Proc. IS&T's Archiving Conference, IS&T, 96-100, June 2010

Capturing the Color of Black and White

Don Williams, Image Science Associates and Peter D. Burns*, Carestream Health Inc.

Abstract

One of the first and iconic imaging targets to be developed for color photography was described by McCamy, Marcus, and Davidson. That was more than thirty years ago and, for better or for worse, the Macbeth ColorChecker™ continues to be a de-facto target standard in imaging circles. The target was designed to simulate the spectral nature of critical or problematic colors for consumer photography so that color calibrations and color reproduction evaluations could be made that were consistent with scene content. However, for digital collections with limited color-gamut, or where detection of small differences in material properties is important, adopting collection-specific color test targets is often advisable. This can reduce metamerism, and sample the device signal space in the most important regions. As an example of such content, the spectral-reflectance characteristics of a collection of early photographic prints were measured. The underlying structure of the data set was investigated in terms of principal components, and a spectral reconstruction based on two principal components was demonstrated.

Introduction

Color test objects such as the ColorChecker™ [1] are used in two principal ways to; (1) establish accurate color image capture (*calibration function*), and (2) measure deviation from the desired capture of color image information (*evaluation function*). An example of the first approach is calibration of photofinishing print scanners using test targets printed on the photographic or inked paper that is similar to that used for the consumer prints being scanned. The scanner calibration targets have similar colorants (dyes or pigments) to those of the print, reducing metamerism. The design of color test targets can also favor important colors during calibration. This is done with the ColorChecker SG target. To weight color-profiling in favor of skin tones, fourteen flesh-tone patches are provided in the color target, compared with two in the original ColorChecker.

In this paper we focus on the second use of color test targets, as part of image quality assurance operations, rather than color calibration. The control of the imaging performance during digital conversion requires that we observe the important characteristics of our image capture operations. A key part of such initiatives is the selection of test methods and analysis. Although we would generally suggest using established standard methods where applicable, there can be benefits from the adaptation of the testing procedures to the needs of particular collection materials.

In a sense our approach to testing is consistent with color calibration. If the test objects are chosen so they share key characteristics with the collection content, the quality assurance program will focus on measuring what is most important. A world map can locate the geographical position of a major city, but if a traveler plans to take full advantage of what any specific city offers, a more detailed street map is required. The same applies to color imaging of near-neutral image content. In order to create that detailed color map, the color 'demographics' of near-neutral objects need to be characterized.

Current digitizing guidelines for testing color image fidelity tend to 'teach to the test' by using the same imaging target for both calibration and color fidelity evaluation. This approach can yield unrealistic results, and may not measure important color regions. We suggest adopting collection-specific color test targets for digital collections, where detection of small differences in material properties is important. Adopting collection-specific color test targets is useful. In addition, it is advisable to perform imaging quality analysis using (software) methods that are independent of color-management software that is used to produce the color-image content.

In this study we investigated collection materials that populate an important but limited region in color space, 19th and early 20th century black and white photographic prints. We start by describing the color characteristics of the commonly used Macbeth ColorChecker test target. Corresponding measured results for the early print collection will then be described. This is followed by a discussion of content-specific test objects and their use in imaging performance quality assurance programs.

Imaging Performance Measurement

Recent developments of standard imaging performance evaluation methods have received broad exposure in the Archiving and Museum communities. [2, 3] Several authors have described the use of established methods that have been adapted to equipment evaluation and image quality-assurance programs for digital conversion projects. [4-6]

A key element of such efforts is the use of well-specified test targets that are tied to analysis software. Object-level and device-level targets [7] are currently being used with three-color and monochrome image acquisition of reflective and transmissive (film) media.

An example of an object-level target is shown in Fig. 1, captured alongside the print. This particular target provides information which not only relates the captured image to object material and colorants, but also image detail, distortion and sampling. The color patches used are those of the Macbeth ColorChecker. When captured as part of routine digitization, data gathered from this target can be used to establish limits for acceptable, in-control, variation and identify special-cause deviation. An example data set is shown in Fig. 2, derived from four neutral patches from the target.

Most of the discussion of test targets has been in the context of three-color image acquisition. However image quality evaluation, quality assurance and the connection between image and object is equally important to multispectral image acquisition projects. Such systems usually operate with integrated illumination and acquisition functions. [8] Tight system control is often

* Peter Burns is now with Burns Digital Imaging. Contact him at pdburns@ieee.org

demonstrated during system development and initial case studies. However, as such systems reach broader use it is anticipated that standard efficient imaging performance methods will be developed and applied, as they are in the rest of the cultural heritage imaging community.

Figure 1: *Example of an object-level target with wide range of color patches from Ref. 4. (Courtesy of Cornell University, USA)*

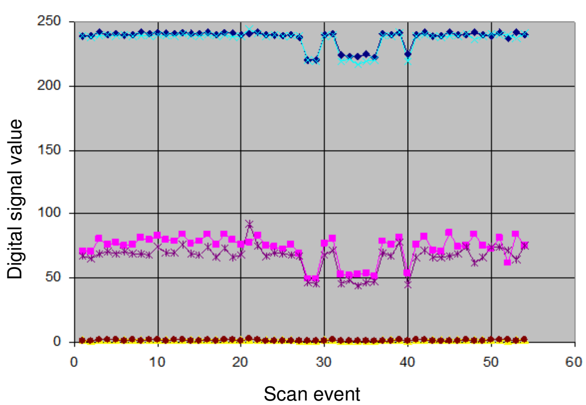

Figure 2: *Example control-chart data derived from the test target of Fig. 1, plotting results from four grey patches*

An experimental object-level target was included as part of a recently reported project to image the Dead Sea Scrolls. [9] The imaging included visible color and infra-red image capture. The experimental test object (see Fig. 2 of Ref. 10) included color patch and ruler elements, similar to those of the commercially available target of Fig. 1, but without the edge feature and registration marks used by automated analysis software.

Early Print Collection

Adopting the same approach as McCamy *et al*. did in 1976, we have characterized the spectral content of samples from a popular class of image types often found in archive environments: the 19th and early 20th century black and white photographic prints. Though not spectrally neutral, this class of objects tends to be near-neutral with predominant brown or yellow hues. By characterizing the near neutrals of the cited content we investigated the selection and development of content-specific imaging targets for this and other classes of collections frequently found in archives.

It is worthwhile noting that there is an increasing awareness of color-gamut moderation in digitizing collections for cultural heritage applications. While adopting a wide color gamut appears *de rigueur*, this may actually serve to diminish color fidelity by increasing the encoded quantization interval between neighboring colors. Larger quantization intervals can reduce color and tone discrimination, which makes it difficult to render subtle near neutral tones in slowly varying image areas.

Sample and Data Collection

A combination of thirty samples of both print out paper (POP) and developed out paper (DOP) prints, largely from the late 19th and early 20th century vintage were provided by the Institute for Image Permanence (IPI). These consisted of prints from albumen, platinum, silver gelatin, collodion and carbon print processes with a variety of toning effects. No cyanotypes were included.

Approximately thirty spectral measurements were made over the entire area of each print. An attempt was made to sample highlights, mid-tones, and shadows equally. After a transparent photocopy was made of each print, areas of interest were identified and circular punches in these areas were made on the transparency. This transparency mask was than overlaid on the accompanying print and spectral measurements made through the punched holes This punched transparency acted as a convenient mask and spatial record of where the spectral sampling was done.

The measurements were performed using an X-Rite Model 530 Spectral densitometer between 400 and 700 nm. at 10 nm. intervals. Example locations for the spectral measurements, superimposed on the example print, are shown in Fig. 3.

Figure 3: *Example print with measurement locations indicated*

An examination of measured spectral reflectance data, plotted vs. wavelength reveals similar smooth curves that we would expect from near-neutral samples. Figure 4 shows results drawn from a set of 300 measurements. We observe an almost flat curve for the low (dark) samples, but note the rising values with increasing wavelength for several of the lighter samples. These characteristics, indicating an increasing chroma with lightness can be seen when we transform the spectral reflectance into an approximately visual color space, CIELAB. Figure 5 shows the results for a single print, where the deviation from the neutral axis where $a^* = b^* = 0$ is observed. Projecting CIELAB data onto the $a^* - b^*$ plane removes the lightness component. Figure 6 compares results of 570 near-neutral print coordinates with those of the ColorChecker color chart. [10]

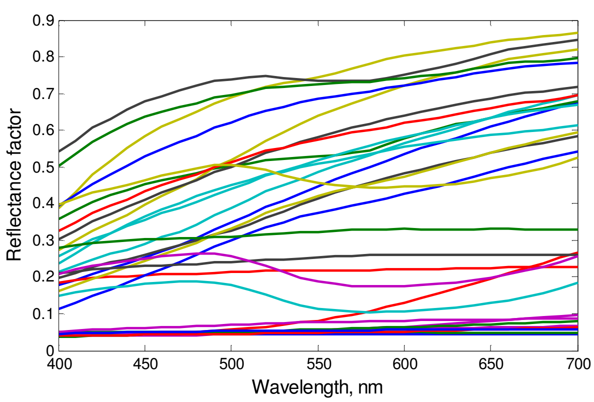

Figure 4: *Example spectral reflectance factor data from the print collection*

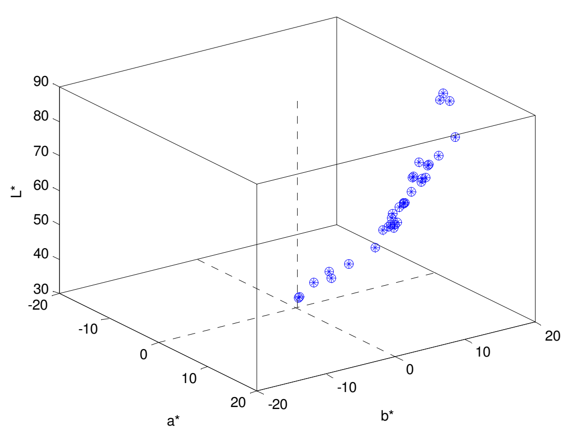

Figure 5: *CIELAB coordinates for measurements from a single print*

Principal Component Analysis

Principle component analysis (PCA) is a statistical technique that seeks to find underlying patterns in sets of observed or measured data. Once we have found these patterns, they can often be used to describe the data in terms of a small number of underlying functions, or relationships. PCA is often used in color imaging to capture spectral reflectance images with far fewer camera signals than would be needed using spectrophotometer. This multispectral imaging is often based on 5-8 color-image records. [11, 12] Here we apply PCA to the measured spectral reflectance data for the print collection.

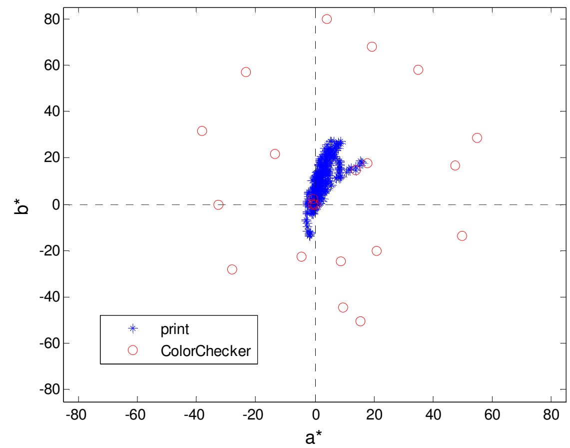

Figure 6: *CIELAB color coordinates (a*, b*) for the Macbeth ColorChecker test target and 570 print collection measurements, for illuminant D50*

To describe how this is done, we use common matrix-vector notation, where the super-scripts, T and -1, indicate the matrix transpose and inverse, respectively. The spectral reconstruction for a given sample is computed by

$$\mathbf{f} = \Phi\alpha + \mu_{\mathbf{f}} \,, \tag{1}$$

where,

$\Phi = [e_1, e_2, ..., e_N]$ is a matrix of principal component vectors, or eigenvectors,

$\mu_{\mathbf{f}}$, the mean spectral vector

$\alpha^T = [a_1, a_2, ..., a_N]$, the set of scalar weights (signals) associated with the sample. In other words, a small number of 'signal' values can be used to compute (reconstruct) the spectral reflectance curve. These signals, which characterize the sample being measured, can be found by

$$\alpha = \left(\Phi^{\mathrm{T}}\Phi\right)^{-1} - \Phi^{\mathrm{T}}\left(\mathbf{f} - \mu_{\mathbf{f}}\right). \tag{2}$$

The term $\left(\Phi^{\mathrm{T}}\Phi\right)^{-1} - \Phi^{\mathrm{T}}$ can be interpreted as a matrix of spectral sensitivity functions that could be used to analyze a sample, **f**, for subsequent spectral reconstruction. In reality these spectral sensitivity functions may not correspond directly to actual camera sensitivities, but often they can be approximated. [11] The spectral reconstruction is then given by,

$$\hat{\mathbf{f}} = \Phi\alpha + \mu_{\mathbf{f}}\Phi \,, \tag{3}$$

where the superscript ^ indicates that the reconstructed spectral reflectance is an approximation of the measured function based on a fewer principal components than the analysis step of Eq.(1).

PCA was performed for a set of 570 spectral reflectance factor measurements of the print collection. The mean vector, or average profile, and first four principal components are shown Fig. 7. Each principal component function does not necessarily correspond to particular colorant material in the prints. However, the relative importance of each component can be determined via the PCA. The magnitude of the eigenvalues associated with each principal component indicates the amount of variation in the data explained by each one.

Figure 8 indicates that most of the variation is explained by 2-3 principal components. The practical use of this is shown in Fig. 9, where results of spectral reconstruction based on two components is compared with the direct measurement of 31 values (10 nm. interval).

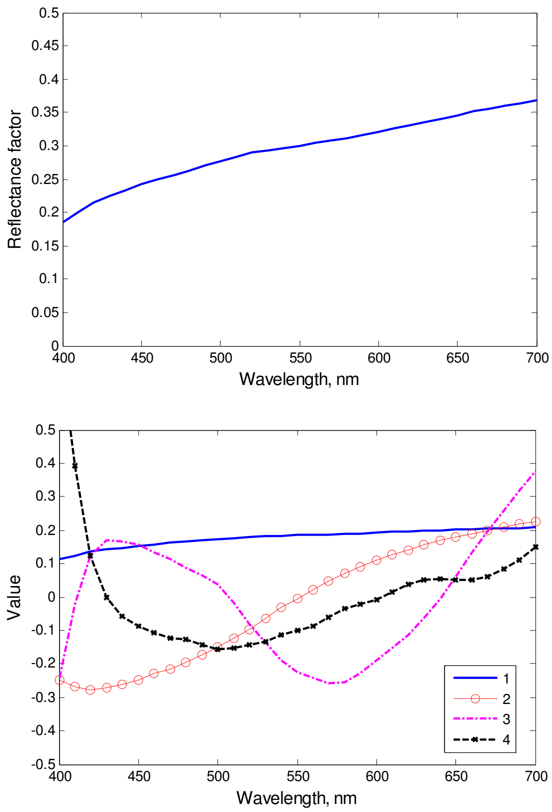

Figure 7: *Mean vector (above) and first four principal components for print spectral reflectance data for the print collection*

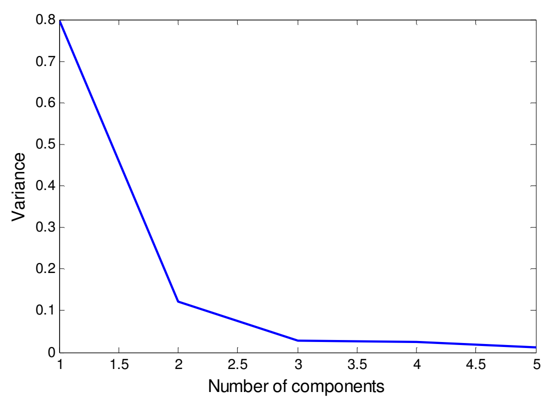

Figure 8: *Normalized variance [0, 1] explained with each of the principal components for the analysis of the print collection*

Adopting collection-specific image quality test objects will usually involve either adding or substituting several color elements (pastels, near-neutrals) in an existing test target. This is usually preferred since object-level or device-level targets are normally designed to evaluate more than color reproduction. Other testing would be unchanged.

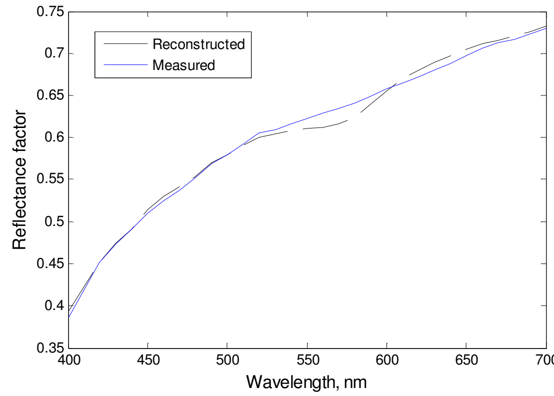

Figure 9: *Spectral reconstruction based on two principal components*

Collection-specific Test Targets

In the above discussion we have focused on color test targets and characteristics of the near-neutral print collection. It is natural to consider adopting collection-specific color test targets for digital collections, where detection of small differences in material properties is important. When considering the number and type of color test objects it is important to consider using test objects that have traceable characteristics and are likely to be in stable supply. One approach is to search a database of manufactured color samples, such as Munsell color samples for which the spectral reflectance data are available. [13]

Flexible Analysis and Tracking Software

Test target selection, is only part of what is needed for imaging performance evaluation and quality-assurance efforts. The companion component involves the selection of software for the analysis of test target elements of captured images, and comparison of results with required performance levels. Ideally the analysis software allows performance criteria (including data and location of custom elements) to be adjusted as needed. In this way results from collection-specific color elements can be reported with the other established performance measures that apply, e.g.,

- spatial frequency response measure of image sharpness
- color registration
- illumination uniformity
- optical distortion
- sampling distortion

without disruption to normal established performance evaluation procedures. The most advanced analysis software products now include functions that help the user interpret results from a collection, and set project-specific control limits.

Conclusions

The elements of a successful imaging performance program should include; establishing of performance goals, efficient test plans, and performance tracking software tools. The first two of these involve the selection and adoption of test objects. In many cases the use of common color elements such as those of the ColorChecker is successful. However, for digital collections with limited color-gamut, or where detection of small differences in material properties is important, adopting collection-specific color test targets is advisable. This can be used to reduce metamerism, and sample the device signal space in the most important regions. As an example of such content, the spectral-reflectance characteristics of a collection of early photographic prints were measured. The underlying structure of the data set was expressed in terms of principal components, and spectral reconstruction based on two principal components was demonstrated. This type of analysis can be used to select color test stimuli. When used with flexible analysis software based on the framework of emerging international standards, these test elements can provide effective imaging performance evaluation and tracking.

Acknowledgements

It is a pleasure to acknowledge the contributions of Natalie Russo who guided and performed the spectral measurements of the print collection in a more organized way than we might have. In addition, thanks to the staff of the Image Permanence Institute at Rochester Institute of Technology for help in the selection and identification of the early print collection.

References

[1] C. S. McCamy, H.Marcus, and J. G. Davidson, A Color Rendition Chart, Jour. *Applied Photo. Engineering*, Vol. 2, 95-99, 1976
[2] Federal (US) Agencies Digital Guidelines Initiative, http://www.digitizationguidelines.gov/
[3] H. van Dormolen, Metamorfoze Preservation Imaging Guidelines, *Proc. Archiving Conf.*, IS&T, 162-165, 2008
[4] M. Stelmach and D. Williams, When Good Scanning Goes Bad: A Case for Enabling Statistical Process Control in Image Digitizing Workflows, *Proc. Archiving Conf.*, IS&T, 237-243, 2006
[5] P. D. Burns and D. Williams, Ten Tips for Maintaining Digital Image Quality, Proc. *Archiving 2007 Conf.*, IS&T, 16-22, 2007
[6] F. B. Wheeler, Exploring the Use of Targets to Study Scanning Exposure Variability, *Proc. Archiving 2008 Conf.*, IS&T, 171-176, 2008
[7] D. Williams and P. D. Burns, Preparing for the Image Literate Decade, *Proc. IS&T Archiving Conf.*, pg. 124-127, IS&T, 2009
[8] D. Emery, F. G. France and M. B. Toth, Management of Spectral Imaging Archives for Scientific Preservation Studies, *Proc. Archiving 2009 Conf.*, IS&T, 137-142, 2009
[9] S. Tanner and G. Bearman, Digitizing the Dead Sea Scrolls, *Proc. Archiving Conf.*, IS&T, 119-123, 2008
[10] D. Pascale, RGB Coordinates of the MacBeth ColorChecker, http://www.babelcolor.com
[11] P. D. Burns and R. S. Berns, Analysis of Multispectral Image Capture, *Proc. Fourth Color Imaging Conf.*, IS&T/SID, 19-22, 1996
[12] P. D. Burns and R. S. Berns, Image Noise and Colorimetric Precision in Multispectral Image Capture, *Proc. Sixth Color Imaging Conf*, IS&T/SID, 83-85. 1998
[13] Spectral Database, University of Joensuu Color Group, http://spectral.joensuu.fi/

Author Biography

Don Williams is founder of Image Science Associates, a digital imaging consulting and software group. Their work focuses on quantitative performance metrics for digital capture imaging devices, and imaging fidelity issues for the cultural heritage community. He has taught short courses for many years, contributes to several imaging standards activities, and is a member of the Advisory Board for the interagency US Federal Agencies Digitization Guidelines Initiative, FADGI.

Peter Burns is with Carestream Health working in image evaluation, modelling, and image processing for medical imaging systems. Previously he worked for Eastman Kodak and Xerox Corp. He has taught imaging courses for many years, as an adjunct faculty member at Rochester Institute of Technology, at Kodak, and at previous Archiving conferences.