Effect of Selected Image Quality Attributes on Overall Preference

Kristen Natale-Hoffman*, Edul Dalal*, René Rasmussen* and Masaaki Sato†
***Xerox Corporation, Webster, NY, USA**
†Fuji-Xerox Co., Ltd., Ebina, Japan

Abstract

This paper describes a relationship between customer preference for color images and more objective image quality attributes. This is just one relationship within a high level image quality hierarchy that needs to be understood in order to completely describe overall image quality. Customer preference was quantified for one set of color images that spanned a wide range of image quality by means of a preference survey. The images were also visually assessed and the levels of five selected image quality attributes were evaluated. Attribute importance was ranked with respect to overall image preference for each of the five image quality attributes. Color rendition was found to have the strongest correlation with average preference values for the image set studied. Results show that subjective preference can be represented in terms of an S-shaped dependence on color rendition and micro uniformity.

Introduction

As discussed by Dalal et al,[1] there are three different domains in which color printer image quality can be represented: image quality metrics, image quality attributes and image preference. Image quality metrics are the objective, typically instrumented, measurements of specific aspects of the quality of an image, and are ideally suited for engineering specifications. Image preference is an overall measure of how well customers like a given image, typically based on expensive surveys involving large numbers of observers. Image quality attributes are used to provide a high-level description of image quality, and also to bridge the gap between the metrics and preference domains. Figure 1 shows an illustration of these three complementary systems for image quality assessment.

An image quality attribute is a high-level image quality descriptor, such as micro uniformity, which describes the perceived uniformity on a small spatial scale in a given image. Several analytical image quality metrics may be associated with a given attribute: e.g., metrics such as solid area graininess, halftone graininess, RMSL mottle, etc. are used to measure different aspects of the micro uniformity attribute.

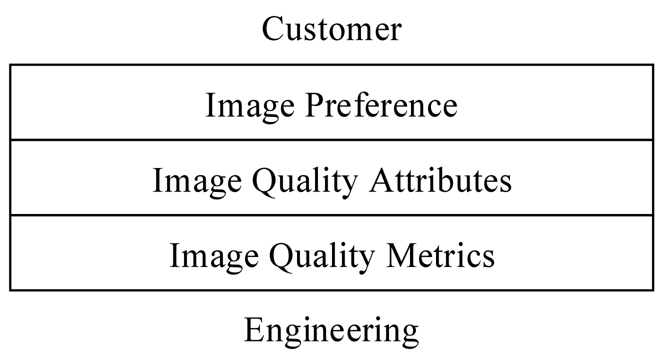

Figure 1. Hierarchy of Three Complementary Image Quality Assessment Systems

Dalal et al[1] define a set of high-level image quality descriptors as one which is suitable for concisely describing the image quality strengths and weaknesses of a given device or technology. The overall image quality can be described with relatively few such descriptors which are orthogonal to (or at least independent of) each other. The DAC (Document Appearance Characterization) system was created and introducted as a means to evaluate or compare the image quality characteristics of a printing system through the use of image quality attributes.

Image preference of a set of selected test images could have been used as an alternative method of evaluating overall image quality. However, not only are preference surveys expensive and time-consuming, they are also highly sensitive to the weakest link in the overall image quality. For these reasons image preference is often not a good way of measuring overall image quality. This is particularly true of immature technologies, where some aspects of image quality still need to be worked on, for example where color management is not yet in place.

Image quality metrics do not have these drawbacks of image preference. They are objective, quantitative, mostly independent of color rendering issues, and can isolate specific areas of image quality for independent evaluation. There are many image quality metrics currently in use, and meaningful engineering conclusions may be drawn from their results. But if one wished to compare the overall image quality of two output devices using image quality metrics, one would need to compare literally thousands of numerical values. This is difficult to do in a meaningful manner. The significance and impact that each of these numbers has on the overall system image quality would be quite difficult to determine.

Image quality attributes are designed to combine some of the best features of both image preference and image quality metrics. Like metrics, they are objective and quantitative. But the strengths and weaknesses of a given output device can be meaningfully represented with a relatively small number of attributes. The DAC system[1] outlines a set of image quality attributes, a standard set of digital images, and a detailed procedure for evaluation of a printer system on the basis of these attributes. The attributes were carefully chosen to be appearance-based and independent of technology. A drawback is that evaluation of most image quality attributes currently must be performed visually, by an experienced panel, although there are plans to eventually relate them directly to image quality metrics.

A long term goal is to be able to predict image preference from measurements on analytical images, using statistical correlations. This is not practicable to do directly with metrics because of the large number of metrics which would need to be taken into account. Image quality attributes provide a reasonable bridge between image preference and image quality metrics, since image quality attributes can be reasonably related to both of them.

This paper presents preliminary results on the relationship between selected image quality attributes of an image and their effect on customer preference.

Evaluation of Image Preference

In order to address the need to measure and understand customer preference for color images, a quantitative preference measurement system was developed. Nine customer-like images were used which included several pictorials, business graphics, text, line art and composites of pictorials, graphics, text and lines. While the images were carefully selected to resemble customer-like documents, they were also chosen with the intent of stressing, in combination, all image quality attributes.

In an effort to simplify the analysis of the preference results for the purposes of this paper, we will focus the analysis on results obtained for just one of the nine image sets. The image content was a pictorial scene of a church surrounded by sky, a grassy field and many red and yellow flowers, thus containing many outdoor memory colors. Of the ten image quality attributes defined by Dalal et al,[1] five were considered potentially important to this image:

- Color Rendition: blue sky, green grass, gray road, red and yellow flowers.
- Micro Uniformity: noise, halftone screen, or other small-scale artifacts are very noticeable in the blue sky and clouds.
- Macro Uniformity: Bands or other large-scale non-uniformities are very noticeable in the blue sky.
- Effective Resolution: Details in trees and structural artifacts of the church building stress the Effective Resolution attribute.
- Gloss Uniformity: The attribute is moderately stressed by the variation in coverage across the image.

Five printing technologies were used to make print samples which cover a range of different image "looks": lithography, electrophotography, ink jet, silver halide and dye diffusion. The images were printed on a total of 21 printing systems using a variety of modes, setup conditions and substrates. The final outcome of all print runs resulted in a total of 48 different color prints within the image set. These images ranged from very high overall image quality (*e.g.*, lithographic images on coated paper, or silver halide prints) to quite low image quality (*e.g,* inkjet on plain paper in draft mode). The image sets were then presented to a large number of observers for preference evaluation.

Observers with both technical and non-technical backgrounds participated in the survey to rank the printed color images according to preference within each of the nine sets. Preference scores were based on a scale of 0 to 100; 0 corresponding to the lowest possible preference for a particular print sample, and 100 corresponding to the highest. A total of 61 observers performed the survey. Average preference values were calculated for each of the 48 images in the set.

Visual assessments of the above mentioned five image quality attributes were completed on each of the print samples in the image set by a qualified expert panel experienced in the field of image quality. Close attention was paid to assessing each attribute individually and care was taken to not intermix the visual effect of each attribute. Each image quality attribute was rated on an arbitrary scale of 0 to 4. A rating of 4 corresponded to a very high quality rating, while a score of 0 corresponded to the lowest level of quality with respect to the attribute.

Results

A linear regression analysis was completed on the effect of the five image quality attributes on overall average image preference. Results show varied degrees of correlation between each of the five image quality attributes and the average preference scores. The following table shows the results from the regression analysis. Each of the image quality attributes are listed with corresponding fitted coefficients (slopes) in order of decreasing magnitude.

Table 1. Fitted coefficients from linear regression of image quality attributes on average image preference

Image Quality Attribute	Coefficient
Color Rendition (CR)	8.8
Micro Uniformity (Micro)	5.5
Effective Resolution (ER)	4.4
Macro Uniformity (Macro)	3.5
Gloss Uniformity (GU)	1.9

The coefficients from the regression analysis may be used to predict the average preference score of a given image if the visually assessed level of each of the five image quality attributes are known:

$$\text{Average Preference} = 8.8 \cdot CR + 5.5 \cdot Micro + 4.4 \cdot ER + 3.5 \cdot Macro + 1.9 \cdot GU \quad (1)$$

Results in Table 1 imply that the average preference for this image set is most highly associated with the level of color rendition. The remaining attributes were correlated in the following order of decreasing importance: micro uniformity, effective resolution, macro uniformity and gloss uniformity. Gloss uniformity appeared to have very little correlation with the average image preference for this particular set of samples.

The relationship between the two highest correlated image quality attributes, color rendition and micro uniformity, and the average preference scores were further examined. In order to examine *only* the effects of these two attributes, the average preference scores were scaled to be approximately independent of the remaining attributes by rearrangement of the above relationship (equation 1). This allowed for the average preference to be tracked with respect to the two most correlated image quality attributes.

A two dimensional non-linear Logistic Dose Response Function was fitted to the color rendition level, micro uniformity level and scaled average preference results. It is a six parameter equation describing the scaled average preference *(z)* with respect to the color rendition level *(x)* and the micro uniformity level *(y)*.

$$z = \frac{a}{1 + \left(\frac{x}{b}\right)^{-c}} + \frac{d}{1 + \left(\frac{y}{e}\right)^{-f}} \quad (2)$$

This equation is a sum of two Logistic Dose Response functions, one for each axis. The Logistic Dose Response is an S-shaped transition function. The parameters *a* and *d* represent the transition height of the curve for the color rendition and micro uniformity axes respectively, parameters *b* and *e* equal the transition height of the curve for each of the attribute axes and parameters *c* and *f* determine the transition width of the function for each of the attribute axes. The values of these parameters fitted to the data are listed in Table 2. Curve fitting was done using TableCurve 3D software. Figure 2 is a plot of the results.

Table2. Parameters used to define preference with respect to Color Rendition and Microuniformity Levels

Fitted Constant	Value
a	66.4
b	3.6
c	0.6
d	14.8
e	2.4
f	9.6

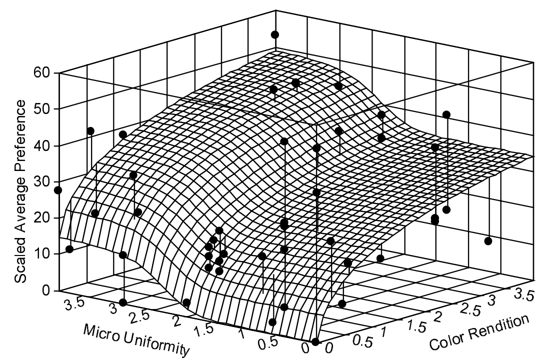

Figure 2. Relationship between Color Rendition, Micro Uniformity and Scaled Preference

Conclusions

Average preference values were examined for a wide range of image quality samples within a set containing a pictorial scene of a church surrounded by blue sky, green grass and several colored flowers, in an effort to rank the importance of selected image quality attributes. Color rendition was found to have the strongest correlation with average preference scores for the color image sample set analyzed. Micro uniformity, macro uniformity and effective resolution demonstrated weaker yet significant correlations to the average preference scores. Gloss uniformity showed little correlation with average preference values for this image set.

Figure 2 shows that both the color rendition and micro uniformity level significantly affect the average preference scores. Improvements in each of the attributes tend to increase overall preference of the color print within a certain range if the effect of all other image quality attributes is ignored. Although improvements in the two attributes track well with increased values of preference, there is a certain plateau that is reached for both beyond which additional improvements fail to provide any greater preference for the print samples.

Clearly this study is only a first step in the effort to bridge the gap between subjective image quality preference and analytical metric analysis of images through the use of selected image quality attributes. The results of this study represent a way to determine the correlations and relationships between image quality attributes and overall image preference. These same attributes may be directly related to analytical metrics and eventually models could be set in place to predict image preference from instrumentally measured metrics, without dependence on human observers.

References

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-98 Conference*, 169-173, Portland (1998).