Objective Image Quality Assessment of Color Prints

Marius Pedersen

Gjøvik University College, The Norwegian Color Research Laboratory, Gjøvik, Norway
Océ Print Logic Technologies S.A., Créteil, France

ABSTRACT: Measuring the perceived quality of printed images are important to assess the performance of printers and to evaluate technology advancements. Image quality metrics have been proposed to objectively assess the quality of images, and new metrics are continuously proposed. However, applying these metrics to printed images are not straightforward, since they require a digital input. We present a new framework for applying image quality metrics to printed images, including the transformation to a digital format, image registration, and the application of image quality metrics. Evaluation of image quality metrics in the new framework showed that some metrics provide better results for certain quality attributes, which lead to an investigation of the different quality attributes used in the evaluation of color prints. Based on a survey of the existing literature and a psychophysical experiment, we identify and categorize existing image quality attributes to propose a refined selection of meaningful ones for the evaluation of color prints.

1 INTRODUCTION: When we print a digital image we get a physical copy of it, and this copy differs from the digital original due to the limitations of the printing system. Furthermore, these differences can contribute to the loss of Image Quality (IQ). One way to assess loss of IQ is by using human observers. However, subjective evaluation is often time-consuming, inconvenient, resource demanding, and even expensive. In addition, observers are not objective, and their preference of IQ may change over time. Objective evaluation of IQ can be used to avoid subjectivity and decrease the other drawbacks of subjective evaluation. Many methods for objective IQ evaluation have been proposed, one of these is commonly referred to as IQ metrics. Their goal is to automatically predict IQ, usually by incorporating several stages of processing to account for specific issues. These metrics have been made for different purposes, such as to quantify a distortion or Quality Attribute (QA) (for example sharpness or contrast), optimize a process or to indicate problem areas. An extensive number of metrics have been proposed in the literature [1], and new metrics are introduced all the time. The goal of this paper is to propose a method to use IQ metrics to evaluate color prints, and to investigate the QAs used in the evaluation of IQ.

This paper is organized as follows. First we propose a framework for using IQ metrics with printed images. Then we discuss the use of QAs in the assessment of IQ, at last we conclude.

2 USING IMAGE QUALITY METRICS TO MEASURE THE QUALITY OF PRINTED IMAGES: Subjective assessment of print quality is rather straightforward, where a group of observers can be asked about the quality of the printed image. However, assessment of printed images by IQ metrics is not straightforward. The original image is of a digital format and the printed image is of an analog format, because of this the printed image must be digitalized before we can carry out IQ assessment with IQ metrics. In this section we discuss the transformation from a physical reproduction to a digital reproduction with the goal of proposing a framework for using IQ metrics to evaluate the quality of color prints.

A few frameworks have been proposed in the literature for using IQ metrics on printed images. These frameworks follow the same procedure; first the printed image is scanned, sometimes followed by a descreening procedure to remove halftoning patterns. Then image registration is performed to match the scanned image with the original. Finally, IQ metrics are applied.

The first framework was proposed by Zhang et al. [2]. To start with the image is scanned, and then three additional scans are performed, each with a different color filter. This results in enough information to transform the images correctly to

CIEXYZ. No information about the image registration was given, nor on the descreening procedure. The applied IQ metric was S-CIELAB [3], and the printed samples were color patches.

Another framework was proposed by Yanfang et al. [4]. Two control points are applied to the image before printing to help in the registration process, one point to the upper left corner and one to the upper center. The images were scanned at 300 dpi before registration, where the control points are used for matching the printed image with the original. Descreening was performed by the scanner at 230 lpi. No information was given regarding the scaling of the image. The applied IQ metric was S-CIELAB [3].

Recently, Eerola et al. [5] proposed a new framework using local features instead of control points. The printed reproduction is scanned, and then both the original and the reproduction go through a descreening procedure, which is performed using a Gaussian low-pass filter. Further, image registration is carried out, where local features are used with a Scale-Invariant Feature Transform (SIFT). A rand sample consensus principle (RANSAC) was used to find the best homography. Scaling was performed using bicubic interpolation. The images were scanned at 1250 dpi, and the applied IQ metric was LABMSE.

2.1 A framework based on control points: We modify and propose a framework similar to the framework by Yanfang et al. [4], which performs image registration based on control points. These control points are used in the image registration to perform different transformation procedures. First the image is padded with a white border and equipped with four control points, which are placed outside the corners. Then the image is printed before being scanned, and the profile of the scanner is assigned in order to achieve a correct description of the colors. Analysis of different scanning resolutions show that 600 dpi is a good trade-off between accuracy and computational time. The next step in the framework is to find the coordinates of the center of the control points in both the original image and the scanned image. This is done by a simple automatic routine based on the detection of squares.

Image registration must be carried since the scanned image can be affected by several geometric distortions, such as translation, scaling, and rotation. The coordinates of the control points, in both images, are used to create a transformation for the registration. There are several possible transformation types for doing the registration, experimental results indicate that a simple transformation correcting for translation, rotation, and scaling is the best. In addition, the interpolation method for scaling also have several possible methods, and the results show that ”bilinear” interpolation is the best. After the scanned image has been registered to match the original, a simple procedure is applied to remove the white padding and the control points. Finally, an IQ metric can be used to calculate the quality of the printed image. An overview of the framework is shown in Figure 1. In our modified framework we do not perform descreening, but we leave this to the IQ metrics in order to avoid a double filtering of the image. This requires the IQ metrics to perform some kind of simulation of the HVS, for example spatial filtering based on contrast sensitivity.

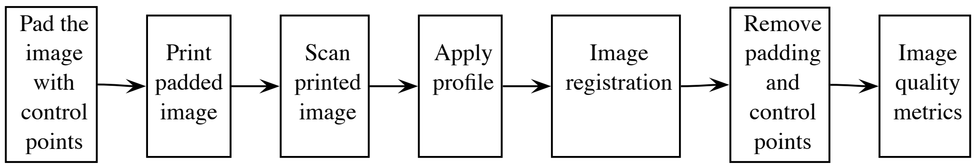

Fig. 1 Overview of the proposed framework for using IQ metrics with printed images.

In order assess the performance of our framework we compare it to the one proposed by Eerola et al. [5]. Three different images were used in the comparison, and the results show that the proposed framework based on control points introduces less error than the framework by Eerola et al. [5] based on local features. In addition the proposed framework is significantly faster.

2.2 Using the framework with image quality metrics: We have used the framework explained above to evaluate a set of IQ metrics on images from a color workflow. 15 images were processed with two different source profiles, the sRGB v2 perceptual transform and the sRGB v4 perceptual transform. These were further processed with four different softwares for obtaining the destination profile. The images and the subjective results were obtained from Cardin [6], where 30 observers participated in the experiment.

A set of IQ metrics, S-CIELAB [3], S-CIELAB with the improved contrast sensitivity function from Johnson and Fairchild [7], S-DEE [8], and SHAME [9], were applied to evaluate the IQ of the printed images. Evaluation of the performance is done by calculating the Pearson correlation coefficient between the subjective score and the objective score. The results show that all metrics have a very low correlation, approximately around zero, indicating that the IQ metrics cannot predict perceived IQ. To verify these results we also used the evaluation method proposed by Pedersen and Hardeberg [10], where the rank of each IQ metric is used as the basis for the analysis. The results from this method support the findings from the evaluation by correlation.

Image wise evaluation by using correlation showed that some IQ metrics had a higher performance for images where certain QAs occurred. Based on this the next natural step is to investigate QAs.

3 IMAGE QUALITY ATTRIBUTES: Evaluation of perceived IQ in color prints is a complex task, due to its subjectivity and dimensionality. The perceived quality of an image is influenced by a number of different QAs, such as sharpness and color. It is difficult and complicated to evaluate the influence of all attributes on overall IQ, and their influence on other attributes. Because of this the most important attributes should be identified in order to achieve a more efficient and manageable evaluation of IQ. Based on a survey of the existing literature and a psychophysical experiment, we identify and categorize existing IQ attributes to propose a refined selection of meaningful ones for the evaluation of color prints.

As a first step towards a subset of the most relevant and important QAs, existing QAs must be identified. In order to do this we have taken the approach of doing a survey of the existing literature. This survey resulted in a list of more than 45 different QAs considered to influence IQ, such as sharpness, contrast, color, and artifacts. All of these QAs cannot be evaluated, and therefore it is required to reduce them to a more manageable set. This was done based on the following criteria:

- they should be based on perception,
- they should account for technological printing issues,
- they should be general, not to exclude novice observers,
- they should be suitable for IQ metrics,
- they should create a link between objective and subjective IQ.

The existing sets of QAs do not fulfill all of these requirements, and therefore a new set of QAs is needed. Based on the criteria above we have reduced the QAs found in the literature to the following six:

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

- **Lightness** is considered so perceptually important that it is beneficial to separate it from color. Lightness will range from ”light” to ”dark”.
- **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 and definition of edges.
- **Artifacts**, like noise and banding, contribute to degrading the quality of an image if detectable.
- The **physical** QA contains physical parameters that affect quality, such as paper properties and gloss.

We have turned to Venn diagrams to create a simple and intuitive illustration of the QAs and their influence on overall IQ (Figure 2). 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 [11]. 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 other QAs.

3.1 Verification of the quality attributes: A set of images was reproduced using the ICC perceptual rendering intent and investigated by 15 observers to verify the proposed QAs, and to learn which QAs that observers use in the IQ evaluation of a color workflow. In order for the observers to use a sufficiently large set of QAs, a broad range of images, natural as well as test charts, were used in order to reveal different quality issues. The instructions given to the observers focused on the overall IQ rating of the reproduction, and which QAs the observers used in their evaluation.

The results show that the observers used more than 50 QAs in the evaluation, with an average of 10 different QAs for each observer. For each image an average of 2.95 QAs were used, with a minimum of one and a maximum of eight. All the QAs used by the observers were fitted to the QAs proposed above by using their definitions, as given in the bullet point list. The results show that almost all of the QAs used by the observers can be fitted within the proposed QAs, where color is the most used QAs, followed by sharpness, contrast, artifacts, and lightness.

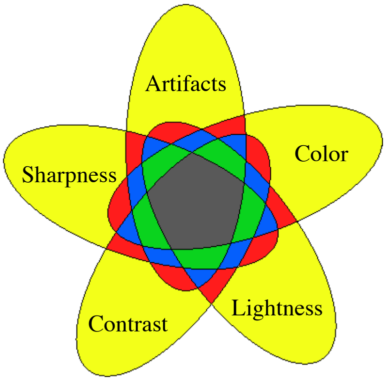

Fig. 2 Simple Venn ellipse diagram with five folds used for an abstract illustration of the QAs. Five different QAs and the interaction between these are shown. Overall IQ can be influenced by one, two, three, four, or five of the QAs.

4 CONCLUSION: We have investigated the use of IQ metrics to evaluate IQ of color prints. In order to do this we have proposed a framework for the transformation of a printed image into a digital format for the application of IQ metrics. The framework is based on control points, and outperforms a state of the art framework. Evaluation of IQ metrics with this framework showed that they cannot predict perceived IQ, but that some metrics perform better for certain QAs. This led to an investigation of the existing QAs in the literature, which was further used to propose a refined set of meaningful QAs to evaluating IQ of color prints.

Future work includes selection of appropriate IQ metrics for the different QAs, and evaluation of these IQ metrics.

ACKNOWLEDGEMENTS: The work in article has been carried out in collaboration with Jon Yngve Hardeberg, Fritz Albregtsen, Nicolas Bonnier, and Seyed Ali Amirshahi. The content has been previously published [12–14].

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.

REFERENCES:

[1] 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.

[2] X. Zhang, D.A. Silverstein, J.E. Farrell, and B.A. Wandell. Color image quality metric S-CIELAB and its application on halftone texture visibility. In *COMPCON97 Digest of Papers*, pages 44–48, Washington, DC, USA, 1997. IEEE Computer Society.

[3] 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.

[4] X. Yanfang, W. Yu, and Z. Ming. Color reproduction quality metric on printing images based on the s-cielab model. In *2008 International Conference on Computer Science and Software Engineering*, pages 294–297, 2008.

[5] T. Eerola, J-K. Kamarainen, L. Lensu, and H. Kalviainen. Framework for applying full reference digital image quality measures to printed images. In *Scandinavian Conference on Image Analysis*, Oslo, Norway, June 2009.

[6] N. Cardin. L'utilisation du perceptual reference medium gamut dans la gestion des coleours améliore-t-elle la qualité d'images produites par impression jet d'encre? Master's thesis, École Nationale Supérieure Louis-Lumière Promotion Photographie, 2009.

[7] G. M. Johnson and M. D. Fairchild. Darwinism of color image difference models. In *The 9th Color Imaging Conference: Color Science and Engineering: Systems, Technologies, Applications*, pages 108–112, 2001.

[8] G. Simone, C. Oleari, and I. Farup. Performance of the euclidean color-difference formula in log-compressed osa-ucs space applied to modified-image-difference metrics. In *11th Congress of the International Colour Association (AIC)*, Sydney, Australia, Oct 2009.

[9] M. Pedersen and J. Y. Hardeberg. A new spatial hue angle metric for perceptual image difference. In *Computational Color Imaging*, volume 5646 of *Lecture Notes in Computer Science*, pages 81–90, Saint Etienne, France, Mar 2009. Springer Berlin / Heidelberg. ISBN: 978-3-642-03264-6.

[10] M. Pedersen and J. Y. Hardeberg. Rank order and image difference metrics. In *CGIV 2008 Fourth European Conference on Color in Graphics, Imaging and Vision*, pages 120–125, Terrassa, Spain, Jun 2008. IS&T.

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

[12] M. Pedersen and S.A. Amirshahi. A modified framework the evaluation of color prints using image quality metrics. In *CGIV - Color in Graphics, Imaging, and Vision*, Joensuu, Finland, Jun. 2010.

[13] M. Pedersen, N. Bonnier, J. Y. Hardeberg, and F. Albregtsen. Attributes of image quality for color prints. *Journal of Electronic Imaging*, 19(1):011016–1 – 011016–13, Jan 2010.

[14] M. Pedersen, N. Bonnier, J. Y. Hardeberg, and F. Albregtsen. Attributes of a new image quality model for color prints. In *Color Imaging Conference*, pages 204–209, Albuquerque, New Mexico, USA, Nov 2009.