

Update on the INCITS W1.1 Standard for Evaluating the Color Rendition of Printing Systems

Susan Farnand and Karin Töpfer
Eastman Kodak Company
Rochester, NY USA

William Kress
Toshiba America Business Solutions
Irvine, CA USA

Óscar Martínez
Hewlett-Packard Company
Barcelona, Spain

Ann McCarthy and Helen Shin
Xerox Corporation
Rochester, NY USA

Eric Zeise and Dmitri Gusev
NexPress Solutions
Rochester, NY USA

Abstract

The color rendition *ad hoc* team of INCITS W1.1 is working to address issues related to color and tone reproduction for printed output and its perceptual impact on color image quality. The scope of the work includes accuracy of specified colors with an emphasis on memory colors, color gamut, and the effective use of tone levels, including issues related to contouring. The team has identified three sub-attributes of color rendition: 1) color quantization, defined as the ability to merge colors where needed; 2) color scale, defined as the ability to distinguish color where needed; and 3) color fidelity, defined as a balance of colorimetric

accuracy, in cases where a reference exists, and pleasing overall color appearance. Visual definitions and descriptions of how these sub-attributes are perceived have been developed. The team is presently working to define measurement methods for the sub-attributes, with the focus in 2004 being on color fidelity. This presentation will briefly review the definitions and appearance of the proposed sub-attributes and the progress to date of developing test targets and associated measurement methods to quantify the color quantization sub-attribute. The remainder of the discussion will focus on the recent progress made in developing measurement methods for the color fidelity sub-attribute.

Introduction

This paper addresses the work of the W1.1 Color Rendition *ad hoc* team of the INCITS W1 committee on the standardization of office equipment. In September 2000, INCITS W1 was chartered to develop an appearance-based image quality standard.^{1,2} The resulting W1.1 project is based on a proposal,³ that perceived image quality could be described by a small set of broad-based attributes. Five *ad hoc* W1.1 teams were established to generate standards for one or more of these image quality attributes. Color rendition, one of the five *ad hoc* teams, is working to create a standard for evaluating color rendition of hard copy output generated from digital input.

When evaluating overall quality of color images in hard copy format, color appearance is an integral part of image quality. If the color does not look right, other image quality parameters might not matter; the print is considered to be of low quality. What does it mean for the color to “look right,” and how can the ability of any given printer to properly render color be measured? These are the questions that the color rendition *ad hoc* team of INCITS W1.1 is working to understand and answer. To date, the first of these issues has been addressed. Three important parameters or sub-attributes were identified as being necessary to describe what it means for color to be well rendered: 1) color quantization, defined as the ability to merge colors where needed; 2) color scale, defined as the ability to distinguish color where needed; and 3) color fidelity, defined as a balance of colorimetric accuracy, in cases where a reference exists, and pleasing overall color appearance. Along with the definitions, descriptions of how these attributes are perceived have been developed. With the definition complete, work has begun to address the second question of how to measure color rendition. This work has been summarized periodically as it has progressed, most recently at the 2004 IQSP Conference.⁴ Plans are in place to develop a measurement methodology for each sub-attribute.

Defining Color Rendition

Before a standard for measuring color rendition of printing systems could be established, definitions of color rendition and printing systems were necessary. Knowing what is and, just as importantly, what is not included in their definitions, when attempting to evaluate how well color is rendered, was considered

essential to determining how color rendition would be measured. First, an understanding of what constituted “the printer” was needed. It was decided that the printer would be defined as “everything that occurred after hitting the ‘print’ button.” As the definition of color rendition was discussed, it was decided that variations in color rendition across an image where a constant code value is specified is not part of color rendition, but would come under the auspices of the macro-uniformity, micro-uniformity, or adjacency attribute *ad hoc* teams (depending on the nature of the variation). Conversely, tonal rendition was considered to be inextricably linked to color rendition and was, therefore, subsumed by the color rendition *ad hoc* team and included in evaluating color rendition. Finally, it was agreed that different metrics might be necessary for natural images and vector graphic images because of the adaptation effects and differences in the image content and intended use of vector graphic images.

With these considerations in place, the color rendition *ad hoc* team determined that “color rendition” could be defined in terms of three unique components or sub-attributes: color quantization, color scale, and color fidelity. A summary of these three sub-attributes of color rendition, their definitions, their correlations to classical color parameters, and examples are listed in Table 1. These sub-attributes are different from the traditional image quality color attributes; however, they do appear to form a complete perceptual set of attributes and were chosen because they emphasize the perceptual nature of the attributes.

Color quantization is defined as “the ability to merge colors where needed.” A failure in color quantization would manifest itself in images primarily as contouring. Images with adequate color quantization would be free of the contouring that can appear in gradient areas, such as sky, walls, or faces, in pictorial images, or backgrounds in graphic images.

Color scale is a term developed to be analogous to tone scale but which is meant to include not just changes in lightness, as might be thought of for tone scale, but changes in chroma and hue as well. With this in mind, color scale is defined as “the ability to distinguish colors where needed.” Images having proper color scale would not suffer from loss of information in shadow or highlight regions but would have adequate separation of colors to provide local color contrast where needed.

The third attribute of color rendition, color fidelity, is defined as a balance of colorimetric accuracy (hue, lightness, and chroma) in cases where a reference exists and pleasing, overall color appearance, including

Table 1: The components or sub-attributes of color rendition.

Color Quantization	
Definition	<ul style="list-style-type: none"> • The ability to merge colors where needed • Colors that should be perceived as smoothly varying are free of contouring
Correlation with classical color sub-attributes	<ul style="list-style-type: none"> • This attribute would include contouring
Examples (Appearance of sub-attribute in images)	<ul style="list-style-type: none"> • Freedom from contouring in smooth sweeps (human faces, sky, differentially lit backgrounds, walls, etc., in pictorial images or vignettes, fades, or graduated backgrounds, etc., in graphical images)
Physical measurement, test target	<ul style="list-style-type: none"> • Printed RGBCMYK sweeps; perhaps skin tone and blue sky sweeps, and the $\Delta E(\Delta L^*, \Delta C^*, \Delta h^*)$ calculated from a digital scan of the sweep. Minimum slopes must be defined such that anything greater would result in contouring, (difficult due to image dependence • Could be 2-D sheets or simulated 3-D objects
Color Scale	
Definition	<ul style="list-style-type: none"> • The ability to distinguish colors where needed • Colors that should be perceived as separate are distinguishable
Correlation with classical color sub-attributes	<ul style="list-style-type: none"> • This attribute would include tone scale (lightness, contrast, shadow and highlight detail, and tonal clipping), hue “scale,” and color saturation (chroma) • The equivalent of “tone scale,” but extended to cover C^* and h in addition to L^*
Examples (Appearance of sub-attribute in images)	<ul style="list-style-type: none"> • Distinguish between dark colors to avoid blocked shadows • Distinguish between light or pale colors to avoid blown highlights • Adequate separation of colors to provide contrast in pictorial images and colors that are perceptually distinguishable in graphics • Are there enough colors and are they distinct where they should be? • Contrast, color saturation, tonal detail in light and dark colors, detail in saturated colors, ability to distinguish hues
Physical measurement, test target	<ul style="list-style-type: none"> • Printed RGBCMYK sweeps; maybe skin tone and blue-sky sweeps. Minimum slopes would have to be defined such that anything less would result in colors that are blocked, etc.
Color Fidelity	
Definition	<ul style="list-style-type: none"> • Colors look correct • Achieve a pleasing balance between colorimetric accuracy (hue, lightness, and chroma) in cases where a reference exists, and pleasing overall color appearance. • Provide preferred memory color reproduction. • The ability to match colors where needed. Does not require matching to a target print but can include matching to target colors.
Correlation with classical color sub-attributes	<ul style="list-style-type: none"> • This attribute would include hue accuracy (lightness and chroma accuracy as well), color balance, and memory color reproduction
Examples (Appearance of sub-attribute in images)	<ul style="list-style-type: none"> • Skin tones, grass, and sky look realistic • Image is not too saturated or too washed out or too dark, etc. • Image looks natural, color rendition meets expectations • Color match a target image when needed
Physical measurement, test target	<ul style="list-style-type: none"> • To be determined

reasonable reproduction of memory colors. This attribute will be a challenge and will likely require separate test targets and metrics to characterize images with text and graphical content and others with pictorial content. Although the ability to match colors is included in color fidelity, it is not meant to imply that matching to a target or original is always necessary or preferred. While in graphical and text images a colorimetric match may often be desirable, a pictorial image with good color fidelity should have, skin tones, grass, and sky areas that look realistic. Pictorial images, in general, should look natural.

Measuring Color Rendition

Once color rendition was defined, the next step was to establish methodology for measuring the three sub-attributes of color rendition. Initially, efforts focused on the color quantization sub-attribute because agreement of a general approach for color quantization came most easily. It seemed that because color quantization and color scale were related and inverted problems, they could be both evaluated using color sweeps. Establishing a test target and measurement methodology for one could very well be suitable for the other. However, while it was difficult to define exactly when a problem occurred in the case of color scale (What exactly is it *not* to have blocked shadows?) for color quantization, it was clear that perceptible contours constituted a problem. It was, therefore, decided to first work on the problem of measuring color quantization.

A strategy of first developing a test target, followed by a method of measuring the target, was adopted because of the difficulty of having any meaningful discussion without data or, at least, an attempt at obtaining data. After much discussion and a few trial runs, it was decided that the test target should comprise color gradients in a serpentine format, Figure 1. Five monochrome black targets were printed, one each having steps between consecutive patches of one, two, and four printer counts and two having three printer count steps, one ranging from 1 to 241 printer code values and one from 16 to 256. Prints of these targets were made on photographic, electrophotographic, and inkjet printers and were evaluated for contours by several members of the group. The results of these initial evaluations, which were well correlated, indicated that the visibility of contours was related to the ΔL (or ΔE) between two adjacent patches along

with the density of those patches, with the contours being more apparent in the mid-density range.

With the development of the color quantization measurement methodology well underway, it was decided that next sub-attribute to be addressed would be color fidelity. This effort was initiated by conducting psychophysical experimentation using a set of images printed on equipment having a wide range of image quality capability in an attempt to establish an initial scale of color fidelity. Eight pictorial images was selected for printing from a set of images purchased from an image "bank" for general testing purposes. The images were transformed from their original color spaces using their source profile, the sRGB destination profile in Adobe Photoshop™ 6.0, and Adobe's Color Management Module. Settings included Relative Colorimetric rendering intent and black point compensation. The chosen scenes include four images containing people, two scenic images, and two images depicting building interiors, Figure 2. A custom color chart, consisting of 64 important colors, was extracted for each image using the "Color Table" feature included when the "Save for Web" function is selected in Adobe Photoshop™. Each custom color chart was printed along side its corresponding image on eight different printers that spanned a range of image quality capability and included photographic, electrophotographic, and inkjet printers. An experiment was conducted in which observers were asked to scale the eight prints for each of the eight scenes in terms of color fidelity, where color fidelity was described as "essentially that the colors look correct and pleasing." A total of 25 observers participated in the experiment at four different locations. The experimental results are currently being examined to determine whether there is some perceptual attribute or attributes associated with the average perception of color fidelity that can be correlated with physical measurements of the accompanying color charts, looking first for the patches from the custom color chart that show reasonable correlation between ΔE_{2000} values (calculated using the highest rated image as a reference) and the observer scale values. Once the patches having ΔE_{2000} values that were well correlated with the observer results were determined, the L^* , C^* , and h values of these patches were plotted relative to the scale values. For some of the patches in each scene, these plots show reasonable trending of the colorimetric value against the observer scale values, especially for the C^* values. One other interesting bit

of information determined from the data was that the standard deviation of the observer scale values was higher across the individual printers than it was over the individual scenes, suggesting more variability, which is due more to individual printer than the individual scene, which is a useful characteristic of a scene set, when trying to distinguish printer performance.

Several issues arose when examining the color fidelity scaling experimental results, foremost of which was that, for a couple of patches for certain prints, the best and worst rated prints had nearly identical colorimetry, indicating that color fidelity quality is driven by more than just color change in isolation. To attempt to examine this effect further, patches that form a "tone scale," of sorts, will be evaluated where possible. Another issue was that the printers and scenes included in the experiment might not have been an adequate representation of printer image quality. It was felt that the eight printers that were used might not have been enough images to evaluate color fidelity quality and that it might be helpful to increase the number of samples of each scene by making deliberate changes to the color in the digital source files using image-manipulation software, such as Adobe™. And it was discovered that the colors found to be important in each image, taken as a whole, did not cover color space; with cyans and purples, in particular, determined to be missing.

Next Steps

In the coming year, attention will continue to be focused on the color fidelity sub-attribute. Further experimentation is planned to investigate possible relationships between perceived color fidelity and physical attributes in images. Evaluation is also planned for the color quantization experimentation examining contouring in the serpentine targets. Having worked through the issues associated with test targets for the color quantization sub-attribute, the team will continue the process of evaluating the results of that exercise and consider what can now be applied to the color scale sub-attribute.

Conclusion

A method of quantifying the perceived image quality of the color rendition of a printer has been defined as comprising three basic components: color fidelity, color scale, and color quantization. These sub-attributes have been defined, related to classic color components, and descriptions of how these attributes might be perceived in images have been developed. Development of a measurement methodology for the color quantization sub-attribute has been initiated. Potential test targets have been designed and generated. Prints have been made of these targets on photographic, electrophotographic, and inkjet printers, and initial attempts of measuring contours have been performed. A color fidelity strategy has been proposed and, again, prints have been made on photographic, electrophotographic, and inkjet printers. Further work is underway and will be reported in a formal presentation, as it is available. Also planned is a review of future work, including planning for fundamental research that will be needed to fully answer the question regarding measuring the ability of a printer to render color.

References

1. E. K. Zeise and N. W. Burningham, Standardization of Perceptually Based Image Quality for Printing Systems (ISO/IEC JTC1 SC28 and INCITS W1.1), *Proc. NIP18*, pp. 699–702 (2002).
2. N. W. Burningham and E. N. Dalal, Status of the Development of International Standards of Image Quality, *Proc. PICS*, pp. 121–123 (1998).
3. E. N. Dalal, D. R. Rasmussen, F. Nakaya, P. A. Crean, and M. Sato, Evaluating the Overall Image Quality of Hardcopy Output, *Proc. PICS*, pp. 169–173 (1998).
4. S. P. Farnand, R. E. Cookingham, K. Topfer, E. N. Dalal, A. McCarthy, W. Kress, O. Martinez, E. K. Zeise, and D. Gusev, Update on the INCITS W1.1 Standard for Evaluating the Color Rendition of Printing Systems, *Proc. IQSP* pp. 55-59 (2004).
5. H. Mizes, Precision Tone Reproduction Curve Measurement Technique for High Image Quality and Halftone Design Applications, *Proc. NIP18*, pp. 486–489 (2002).
6. W. Kress, Digitization and Metric Conversion for Image Quality Test Targets, *Proc. PICS*, (2003).

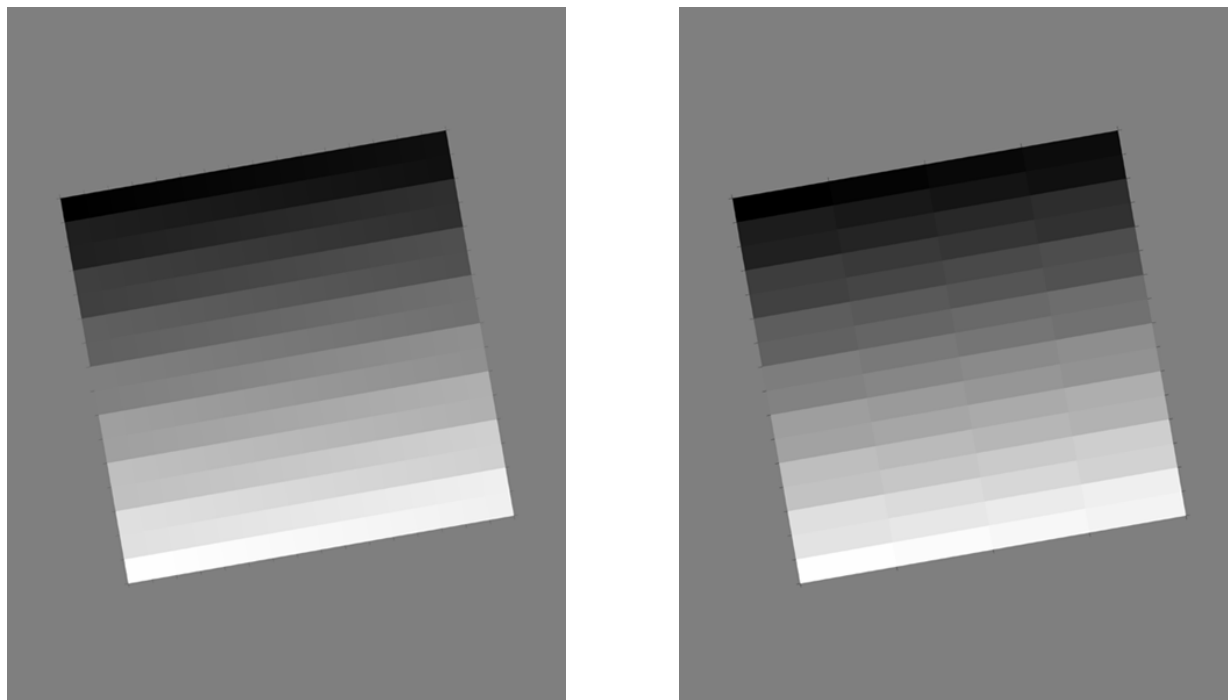


Fig. 1: The serpentine test targets in steps of one printer code value (left) and four printer code values (right) for the assessment of color quantization.

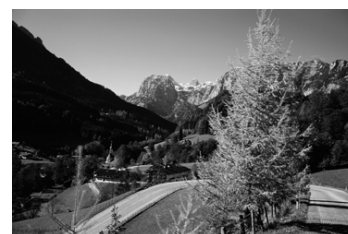
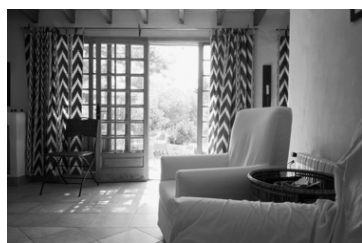


Fig. 2: The eight scenes included in the initial color fidelity experiment.

