INCITS W1.1 Standardization for Evaluation of Perceptual Macro-Uniformity for Printing Systems

D. René Rasmussen, Xerox Corp., Webster, NY
William C. Kress, Toshiba, Irvine, CA
Marguerite Doyle, Lexmark Intl., Lexington, KY
Yee Ng, NexPress Solutions, Rochester, NY
David Wolin, ImageXpert Inc., Nashua, NH
Steven Korol, Xerox Corp., Williamsville, OR

Abstract

This paper describes the status and progress of the INCITS W1.1 Macro Uniformity ad hoc team. The team has defined the Macro Uniformity attribute, has developed several test patterns to be used for subjective and objective evaluations and has defined test patterns and methods to address color conversions of digitizing devices. These activities are reviewed in this paper.

Introduction

INCITS W1 is the U.S. representative of ISO/IEC JTC1/SC28, the standardization committee for office equipment. At PICS 1998 a methodology of describing perceived image quality, by a small set of broad-based attributes, was presented[1] and used as the basis for the current W1.1 project for development of an appearance-based image quality standard.[2],[3] There are currently five ad hoc W1.1 teams, each working on one or more of these image quality attributes. This paper summarizes the work of the W1.1 Macro Uniformity ad hoc team. The active members of the team are:

Steven Korol, Xerox Corp.
William Kress, Toshiba
Yee Ng, NexPress Solutions
René Rasmussen, Xerox Corp., Chair
David Wolin, ImageXpert Inc.

and until recently also Marguerite Doyle of Lexmark International.

In order to secure that the resulting standard becomes appearance-based, and can be applied on equal terms across marking technologies, the guidelines of W1.1 specifies the following steps for development of the standards[4]:

A. Define the attribute based on appearance.
B. Design digital test targets.
C. Create a collection of hardcopy test samples spanning a diverse range of marking technologies, image quality levels, and defect types (within the given attribute).
D. Digitize the hardcopy sample images (e.g. by scanning).
E. Perform surveys to obtain subjective ratings of the hardcopy samples.
F. Develop objective metrics for quantification of the attribute. These metrics must be appearance-based, in the sense that they scale with human visual perception of the attribute.

These steps are followed by several more steps in order to test and establish correlations between the objective and subjective measures of the attribute. In this paper we will discuss progress on each of the above steps, focusing on steps A through E, where most progress has been made. We will then discuss next steps and opportunities to contribute to this activity.

Macro Uniformity Attribute Definition

The team has reached a tentative attribute definition.[5],[6] It may be subject to revision when the method is tested in practice and after further coordination with the Micro Uniformity team. The definition is stated here with comments in cursive type:

1. Macro-uniformity refers to the <u>subjective</u> impression of color uniformity across a <u>large</u> image area that is intended to have uniform color (also know as a "flat field"), as determined from surveys which involve a large population that is representative of document end-users (rather than image quality experts). *A complete definition of the subjective evaluation method and viewing conditions will be defined later (step E).*
2. <u>Color uniformity</u> here refers to all forms of color variations from the average, be it in lightness, hue, saturation or combinations thereof.
3. <u>All forms of spatial</u> color variations are taken into account, including for example 1-dimensional, 2-dimensional, periodic, non-periodic, localized, large-scale as well as small-scale. The subjective evaluation addresses only the "overall uniformity". *The objective evaluation may consist of several measurements of specific sub-attributes (e.g. 1-dimensional banding), but a suitable combination of those measurements must be established*

which scales with the subjective assessment of overall uniformity.

4. The image size and viewing distance must be such that attention is naturally drawn towards the "macro" aspects of uniformity. This can be obtained by evaluation of large (tentatively 170mm by 170mm) flat field areas at a specified normal viewing distance of approximately 40cm. *This part of the definition represents a compromise between (i) the need to ensure that the subjective evaluation can be performed in an easy, natural manner, and (ii) the desire to minimize the overlap between the Macro and Micro Uniformity attributes. Instructions to the subjects such as to "ignore high-frequency non-uniformities" would likely lead to much noise in the subjective evaluation and are not desirable. If necessary, hardcopy samples with little or no Micro Uniformity defects may be used to help obtain a better distinction between the two attributes.*
5. Images must be viewed in a manner such that only diffusely (non-specularly) scattered light is taken into account. The Gloss and Gloss Uniformity attributes will address image quality issues related to the specularly scattered light.

Notice that this definition does not incorporate a lower limit on the scale (that is, an upper frequency limit), and therefore will overlap with the Micro-uniformity attribute. The Macro (or Micro) aspect comes in only through the selection of the size of the print sample that is used for the surveys. This makes it significantly easier to design a good subjective survey, than if the effect of small-scale color variations had to be eliminated by technical instructions to the observers. With such an overlap between Macro Uniformity and Micro Uniformity, these attributes will not be orthogonal in a mathematical sense, but they will be independent and span all aspects of color uniformity.

Given the challenges of defining a suitable boundary between Macro and Micro Uniformity, the teams considered whether the two attributes should be replaced by a single Uniformity attribute. However, it was agreed that with only a single uniformity attribute, the full set of W1.1 attributes would not be able to adequately characterize the image quality of a printer system. The type of documents printed by the end-user, has a significant impact on the relative importance of Macro and Micro Uniformity. For documents that contain no large-size nearly flat-field regions, the Macro Uniformity attribute will have very little consequence for the perceived quality, while Micro Uniformity may still play a significant role. Therefore (at least) two attributes, along the continuum of spatial scales encountered in end-user documents, should be used.

The stated definition deliberately avoids reference to specific defects known to exist for the various marking technologies. Examples of defects that would be in the Macro Uniformity category are:

- Banding (1-dimensional, periodic lightness and/or chromatic variations)
- Streaks (1-dimensional, isolated lightness and/or chromatic variations).
- Mottle (2-dimensional, random lightness and/or chromatic variations).
- Gradients (could be a special case of banding where the spatial period is larger than the image size).
- Moiré patterns (1- or 2- dimensional regular patterns).

Most or all of these defects are also candidates for Micro Uniformity, but as the spatial scale of a defect increases, it will assume a relatively higher weight in the Macro Uniformity attribute.

Test Targets and Hardcopy Samples

The W1.1 project addresses both black & white as well as full-color printing systems. This means, that to characterize a printing system in terms of Macro Uniformity, evaluations must be performed on a suitably large sample of colors within the color gamut of the system. In particular, it is not adequate to evaluate cyan, magenta, yellow, and black separations only. Firstly, the end user rarely prints cyan, magenta, and yellow separations, and secondly, some marking technologies may introduce interactions between the separations, such that Macro Uniformity of process colors cannot be predicted from the Macro Uniformity of separations.

For Macro Uniformity, where a large image (say 170mm by 170mm) must be printed in order to evaluate a single color, this poses a potential issue with respect to practical evaluations that cover the color gamut adequately. This issue was temporarily bypassed by distinguishing between two phases of the standard development:

- Phase I—Establishing the objective measurement method. This covers how to measure a single flat field print sample of a given color, not how to characterize the printer system with respect to all colors.
- Phase II—Specification of how to sample the color gamut of a printer system and combine the Macro Uniformity measurements of individual colors into an overall Macro Uniformity assessment of the printer system.

We are currently focusing on phase I. Here the main challenges lie in the subjective survey and in the development of measurements that correlate with human visual perception of Macro Uniformity as defined by that survey. While human spatial visual perception does have significantly different properties with respect to lightness variations versus chromatic variations, correlations between measurements and perception can be developed and tested using relatively few, selected colors and types of color variations. Therefore, in phase I, we are primarily concerned with selecting a set of colors and color variations that will allow development of measurements and establish correlation to subjective ratings. Phase II will then address how to apply those measurements to fully characterize the Macro Uniformity of a printing system across its color gamut.

As a start, the team has developed a suite of test patterns that will allow testing at 7 different base colors:

- Black only (K) test patterns, with target CIE L* values of 40, 60 and 80, as well as 100% K.

- Blue test patterns defined as equal amount of cyan (C) and magenta (M), with target CIE L* values of 60 and 80, as well as 100% CM. The hue and exact ratio of C to M percentages are not important for this purpose, so although the test pattern is defined in terms of CMYK, it can be used also for “RGB printers” where the color is intermediately represented as RGB and may not produce prints with the originally specified C to M ratio.

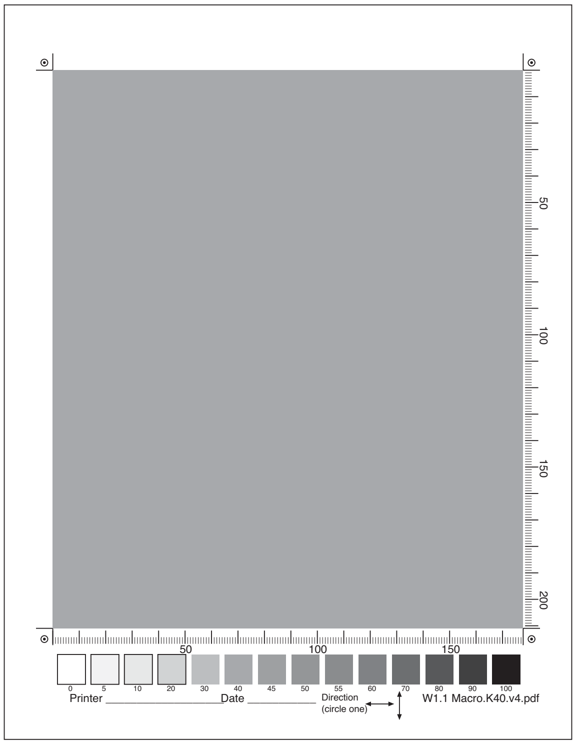

Figure 1. The “W1.1.Macro.K40” (version 4) test pattern, reduced to approximately 35% size.

It is well known that human visual sensitivity to spatial L* variations depends on the base lightness, and therefore this set of test patterns covers a range of base L* values. The test patterns are identical in terms of spatial content, and differ only in the base color. Figure 1 shows one of the test patterns in reduced size. The primary content is a 178mm by 211mm flat field region, part of which will be used for the subjective survey.

The procedure to be used when printing the test patterns has been defined,[7] and specifies how to obtain the target L* values with the required accuracy. Notice, that for phase I these is no need for the base color to very accurately match a target color; however, phase II may need to involve a printer calibration procedure, which will ensure that the printed hardcopy samples are within specified limits of a specified base color.

Several team members, as well as members of the Micro Uniformity team, have already printed these test patterns, covering several different marking technologies. However, the process of producing the hardcopy samples were halted until issues related to color calibration of digitizing devices had been resolved. This is explained in more detail in the next section.

In addition to generating hardcopy samples directly via printing the test patterns on existing printers, the team is exploring the possibility of producing prints with simulated, controlled levels of various Macro Uniformity defects. Photo-quality ink jet printers and image setters are considered as output devices for such simulations. One advantage of simulations is that, if deemed necessary, the effect of high-frequency non-uniformities, which are addressed by Micro Uniformity, can be reduced or eliminated.

Digitization

Step D in the process calls for digitization of the hardcopy sample images, representing the images with sufficient accuracy both spatially and colorimetrically. Purely from the perspective of visual perception, we make the following observations regarding accuracy requirements for the Macro Uniformity attribute:

- Spatial resolution. The human visual contrast sensitivity function for lightness falls off rapidly when the frequency is higher than about 3 cycles per degree[8] (equivalent to about 0.5 cycles per millimeter at 40cm viewing distance). Sensitivity to chromatic variations is poorer than this. Digitization with an MTF equivalent to this is therefore more than adequate for Macro Uniformity.
- Color accuracy and resolution (bits). For a uniformity metric, the absolute accuracy in CIELab color space is relatively insignificant, but the accuracy of measuring color differences is very important. The human visual system is extremely sensitive to periodic lightness variations at the “right” frequency, capable of detecting and objecting to L* amplitudes as small as 0.15.[9]

For the standard to provide practical value it should be easy and cost-effective to perform the measurements. The team is currently exploring commercial “graphics arts” scanners as digitizing devices. Such scanners are not designed as physical measurement devices, and while they may be able to provide sufficient accuracy, it would be ideal to develop the standard based on measurements performed with established physical measurement devices (e.g. a scanning micro-colorimeter). However, the team does not currently have access to such a device. The team, jointly with the Micro Uniformity team, has currently identified 3 scanners, which will be used and tested initially for digitization:

- ScanView ScanMate 4000 drum scanner. Optical resolution 4000dpi. Three photomultiplier sensors (RGB). 36 bits. Scanner will be operated at 600dpi with a corresponding aperture size to avoid aliasing.
- Epson Expression 1680 flatbed scanner. 1600 by 3200 dpi optical resolution. CCD sensors. 48 bits.
- UMAX Powerlook III flatbed scanner. 1200 by 2400 dpi optical resolution. CCD sensors. 42 bits.

Device spatial resolution

The spatial resolutions of all these scanners are expected to be sufficient to meet the requirements. Prior characterizations of these scanners have found MTFs that are better than the human contrast sensitivity function at normal viewing distance. A larger issue may be potential sampling artifacts. Whereas most of the print samples, which we have used so far in testing, have been contone, the majority of hardcopy samples that will be used for the standard development and application of the standard will be halftoned. To avoid moiré pattern artifacts in the scanned images caused by aliasing, the sampling resolution must be carefully adjusted to the optical MTF of the scanner. This means that scanners with a "good" optical MTF such as for example the Epson 1680 may have to be operated at a higher sampling resolution, than a scanner with a poorer optical MTF. After acquiring the image data at a resolution that is sufficiently high relative to the optical MTF, the data may be converted to a lower resolution (e.g., by averaging pixels in reflectance space), which would lead to smaller data storage requirements. The team plans to perform scans at 600dpi and higher, depending on the scanner, possibly with subsequent resolution conversion.

Device Uniformity

Flatbed scanners typically exhibit spatial non-uniformities, especially low-frequency variations across the platen, for example caused by non-uniform illumination. More complicated non-uniformities can be induced by the image that is being scanned, through an integrating cavity effect, whereby the light scattered off one position of the image, through multiple scattering, contributes to the effective illumination of another part of the image. The drum scanner performance is typically much better in this regard, and it will be used as a tool to check scanner uniformity.

Device color accuracy and resolution

To develop measurements that will correlate with human perception of Macro Uniformity it is critical that any image analysis is performed on colorimetrically calibrated images, rather than on device dependent RGB images. However, the scanners being used all acquire the image as device dependent RGB, such that it is necessary to establish a transformation from the scanner RGB to colorimetric space, e.g. to CIELab. Furthermore, such a transform will depend on the materials of the printed sample that has been scanned, most notably on the ink or toner colorants. For this reason, it was decided to halt the printing of hardcopy samples, until a viable means of scanner calibration was determined.

To address this issue, the team has designed two digital test patterns which can be used to calculate the RGB to CIELab transformation for the given scanner / printer / materials combination. Several existing test patterns for scanner calibration were considered, but were either proprietary or used color swatches too small to be reproduced uniformly by all marking technologies. The test pattern shown in Figure 2, "W1.1.SCC.Macro.CMYK," will

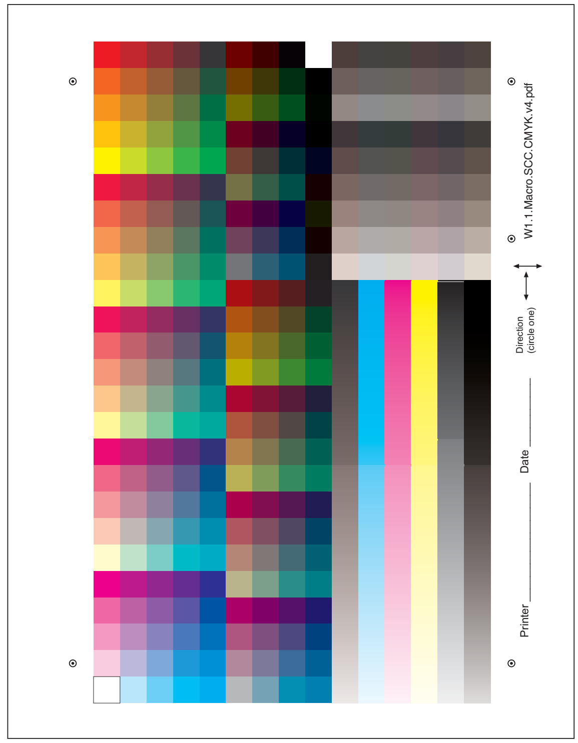

Figure 2. The "W1.1.Macro.SCC.CMYK" (version 4) test pattern, reduced to approximately 35% size.

be used for scanner calibration for CMYK printers. The primary content of the test pattern is as follows:

- 5 by 5 by 5 matrices of (C,M,Y) at K=0%.
- 4 by 4 by 4 matrices of (C,M,Y) at K=33%.
- 3 by 3 by 3 matrices of (C,M,Y) at K=67%.
- Step wedges of K, C, M, and Y, with step size 6-7%.
- Step wedges of (C=M=Y) with step size 6-7% and K=0%.
- Step wedges of (C=M=Y=K) with step size 6-7%.
- Near neutral colors.

A test pattern similar to the one shown in Figure 2 has been designed for "RGB printers", in order to ensure that the full gamut of such printers is covered.

A third test pattern, "W1.1.Macro.SCCTest," has been designed to allow independent testing of the accuracy of the color transformation.[10] This test pattern is shown in Figure 3. The test pattern has been designed with these points in mind:

- The majority of colors should not be identical to the "training colors" which were used to generate the scanner color calibration.
- Since the ultimate purpose is to measure image quality, especially the spatial variation of color, with a calibrated scanner, it is important that we can assess the scanner's capability of accurately measuring small color differences over short distances, or similarly, the scan uniformity over small distances.

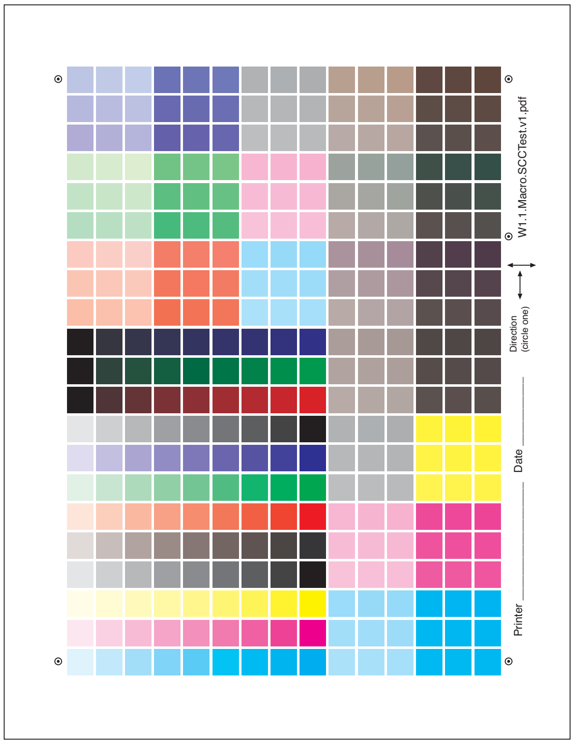

Figure 3. The "W1.1.Macro.SCCTest" (version 2) test pattern, reduced to approximately 35% size.

The test pattern contains twelve step wedges, five of which are in CMYK color space, and seven in RGB color space. These are meant to check the color calibration for primary and secondary colors. The test pattern also contains 23 blocks of 3x3, which are meant to check accuracy of small color differences over relatively small distances (that is, distances of 1-2cm, as opposed to distances from one end of the paper to the other). Many scanner non-uniformities vary slowly across the platen, and therefore scanners can more accurately measure color differences over such small distances, and that may be sufficient for many measurements—especially for Macro Uniformity. Within each 3x3 block the C, M, Y, or K (or R, G, or B) levels vary in steps of approximately 1-2%. Only one or two of the color components vary. For example, for the block at the top-left corner, RGB colors vary from (179, 179, 255) to (199, 219, 255) with R in steps of approximately 1% and G in steps of approximately 2%.

Extensive analyses were done[11] using these and other test patterns and the results are reported elsewhere.[12] Based on those results the team has concluded that we can proceed with generation of hardcopy samples of the Macro Uniformity test patterns, and that one of the scanner calibration test patterns must be printed at the same time and under the same printing conditions.

Next Steps

The test patterns for scanner color calibration and calibration test will be modified slightly to accommodate requests from other W1.1 teams. After those test patterns are finalized the Macro Uniformity team is ready to proceed with generating more hardcopy samples and with scanning the samples. This will be done jointly with the Micro Uniformity team, which at least initially will use the same test pattern. We expect to need a substantial number of hardcopy samples, so the practical issue of how to store and share the Giga- or Tera-bytes of scan data will have to be resolved.

After acquiring the hardcopy and digitized images, the team will address subjective surveys (step E) and development of objective measurements (step F).

The team will also address phase II, the application of the standard, including questions such as which base colors must be sampled, tolerances on the base colors, and potential procedures to ensure that those tolerances are met.

Invitation to Contribute

The reader is invited to contribute to this activity! There are several ways to contribute:

- Submit hardcopy print samples that illustrate various aspects of Macro Uniformity defects. Contact any of the authors for further instructions.
- Become an active team member.
- And finally, if you have an accurate scanning micro-colorimeter to offer for this activity you will be more than welcome!

Conclusion

The W1.1 Macro Uniformity team has made progress along the W1.1 project guidelines. The attribute has been defined in terms of appearance, and digital test patterns for both subjective and objective evaluations have been developed. Issues related to measurements of hardcopies have been addressed, specifically the color calibration of scanners. The team expects that further work will be necessary to address the digitization process, but at this point generation of hardcopy samples can continue. The significant tasks of subjective surveys (step E) and development of objective measurements (step F) lie ahead of us.

References

Documents of the W1.1 project are herein referred to with the format "*W1.1 2002-001*" and can be found at the web site http://www.incits.org/tc_home/w11htm/incits_w11.htm.

1 E. N. Dalal, D. R. Rasmussen, F. Nakaya, P. A. Crean and M. Sato, Evaluating the Overall Image Quality of Hardcopy Output, *Proc. PICS,* pg. 169. (1998).

2 N. W. Burningham and E. N. Dalal, Status of the Development of International Standards of Image Quality, *Proc. PICS,* pg. 121. (1998).

3 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,* pg. 699. (2002).
4 *W1.1 2000-003*, W1.1 Image Quality of Printers. Guide for the Development of Standards. (2000).
5 *W1.1 2001-022*, Proposal regarding attribute definition for Macro uniformity. (2001).
6 *W1.1 2001-023*, Teleconference meeting minutes. (2001).
7 *W1.1 2002-008*, Macro uniformity print procedure. (2002).
8 Brian A. Wandell, *Foundations of Vision*, Sinauer, MA,1995.
9 H. Mizes, N. Goodman and P. Butterfield, The perceptibility of random streaking, *Proc. PICS,* pg. 89. (2000).
10 *W1.1 2002-038*, Proposal for test pattern to test scanner color calibration. (2002).
11 *W1.1 2002-040*, Preliminary analysis of scanner RGB conversion to RGB. (2002).
12 W. C. Kress, Digitization and Metric Conversion for Image Quality Test Targets, *Proc. PICS*. (2003).

Biography

René Rasmussen received his Ph.D. degree in Physics from the Niels Bohr Institute, Copenhagen, in 1990. Since 1992 he has worked in the Wilson Center for Research and Technology at Xerox Corporation in Webster, NY. His work has primarily focused on the development of methods for subjective and objective image quality measurements with applications to product development and manufacturing. He is a member of the IS&T and SPIE.