Attributes of a New Image Quality Model for Color Prints

Marius Pedersen*, Nicolas Bonnier**, Jon Y. Hardeberg* and Fritz Albregtsen.***
**** Gjøvik University College (Norway). ** Océ Print Logic Technologies S.A. (France). *** University of Oslo (Norway).***

Abstract

Evaluation of perceived image quality in color prints is a complex task, due to its subjectivity and dimensionality. Perceived image quality is influenced by a number of quality attributes. Evaluation of all attributes influence, and their influence on other attributes, is is a difficult and complex task. Because of this the most important attributes should be identified in order to achieve a more effective and manageable evaluation of image quality. In this paper we identify and categorize existing quality attributes, and we propose a refined selection of meaningful quality attributes for the evaluation of color prints. Experimental data verify the proposed set of quality attributes.

Introduction

Technology advancements are rapid in the printing industry. One of the goals for the printing industry, and a motivation for the advancements, is to produce high quality prints fast and economically. With this goal and motivation, new and more refined ways to deal with the limitations of printing systems are proposed continuously to achieve high quality prints. In order to show if technology advancements increase the quality of a print the industry performs quality assessment.

There are basically two ways to judge Image Quality (IQ): subjectively or objectively. Subjective evaluation is carried out by observers, and is therefore influenced by the Human Visual System (HVS). There are several possible ways to carry out objective evaluation. One typical way is to use measurement devices gathering numerical values. Another way is to use algorithms, commonly known as IQ metrics, in an attempt to quantify IQ. IQ metrics are usually developed to take into account properties of the HVS, and thus with the goal of being well correlated with subjective evaluations.

Both objective and subjective evaluation of IQ are dependent on a number of Quality Attributes (QAs), which are terms of perception [1], such as colorfulness, contrast, and sharpness. These QAs influence the overall IQ differently, and knowledge about their importance can be used to achieve an optimal reproduction of an image [2]. Many researchers have investigated and acknowledged the importance of different QAs [2–9], but there is, so far, no overall agreement on which QAs are most important. In this paper we focus on QAs for the evaluation of color prints.

For many years, a goal in IQ evaluation has been to develop objective measures that are correlated with subjective quality. The advancements in this field have been driven by the desire to reduce the dependence on human observers and minimize both time and resources needed to quantify IQ. To achieve this, the QAs used in subjective evaluation of IQ need to be identified and their importance need to be assessed. These objective measures, such as IQ metrics, can be used to help observers detect IQ issues, identify loss of IQ in printing workflows, optimize a process, or compare the IQ of different printing systems.

IQ models have been created to establish a link between subjective and objective IQ. These models are theories of perception that enable the prediction of IQ [10]. They are intended to describe the overall IQ and to help researchers in the evaluation of IQ. IQ models are composed of QAs, and they show how QAs relate to each other and their influence on overall IQ. The goal for IQ models is to evaluate all QAs and their relationships, yet, this is a very difficult and complex task. In order to reduce the complexity most IQ models use a subset of QAs composed of the most important QAs. A subset of QAs can be defined based on technological issues, or by asking observers, or by a combination of the two. By defining a subset of QAs, strengths and weaknesses of a given system can be meaningfully represented with a relatively small number of QAs [11]. Several IQ models have been proposed [3, 8, 12, 13], but the search for better and improved IQ models is still ongoing.

The goal of this paper is to identify and categorize existing QAs to propose a refined selection of meaningful QAs for the evaluation of color prints. These QAs can further be used to create a link between subjective and objective IQ in an IQ model. Identification and categorization of QAs can be used as assistance in the evaluation of color prints, as well as to improve or develop new evaluation methods.

This paper is organized as follows: First a survey of QAs and IQ models, then a discussion about the selection of important QAs for the evaluation of color prints. Next, an experiment investigating a color workflow is carried out, where QAs are evaluated by observers. Finally we conclude and suggest directions for further research in this field.

Image Quality Attributes - State of the Art

Norberg et al. [5] evaluated overall quality, as well as eight QAs, in a comparison of digital and traditional print technologies. In a study by Lindberg [4] overall IQ and 12 different QAs were used for the evaluation of color prints. Based on the evaluation performed by observers these 12 QAs were reduced to two orthogonal dimensions; print mottle and color gamut. In addition several researchers have investigated the importance of QAs like sharpness [6], contrast [14], artifacts (for example noise [7, 8] and banding [15]), naturalness [2], and color [3, 9, 14, 16].

Research on the combined influence of QAs has been carried out as well. Sawyer [7] investigated the influence of sharpness and graininess on perceived IQ, and their combined influence. Some years later Bartleson [8] investigated the combined influence of sharpness and graininess on color prints. Both Sawyer and Bartleson showed results where the worst QA tended to determine the quality, and a change in the least worse QA would not increase quality. Natale-Hoffman et al. [11] investigated the relationship between color rendition and micro uniformity on preference. This was considered by the authors as a step towards predicting preference, without dependence on human observers.

Identification of QAs has also been recognized as important for IQ metrics. Morovic and Sun [9] based an IQ metric on perceptual QAs, where the QAs were determined based on answers from observers. Lightness, hue, chromaticity, details, and contrast were found to be important. Only the three first, being the most important according to the authors, were incorporated in the metric. Later Wang and Shang [17] showed that defined QAs were beneficial for training IQ metrics.

Image Quality Models - a Brief Survey

A framework for IQ models was proposed by Bartleson [8] in 1982. His approach was divided into three parts:

1. identification of important QAs,
2. determination of relationships between scale values and objective measures,
3. combination of QA scale values to predict overall IQ.

Bartleson used this framework to investigate the combined influence of sharpness and graininess on the quality of color prints. This framework has the advantage of representing strengths and weaknesses of a given system by a relatively small number of QAs. Because of this, and its perceptual considerations, the framework has been adopted by several researchers [3, 10, 12]. We also adopt this framework, where we discuss the first part, identification of important QAs.

Dalal et al. [3] followed the framework in the creation of the Document Appearance Characterization system, which is a two-sided system composed of QAs: one part for the printer and one for materials and stability. For most QAs in the system, evaluation is performed by experts. The basic IQ is given by 10 QAs for both sides. These describe different aspects of the system, such as color rendition, uniformity, and stability. The system has several advantages. It uses high level descriptors, which cover a wide range of IQ issues. The printer is also separated from materials and stability, allowing for separate analysis. Being technology independent is an advantage as well. But this system has some drawbacks as well, since the evaluation is mostly carried out by experts the results will be influenced by the subjectivity of the expert. It might be unsuitable for non-experts due to its complexity. The model is problematic to use with IQ metrics, since it is difficult to extract information with IQ metrics for all QAs.

Keelan [12] also followed the same framework. First important QAs are identified, then the relationship between a subjective scale (based on just noticeable differences) and an objective metric is found. In the case where multiple QAs influence the quality of an image, the influence of each QA to overall IQ is found. Keelan adopted a multivariate formalism as a tool to combine the influence of each QA in order to obtain a value for overall IQ.

Engeldrum [10] focused on the building of an IQ model, and partially adopted Bartleson's framework. He proposed the IQ circle, which is based on four elements; customer quality preference, technology variables, physical image parameters, and customer perceptions. The last element, customer perceptions, contains the perceptual QAs (or "nesses"), being the topic for this study. The quality circle shows the relationship between objective and subjective quality, but does not include which QAs are important, nor how they should be quantified. Engeldrum also states that observers most likely will not be able to perceive more than five QAs simultaneously.

Many other IQ models have been proposed as well. Some of these are IQ metrics, which accept an image or a set of images as input and the output is one or several values representing IQ. These models are most often constructed to quantify either overall IQ or the quality of specific QAs. They usually incorporate several stages of processing, where characteristics of different QAs are taken into account. For an overview of different models we refer to Pedersen and Hardeberg [18].

Investigation and Selection of Quality Attributes

As a first step towards an IQ model important QAs must be identified. In order to do this we have performed a survey of the existing literature. Numerous QAs have been considered as important and evaluated by researchers to quantify IQ. These QAs include for example lightness [9, 12], sharpness [4–6, 19, 20], contrast [4, 5, 9, 14], noise/graininess [7, 8, 19, 21, 22], banding [15], details [5, 9, 14, 16, 19], naturalness [2] , color [3, 16], hue [9], chroma [9], saturation [14], color rendition [3], process color gamut [3], artifacts [14], mottle [4, 20], gloss [4, 5], color reproduction [22], tone reproduction [22], color shift [5, 20], ordered noise [5], patchiness [5], line quality [3, 23], text quality [3], gamut size [24], adjacency [3], effective resolution [3], effective tone levels [3], gloss uniformity [3], skin color [16], paper roughness [20], paper flatness [3], paper whiteness [20], perceived gray value [19], structure changes [19], micro uniformity [3], macro uniformity [3], structure properties [19], color gamut [20], correctness of hue [25], correctness of lightness [25], contouring [26], colorfulness proportional to the original [25], edge sharpness [23], and edge raggedness [23].

When reducing these QAs found in the literature, there are several important issues to consider, such as the intention of how QAs should be used, and their origin. A long term goal of this research is to create a link between subjective and objective IQ of color prints. With this intention the QAs should be based on perception and account for technological printing issues. The QAs should be general enough to be evaluated by observers, and in order not to exclude novice observers the QAs should be somewhat straightforward to evaluate. In addition, the QAs should be suitable for IQ metrics, being the intended objective method. The existing sets of QAs and models do not fulfill all of these requirements, and therefore a new set of QAs is needed.

Many of the QAs listed above are similar and have common denominators, which enables them to be grouped within more general QAs in order to reduce the dimensionality and create a more manageable evaluation of IQ. There is usually a compromise between generality and accuracy when it comes to dimensionality. A small set of general QAs results in lower accuracy, but low complexity, while a higher dimensionality offers accuracy, but higher complexity. We have linked most of the above QAs to six different dimensions, considered as important for the evaluation of IQ. This results in a reasonable compromise between accuracy and complexity, as well as being close to the statement by Engeldrum [10] that observers will not perceive more than five QAs simultaneously. We have reduced the QAs found in the literature to the following six:

- **Color** contains aspects related to color, such as hue, saturation, and color rendition, except lightness.

- **Lightness** is considered so perceptually important that it is beneficial to separate it from the color QA [12]. Lightness will range from ”light” to ”dark” [1].
- **Contrast** can be described as the perceived magnitude of visually meaningful differences, global and local, in lightness and chromaticity within the image.
- **Sharpness** is related to the clarity of details [6] and definition of edges [27, 28].
- In color printing some **artifacts** can be perceived in the image. These artifacts, like noise and banding, contribute to degrading the quality of an image if detectable [29].
- The **physical** QA contains all physical parameters that affect quality, such as paper properties and gloss.

The six dimensions are general high-level descriptors, either artifactural, i.e., those which degrade the quality if detectable [29], or preferential, i.e., those which are always visible in an image and have preferred positions [29]. Most of the QAs found in the literature can be linked with these six QAs. As an example, lightness can be linked to tone reproduction [22], perceived gray value [19], correctness of lightness [25], and lightness [9, 12]. Other QAs are more difficult to link with just one QA, such as color shift [5, 20], gamut size [24], process color gamut [3], and color rendition [3]. All of these have ties to lightness, but also to the color QA. QAs, like gloss level and paper flatness, can influence both lightness and color, but cannot be accounted for with the first five QAs. Therefore, a grouping of these physical QAs is needed.

In order to create a simple and intuitive illustration of the QAs and their influence on overall IQ we have turned to Venn diagrams. Venn diagrams may be used to show possible logical relations between a set of attributes. However, it is not possible to create a simple Venn diagram with a six fold symmetry [30]. Therefore we illustrate the QAs using only five folds, leaving the physical QA out. This does not mean that the physical QA is less important than the others.

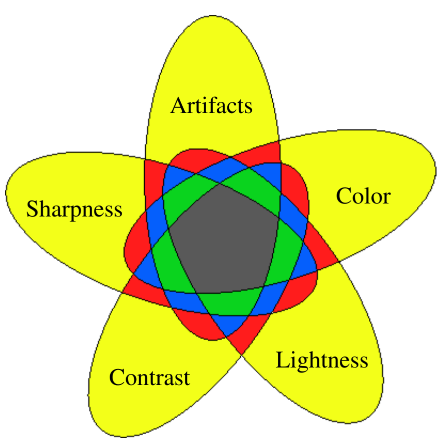

Figure 1. Simple Venn ellipse diagram with five folds used for an abstract illustration of the QAs. Five different QAs and the interaction between them are shown. Overall IQ can be influenced by one (yellow), two (red), three (blue), four (green), or five (gray) of the QAs.

The Venn diagram of Figure 1 illustrates how the overall IQ is influenced by one, two, three, four, or five of the QAs. Many of the QAs are interdependent [31], making IQ a multidimensional issue [32], in this case five dimensions. These QAs can influence overall IQ differently, and therefore the ellipses may not have equal sizes or the same positions in all situations. Influence of different QAs on IQ and how they can be used in an IQ model for prediction of perceived quality will be dealt with in future work.

Each of these QAs can be divided into sub-QAs for adaptation to specific issues. The artifact QA can, for example, be divided into three sub-QAs; noise, contouring, and banding. Separate analysis of these can be advantageous since it allows for specific analysis either by experts or IQ metrics. Some sub-QAs can be placed under several main QAs, such as uniformity. This sub-QA can be placed under color, but also under artifacts since lack of uniformity can be thought of as an artifact. The placement of these sub-QAs must be done where it is most appropriate. Furthermore, all QAs might not be used in the evaluation of IQ, in this case QAs can be excluded.

In the following section we will take a closer look at the three first QAs; color, lightness, and contrast as seen on Figure 2. A psychophysical experiment was carried out to investigate these QAs. QAs used by observers to describe quality changes were recorded and analyzed.

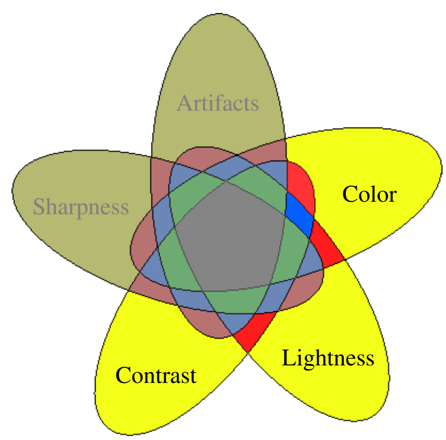

Figure 2. A psychophysical experiment was carried out to investigate QAs in a color workflow, where a subset of the QAs is considered to affect overall IQ; color, lightness, and contrast.

Investigation of Quality Attributes in a Color Workflow

In order to investigate quality issues in a color workflow, and to confirm the proposed QAs in the previous section, an experiment was carried out.

The images were reproduced using the ICC perceptual rendering intent, which adjusts color appearance to achieve the most attractive result on a medium different from the original [33]. In the evaluation of this color workflow, observers evaluate different QAs, and the influence these QAs have on overall IQ affect the observer’s judgment of IQ. For some QAs the quality decreases, for other QAs the quality might increase, while some QAs neither increase nor decrease quality. Investigating only the QAs that influence quality might not give a correct representation of which QAs are important. All QAs being evaluated by observers should be investigated, because of this the instructions given to the observers are crucial for correct results.

Experimental Setup

Test Images

In order for the observers to use a sufficiently large set of QAs, a broad range of images should be used in order to reveal different quality issues [34]. To achieve this we followed the recommendations of Field [35] and CIE [36], where the images were chosen based on the following criteria:

- Low, medium, and high levels of **lightness**,
- Low, medium, and high levels of **saturation**,
- **Hue** primaries,
- Low, medium, and high **contrast** ,
- **Larger areas of the same color**,
- **Fine details**,
- **Memory colors** as skin tones, grass, and sky blue,
- Color **transitions**,
- **Neutral gray**.

Most of the images were pictorial with a wide range of scenes like landscapes, portraits, and personal items (such as jewelry, books and clothes). This helps to characterize the impacts for QAs [12], and ensures that the observers examine a wide variety of QAs. In addition to the pictorial images a set of test charts were included, since these are content-free and have a selection of "interest area" colors suitable for evaluation of different aspects of IQ [35].

A total of 56 images, as seen in Figure 3, were used in this experiment. 7 images from ISO [37], 2 images from CIE [36], 3 test charts, and 44 other images captured by the authors. These 44 images were RAW format, which were converted to sRGB using Camera Raw 5.0 in Adobe PhotoShop CS4 with a resolution of 150 dpi and a 16-bits encoding. The images were printed at a resolution of 150 pixels per inch resulting in a reproduction of approximately 8 by 10 inches.

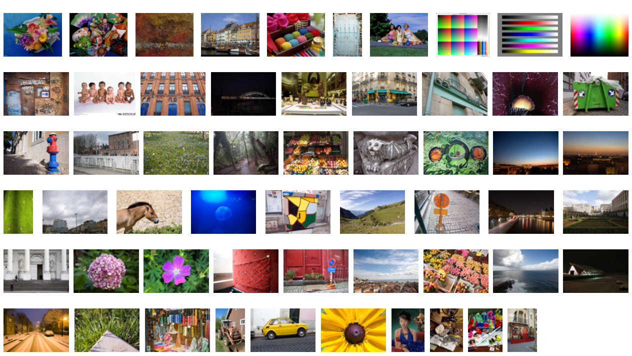

Figure 3. Images used in the experiment.

Color Workflow

The first step in the color workflow was to re-render the set of images from sRGB to the perceptual reference medium gamut [38] using the sRGB v4 perceptual transform. The output profile was generated using the TC3.5 CMYK test target, measured using a GretagMacbeth Eye-One Pro spectrophotometer and ProfileMaker Pro 5.0.8. Then, as a second step, linear lightness scaling compensation in the CIE XYZ color space plus the hue preserving minimum ΔE clipping gamut mapping algorithm [36] was applied to the image to re-render from the perceptual reference medium gamut to the gamut of the printing system. The linear lightness scaling were made between the black point CIELAB coordinates of each images to the black point CIELAB coordinates of the printing system contained in the output profile. The third and last step was to convert the color data from the profile connection space values to CMYK values of the printing system by a relative colorimetric transform. The images were then printed with the Océ ColorWave 600 wide format CMYK printer on Océ Red Label paper.

Viewing Conditions

The observers were presented with a reference image on an EIZO ColorEdge CG224 (some observers on an EIZO ColorEdge CG221 since the experiment was carried out in two locations) display at a color temperature of 6500 Kelvins and luminance level of 80 cd/m^2. This set was rendered for sRGB display, and therefore a monitor capable of displaying the sRGB gamut was the most adapted reproduction device for this set of images. The printed images were presented randomly in a controlled viewing room at a color temperature of 5200 Kelvins, an illuminance level of 450 $\pm$75 lux and a color rendering index of 96. The observers viewed the reference image and the printed image simultaneously from a distance of approximately 60 cm. The experiment followed the CIE guidelines [36] as closely as possible.

Instructions Given to the Observers

The instructions given to the observers focused on the overall quality rating of the reproduction and which QAs the observer used in the evaluation. Instructions specified that all QAs used in the evaluation should be stated, even if they did not influence IQ. The two following questions were given to the observers:

- *Is the printed image a pleasing reproduction?*
- *According to you, which quality attributes influence the quality of the reproduction?*

For the first question a scale from 1 (most pleasing) to 7 (least pleasing) was given to the observer. QAs used by the observers were noted on a form, where QAs decreasing, increasing or not influencing IQ was marked with different symbols.

Experimental Results

15 observers participated in the experiment, with mixed expertise and of both genders. Five observers rated the whole data set, and 10 observers rated parts of the data set. A total of 452 evaluations were carried out by the observers, where a scale value was given to each image and the observers described the QAs they used in their evaluation. The evaluation was carried out in several sessions to prevent observer fatigue.

The average pleasantness of the images, based on the seven step scale, has been found to be between fairly and very pleasing. Analysis of the ratings given by the observers indicates that images with a majority of shadow areas and images with color transitions (both test charts and natural images) are rated as least pleasing. In some images with shadow areas details are lost due to the difference between the input and output gamut (output gamut volume is 56% of the input gamut volume). In images with color transitions color breaks occur because of the gamut clipping algorithm. The images rated to be most pleasant are according to

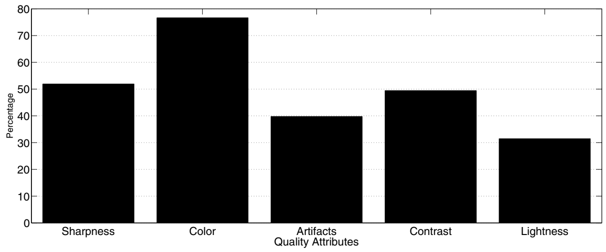

Figure 4. *Frequency of QAs used by the observers in the experiment. All QAs used by the observers have been fitted to one of the QA proposed in the previous section.*

the observers equally or more colorful than the original and have equal or better contrast than the original.

The observers used more than 50 different QAs in their evaluation, with an average of 10 different QAs for each observer. An average of 2.95 QAs have been used for each image, with a maximum of eight QAs and a minimum of one QA. This indicates that a high number of QAs are not considered by the observers in the evaluation of IQ. Many of the QAs used in the experiment overlap, such as lightness, brightness, luminance, and darkness. All of these are connected to the lightness of the image, and have been grouped within the QA lightness. Similar grouping is done for other QAs to fit within the proposed QAs (Figure 1).

Color is the most frequently used QA by the observers, as seen in Figure 4 it has been used to describe the IQ of more than 70 percent of the images. This is not surprising since a color workflow was investigated and the color QA is fairly large, containing sub-QAs as hue, saturation, and colorfulness. These three sub-QA are commonly used, and often used together. The second most used QA, sharpness, mainly contains two sub-QAs; edges and details. Details both in highlights and shadows have been frequently used by the observers. This is not surprising since a gamut clipping algorithm was used. Some observers also commented that loss of contrast lead to a loss of perceived sharpness, since edges and details were less prominent. It is interesting to notice the frequent use of contrast ,50 percent, in evaluation of color prints. The term artifacts or sub-QAs of this were used in approximately 40 percent of the images, mostly because sub-QAs as noise, contouring, and banding could be perceived in the images. Lightness is considered in more than 30 percent of the images. Even though this is the least frequently used QA it should be noted that some observers used the more general term color rather than separating lightness and chromaticity.

Analysis of the relations between QAs has also been carried out, using cross-tabulation and chi-square tests. Analysis for these results indicates a dependence between color and lightness, but also between lightness and sharpness, contrast and artifacts, and artifacts and lightness. This indicates that the use of these QAs occur simultaneously, but not how they affect IQ. This issue will be dealt with in future work.

In the experiment observers distinguished between QAs that decreased IQ, did not influence IQ, or increased IQ. Observers have marked more QAs to decrease IQ than the two other groups, and more QAs to increase IQ than QAs that do not influence IQ. For the artifacts attribute, some observers stated that lack of artifacts increased IQ. Observers have not considered artifacts where it does not influence IQ, indicating that artifacts only are considered when they are perceivable, or not present in areas where observers expect to find artifacts. In the sharpness QA lack of details in several of the reproductions contribute to decrease IQ.

In this experiment the physical QA was not considered, and therefore not a part of this analysis. There are also some QA used by the observers that are difficult to link with one of the six QA, for example naturalness and warmness. These can be linked with changes in other QAs, such as color and lightness.

Conclusion and Future Directions

In this paper we have identified and categorized existing QAs, and proposed a refined set of selection of the most meaningful QAs for the evaluation color prints. The number of QA considered to be important in IQ evaluation has been reduced to a set of six QAs; color, lightness, sharpness, contrast, physical, and artifacts. These QAs present a good starting point to describe overall IQ, and they can be considered as a step towards achieving a link between objective and subjective IQ. A psychophysical experiment has been carried out to evaluate a color workflow, where QAs used by observers were recorded and analyzed. Results obtained from this experiment support the proposed set of QAs.

Future work includes investigation of interactions between different QAs, locally and globally, and their influence on overall IQ. Incorporation of IQ metrics should be considered in order to achieve objective evaluation of IQ.

References

[1] G. Wyszecki and W.S. Styles. *Color Science, Concepts and Methods, Quantitative Data and Formulae*. Wiley Interscience, Derby, UK, 2nd edition, Aug 2000.

[2] E. A. Fedorovskaya, F.J.J. Blommaert, and H. de Ridder. Perceptual quality of color images of natural scenes transformed in CIELUV color space. In *Color Imaging Conference*, pages 37–40. IS&T/SID, 1993.

[3] E. N. Dalal, D. R. Rasmussen, F. Nakaya, P. A. Crean, and M. Sato. Evaluating the overall image quality of hardcopy output. In *Image Processing, Image Quality, Image Capture, Systems Conference*, pages 169–173, Portland, OR, May 1998. IS&T.

[4] S. Lindberg. *Perceptual determinants of print quality*. PhD thesis, Stockholm University, 2004.

[5] O. Norberg, P. Westin, S. Lindberg, M. Klaman, and L. Eidenvall. A comparison of print quality between digital, offset and flexographic printing presses performed on different paper qualities. In *International Conference on Digital Production Printing and Industrial Applications*, pages 380–385. IS&Ts, May 2001.

[6] S. Bouzit and L. MacDonald. Colour difference metrics and image sharpness. In *8th Color Imaging Conference*, pages 262–267, Scottsdale, AZ, 2000. IS&T/SID. ISBN / ISSN: 0-89208-231-3.

[7] J.F. Sawyer. Effect of graininess and sharpness on perceived print quality. In *Photographic Image Quality Symposium*, pages 222–231. Royal Photographic Society, Sep 1980.

[8] C.J. Bartleson. The combined influence of sharpness and graininess of the quality of color prints. *J. Photogr. Sci.*, 30: 33–38, 1982.

[9] J. Morovic and P. Sun. Visual differences in colour reproduction and their colorimetric correlates. In *10th Color*

Imaging Conference, pages 292–297, Scottsdale, AZ, 2002. IS&T/SID.

[10] P. G. Engeldrum. Image quality modeling: Where are we? In *Image Processing, Image Quality, Image Capture, Systems Conference*, pages 251–255. IS&T, 1999.

[11] K. Natale-Hoffman, E. Dalal, R. Rasmussen, and M. Sato. Effect of selected image quality on overall preference. In *Image Processing, Image Quality, Image Capture, Systems Conference*, pages 266–269, Savannah, GA, Apr 1999.

[12] B. W. Keelan. *Handbook of Image Quality: Characterization and Prediction*. Marcel Dekker, 2002.

[13] X. Zhang and B.A. Wandell. A spatial extension of CIELAB for digital color image reproduction. In *Soc. Inform. Display 96 Digest*, pages 731–734, San Diego, 1996.

[14] N. Bonnier, F. Schmitt, H. Brettel, and S. Berche. Evaluation of spatial gamut mapping algorithms. In *14th Color Imaging Conference*, volume 14, pages 56–61. IS&T/SID, Nov 2006.

[15] Y. Bang, Z. Pizlo, and J. P. Allebach. Discrimination based banding assessment. In *IS&T's NIP19: International Conference on Digital Printing Technologies*, pages 745–750, New Orleans, LA, Sep 2003.

[16] J. Y. Hardeberg, E. Bando, and M. Pedersen. Evaluating colour image difference metrics for gamut-mapped images. *Coloration Technology*, 124(4):243–253, Aug 2008.

[17] Z. Wang and X. Shang. Spatial pooling strategies for perceptial image quality assessment. In *IEEE International Conference on Image Processing*, pages 2945–2948, Atlanta, Oct 2006.

[18] M. Pedersen and J.Y. Hardeberg. Survey of full-reference image quality metrics. Høgskolen i Gjøviks rapportserie 5, The Norwegian Color Research Laboratory (Gjøvik University College), Jun 2009. ISSN: 1890-520X.

[19] F. Nilsson and B. Kruse. Objective quality measures of halftones images. In *IS&T's NIP 13: International Conference on Digital Printing Conference*, pages 353–357, Mar 1997.

[20] M. Klaman. *Aspects on Colour Rendering, Colour Prediction and Colour Control in Printed Media*. PhD thesis, Stockholm University, 2002.

[21] P. Oittinen, I. Huttu, and A. Mensonen. End user preferences for quality of on-demand printed products. In *International Conference on Digital Printing Technologies*, pages 147 – 152, 1996.

[22] K. Miyata, N. Tsumura, H. Haneishi, and Y. Miyake. Subjective image quality for multi-level error diffusion and its objective evaluation method. *Journal of Imaging Science and Technology*, 43:170–177, 1999.

[23] M. Tse and A. H. Klein. Automated print quality analysis in inkjet printing: Case study using commercially available media. In *Recent Progress in Ink Jet Technologies II*, pages 515–519, 1999.

[24] M. Andersson and O. Norberg. Color gamut - is size the only thing that matters? In *2006 TAGA conference*, pages 273–291, Vancouver, Canada, Mar 2006.

[25] R. W. G. Hunt. *The reproduction of Color in Photography, Printing and Television*. Tolsworth: Fontain., 1987.

[26] P. Green. *Gamut mapping and appearance models in graphic arts colour management*. PhD thesis, University of Derby, 2003.

[27] J. Caviedes and F. Oberti. A new sharpness metric based on local kurtosis, edge and energy information. *Signal Processing: Image Communication*, 19:147–161, 2004.

[28] E. A. Fedorovskaya. Image quality as a problem of computational vision. In *Image Processing, Image Quality, Image Capture, Systems Conference*, pages 22–28, Rochester, NY, May 2003. IS&T.

[29] K. Topfer, B. Keelan, S. O'Dell, and R. Cookingham. Preference in image quality modelling. In *Image Processing, Image Quality, Image Capture, Systems Conference*, pages 60–64, Apr, Portland, OR 2002. IS&T.

[30] B. Grunbaum. The search for symmetric venn diagrams. *Geombinatorics*, 8:104 – 109, 1999.

[31] T. N. Pappas and D. L. Neuhoff. Least-squares model-based halftoning. *IEEE Transactions on Image Processing*, 8(8): 1102–1116, Aug 1999.

[32] F. Nilsson. Objective quality measures for halftoned images. *Journal of Optical Society of America A*, 16:2151–2162, 1999.

[33] ICC. White paper number 2: Perceptual rendering intent use case issues. 16/03/09: www.color.org, Jan 2005.

[34] J. Holm, J. Tastl, and T. Johnson. Definition & use of the ISO 12640-3 reference color gamut. In *14th Color Imaging Conference*, pages 62–68, Scottsdale, AZ, 2006. IS&T/SID.

[35] G. G. Field. Test image design guidelines for color quality evaluation. In *The Seventh Color Imaging Conference: Color Science, Systems and Applications*, pages 194–196, Scottsdale, AZ, Nov 1999. IS&T.

[36] CIE. Guidelines for the evaluation of gamut mapping algorithms. Technical Report ISBN: 3-901-906-26-6, CIE TC8-03, 156:2004.

[37] ISO. ISO 12640-2: Graphic technology - prepress digital data exchange - part 2: XYZ/sRGB encoded standard colour image data (XYZ/SCID), 2004.

[38] ISO. ISO 12640-3 graphic technology - prepress digital data exchange - part 3: CIELAB standard colour image data (CIELAB/SCID), 2007.

Acknowledgments

We would like to thank Nicolas Cardin for his assistance with the experiment, and to Alessandro Rizzi and Gabriele Simone for their enlightening discussions regarding contrast.

The author hereof has been enabled by Océ-Technologies B.V. to perform research activities which underlies this document. This document has been written in a personal capacity. Océ-Technologies B.V. disclaims any liability for the correctness of the data, considerations and conclusions contained in this document.

Author Biography

Marius Pedersen *received his BsC in Computer Engineering in 2006, and MiT in Media Technology in 2007, both from Gjøvik University College, Norway. He is currently pursuing a PhD in Color Imaging, under the supervision of Pr. Hardeberg and Pr. Albregtsen, sponsored by Océ. He is also a member of the Norwegian Color Research Laboratory at Gjøvik University College. His work is centered on image quality metrics for color prints.*