printer characterization target is shown in Fig 5, and will be referred to in the ensuing discussion.

The target patches are oriented along a rectangular grid (this is usually the case, though not a requirement) over the page. Typical automated color/spectral measurement stages provide methods for readily measuring such targets by placing the target on the stage and then specifying, by means of a control file, a measurement sequence for the patches as they appear on the stage. Examples of measurement sequences are: left to right across rows, moving among the rows from top to bottom; or top to bottom across the columns moving across columns from right to left. The target orientation on the measurement stage could potentially be varied in several different ways by flipping or rotating the target and the measurement stage control file could vary the order in which the patches are measured. Due to these various degrees of freedom in the target placement and in the control files, there are several possible correspondences between the sequence of target measurements and the physical placement of the target patches on the page. In existing systems, usually only one of these correspondences is allowed as a correct one; and others result in incorrect characterization or request a re-measurement by the user. In order to remedy this problem, we propose trying out the different potential correspondences and determining the one that is correct. For a rectangular target layout like the one shown in Fig 5, it is highly likely that the target region of the page will be measured in one of the following 8 orders (specified, for instance, when the target is oriented with the text upright): 1) left-to-right across rows, moving among the rows from top to bottom; 2) left-to-right across rows, moving among the rows from bottom to top; 3) right-to-left across rows, moving among the rows from top to bottom; 4) right-to-left across rows, moving among the rows from bottom to top; 5) top to bottom across the columns moving across columns from right to left; 6) bottom to top across the columns moving across columns from right to left; 7) top to bottom across the columns moving across columns from left to right; and 8) bottom to top across the columns moving across columns right to left. These measurement orders are shown graphically in Fig. 6. Note additional orders, such as for example a serpentine scan could be considered similarly.

3.1 Visual Pattern Matching

In this approach, the characterization software re-generates 8 different rectangular-grids of patches having the same number of rows and columns as the characterization target (this is known since the target itself was generated by the characterization software). The colors of the patches in each of the grids are set by arranging the sequence of data from the measurement device in each of the aforementioned 8 measurement orders. The user is then presented with these different rectangular-grids on a display and asked to pick the one whose orientation visually matches the orientation of the physical printed target held in its correct orientation. Usually the proper orientation of the physical target is easily identified in the presence of text labels on the target (i.e. the text must be upright). Alternate methods may be used for specifying the orientation of the physical target if no text is present – for example, by indicating that certain colors should appear at certain locations. Note that since the measurement device will usually report spectral reflectance or colorimetric data such as CIEXYZ or CIELAB, this must be converted to an additive RGB space such as sRGB[4] for approximately correct rendition on a nominal display device. An example of such a display is shown in Figure 7, where the 8 possible orientations of the target are shown using a set of measurements corresponding to the target in Figure 5. Note that in this case, the orientation 4 matches the orientation of the physical target of Figure 5, when it is held with the text upright. Once the user enters the orientation that matches the printed target, the correspondence between the physical target patches and the measurement data is unambiguously established. Since the physical target layout is generated by the characterization software, the latter can use this correspondence to obtain the correspondence between device control values and measurement values for each of the target patches and the characterization can then be performed using the measured data. For testing a large number of potential measurement orders (for instance with serpentine scans of the patches), multiple display screens of orientations may be used until the user indicates a match. Organizing the most common "mis-orientations" in the earlier screens would make this more efficient. In addition, as a first step, a single rectangular-grid layout corresponding to a default assumed measurement order could be presented to the user asking for visual validation with respect to the physical target. Subsequent presentations of different orderings can then occur only if the user indicates that this order is incorrect.

3.2 Automatic Pattern Matching

The process can be further automated by using *a priori* knowledge of the printer characteristics. For instance, if a profile or characterization function for the printer or a similar class of printers is available (e.g., from a previous characterization) the device control values used in generating the target can be used to predict the color values of the target patches. The measured target color values could then be arranged in different orders, as in Sec. 3.1, and those values compared with the predicted values obtained from the CMYK values used to generate the target, the printer characterization function, and the correct order of anticipated measurements. Color differences (e.g. ΔE_{ab}^* or ΔE_{94}^*) between corresponding color values in these predictions and the actual measurements can then be computed for each of the assumed measurement orders and these can be used to determine the most likely measurement order. For instance, the most likely measurement order may be determined as the case which yields the smallest average color difference.

For the example of Figs. 5-7, the results of automated pattern matching are illustrated in Table 2. The characterization software anticipates target measurements according to the "Order 1" shown in Fig. 6. The target was however measured in the sequence corresponding to "Order 4" (as already established by the visual matching results). In order to perform the automated matching, the CMYK values used to generate the characterization target were arranged in a sequence corresponding to "Order 1". These were then used along with a CMYK to CIELAB forward response model for a CMYK printer to predict the CIELAB values expected from the target measurement (in correct order). The measured colorimetry for the target was then reordered in 8 different ways corresponding to the hypotheses of the 8 measurement orders shown in Fig. 6. The mean color difference between the predicted sequence of target measurements and the re-ordered sequence of target measurements was computed for each of these 8 re-orderings. These mean color differences are listed in Table 2. From the values, it is apparent that the correct hypothesis for the measurement order, i.e. Order 4, has a much lower (≈11) average color error than all the other hypothesized orderings (>47). Thus the actual measurement order and the corresponding re-ordering needed for correction, can be selected as the one corresponding to the smallest average color error in the prediction. The large difference between the correct and incorrect hypotheses in Table 2 indicates that this process has little margin for error. Note also that the forward response model used in this example was from a printer in an entirely different family. An even sharper minimum can be achieved if a response model for the same printer/printer family is available.

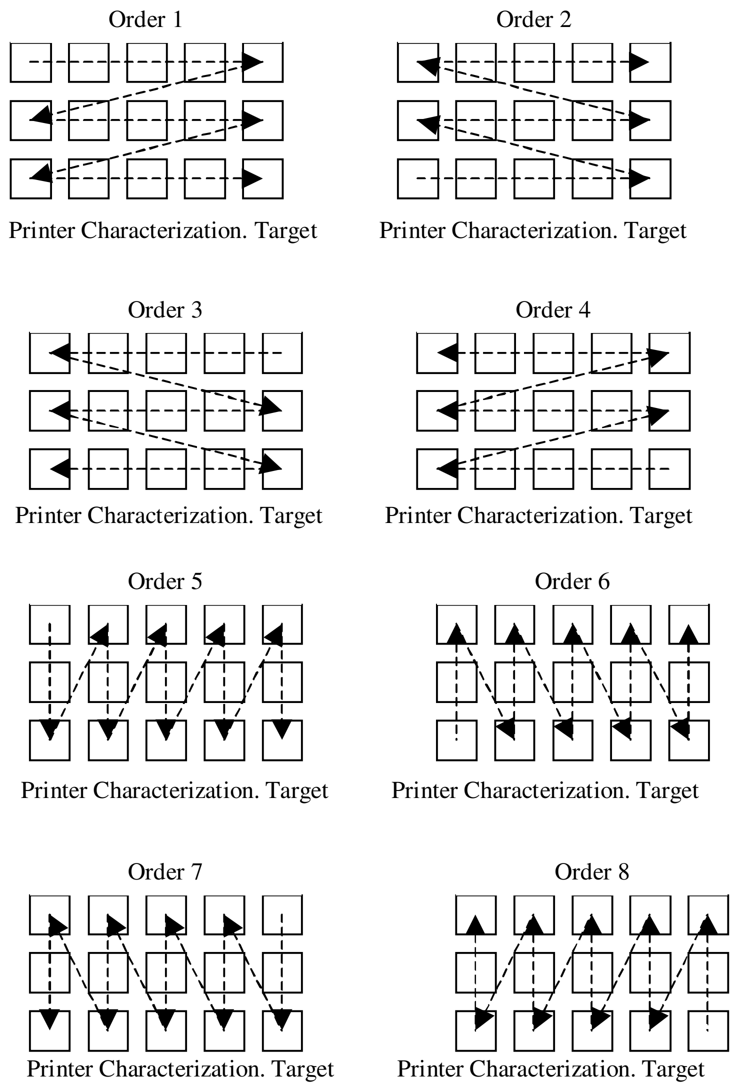

Figure 6. Common possible measurement orders for a rectangular layout target

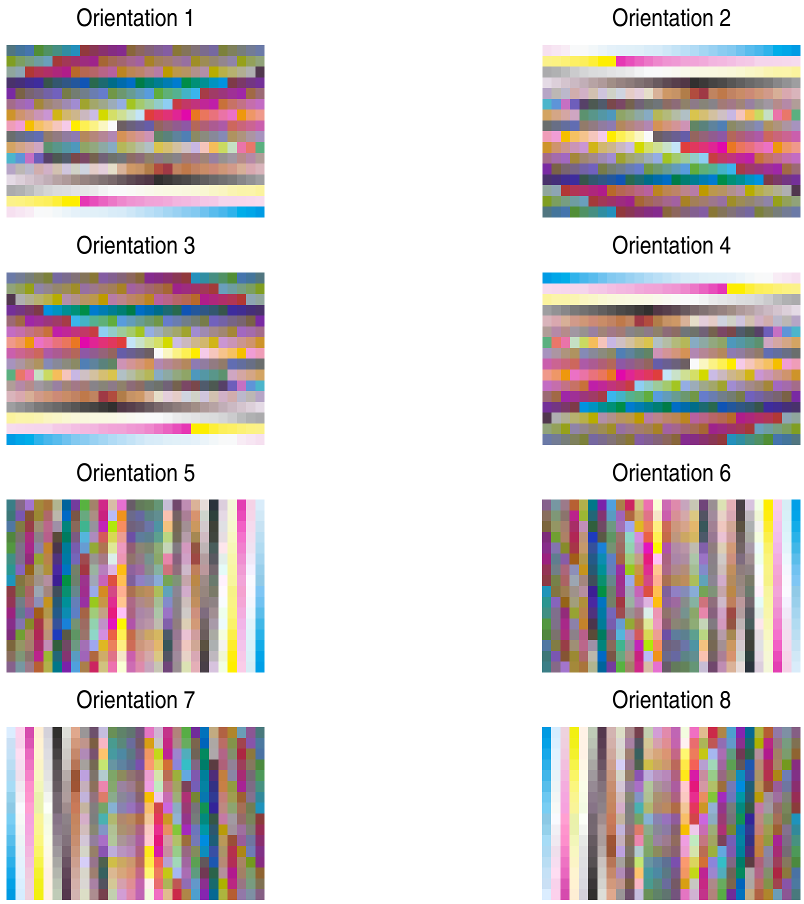

Figure 7. Example of layout validation screen presented to a user.

Table 2. Average color differences (ΔE^*_{ab} units) between the colorimetry for the target patches in measured sequence order and the predicted colorimetry for the same patches in the anticipated sequence for correct measurement order.

Order No.	1	2	3	4	5	6	7	8
Avg. color difference wrt prediction (ΔE^*_{ab})	58.0	47.1	59.1	10.9	57.5	55.6	56.7	56.1

Note that the automated system can readily search through a much larger number of "mis-orientations" than can be reasonably presented to a user for visual validation. Note also that the visual and automatic methods can be combined with the average color difference from the automatic approach being used to winnow down the number of orderings to be presented for visual matching to the physical target.

Conclusions

Two approaches have been presented for increasing the robustness of measurement processing in device characterization applications. The first technique utilizes patch codes to encode information about the target, including its layout and orientation. A preliminary analysis has shown that patches that take on one of 8 colors (white, C, M, Y, R, G, B, black), can encode 3 bits of information with reliable decoding accuracy. We demonstrate with a simple example that even a relatively small sequence of such patches can convey a lot of useful information within the target. The second technique employs re-ordering and pattern matching to correct for common mis-orientations. Both techniques avoid erroneous characterizations as well as the need to re-measure the target.

As a final note, the methods described here are particularly useful in a distributed (e.g. web-based) color management system where there is little control over the user's measurement environment.

5. References

1. R. Bala, "Device Characterization", Chapter 5 in *Digital Color Imaging Handbook,* G. Sharma, Ed., CRC Press, 2003.
2. D. Venable, R. Bala, "Patch codes for color calibration job identification encoding", U. S. Patent filed, November 2000.
3. G. Sharma, R. Bala, "Visual layout validation for color target measurements", U. S. Patent filed, November 2000.
4. See www.srgb.com for details of the sRGB standard.

Biography

Raja Bala received his Ph.D. degree from Purdue University in 1992 in Electrical Engineering. Since then he has been employed at Xerox Corporation, where he is currently a Principal Scientist in the Solutions and Services Technology Center. His research interests include color science, color management and color image processing. Raja holds over 20 patents and over 35 publications in the field of color imaging. He is a member of IS&T.