https://doi.org/10.2352/EI.2025.37.9.IQSP-257

Color and Spectral Matching in Imaging Performance Evaluation

Peter D. Burns[1] and Don Williams[2]; [1]Burns Digital Imaging, [2]Image Science Associates

Abstract

Evaluation of camera or scanner image capture usually includes the selection of reference color test objects. For image capture, since the optical characteristics are known, the intended camera color-encoded image data are compared with the corresponding idea values. However, for critical scene content, such as human skin, the color patches may not sample the encoded color space efficiently. In addition, the colorimetric sampling by the target, its spectral-reflectance characteristics may differ from those of important scene or object elements. We address the selection of color test patches for both colorimetric and spectral criteria. Results are shown for the imaging of skin tones and include an analysis of likely variability.

Introduction

Most programs for imaging performance evaluation include a selection of reference test objects. Since the optical characteristics are known, for image capture, the intended camera color-encoded image data are compared with the corresponding idea values. This is usually done for a set of uniform color patches on a standard test chart.

However, for critical scene content, the color patches may not sample the encoded color space efficiently. This can be due to the original selection having been made for a different imaging sector. For example, the ColorChecker® patches [1] were selected for consumer and professional (film-based) photography. The 24 colors sampled largely outdoor objects, with only two skin colors. The ColorChecker® Digital SG replaces the original, has 140 colors, and is better suited for creating ICC color profiles, used by color management software.

In addition to this colorimetric sampling by the test target, its spectral-reflectance characteristics may not reflect those of important scene or object elements. In this situation, measurements of color (encoding) accuracy can underestimate the actual errors, and variability in, e.g., unconstrained and mixed lighting conditions.

We address the selection of color test patches for image capture evaluation and benchmarking. Here, the color encoding will be input-referred, where color information describes the scene rather than the intended viewed image. Both colorimetric and spectral criteria are considered. After discussing previous results for cultural heritage imaging, we turn to the capture of human skin tones.

In cases where the evaluation of printing systems is the objective, the intended or ideal color characteristics are also selected. These are then printed, and measured, e.g., for image permanence. However, in image capture evaluation the (reference) color test patches are often chosen as a commercially available standard set.

Imaging Important Scene Content

Color imaging performance evaluation for cultural heritage (museums, libraries, and archives) has often been addressed by selecting test objects whose spectral reflectance characteristics resemble the ensemble of collection materials [2]. For example, we previously reported on studies of black-and-white photographic prints, [3] and pastels [4]. In both cases, the spectral reflectance factor characteristics were directly measured. These results were then used to guide color patch selection from a commercially available set of candidate materials. Figure 1, from [4] outlines the selection method.

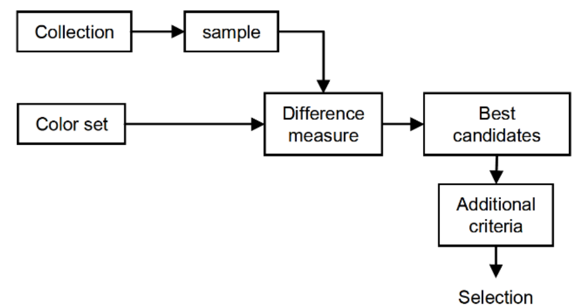


Figure 1: Selection from a set of available color samples based on object characteristics [4]

Following this approach, Fig. 2 shows the results based on a minimal absolute spectral error. The selection was made from the set of Munsell colors [5]. The colored circles are rendered as corresponding to the *sRGB* values. In this case, the aim spectrum was based on a measurement. An alternative approach would be to generate potential members of the target population based on a principal component analysis of the population [3].

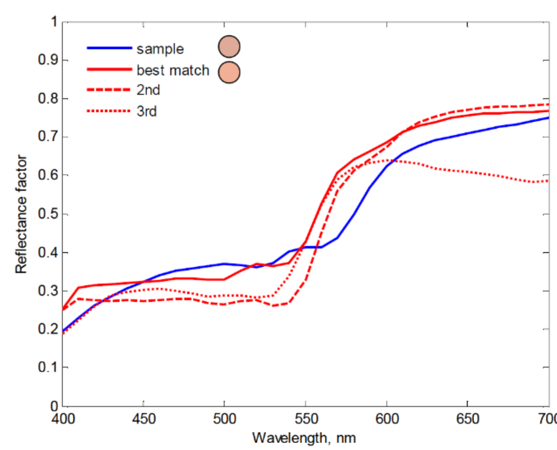
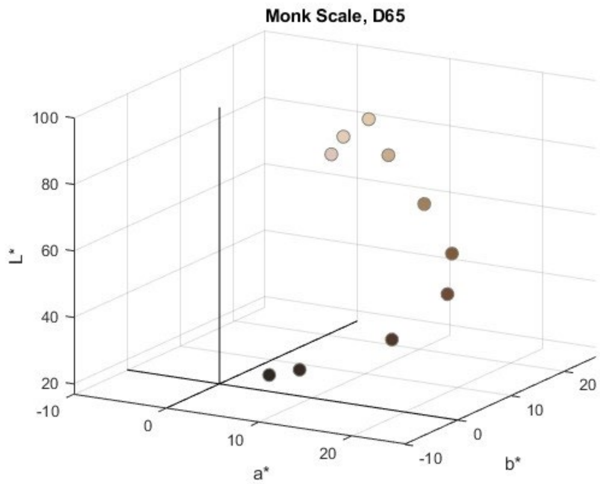


Figure 2: Results of spectral matching to a set of Munsell color samples [4]

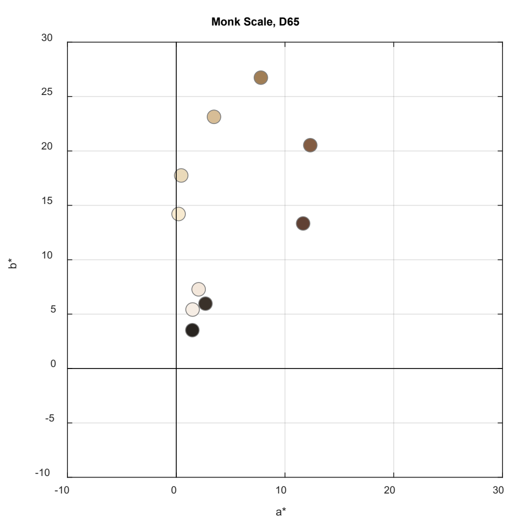
Skin Tone Test Target

For system evaluation, several commercially available color test targets are commonly used. For example, the extensive ColorChecker® family of targets for photography and filmmaking [5], DT Next Generation [6], and GoldenThread FADGI 19264 targets [7] for cultural heritage imaging.

Currently, there are several efforts aimed at improving the capture and display of the range of human skin. For example, for video and consumer cameras [8]. This study used the Monk skin tone scale [9]. The scale has ten steps, whose characteristics are defined by colorimetric (*sRGB* and *CIELAB*) values. [9] Figure 3 shows the Monk scale plotted as CIELAB coordinates, for a D65 illuminant. For display here, the circles are filled as *sRGB* values.



(a)



(b)

Figure 3: (a). CIELAB values for Monk Scale under D65 illuminant. (b) the a-b* plane*

The Monk scale is presented in two ways; as uniform patches and as 'orbs'. Each orb is a non-uniform circular (*sRGB*) plane, as shown in Fig. 4. The intent is to indicate the differences in skin tone that can occur, e.g., in common scene illumination, and ordinary images. It is possible to produce a camera target based on a printed version of the Monk scale. Note that the range of *sRGB* values across each orb step may infer a corresponding perceptual range for skin tones, from which target colorimetric tolerances could be established.

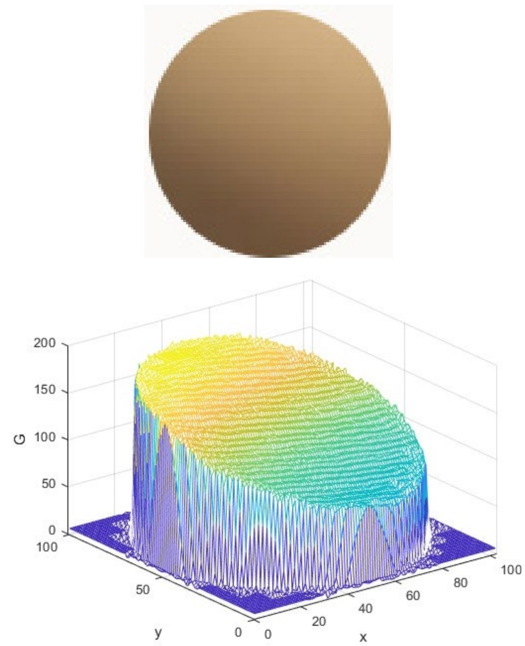


Figure 4: A single orb step of the Monk Scale (top), and corresponding green pixel values

Spectral Skin Measurements

A set of 100 spectral measurements of human skin is available from the US National Institute of Standards and Technology, described in Cooksey, *et al*. [10, 11]. These are shown in Fig. 5. The spectral reflectance values range from 250 – 2500 nm, in 3 nm intervals. This wide wavelength range was based on the intended span of potential applications; medical, therapeutic, etc., well beyond our area of focus. Here we display the visible wavelength range.

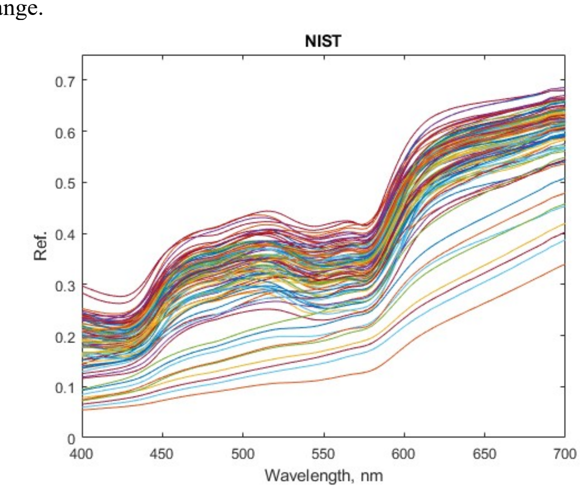


Figure 5: NIST 100 reflectance spectra and for visible wavelengths

These data are traceable, and the spectral values could be used to investigate potential test colors under a variety of illumination conditions. The original 51-subject set was expanded to include 100. Subjects were not selected to span the full range of skin color, so the range was lower than that of the Monk or other scales, such as that developed by L'Oréal Company [12].

Pantone SkinTone Guide

As an example of spectral selection from a set of commercially available colors, we chose the *Pantone SkinTone Guide.* [13] This collection is based on measurements of human skin and contains 138 color samples. Figure 6 shows the color patches of the guide (uncorrected).

Figure 6: Pantone SkinTone Guide color patches, a 'SkinTone Genome'

Each of the color samples of a physical guide was measured, and the spectral reflectances for half of them are plotted in Fig. 7. The measurements were made using an *Xrite iPro2* spectrophotometer, and *BabelColor PatchTool* software. The CIELAB values for the guide are shown in Fig. 7. The corresponding CIELAB values are shown in Fig. 8

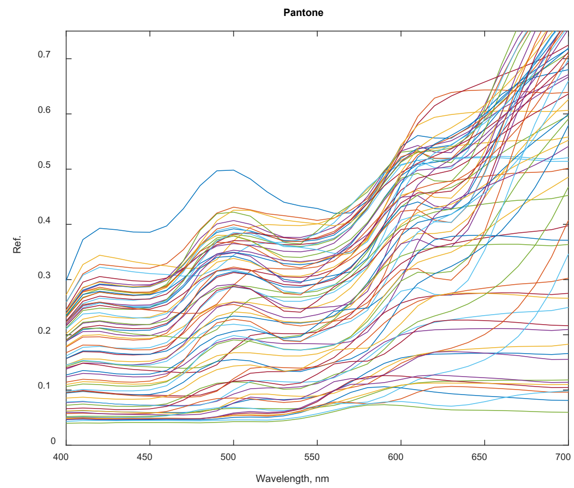


Figure 7: A sampling of the Pantone SkinTone Guide reflectance spectra for visible wavelengths

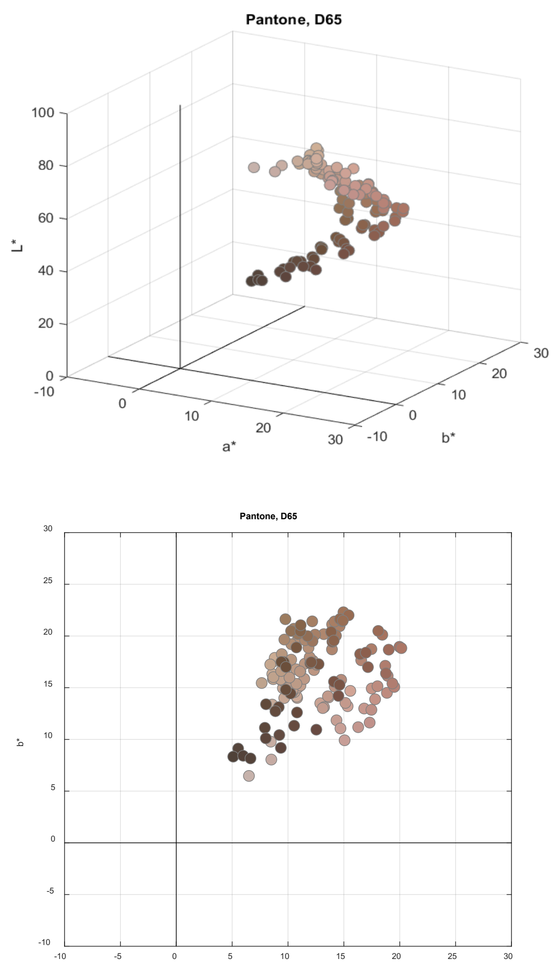


Figure 8: CIELAB coordinated for the Pantone SkinTone Guide patches

Patch Selection

Matching skin measurements to color test patches from the above collection was approached in two ways; colorimetric and spectral. The aim sample was a direct skin measurement. The colorimetric match was achieved by minimizing a computed color-difference measure. The commonly used, ΔE_{00} ('Delta E 2000') measure is a modified Euclidean (visual) distance in CIELAB space.

The minimum ΔE_{2000} was used to select the best match under illuminant D65. Based on CIE color-matching functions, the result shows a close correspondence in the middle wavelengths. The second selection was made by minimizing the mean-squared spectral reflectance error, as was done for the results of Fig. 2. Fig. 9 shows the resulting matches to a member of the NIST skin measurement set.

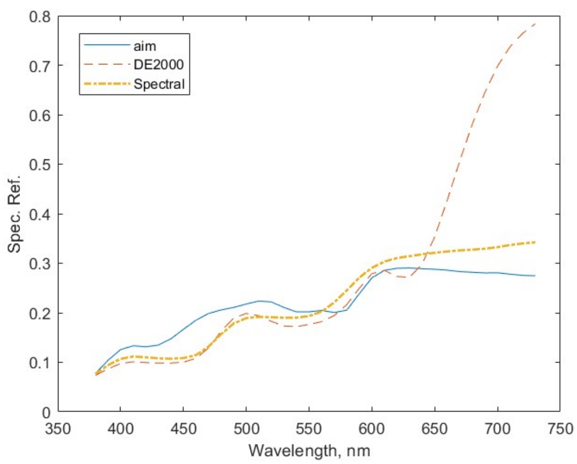


Figure 9: Spectral selection from a set of Pantone SkinTone Guide to match a skin measurement.

Visual Differences

For visually displayed and interpreted imaging systems, e.g., portrait photography, we address visual color differences. This is one way to address both the color-sampling and testing tolerance requirements for both design and evaluation. Most studies of (minimal) perceivable differences rely on the CIELB color space. We again use the ΔE_{00} measure.

We can directly compute the regions corresponding to visual color differences, as estimated by the ΔE_{00}. This was done in a way similar to that reported by Urban, *et al*. [14]. These regions do not vary with L^* level and are shown in Fig. 10. Note also that, although these regions appear as ellipses, these results are not constrained to this form. We can interpret the size of these regions as setting the sampling of color-testing tests. The contour level (in this case 1.0) can be chosen based on the imaging application.

When selecting the number and colors of patches in skin tone test charts, we can be guided by these results. For example, where the contours are smaller, the target color intervals would be smaller, to capture imaging performance when color vision sensitivity is more sensitive. This is particularly useful when the complexity and cost of manufacturing test charts are considered.

Measurement Variability

A second consideration when designing physical color targets is the influence of measurement variability (error). To evaluate this, we used a direct method based on replicate measurements. For all patches of the Pantone SkinTone Guide, three measurements were recorded. Between each measurement, the instrument was removed, so the location was intentionally varied. The CIELAB (L^*, a^*, b^*) values, under D50 illumination. These values were then transformed to the D65 CIE illuminant using the simplified Bradford Colour Adaptation Transform [15]. Two measures were then computed for each set of measurements, the range and standard deviation. Figures 11 and 12 show the results for the L^* and a^*-b^* values, respectively.

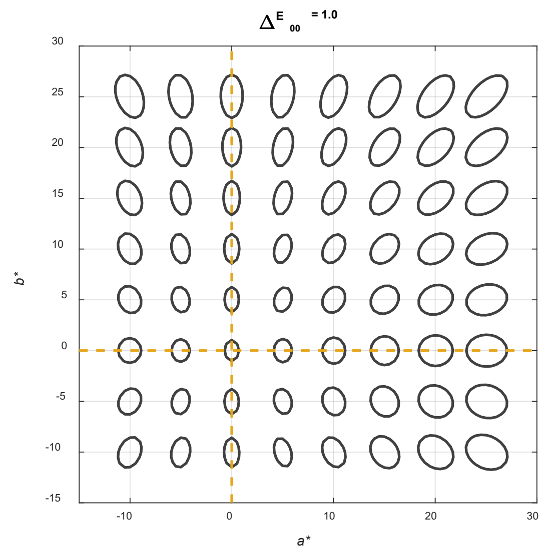


Figure 10: Contours for ΔE_{00}= 1.0 for the a-b* plane range that is important for skin tones*

The variability observed for the repeated colorimetric measurements was generally less than 0.2. As expected, perhaps, this low level compares well with testing tolerances based on visual differences (Fig. 10).

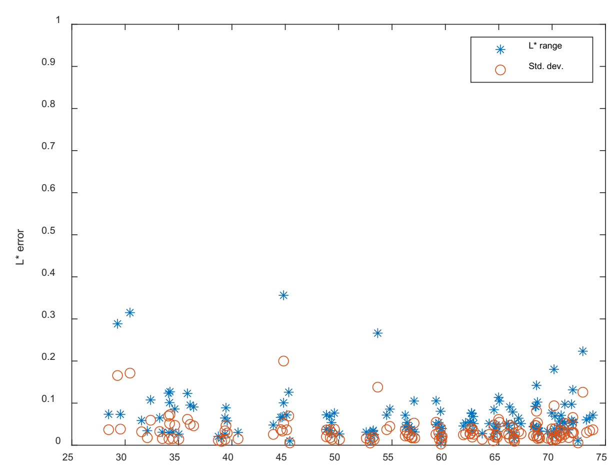


Figure 11: L measurement variability for the Pantone SkinTone Guide samples*

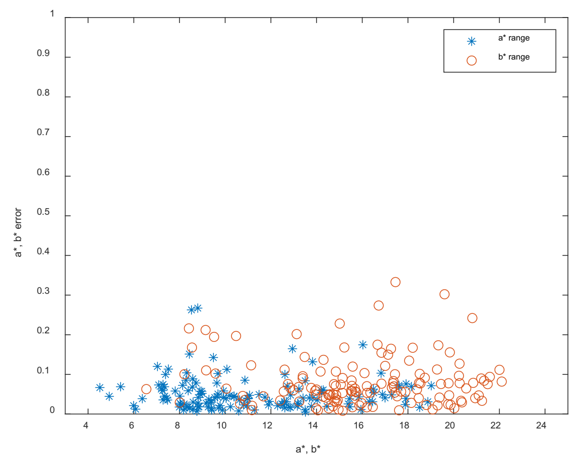


Figure 12: a and b** measurement variability range for the Pantone SkinTone Guide samples*

Discussion

We have addressed imaging performance evaluation color capture. For this application, we consider input-referred color encoding, where information about the scene content is the priority. This would apply, e.g., to archival, scientific, and medical applications. However, color selection for evaluation of image (print) permanence, rendering, and display, out-referred encoding is often used.

While the number of colors depends on the evaluation objective, e.g., system benchmarking usually requires fewer than ICC color profiling. The optical characteristics, however, can be specified based on the anticipated scene content. In many cases, the spectral reflectance characteristics of test charts are chosen for this purpose. In addition, sampling the range, or volume, of the signal space needs is important. For the imaging of human skin, we have considered the color sampling of both a current (colorimetric) scale and a (physical) commercially available set.

Starting with the end in mind, the color sampling can be based on visual color differences, as expressed in CIELAB. When setting system performance levels, however, accommodation of likely measurement variability is also helpful. Our results include both the instrument and test object variability due to the texture or microstructure of the test target.

Acknowledgment

Many thanks to the ISO TC 42 Skin Tone Ad Hoc Group for their helpful discussions.

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] R. S. Berns and M. I. Haddock., A Color Target for Museum Applications, *Proc. Color Imaging Conf.*, 27-30, IS&T/SID, 2010

[3] D. Williams and P. D. Burns, Capturing the Color of Black and White, *Proc. IS&T Archiving Conf.*, 96-100, IS&T, 2010

[4] D. Williams and P. D. Burns, Targeting for Important Color Content: Near Neutrals and Pastels, *Proc. IS&T Archiving Conf.*, 190-194, IS&T, 2012

[5] ColorChecker Target Family, xrite.com, https://www.xrite.com/categories/calibration-profiling/colorchecker-targets

[6] Next Generation Camera Characterization Target, Avian Rochester LLC, https://www.avianrochester.com/nextgen/

[7] FADGI Resources, Image Science Associates, https://www.imagescienceassociates.com/fadgi-resources.html

[8] Megan Borek, Alexander Schwartz, and Amelia Spooner, Evaluating Camera Performance in Face-Present Scenes with Diverse Skin Tones, *Proc. Sym. on Electronic Imaging,* vol. 36., pp 260-1 - 260-5, IS&T, 2024

[9] E. Monk, The Monk Skin Tone Scale (MST), *Open Archive of the Social Sciences*, 2023. doi:10.31235/osf.io/pdf4c

[10] Cooksey, C., Tsai, B. and Allen, D., Spectral reflectance variability of skin and attributing factors, Proceedings of SPIE Defense and Security Symposium, 2015

[11] NIST data set for Ref. 1. Original 51 samples expanded to 100 in July 2022, Reference Data Set of Human Skin Reflectance

[12] L'Oréal Company, A New Geography of Skin Color, posted at https://www.loreal.com/en/articles/science-and-technology/expert-inskin/, (retrieved 25 Jan. 2025)

[13] Pantone Limited Edition SkinTone Guide, https://www.pantone.com/products/fashion-home-interiors/skintone-guide-limited-edition,

[14] P. Urban, M Rosen, R. Berns, Embedding non-Euclidean Color Spaces into Euclidean Color Spaces with Minimal Isometric Disagreement, *J. Opt. Am. A*, vol. 24, pp. 1516-1526, 2007

[15] G. Finlayson, S. Susstrunk, Performance of a Chromatic Adaptation Transform based on Spectral Sharpening, *Proc. Eight Color Imaging Conference*, IS&T, pp. 49-55, 2000

Author Biography

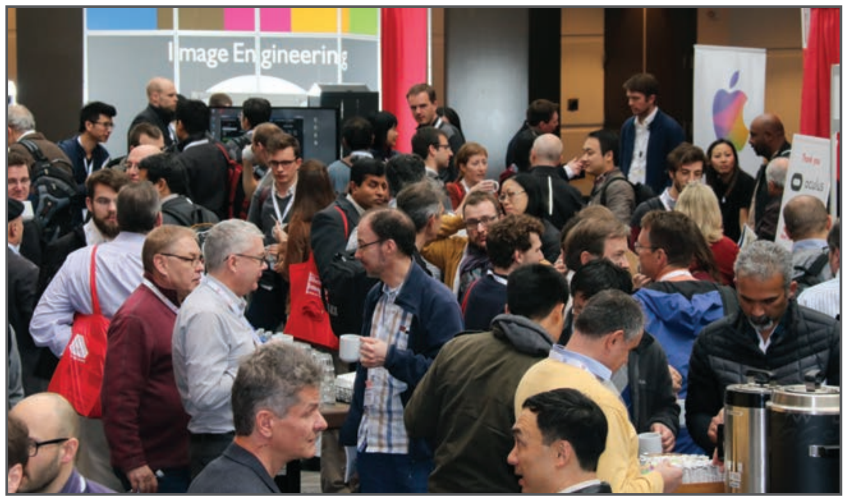
Peter Burns is a consultant for imaging system evaluation, modeling, and design. Previously he worked for Carestream Health, Xerox, and Eastman Kodak. A frequent speaker at technical conferences, he has taught imaging courses for clients and universities for many years. Dr. Burns has authored over 80 technical publications and 23 US patents and is a fellow of IS&T and senior member of IEEE.

Don Williams is the founder of Image Science Associates, which focuses on quantitative imaging performance/fidelity evaluation for digital capture systems. His clients include national libraries, museums, and those with dental, mobile, and automotive applications. Also a fellow of IS&T, Don has taught short courses for many years and contributes to several imaging standards activities

JOIN US AT THE NEXT EI!

electronic IMAGING

Imaging across applications . . . Where industry and academia meet!



• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org