Flexible and Robust Calibration of the Yule-Nielsen Model for CMYK Prints

THÈSE N^{O} 5049 (2011)

PRÉSENTÉE LE 17 JUIN 2011

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE DE SYSTÈMES PÉRIPHÉRIQUES
PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Thomas BUGNON

acceptée sur proposition du jury:

Prof. P. Fua, président du jury
Prof. R. Hersch, directeur de thèse
Dr Ph. Green, rapporteur
Dr H. Hoffstadt, rapporteur
Prof. S. Süsstrunk, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2011

Abstract

Spectral reflection prediction models, although effective, are impractical for certain industrial applications such as self-calibrating devices and online monitoring because of the requirements imposed by their calibration. The idea emerged to make the calibration more flexible. Instead of requiring specific color-constant calibration patches, the calibration would rely on the information contained in regular prints, e.g. on information found in printed color images. Using the CMYK Ink Spreading enhanced Yule-Nielsen modified Spectral Neugebauer model (IS-YNSN), the objective of this dissertation is to recover the Neugebauer primaries and ink spreading curves from image tiles extracted from printed color images.

The IS-YNSN is first reviewed in the context of CMYK prints. Two sources of ambiguity are identified and removed, yielding a more robust model better suited for a flexible calibration. We then propose a gradient-descent method to acquire the ink spreading curves from image tiles by relying on constraints based on a metric evaluating the relevance of each ink spreading curve to the set of image calibration tiles. We optimize the algorithm which automatically selects the image tiles to be extracted and show that 5 to 10 well-chosen image tiles are sufficient to accurately acquire all the ink spreading curves.

The flexible calibration is then extended to recover the Neugebauer primaries from printed color images. Again, a simple gradient-descent algorithm is not sufficient. Thanks to a set of constraints based on Principal Component Analysis (PCA) and the relationships between composed Neugebauer primaries and the ink transmittances, good approximations of the Neugebauer primaries are achieved. These approximations are then optimized, yielding an accurately calibrated IS-YNSN model comparable to one obtained by classical calibrations. A detailed analysis of these calibrations shows that 25 well-chosen CMYK image calibration tiles are sufficient to accurately recover both the Neugebauer primaries and the ink spreading curves.

Keywords : color prints, color reproduction, dot gain, halftones, ink spreading, prediction model calibration, spectral reflection prediction, automatic selection, image tiles

Résumé

Les modèles de prédiction de réflectances spectrales, bien qu'efficaces, ne sont pas utilisables pour certaines applications industrielles, telles que les systèmes auto-calibrants ou le contrôle continu de systèmes de reproduction, à cause des restrictions imposées par leur calibration. L'idée de rendre cette calibration plus flexible a donc germé. Au lieu d'utiliser des échantillons de calibration de couleur constante, la calibration utiliserait l'information contenue dans les pages imprimées, comme par exemple l'information contenue dans les images. Basé sur le modèle spectral CMJN de Neugebauer modifié par Yule et Nielsen et étendu par un modèle d'accroissement du point de trame (modèle IS-YNSN), la présente thèse a pour objectif de déterminer les primaires de Neugebauer et les courbes d'accroissement du point de trame à partir de tuiles image extraites d'images couleur imprimées et sans connaissances préalables du système de reproduction.

Le modèle IS-YNSN est d'abord réexaminé dans le contexte des quatre encres cyan, magenta, jaune et noire (CMJN). Deux sources d'ambiguïté sont identifiées et supprimées, rendant le modèle plus robuste et mieux adapté à une calibration flexible. Nous proposons ensuite une méthode de descente de gradient afin d'acquérir les courbes d'accroissement du point à partir de tuiles image. Cette méthode utilise des contraintes basées sur une métrique évaluant la pertinence des courbes d'accroissement du point de trame par rapport à l'ensemble des tuiles de calibration. L'algorithme de sélection de tuiles à extraire des images est ensuite optimisé et nous montrons que 5 à 10 tuiles de calibration soigneusement choisies sont suffisantes pour déterminer les courbes d'accroissement du point de trame avec exactitude.

La calibration flexible est ensuite étendue à la détermination des spectres des primaires de Neugebauer à partir d'images couleur imprimées. À nouveau, une simple descente de gradient n'est pas suffisante. Grâce à plusieurs contraintes basées sur une analyse en composantes principales (PCA) et à la relation entre les primaires de Neugebauer composées et la transmittance des encres, de bonnes approximations des primaires de Neugebauer sont obtenues. Ces approximations sont ensuite optimisées afin d'obtenir une calibration du modèle IS-YNSN pré-

cise et comparable à une calibration classique. Une analyse détaillée de ces calibrations montre que 25 tuiles de calibration CMJN soigneusement choisies sont suffisantes pour déterminer précisément les primaires de Neugebauer et les courbes d'accroissement du point de trame.

Mots-clés : impressions couleur, reproduction couleur, accroissement du point de trame, demi-tons, étalement d'encre, calibration de modèles de prédiction, prédiction de réflectances spectrales, sélection automatique, tuiles image

Acknowledgments

Beginning a PhD is not an easy task. I was new to the field when I arrived at the Peripheral Systems Laboratory (LSP) and the two years I spent as an assistant helped me greatly to find the track I wanted to follow. Thank you Roger for giving me the time and the opportunity to find my way. Thank you also for the numerous discussions about my work, my progress, and for providing a never-ending stream of interesting problems to solve. Your support and guidance are invaluable.

Completing my PhD would have been far more difficult if the LSP had not been such an inspiring place to work in. A friendly atmosphere, dedicated people, and interesting weekly meetings tremendously help overcoming the many obstacles. A big thank you to Mathieu Brichon, Basile Schaeli, Sebastian Gerlach, Isaac Amidror, Fabienne Allaire, Maria Anitua, Sylvain Chosson, Romain Rossier, Mathieu Hébert, Fabrice Rousselle, Andrea Maesani, Zhe Wei, Arunanshu Pandey, Jean-Marc Comby, Julien Andrès, Petar Pjanic, and Vahid Babaei for sharing at least part of this time with me. I also thank Xavier Jimenez, Rafik Chaabouni and Florent Garcin who regularly joined our team and contributed to making a great time out of my PhD.

I am also greatly indebted to the people I have met on the field: Herbert Janser, Jiang Yufan, Matthias Rippenhoff, and the entire team at Wifag AG; Sergio Lizzola; Pascal Fehr; Dominique Ackermann and the entire team at Perfect SA. Your experience, skills and patience greatly help bridge the gap between academia and industry and make research all the more relevant.

The support of the Swiss National Science Foundation to research helps many projects and contributes to keeping Switzerland one of the most innovative countries. Thank you for supporting my thesis under grant n° 200020–126757/1.

Finally, life would be meaningless without the joys of friends and family. Thank you all for making mine so colorful, and thank you Michèle for making it so special.

Contents

“The pursuit of truth and beauty is a sphere of activity in which we are permitted to remain children all our lives.”
– Albert Einstein

Chapter 1

Introduction

1.1 Motivations

Printing technologies used to be a luxury only accessible to professionals. Since the 21st century, advances in miniaturization and manufacturing processes have made devices such as digital cameras, video cameras and desktop printers cheap and ubiquitous. By combining these technologies with personal computers, anyone has access to powerful tools and can become an artist, a designer, a publisher, etc. However, it is a difficult task to master these tools in order to produce the desired results and non-professionals do not have the required knowledge to overcome all the obstacles. Technology must be robust and flexible in order to assist them in their tasks.

Another aspect of modern society having an impact on how technology evolves is that everything must go fast and costs be reduced to the minimum. This attitude puts a lot of pressure on professionals, especially since technology keeps getting ever more complex and specialized. It is difficult for them to meet the productivity targets while maintaining the expected high quality. In the printing industry, they increasingly rely on the new printing devices to automatically manage many aspects of the reproduction process in, again, a robust and flexible way.

The need for robust and flexible tools has been called to our attention while working on a project in collaboration with Wifag, a company building offset presses. A collaborator of Wifag remarked that our solution to their problem, although effective, was rather impractical. The solution was based on the Clapper-Yule spectral reflection prediction model. This theoretical model needs to be calibrated, i.e. tuned to the current working environment, and this calibration requires patches of specific colors to be printed and measured. Due to the cost

of printing the required colors, measuring them, and due to the number of different ink and paper combinations, calibrating the theoretical model for all these combinations was not an option. The idea emerged then to rethink the method used to calibrate such spectral reflection prediction models.

1.2 Spectral reflection prediction models

The most important aspect about color printing are the colors. They must be pleasant and faithful. Although faithful colors refer more to what people expect to see rather than what people actually see, it is fundamental to accurately control the printing device in order to deliver the desired colors. To achieve such color fidelity, printing devices must be both calibrated and characterized [Bal03].

Device calibration ensures that the device has a known characteristic color response, i.e. the device always prints the same color when given the same set of control values. Once a device is calibrated, it is possible to control the printing process and to repeat given results. Control and repeatability are two characteristics that are expected of any printing device.

Device characterization relates the characteristic color response achieved through the device calibration to a device independent representation such as the CIE-XYZ or CIE-LAB color spaces. It defines the gamut of the printing device, i.e. the set of printable colors. This step enables the comparison of different devices as well as the conversion of colors from one device to another.

Both device calibration and characterization can benefit from the use of a model predicting the reflectance of color halftones as a function of their nominal surface coverages, i.e. of the control values specifying the amount of inks to be printed. For device calibration, spectral reflection prediction models are helpful in studying the influence of given factors on the range of printable colors: the inks, the type of paper, the illumination conditions, and the halftones. A spectral reflection prediction model is also useful for the characterization of a printing device since it predicts the device-independent color values resulting from a set of nominal ink surface coverages. It enables not only the creation of ICC profiles [GL96], but also new color separation strategies, for example when printing with custom inks [TB01].

These spectral reflection prediction models also need to be calibrated. Classical model calibration is performed by measuring the reflection spectra of specially printed calibration patches. These special calibration patches comprise all the solid colorants (paper, solid inks and superpositions of solid inks) and single ink halftones superposed with all combinations

of solid ink superpositions. For CMYK prints, a minimum of 36 calibration patches are required [BBH08], but that number may be considerably larger to achieve even better accuracy, e.g. several hundreds of patches may be used for the Cellular Yule-Nielsen model [RB93].

1.3 Scope and objective

There exist several different spectral reflection prediction models, but the most widely used is the Yule-Nielsen modified Spectral Neugebauer model (YNSN) because of its simplicity and its relatively good accuracy [WB00]. In order to improve its accuracy, the YNSN model has been enhanced with an ink spreading model accounting for physical dot gain [HC05]. First derived for 3 inks, the extension to 4 inks of this ink spreading model is presented in [HECC05]. The scope of this dissertation is limited to this 4 ink IS-YNSN model applied to CMYK prints, which corresponds to many practical situations such as the one that arose in our collaboration with Wifag. Note that alternative prediction models, e.g. the Clapper-Yule model mentioned above, are discussed in Section 2.5.

The choice to limit ourselves to 4 inks is one often made in the literature [TB00]. The first reason is that the YNSN model requires an exponential number of measurements for its calibration with respect to the number of inks, which is time-consuming and expensive. The second reason is the effect called *ink-trapping limitation*. When the total ink coverage exceeds 400% at a given location, the inks stop adhering to the substrate [TB00]. The paper begins to crinkle and the inks to leak. To avoid this problem, the upper bound for total ink coverage is usually set to 300%. With this limit, it is not possible to print too many different inks at the same location. The third reason is that all the colors in the gamut of a printing device using more than 4 inks can be achieved with a combination of at most 4 of the available inks.

Among all possible combinations of 4 inks, we have chosen to focus on cyan, magenta, yellow, and black inks because this combination generates the largest gamut. It is also the combination that is available in all printing devices providing a minimum of 4 inks. We also consider the simpler CMY case when appropriate.

The objective of this dissertation is to propose a calibration method for the chosen spectral reflection prediction model as flexible and robust as possible, i.e. a calibration method based only on the information that can be obtained from produced prints without relying on specially conceived calibration patches or prior knowledge of the printing device. The main problem of the classical calibration procedure is to print the required and specially conceived calibration patches. If they are too numerous, it is costly and time-consuming to print and measure them all.

In other circumstances, they take valuable space on the printed product and must be removed. For example, on offset presses, the special calibration patches are placed in the margins of the printed pages and are cut off before assembling the final print product. Instead of relying on these special calibration patches, we propose a calibration method based on measurements of image tiles extracted from printed color images. This flexible calibration allows the use of the IS-YNSN model in situations similar to the situation at Wifag described above and any printing device can be monitored online.

1.4 Dissertation outline

Chapter 2 begins by introducing the different parts of the IS-YNSN model and how these parts are calibrated. It then discusses different approaches to the problem of spectral reflection prediction.

In Chapter 3, we focus on the robustness of the IS-YNSN model for CMYK prints. Compared to the CMY case, the addition of the black ink introduces two sources of ambiguity for the model. The first ambiguity follows from the two possibilities to create the color black, either using the black ink or superposing cyan, magenta, and yellow. When the IS-YNSN model is used to deduce the surface coverages of printed color halftone patches from their measured reflectance spectra, it cannot accurately distinguish between the two different blacks, introducing inaccuracies in the deduced coverages. The second ambiguity is that printing solid black along with other inks yields black. There is no possibility to detect the presence of the other inks from measured spectra, generating instabilities when calibrating certain ink spreading curves. We therefore propose solutions that remove these sources of errors. This more robust version of the IS-YNSN model is used as basis for our calibration based on image tiles extracted from printed color images.

In Chapters 2 and 3 and in the literature, the IS-YNSN model is used to predict the reflection spectra of patches of a constant color. In Chapter 4, instead of these color-constant patches, we investigate if tiles extracted from color images can be accurately predicted and how they can be used to facilitate the calibration of the ink spreading curves.

Chapter 5 extends the calibration method described in Chapter 4 from the ink spreading model to the entire IS-YNSN model. The proposed calibration method is able to accurately calibrate the IS-YNSN model using exclusively tiles extracted from color images. No prior knowledge about the printing device or the inks is necessary.

The conclusions are drawn in Chapter 6.

1.5 Contributions

The detailed list of contributions is as follows:

- We show that the near-infrared wavelength range is useful when a pigment-based black ink is used, removing the ambiguity between the black ink and the superposition of the cyan, magenta, and yellow inks;
- We describe a new set of ink spreading equations specific to CMYK prints. We also discuss how this new set of equations is derived and how the same reasoning can be applied to other ink combinations;
- We show that tiles extracted from color images can be measured and accurately predicted if the colors contained in the tiles are uniform enough. We also show that the accuracy of the predicted spectra of color image tiles is well correlated with the uniformity of these tiles;
- We propose an algorithm to calibrate the ink spreading curves based on tiles extracted from color images. This calibration algorithm relies on a metric evaluating the relevance of each ink spreading curve for predicting a given set of color image tiles;
- Instead of measuring the reflectances of the Neugebauer primaries, we describe a set of constraints imposed on these Neugebauer primaries that enable their recovery based on color image tiles;
- We propose a least-squares minimization algorithm to optimize the reflectances of the Neugebauer primaries calibrated from color image tiles. This optimization relies on a metric evaluating the relevance of each Neugebauer primary for predicting a given set of color image tiles; and
- We describe algorithms that automatically select color image tiles best suited to the proposed calibrations based on color image tiles.

Chapter 2

The Ink-Spreading Enhanced Yule-Nielsen modified Spectral Neugebauer model

2.1 Introduction

The Yule-Nielsen model is the most widely used spectral reflection prediction model for its simplicity and accuracy. Based on the Neugebauer model, the Yule-Nielsen modified Spectral Neugebauer model (YNSN) predicts the spectral reflectance of printed samples given their nominal ink coverages, i.e. the relative surface covered by the inks and derived from the device control values used to print the samples. However, predictions performed using nominal ink coverages are not the most accurate since the YNSN model does not account for physical dot gain.

When inks are printed on a substrate, they tend to spread out and cover an effective surface larger than the intended nominal surface. In order to account for this phenomenon, an ink spreading model must be used. Such a model computes effective ink surface coverages from nominal ink surface coverages. The effective ink surface coverages are then used in the YNSN model instead of the nominal ink surface coverages to improve the accuracy of the predictions.

This chapter first describes the Yule-Nielsen modified Spectral Neugebauer model (YNSN). It follows by describing the ink spreading model on which this dissertation is based and combines both the YNSN and ink spreading models to form the IS-YNSN model. The different assumptions made by the two models are finally discussed and compared to the literature.

2.2 The Yule-Nielsen modified Spectral Neugebauer model (YNSN)

One of the first color prediction models is the Neugebauer model ([Neu37], translated into English in [WK05]). In its original form, it predicts the RGB values of a color halftone patch as the sum of the RGB values of their individual colorants weighted by their fractional area coverages a_i. By considering the reflection spectra R_i of colorants instead of their respective RGB values, one obtains the spectral Neugebauer equations [Vig90]. They predict the reflection spectrum R of a printed color halftone patch as a function of the reflection spectra of its individual colorants (also called Neugebauer primaries):

$$R(\lambda) = \sum_i a_i \cdot R_i(\lambda) \tag{2.1}$$

With k inks, there are 2^k colorants: white, the k single ink colorants and all the different superpositions of solid inks. For example, the red colorant is the superposition of the magenta and yellow inks. When the ink layers are printed independently one from another, the fractional area coverages of the individual colorants are calculated from the surface coverages of the inks by the Demichel equations ([WB00], originally published in [Dem24]). For 4 inks, the Demichel equations are the following:

$$\begin{array}{llll}
a_w &= (1-c)\cdot(1-m)\cdot(1-y)\cdot(1-k) & a_k &= (1-c)\cdot(1-m)\cdot(1-y)\cdot k \\
a_c &= c\cdot(1-m)\cdot(1-y)\cdot(1-k) & a_{ck} &= c\cdot(1-m)\cdot(1-y)\cdot k \\
a_m &= (1-c)\cdot m\cdot(1-y)\cdot(1-k) & a_{mk} &= (1-c)\cdot m\cdot(1-y)\cdot k \\
a_{cm} &= c\cdot m\cdot(1-y)\cdot(1-k) & a_{cmk} &= c\cdot m\cdot(1-y)\cdot k \\
a_y &= (1-c)\cdot(1-m)\cdot y\cdot(1-k) & a_{yk} &= (1-c)\cdot(1-m)\cdot y\cdot k \\
a_{cy} &= c\cdot(1-m)\cdot y\cdot(1-k) & a_{cyk} &= c\cdot(1-m)\cdot y\cdot k \\
a_{my} &= (1-c)\cdot m\cdot y\cdot(1-k) & a_{myk} &= (1-c)\cdot m\cdot y\cdot k \\
a_{cmy} &= c\cdot m\cdot y\cdot(1-k) & a_{cmyk} &= c\cdot m\cdot y\cdot k
\end{array} \tag{2.2}$$

Since the Neugebauer model neither takes explicitly into account the lateral propagation of light within the paper bulk nor the internal reflections (Fresnel reflections) at the paper-air interface, its predictions are not accurate [Kan94]. Yule and Nielsen [YN51] modeled the non-linear relationship between the reflectances of paper, single ink halftones, and the corresponding solid ink prints by a power function whose exponent n can be optimized according to the reflectances

of a limited set of halftone patches. Viggiano [Vig90] applied the Yule-Nielsen relationship to the spectral Neugebauer equations, yielding the Yule-Nielsen modified Spectral Neugebauer model (YNSN):

$$R(\lambda) = \left(\sum_i a_i \cdot R_i(\lambda)^{1/n} \right)^n \tag{2.3}$$

When calibrated, i.e. when the reflection spectra of the colorants R_i and the Yule-Nielsen n-value are known, the YNSN can be seen as a function taking the ink surface coverages of a color halftone as input and yielding its reflection spectrum. It is possible to deduce ink coverages from a given spectrum by minimizing a difference metric between predicted and measured reflection spectra as follows:

$$covs_{opt} = \arg\min_{covs} \left[\sum_k \left(predSpectrum(\lambda_k, covs) - R(\lambda_k) \right)^2 \right] \tag{2.4}$$

where $covs = \{c, m, y, k\}$ are the nominal ink coverages, $R(\lambda)$ the measured spectrum, and $predSpectrum$ the YNSN model predicting function combining Eqs. (2.2) and (2.3).

Ink coverages deduced using Eq. (2.4) are usually larger than the corresponding nominal ink coverages because physical dot gain due to ink spreading is not accounted for in the YNSN model. As a consequence, spectral predictions made from nominal ink coverages are not accurate. We therefore enhance the YNSN model by adding an ink spreading model.

2.3 The Ink Spreading model

When a reproduction device reproduces a color with given nominal ink coverages by depositing inks on paper, the inks spread out on the paper, resulting in *effective ink surface coverages* usually greater than the requested *nominal ink surface coverages*. This phenomenon is known as *physical* or *mechanical* dot gain, as well as *ink spreading*. The amount of dot gain of an ink depends on whether the ink halftone is printed alone on paper or in superposition with one or more other inks [YC51, HECC05]. The goal of the ink spreading model is therefore to compute the effective ink surface coverages from the nominal ink surface coverages in order to use the effective ink surface coverages as input to the YNSN model.

Since the ink spreading model is based on ink spreading curves, this section first presents the available ink spreading curves and the method used to calibrate them. It then describes how the ink spreading model computes the effective surface coverages using these ink spreading

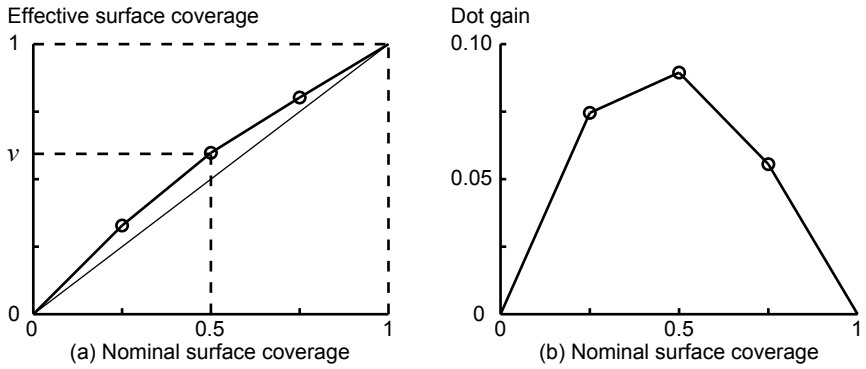


Figure 2.1: (a) Example of a linearly interpolated ink spreading curve (thick solid line) characterized by three calibration points (circles). The thin solid line indicates when there is no ink spreading. (b) Corresponding dot gain curve, defined as the difference between effective and nominal surface coverages.

Table 2.1: List of the considered ink spreading curve indicia.

Cyan		Magenta		Yellow		Black	
c	c/k	m	m/k	y	y/k	k	k/y
c/m	c/mk	m/c	m/ck	y/c	y/ck	k/c	k/cy
c/y	c/yk	m/y	m/yk	y/m	y/mk	k/m	k/my
c/my	c/myk	m/cy	m/cyk	y/cm	y/cmk	k/cm	k/cmy

curves. Three different methods are proposed, further referred to as *ink spreading directives*, each directive corresponding to given assumptions about how the inks influence each other.

2.3.1 Ink Spreading Curves

The ink spreading model relies on ink spreading curves. An ink spreading curve maps the nominal surface coverages of an ink halftone into its effective surface coverages, i.e. to the surface that the ink halftone effectively covers after being printed. For example in Figure 2.1a, the 50% nominal surface coverage is mapped to effective surface coverage v. There is one ink spreading curve for each ink halftone in each superposition condition. For example, a cyan halftone may be printed alone, c; superposed with solid magenta, c/m; with solid yellow, c/y; with solid black, c/k; with solid magenta and solid yellow, c/my; with solid magenta and solid black, c/mk; with solid yellow and solid black, c/yk; and with solid magenta, yellow and black, c/myk. There are 8 different ink spreading curves for each ink, yielding a total of 32 ink spreading curves. Table 2.1 lists the 32 possible ink spreading curves for CMYK prints.

Figure 2.1b shows the dot gain curve corresponding to the ink spreading curve of Figure 2.1a, where dot gain is defined as the difference between effective and nominal coverages.

Figure 2.1 also shows that three calibration patches (the three circles) have been used to calibrate the ink spreading curve. The nominal ink coverages of these calibration patches are 0.25, 0.5, and 0.75, respectively. These nominal ink surface coverages determine the position on the x-axis of the ink spreading curve. The corresponding effective surface coverages are fitted using Eq. (2.4) and determine the position on the y-axis. The ink spreading curve is then linearly interpolated. Any number of calibration patches can be used. Each additional patch adds another circle and refines the ink spreading curve. Ink spreading curves calibrated using this method are further referred to as *linearly interpolated ink spreading curves*.

2.3.2 Ink Spreading Directives

When using more than one ink spreading curve per ink, there are many possible combinations to compute the effective ink coverages. Each directive of the ink spreading model specifies a different combination based on certain assumptions. Below, we describe the three directives presented in [HECC05].

In the following equations, the cyan, magenta, yellow, and black nominal coverages are designated using the variables c, m, y, and k, respectively, whereas the effective coverages are designated using the variables c', m', y', and k'.

Single ink dot gain

In the literature, ink spreading is often characterized by one dot gain curve per ink and does not take superposition conditions into account [RB93, WB00]. In that case, the effective coverage of an ink depends only on the nominal coverage of the ink and is computed from the dot gain curve of the ink alone. This ink spreading directive is further referred to as the *single* directive and yields the following equations:

$$c = f_c(c) \qquad m = f_m(m) \qquad y = f_y(y) \qquad k = f_k(k) \tag{2.5}$$

The problem with the single directive is that it does not take superposition conditions into consideration.

Ink spreading on top of solid inks

The next directive, further referred to as the *top* directive, assumes that an ink halftone is influenced by the inks that are already printed. It is influenced by the underlying inks, but not

by new ink layers printed on top of it.

The effective coverages are computed in order in which the inks are printed. Assuming that the inks are printed in the order cyan, magenta, yellow, and black, we first compute the effective coverage of cyan. As no other ink is present when cyan is printed, the effective coverage of cyan is computed using the ink spreading curve of cyan alone, as in the *single* directive. Then, we compute the effective coverage of magenta. Since the cyan ink covers a percentage c' of the surface, a percentage c' of magenta is printed over solid cyan and a percentage $(1 - c')$ of magenta is printed alone on paper. We therefore use these percentages to weight the effective coverages returned by the ink spreading curve of magenta over solid cyan and of magenta alone. For yellow, we weight the ink spreading curves of yellow alone, yellow on cyan, yellow on magenta, and yellow on cyan and magenta according to the percentages of surface covered respectively by paper, $(1 - c') \cdot (1 - m')$; cyan only, $c' \cdot (1 - m')$; magenta only, $(1 - c') \cdot m'$; and cyan and magenta, $c' \cdot m'$.

We apply the same logic for the black ink halftone. We obtain a total of 15 ink spreading curves weighted according to the following equations:

$$
\begin{aligned}
c' &= f_c(c) \\
m' &= (1-c')\cdot f_m(m) \\
&\quad + c' \cdot f_{m/c}(m) \\
y' &= (1-c')\cdot(1-m')\cdot f_y(y) \\
&\quad + c' \cdot(1-m')\cdot f_{y/c}(y) \\
&\quad + (1-c')\cdot m' \cdot f_{y/m}(y) \\
&\quad + c' \cdot m' \cdot f_{y/cm}(y)
\end{aligned}
\qquad
\begin{aligned}
k' &= (1-c')\cdot(1-m')\cdot(1-y')\cdot f_k(k) \\
&\quad + c' \cdot(1-m')\cdot(1-y')\cdot f_{k/c}(k) \\
&\quad + (1-c')\cdot m' \cdot(1-y')\cdot f_{k/m}(k) \\
&\quad + c' \cdot m' \cdot(1-y')\cdot f_{k/cm}(k) \\
&\quad + (1-c')\cdot(1-m')\cdot y' \cdot f_{k/y}(k) \\
&\quad + c' \cdot(1-m')\cdot y' \cdot f_{k/cy}(k) \\
&\quad + (1-c')\cdot m' \cdot y' \cdot f_{k/my}(k) \\
&\quad + c' \cdot m' \cdot y' \cdot f_{k/cmy}(k)
\end{aligned}
\tag{2.6}
$$

The *top* directive improves the predictions, but because of the speed of printing devices, the inks that are already printed are not dry when the next layer is printed.

Ink spreading depending on solid inks printed on top or below the considered ink halftone

Since a given ink halftone is influenced both by the inks already printed and by the inks that are printed on top of it, a new directive, further referred to as the *top or below* directive, is

required. The set of equations corresponding to this assumption is the following:

$$
\begin{aligned}
c' &= (1-m')(1-y')(1-k')\,f_c(c) \\
&+ \quad m' \quad (1-y')(1-k')\,f_{c/m}(c) \\
&+ (1-m') \quad y' \quad (1-k')\,f_{c/y}(c) \\
&+ \quad m' \quad\quad y' \quad (1-k')\,f_{c/my}(c) \\
&+ (1-m')(1-y') \quad k' \quad f_{c/k}(c) \\
&+ \quad m' \quad (1-y') \quad k' \quad f_{c/mk}(c) \\
&+ (1-m') \quad y' \quad\quad k' \quad f_{c/yk}(c) \\
&+ \quad m' \quad\quad y' \quad\quad k' \quad f_{c/myk}(c)
\end{aligned}
\qquad
\begin{aligned}
y' &= (1-c')(1-m')(1-k')\,f_y(y) \\
&+ \quad c' \quad (1-m')(1-k')\,f_{y/c}(y) \\
&+ (1-c') \quad m' \quad (1-k')\,f_{y/m}(y) \\
&+ \quad c' \quad\quad m' \quad (1-k')\,f_{y/cm}(y) \\
&+ (1-c')(1-m') \quad k' \quad f_{y/k}(y) \\
&+ \quad c' \quad (1-m') \quad k' \quad f_{y/ck}(y) \\
&+ (1-c') \quad m' \quad\quad k' \quad f_{y/mk}(y) \\
&+ \quad c' \quad\quad m' \quad\quad k' \quad f_{y/cmk}(y)
\end{aligned}
\tag{2.7}
$$

$$
\begin{aligned}
m' &= (1-c')\,(1-y')(1-k')\,f_m(m) \\
&+ \quad c' \quad (1-y')(1-k')\,f_{m/c}(m) \\
&+ (1-c') \quad y' \quad (1-k')\,f_{m/y}(m) \\
&+ \quad c' \quad\quad y' \quad (1-k')\,f_{m/cy}(m) \\
&+ (1-c')\,(1-y') \quad k' \quad f_{m/k}(m) \\
&+ \quad c' \quad (1-y') \quad k' \quad f_{m/ck}(m) \\
&+ (1-c') \quad y' \quad\quad k' \quad f_{m/yk}(m) \\
&+ \quad c' \quad\quad y' \quad\quad k' \quad f_{m/cyk}(m)
\end{aligned}
\qquad
\begin{aligned}
k' &= (1-c')(1-m')(1-y')\,f_k(k) \\
&+ \quad c' \quad (1-m')(1-y')\,f_{k/c}(k) \\
&+ (1-c') \quad m' \quad (1-y')\,f_{k/m}(k) \\
&+ \quad c' \quad\quad m' \quad (1-y')\,f_{k/cm}(k) \\
&+ (1-c')(1-m') \quad y' \quad f_{k/y}(k) \\
&+ \quad c' \quad (1-m') \quad y' \quad f_{k/cy}(k) \\
&+ (1-c') \quad m' \quad\quad y' \quad f_{k/my}(k) \\
&+ \quad c' \quad\quad m' \quad\quad y' \quad f_{k/cmy}(k)
\end{aligned}
$$

As the *top* directive, the *top or below* directive weights the effective coverages returned by the different ink spreading curves according to the colorant surface coverages in the different superposition conditions. However, instead of considering only the inks already printed in the possible superposition conditions, the *top or below* directive accounts for all possible superposition conditions. As there are 8 possible superposition conditions for each ink (see the list of superposition conditions in Table 2.1), the effective coverage of an ink is the weighted average of 8 different ink spreading curves. For example, if a patch with 50% cyan, 50% magenta, and 50% yellow is printed, the superposition dependent colorant surface coverages weighting the ink spreading curves of magenta are 25% unprinted paper, 25% cyan, 25% yellow, 25% green and 0% for the 4 superposition conditions containing solid black. For four inks, 32 ink spreading curves are required.

Eqs. (2.7) weight for each ink halftone its ink spreading curves according to the underlying colorant surface coverages. However, computing the effective coverage of an ink requires the effective coverages of the other inks, which are not available. Because of the non-linearity

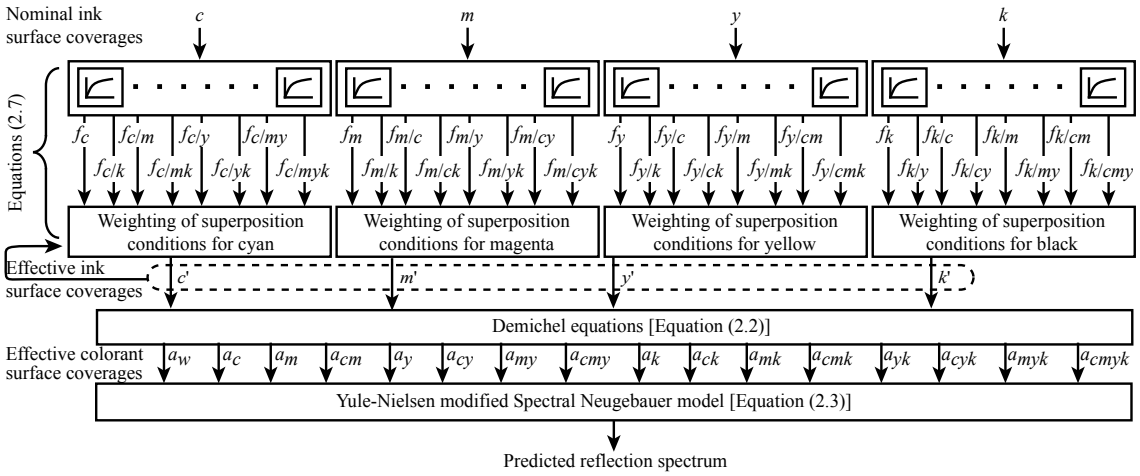


Figure 2.2: The ink spreading enhanced Yule-Nielsen modified Spectral Neugebauer model with nominal ink surface coverages c, m, y, and k; ink spreading curves $f_{i/jk}$ of ink i superposed with solid inks j and k; effective ink surface coverages c', m', y', and k'; and effective colorant surface coverages a_w to a_{cmyk}.

of the equations, there is no analytical solution to the equation system. The effective coverages must therefore be computed iteratively, starting with the nominal coverages. Usually, five iterations ensure sufficient convergence to determine the effective ink coverages.

2.4 The Ink Spreading Enhanced YNSN Model (IS-YNSN)

The spectral reflection prediction model used in the present contribution is the IS-YNSN model. It is the combination of the YNSN model with the ink spreading model presented in the previous section.

Spectral predictions using the IS-YNSN model are performed according to Figure 2.2. For given nominal ink surface coverages c, m, y, k, we obtain effective surface coverages c', m', y', k' by weighting the ink spreading curves $f_c(c)$ to $f_{k/cmy}(k)$ according to the surface coverages of the colorants contributing to that color halftone [Equations (2.7)]. With the Demichel equations, we then compute the corresponding effective surface coverages of the colorants forming that color halftone. With these effective colorant coverages, the YNSN model calculates the predicted reflection spectrum.

2.5 Review of the prior art

The YNSN model has been used by many researchers for the characterization of printing systems. Iino and Berns have used to it to characterize inkjet printers [IB98a] and offset presses [IB98b], while Taplin and Berns [TB01] or Chen et al. [CBT04] used it to characterize a six-color inkjet printer. A cellular version of the Neugebauer model was first proposed by Heuberger [HJP92]. Based on his approach, Agar and Allebach [AA98] proposed the cellular YNSN model (CYNSN) to improve the accuracy of the predictions, but at the cost of an increased complexity and calibration effort. Rolleston [RB93] and Balasubramanian [Bal95] have both reviewed this cellular approach. However, since the CYNSN model can be seen as multiple instances of the basic YNSN model, our focus is on the basic YNSN model.

A few researchers have proposed different approaches to calibrate the YNSN model. Xia et al. proposed a total least-squares optimization to account for the noise in the measurements of the Neugebauer primaries [XSSMT99], but the drawback of this approach is that the Neugebauer primaries must still be initially measured. Bala also proposed a similar approach in which an additional optimization is performed on the Neugebauer primaries after an initial calibration [Bal99]. Bala proposed yet another calibration method in which the Neugebauer primaries are not measured, but obtained by a least-squares optimization [Bal96]. This method requires a training set of measured calibration patches well distributed within the gamut of the printing device. This contrasts to our approach in which no prior knowledge about the Neugebauer primaries is necessary and no constraint on the calibration patches are imposed. Zuffi et al. proposed a calibration method based on neural and genetic algorithms [ZSM03, ZSS06], but they based their solution on an ink spreading model lacking consistency since thresholding is applied to avoid having total surface coverages larger than 1.

Spectral reflection prediction models predict spectra or colors from nominal ink coverages, but the inverse problem is a key component to many color management tasks. Mahy and Delabatista [MD96] or Urban and Grigat [UG06] have both proposed optimization to accurately solve this problem. The cellular approach is widely used in the literature for such color management tasks. Besides the works by Iino and Berns cited above, Tzeng and Berns [TB00] followed by Chen et al. [CBTI03] used the CYNSN model to maximize color constancy and minimize metamerism when printing with multiple inks. Urban later relaxed the metamerism constraints to allow for better reproduction under several illuminants [Urb09].

The Yule-Nielsen n-value has also been the focus of many papers. In their seminal article, Yule and Nielsen showed that the effect of light scattering in paper could be modeled using a

non-linear relationship based on the n-value [YN51]. Ruckdeschel and Hauser demonstrated that an n-value between 1 and 2 made physical sense within their framework and Pearson proposed to use an n-value of 1.7 in general situations. However, Berns et al. [BBT96] showed that large n-values regularly occur. Recently, Lewandowski et al. even proposed the use of negative n-values [LLBD06] and Viggiano evaluated the physical significance of these negative n-values [Vig06, Vig10]. Iino and Berns also proposed spectral dependent n-values, again in an attempt to improve the accuracy of the Yule-Nielsen model by increasing its complexity [IB98a][IB98b]. Considering the n-value to be wavelength dependent or negative has not gained much interest. We therefore only consider positive scalar n-values in this dissertation.

For any variation of the YNSN model to be reasonably accurate, it must take dot gain into account. Most of the time, dot gain is accounted for each ink separately through the use of tone reproduction curves [PP72], but Yule and Colt already mentioned that printing inks on top of each other modifies how the inks spread on paper [YC51]. Hersch et al. proposed an ink spreading model accounting for ink spreading in different superposition conditions [HECC05]. Although this ink spreading model uses more ink spreading curves than traditional approaches based on tone reproduction curves, each ink spreading curve is determined independently. Other approaches exist. For example, Bala et al. proposed 2D transforms to increase the amount of control over given regions of the color space [BSMVdC05]. Instead of interpolating the tone reproduction curves between calibration points or populating lookup tables, Viggiano [Vig83], Zuffi and Schettini [ZS04], or Yang et al. [YLK01] used mathematical functions describing the curves. In our case, starting from Chapter 4, we also describe ink spreading curves using mathematical functions, but we choose simple parabolas having a single degree of freedom. We have found that this approach yields sufficient accuracy and keeps the complexity of the model low, which is critical when introducing flexibility in the calibration process.

The Yule-Nielsen model plays a significant role in building color management systems. Its low complexity and high accuracy make it a highly practical approach in real-world settings. However, it does not try to model physical phenomena as they occur in paper. There exist many models describing these physical and optical phenomena. Kubelka and Munk described how light travels within turbid media [KM31, Kub48]. This Kubelka-Munk theory is extensively used in the textile or painting industry [Nob85]. Yang and Kruse proposed several enhancements to the Kubelka-Munk model to widen its field of use [YK04]. Clapper and Yule [CY53] or Williams and Clapper [WC53] described how light travels within paper and created models tuned to their assumptions. Emmel and Hersch [EH00] or Hébert and Hersch [HH06] created unified models whose particular cases are the Neugebauer, Kubelka-Munk, Clapper-Yule, or

Williams-Clapper models. However, these models are either not as accurate as the IS-YNSN model or too complex for our purpose.

Chapter 3

Improving the robustness of the IS-YNSN model

3.1 Introduction

Before making the calibration of the IS-YNSN model more flexible, the model itself must be stable and robust. If not, the introduced flexibility risks to exacerbate the instabilities of the model. There can be different sources of instabilities in a model. In the case of the IS-YNSN model, the 3 ink version has been extensively studied and no instability found. However, the extension to 4 inks, especially in the case of CMYK, introduces two instabilities that are not readily apparent.

The availability of the black ink creates two possibilities to print the color black, either using the black ink or superposing cyan, magenta, and yellow. This redundancy has an impact on the use of the inverted IS-YNSN model, i.e. when deducing the surface coverages of printed color halftone patches from their measured reflectance spectra. Since the inverted IS-YNSN model cannot accurately distinguish between the two different blacks, the deduced coverages are inaccurate. The ability to accurately deduce coverages from spectra plays an important role in color management. Section 3.2 shows how this first source of inaccuracies can be prevented by extending the measurements from the visible domain to the near-infrared domain (NIR).

The black ink is used because of its strong absorbance of the light in all the visible spectrum. This absorbtion creates a strong black color, whereas the superposition of cyan, magenta, and yellow looks brownish or greyish. The gamut of CMYK printers is greatly enhanced thanks to this new ink. However, printing any ink on top of solid black yields black. The spectral

difference between the two colors is extremely small and the information contained in this difference cannot be distinguished from noise. Since certain ink spreading curves attempt to model this difference, their calibration is not stable. Moreover, since the difference they model is not significant, they do not improve the prediction accuracy of the IS-YNSN model. Section 3.3 proposes a simplified ink spreading directive and shows that the predictions performed with this new directive are at least as accurate as predictions performed with the *top or below* directive (Eq. (2.7)).

The flexible calibration of the IS-YNSN model presented in Chapters 4 and 5 rely on this increased robustness. Thanks to the improved accuracy of the deduced coverages, the information contained in the measurements performed on printed color images is more efficiently extracted; and thanks to the simplified ink spreading directive, there is no useless parameter to calibrate.

3.2 Model-Based Deduction of CMYK Surface Coverages from Visible and Infrared Spectral Measurements of Halftone Prints

The IS-YNSN model allows predicting reflectance spectra in function of the ink surface coverages. Such predictions, further referred to as *forward predictions*, can be accurately performed for four or more inks. Inverting the models, i.e. deducing the ink surface coverages from a given printed color patch, is an important building stone for applications such as optimizing GCR parameters or calibrating the ink spreading model. Working with the four cyan, magenta, yellow, and black inks, the IS-YNSN model cannot easily perform *inverse* or *backward predictions*, i.e. deduce the ink surface coverages from a given printed color patch. We show why it is hard to perform inverse predictions with CMYK inks and how we can improve such predictions using spectra spanning both the visible and infrared domain.

3.2.1 Measurement setup

The measurements used for the experiments are made on color patches printed on newspaper using a web offset press. Classical rotated halftone screens are printed at 100 lpi. The colors printed are all the combinations of cyan, magenta, and yellow inks at surface coverages of 0%, 25%, 50%, 75%, or 100% and black ink at surface coverages of 0%, 10%, 30%, 50%,

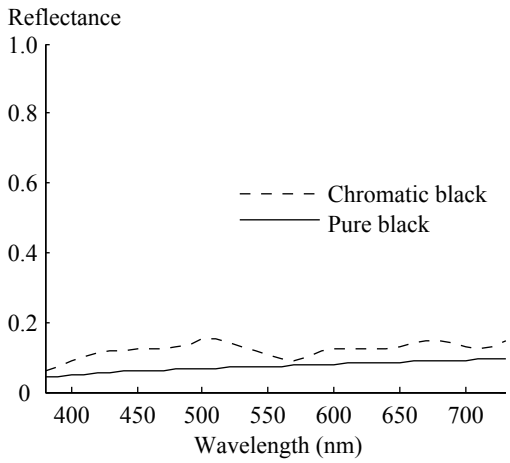


Figure 3.1: Reflection spectra of the chromatic and pure black printed colors.

80%, or 100%, creating a set of 750 different colors. The reflectance spectra are measured with a Datacolor MF-45 spectrophotometer (geometry: 45°d:0°, i.e. 45° directed incident light and capture at 0°) at 10 nm intervals between 380 nm and 850 nm.

3.2.2 Limitations of the visible domain

With the rods and cones, the human eye perceives light whose wavelength ranges from 400 nm to 700 nm. The cones are responsible for the perception of colors. There are three types of cones: the L, M and S cones which have a high sensibility in the red, green and blue wavelength ranges, respectively. Many printing systems therefore use cyan, magenta and yellow inks to reproduce colors because they absorb red, green and blue, respectively. This enables printing systems to reproduce a large part of the colors that the human visual system can see. Since each ink absorbs in a different spectral region, it is relatively easy with the IS-YNSN model to deduce the surface coverages of inks needed to reproduce a given color [Eq. (2.4)].

In printed systems, the black ink is used to increase the gamut and to reduce the amounts of superposed cyan, magenta, and yellow inks. By performing gray component replacement (GCR), a given amount of black replaces the superposition of similar amounts of cyan, magenta, and yellow inks. One may create the black color either by superposing the cyan, magenta, and yellow inks, referred to as *chromatic black*, or by using the black ink, referred to as *pure black*. The spectra of the two black colors are shown in Figure 3.1. We can see that the two spectra do not yield the same color, i.e. the same black. For the considered web offset press technology, pure black is darker than chromatic black. Moreover, its spectrum is flatter, which

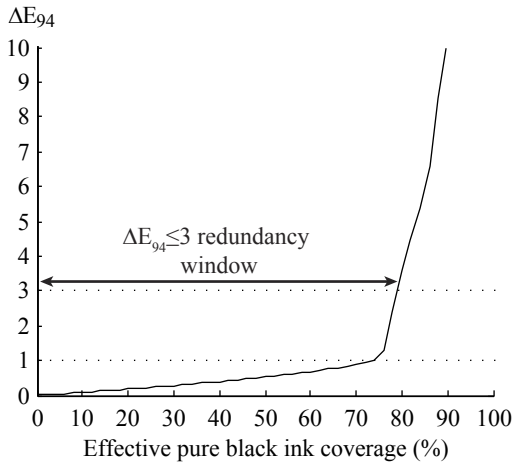


Figure 3.2: ΔE_{94} difference between the target 100% chromatic black and its closest color that can be reproduced using a given amount of pure black and appropriate amounts of chromatic black. The ΔE_{94} difference is in function of the amount of pure black. The $\Delta E_{94} \leq 3$ redundancy window is also drawn.

ensures that it is really black and not slightly colored.

Spectral prediction models can easily predict the reflection spectrum and therefore the resulting color given nominal ink surface coverages of the cyan, magenta, yellow and black inks. On the other hand, reversing the model when cyan, magenta, yellow and black inks are used, i.e. deducing the four effective ink surface coverages for a reflection spectrum, is much harder because of the redundancy between pure and chromatic black. It is indeed possible to achieve similar colors by replacing chromatic black with pure black and vice versa.

Let us show the redundancy between pure and chromatic black. In a first experiment, we try to reproduce 100% chromatic black, given by the superposition of solid cyan, solid magenta, and solid yellow, by using a given amount of pure black and appropriate amounts of chromatic black. We first choose the amount of pure black and then use the IS-YNSN spectral prediction model to deduce the amounts of cyan, magenta, and yellow inks reproducing the closest possible color to the desired chromatic black. We use the ΔE_{94} difference metric to evaluate the match [Gau03]. With a ΔE_{94} difference below 1, there is no noticeable difference for the human visual system and the match is perfect. With a ΔE_{94} difference below 3, the color match is reasonable and corresponds to the current practice in web offset printing. Figure 3.2 shows that chromatic black can be reproduced with up to 75% of pure black, i.e. black ink, while keeping the ΔE_{94} difference below 1, i.e. without any noticeable difference for the human visual system. Moreover, we can use between 0% and 80% of pure black and keep the ΔE_{94} difference below 3. This interval is called the $\Delta E_{94} \leq 3$ redundancy window. In this experiment, the $\Delta E_{94} \leq 3$ redundancy window ranges from 0% to 80% and its width is 80%.

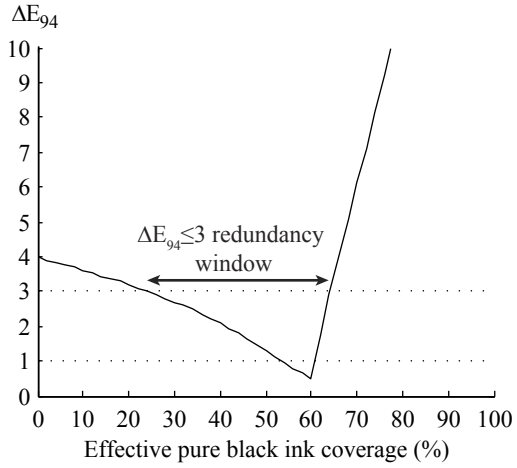


Figure 3.3: ΔE_{94} difference between the target 60% pure black and its closest color that can be reproduced using a given amount of pure black and appropriate amounts of chromatic black. The ΔE_{94} difference is in function of the amount of pure black. The $\Delta E_{94} \leq 3$ redundancy window is also drawn.

The next experiment is performed on a reference color patch composed of 60% pure black. Again, we try to reproduce this color by first choosing the amount of pure black and then using the IS-YNSN spectral prediction model to deduce the amounts of cyan, magenta, and yellow inks that reproduce the closest possible color to 60% pure black. Figure 3.3 shows that we have a $\Delta E_{94} \leq 3$ redundancy window ranging from 25% to 65% pure black. The color composed of 60% pure black can therefore be reproduced using only 25% pure black, which means that 35% pure black has been replaced by chromatic black.

3.2.3 Patches respecting the gray component replacement (GCR) rule

Gray component replacement (GCR) is an important tool for color reproduction. It enables reducing the amount of superposed inks and ensures an improved reproduction of gray shades. Since GCR replaces given amounts of cyan, magenta and yellow inks by the black ink, only a subset of the entire CMYK color space is used for printing. In this section, we therefore perform the experiments both on the entire CMYK color space using the 750 colors selected in Section 3.2.1 and on a subset of these colors respecting the GCR rule. Understanding if the IS-YNSN model behaves differently when GCR is applied is important since the methods described in Chapters 4 and 5 rely on information extracted from printed color images.

Gray component replacement (GCR) is based on the redundancy between pure and chromatic black. It replaces a certain amount of chromatic black by pure black in order to limit the size of the areas covered by three superposed inks. GCR algorithms may replace all the chromatic black by pure black or only part of it. When GCR is applied, a given amount of pure

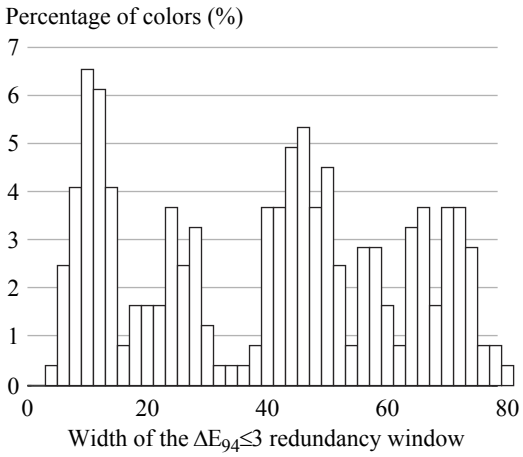


Figure 3.4: Percentage of colors in function of their $\Delta E_{94} \leq 3$ redundancy window.

black replaces a similar amount of chromatic black, i.e. similar amounts of cyan, magenta and yellow inks. Assuming that maximum GCR is applied, the sum of the pure black ink surface coverage and the maximum surface coverage of the cyan, magenta and yellow inks of any given patch does not exceed 100%. Out of the 750 colors selected in Section 3.2.1, 245 colors match this criterion.

Figure 3.4 shows the percentage of these 245 colors in function of the width of the $\Delta E_{94} \leq 3$ redundancy window. We can see that one color out of four have a $\Delta E_{94} \leq 3$ redundancy window greater than 60% and one out of two a $\Delta E_{94} \leq 3$ redundancy window greater than 40%. When the $\Delta E_{94} \leq 3$ redundancy window of a color is wide, the effective ink surface coverages deduced by the IS-YNSN model for the given color are less accurate because the model is not able to distinguish between pure and chromatic black. Even if the deduced ink surface coverages reproduce the desired color for the human visual system, the low prediction accuracy prevents recovering the exact ink surface coverages.

3.2.4 Extending the spectral prediction model to the infrared wavelength range

The experiments performed in Section 3.2.2 show that it is not possible to accurately distinguish between chromatic and pure black in the visible wavelength range only, i.e. the wavelength range between 380 nm and 730 nm. The goal is therefore to find a method that can determine unambiguously the amount of black ink. Once this amount is known, one can then infer the amounts of the cyan, magenta and yellow inks.

Experiments have shown that cyan, magenta, yellow, and dye-based black inks do not ab-

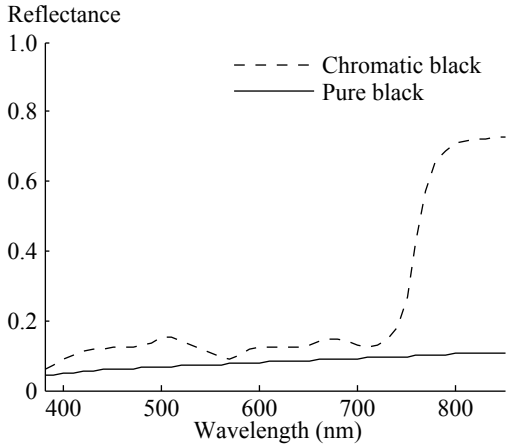


Figure 3.5: Reflection spectra of chromatic black and pure black prints including the near-infrared domain.

sorb outside the visible domain, whereas pigmented black inks do. This is interesting because, in reproduction systems such as offset printing, the cyan, magenta, and yellow inks are based on dyes whereas the black ink is based on pigments. The pigmented black ink is the only ink that absorbs in the near-infrared wavelength range, i.e. from 730 nm to 1100 nm. Since the measured reflection spectra remain flat between 800 nm and 1100 nm, we consider a wavelength range between 380 nm and 850 nm, which includes the specific yellow (380-490 nm), magenta (500-610 nm), cyan (620-730 nm) and pure black (740-850 nm) absorption zones.

Figure 3.1 shows the difference between chromatic and pure black in the visible domain only. Figure 3.5 shows the same spectra, but extended to the near-infrared domain. The difference is obvious and allows us to clearly discriminate the two types of black. It is therefore possible to determine the amount of pure black ink by also considering the near-infrared wavelength region. After recovering the pure black surface coverage, the surface coverages of the cyan, magenta, and yellow inks can be recovered by considering the visible region only.

Extending the model to the near-infrared domain does not require many modifications since the YNSN equation, i.e. Eq. (2.3), is independent of the number of wavelength components. The IS-YNSN model extended to the near-infrared domain is further referred to as the *ink spreading and near-infrared enhanced Yule-Nielsen modified Spectral Neugebauer model* (NIR-YNSN).

We evaluate the benefits brought by the near-infrared wavelength range as follows: Given a measured spectrum, we fit the effective ink surface coverages that reproduce this spectrum as closely as possible. These backward fitting predictions are performed using either the IS-YNSN model or the NIR-YNSN model. Then, using the NIR-YNSN model, we predict the reflection

Table 3.1: Average, maximum and standard deviation of the root mean square errors (RMSE) of the 750 colors selected in Section 3.2.1 when fitting surface coverages in the visible domain only or in the visible and near-infrared domains.

	RMSE		
	Mean	Max	Stdev
Test in visible domain only			
IS-YNSN coverage fit	0.0097	0.0839	0.0101
NIR-YNSN coverage fit	0.0114	0.0788	0.0098
Test in visible and near-infrared domains			
IS-YNSN coverage fit	0.0242	0.1700	0.0234
NIR-YNSN coverage fit	0.0104	0.0689	0.0086

spectrum from the deduced effective coverages. We finally compute two different root mean square errors (RMSE) between the measured and predicted spectra. The first RMSE is computed considering for the test set only the visible wavelength range, i.e. between 380 nm and 730 nm. The second RMSE is computed considering for the test set the visible and near-infrared wavelength range, i.e. between 380 nm and 850 nm. Using this method for the 750 colors selected in Section 3.2.1, we collect two sets of 750 RMSE for each model. The results are summarized in Table 3.1.

When computing the RMSE in the visible domain only, both the IS-YNSN and NIR-YNSN coverage fits provide low errors, i.e. the measured and predicted spectra are closely matched. When computing the RMSE in both the visible and near-infrared domains, the RMSE increases by a factor of 2.4 for the IS-YNSN coverage fit and remains constant for the NIR-YNSN coverage fit. The increase is caused by the lack of accuracy of the IS-YNSN surface coverage fit when performing backward predictions. The effective ink surface coverages that it finds match the measured spectrum in the visible domain, but are not accurate enough to match the measured spectrum in the near-infrared domain. Extending the considered wavelength range to the near-infrared domain by applying the NIR-YNSN surface coverage fit provides a greater accuracy for backward predictions, which is confirmed by the low RMSE both in the visible domain alone and in the visible and near-infrared domains.

We then perform the same experiment using only the 245 colors selected in Section 3.2.3 and summarize the results in Table 3.2. The considerations made for Table 3.1 also apply to Table 3.2. The only difference is that the maximum and standard deviation values drop significantly when considering only 245 colors. We explain this behavior in the next section in the light of another experiment.

Table 3.2: Average, maximum and standard deviation of the root mean square errors of the 245 colors selected in Section 3.2.3 when fitting surface coverages in the visible domain only or in the visible and near-infrared domains

	RMSE		
	Mean	Max	Stdev
Test in visible domain only			
IS-YNSN coverage fit	0.0083	0.0290	0.0038
NIR-YNSN coverage fit	0.0106	0.0325	0.0045
Test in visible and near-infrared domains			
IS-YNSN coverage fit	0.0251	0.1685	0.0243
NIR-YNSN coverage fit	0.0098	0.0300	0.0043

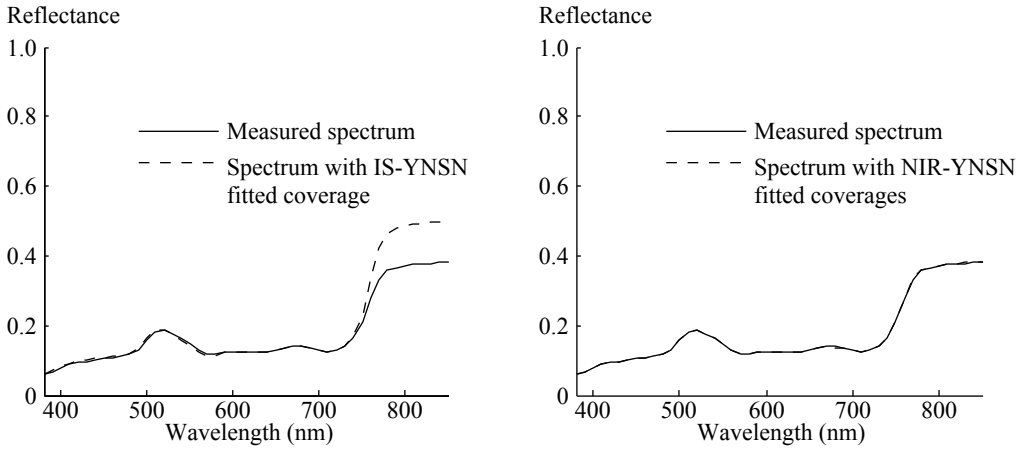


Figure 3.6: Comparison of the spectra resulting from backward coverage predictions using either the IS-YNSN or NIR-YNSN coverage fit. The measured spectrum corresponds to a color composed of 75% cyan, 25% magenta, 75% yellow, and 30% black nominal ink surface coverages.

In Figure 3.6, we show the resulting spectra when performing the above experiment with a color composed of 75% cyan, 25% magenta, 75% yellow and 30% black. On the left-hand side, the backward prediction is performed using the IS-YNSN model. We see that the predicted spectrum matches well the measured spectrum in the visible wavelength range, i.e. from 380 nm up to 730 nm, but diverges in the near-infrared wavelength range. Since only the black ink absorbs in the near-infrared domain, the effective black ink surface coverage is not accurate. On the other hand, if the backward coverage prediction is performed using the NIR-YNSN coverage fit, the predicted spectrum matches well the measured spectrum for the full wavelength range.

Table 3.3: Coverage differences between fitted and effective surface coverages for the IS-YNSN model and the NIR-YNSN model when considering the 750 colors selected in Section 3.2.1.

	RMSE		
	Mean	Max	Stdev
IS-YNSN surface coverage fits			
Δc	0.0254	0.2944	0.0394
Δm	0.0286	0.5227	0.0606
Δy	0.0536	0.7243	0.0813
Δk	0.0274	0.1912	0.0309
NIR-YNSN surface coverage fits			
Δc	0.0112	0.1138	0.0152
Δm	0.0184	0.0961	0.0212
Δy	0.0521	0.7013	0.0906
Δk	0.0162	0.1026	0.0207

3.2.5 Comparing forward and backward predictions

In the preceding section, we have shown that extending the IS-YNSN model to the near-infrared wavelength range yields more accurate backward surface coverage predictions. Such predictions allow the deduction of effective ink surface coverages from measured spectra. We now compare these backward predictions with forward predictions (see Figure 2.2). The comparison between forward and backward predictions is based on the differences between the effective ink surface coverages computed by forward predictions, or simply effective coverages, and the ones deduced by backward predictions, or *fitted* effective coverages.

We proceed as follows: First, the four fitted effective ink surface coverages of a given color halftone patch are deduced from the measured reflection spectrum of this color. These fitted coverages are then compared to the effective ink surface coverages computed by the ink spreading model. Finally, the respective absolute difference between fitted and effective coverages of cyan (Δc), magenta (Δm), yellow (Δy), and black (Δk) are computed. With the 750 colors selected in Section 3.2.1, we have 750 sets of differences. The results are summarized in Table 3.3 for the IS-YNSN and NIR-YNSN surface coverage fits. Note that the ink spreading curves of the IS-YNSN model are calibrated in the visible domain only and the ones of the NIR-YNSN model are calibrated in both the visible and near-infrared domains.

We can see that the effective ink surface coverage differences are lower when using the NIR-YNSN model than when using the IS-YNSN model. The NIR-YNSN model is therefore more consistent than the IS-YNSN model in regards to forward and backward predictions.

Table 3.4: Coverage differences between fitted and effective surface coverages for the IS-YNSN model and the IS-YNSN model when considering the 245 colors selected in Section 3.2.3.

	RMSE		
	Mean	Max	Stdev
IS-YNSN surface coverage fits			
Δc	0.0222	0.1854	0.0274
Δm	0.0191	0.1728	0.0258
Δy	0.0391	0.3363	0.0413
Δk	0.0240	0.1912	0.0330
NIR-YNSN surface coverage fits			
Δc	0.0075	0.0404	0.0083
Δm	0.0133	0.0851	0.0168
Δy	0.0197	0.1522	0.0243
Δk	0.0069	0.0671	0.0094

Moreover, since forward predictions are known to yield good results for both the IS-YNSN and the NIR-YNSN models, the fitted effective coverages deduced by the NIR-YNSN model are more accurate than the fitted effective coverages deduced by the IS-YNSN model.

It is also interesting to note that the differences in yellow effective coverages are significantly higher than the differences of the other inks. They also do not significantly decrease when fitting the surface coverages in the near-infrared domain. Possible reasons include the fact that yellow ink scatters light and that the superposition of yellow and other inks is not consistent across different halftones.

We then compute the coverage differences only for the 245 colors selected in Section 3.2.3 and summarize the results in Table 3.4. Again, we see that the NIR-YNSN model is more consistent than the IS-YNSN model since its coverage differences are lower. The yellow coverage differences are still higher than the differences of the other inks, but extending the IS-YNSN model to the near-infrared domain divides the yellow coverage differences by a factor of two. This tends to confirm the assumption that the superposition of yellow and other inks is not consistent across different halftones. The criterion used to select the 245 colors used in this experiment indeed limits the amount of superposed inks and therefore the inconsistencies caused by such superpositions.

Limiting the experiment to the 245 colors also affects the maximum and standard deviation values, which significantly drop. This is another confirmation of the previous assumption. Moreover, it explains why the maximum and standard deviation values in Table 3.2 are lower

than in Table 3.1. Since the models handle the 245 color set better than the 750 color set, the errors they make are more focused around the mean value.

3.2.6 Summary

Section 3.2 shows the difficulty in trying to deduce ink surface coverages due to the redundancy between pure and chromatic black when using the visible wavelength range only. Since the pigmented black ink absorbs in the near-infrared wavelength range whereas the cyan, magenta, and yellow inks do not, we propose to extend the measurements to the near-infrared domain. We then show that this extension improves the accuracy of the IS-YNSN model when deducing effective ink surface coverages. Finally, we show that the fitted effective ink coverages better agree with the effective ink surface coverages computed by the ink spreading model when the measured spectra are extended to the near-infrared domain.

Although some inconsistencies remain, applications such as the recovery of the gray component replacement strategies from halftone patches highly benefit from accurate backward surface coverage predictions. This accuracy is also important for the flexible calibration proposed in Chapters 4 and 5 in order to extract correct information directly from printed color images. Eschbach et al. also used a similar extension in the near-infrared domain in order to build anti-counterfeiting optical security features by taking advantage of the increased control provided by these measurements on the ambiguity between pure and chromatic black [EBMZ09].

In the remaining part of this dissertation, we always use spectra extended to the near-infrared domain for CMYK measurement sets. However, when working with CMY data, we use spectra limited to the visible domain. In both cases, we refer to the spectral reflection prediction model as the IS-YNSN model. The NIR-YNSN notation is strictly limited to Section 3.2.

3.3 Simplified Ink Spreading Equations for CMYK Halftone Prints

One of the fundamental aspects a spectral reflection prediction model has to consider is how the inks spread on paper, a phenomenon also referred to as physical or mechanical dot gain. With an ink spreading model accounting for physical dot gain, a spectral reflection prediction model is able to accurately predict reflectance spectra in function of ink surface coverages for three or four inks. In order to be effective, such an ink spreading model must take into account

not only the interaction between an ink halftone and paper, but also the interaction between an ink halftone and superposed inks.

One proposed solution is to use multiple ink spreading curves, as shown in Section 2.3, to characterize the physical dot gain of the ink halftones on paper and in all solid ink superposition conditions. For CMYK halftone prints, such an approach requires the characterization of 32 ink spreading curves. In this section, we evaluate the relevance of each ink spreading curve and subsequently simplify the ink spreading model in order to use fewer curves. The relevance of an ink spreading curve depends on its impact on the spectral prediction model accuracy as well as how it is affected by noise.

3.3.1 Measurement setup

We use two different measurement sets for the experiments. The first set, further referred to as the *PIXMA set*, is composed of the measurements of 625 patches printed on Canon MP-101 matte paper using the Canon PIXMA 4000 ink jet printer, classical rotated clustered dot screens at 150 lpi, and a resolution of 600 dpi. The 625 patches are composed of all combinations of cyan, magenta, yellow, and pigmented black at 0%, 25%, 50%, 75%, and 100% nominal surface coverages.

The second set, further referred to as the *CIEL set*, is composed of the measurements of 750 patches printed on newspaper using a web-offset press, classical rotated clustered dot screens at 100 lpi and a resolution of 1200 dpi. The 750 patches are composed of all combinations of cyan, magenta, and yellow at 0%, 25%, 50%, 75%, and 100% nominal surface coverages and black at 0%, 10%, 30%, 50%, 80%, and 100%.

The reflectance spectra are measured with a Datacolor MF-45 spectrophotometer (geometry: 45°d:0°, i.e. 45° directed incident light and capture at 0°) at 10 nm intervals between 380 nm and 850 nm.

3.3.2 Relevance of the different ink spreading curves

Using the *top or below* directive for the ink spreading model requires 32 different ink spreading curves (Section 2.3.2). Calibrating these ink spreading curves requires at least 32 spectral measurements. In order to reduce the number of ink spreading curves, we evaluate their relevance. If a specific ink spreading curve does not improve the accuracy of the spectral prediction model, we try to discard it.

The following experiment is performed both on the PIXMA and CIEL sets. It first calibrates

the spectral prediction model, including the 32 ink spreading curves. Each ink spreading curve is calibrated using a single calibration patch at 50% nominal ink surface coverage. Then, we successively modify each ink spreading curve by setting the effective coverage at 50% to a value between 0% and 100%. For each modification, we predict the spectra of the patches in the measurement set and compute the average ΔE_{94} between the measured and predicted spectra. Each modification is therefore associated to an average ΔE_{94}. The resulting curves are called surface coverage accuracy curves.

Figures 3.7 and 3.8 show the surface coverage accuracy curves for the PIXMA set (ink jet) and CIEL set (offset), respectively. The circles indicate the calibration points, i.e. prior to modifications, and the crosses indicate the effective coverages at 50% yielding the minimum average ΔE_{94}. All the surface coverage accuracy curves have the shape of convex parabolas, but not all parabolas have similar curvatures. Some are nearly flat, especially the ink spreading curves whose superposition condition includes solid black, i.e. the last four surface coverage accuracy curves of the first three lines of Figures 3.7 and 3.8. We further refer to these 12 ink spreading curves as the ink spreading curves on solid black.

When observing the surface coverage accuracy of ink spreading curves which do not include solid black, we make the following observations. The effective coverages at 50% chosen during calibration (circles) agree well with the minima (crosses). Moreover, the surface coverage accuracy curves are highly relevant to the accuracy of the prediction model because of their curvature. If the effective coverage at 50% of one of these ink spreading curves is set too far off the minimum, the average ΔE_{94} becomes large. Finally, we can see that the effective coverage at 50% of a given ink halftone yielding the minimum ΔE_{94} is different depending on the superposition condition. This emphasizes the fact that the ink spreading of an ink halftone depends on the presence of other inks.

The opposite analysis is true for the ink spreading curves on solid black. The calibrated effective surface coverages do not agree well with the computed minima. They are not relevant for the accuracy of the prediction model because the curvature of the parabolas is low, i.e. a modification of an effective halftone surface coverage at 50% nominal coverage on solid black does not deteriorate the accuracy of the prediction model. The system is underdetermined and it seems preferable to discard the ink spreading curves on solid black.

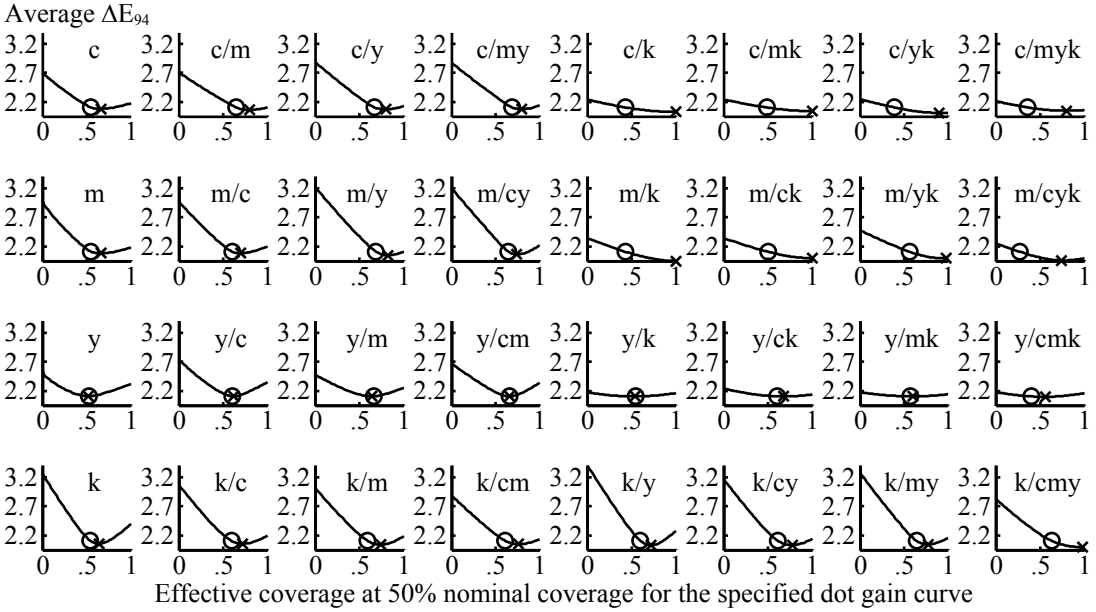


Figure 3.7: Surface coverage accuracy curves showing the impact of dot gain variations computed for the PIXMA measurement set (inkjet, 150 lpi, Yule-Nielsen n-value $= 10$). The horizontal axis denotes the effective surface coverage at 50% nominal surface coverage and the vertical axis the prediction accuracy in terms of average ΔE_{94} difference between measured and predicted spectra.

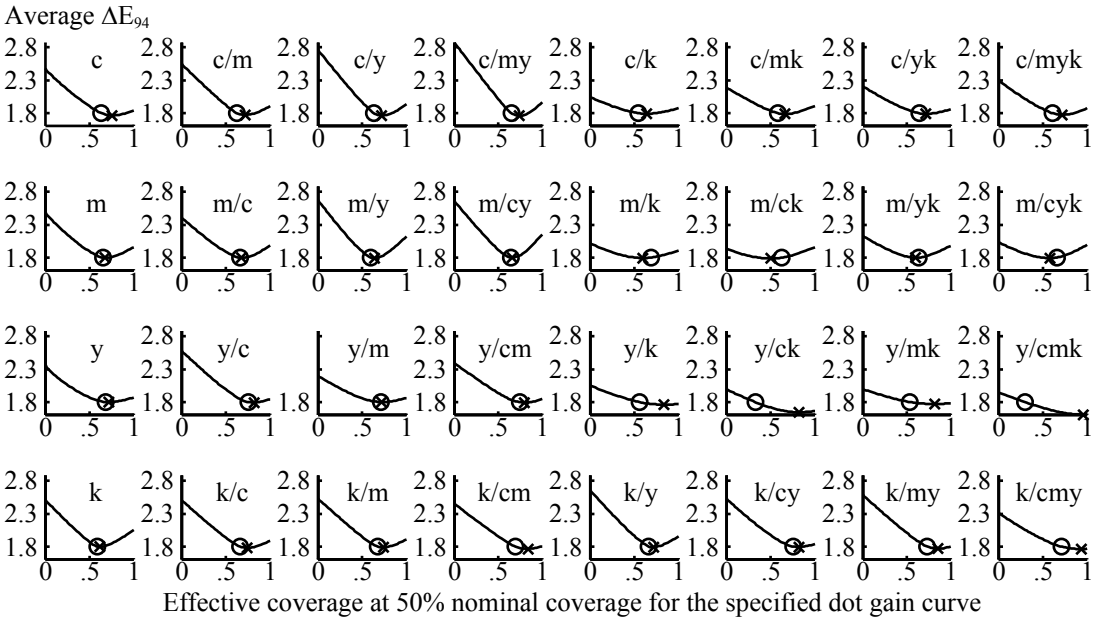


Figure 3.8: Surface coverage accuracy curves showing the impact of dot gain variations computed on the CIEL measurement set (offset, 100 lpi, Yule-Nielsen n-value $= 7.8$). The horizontal axis denotes the effective surface coverage at 50% nominal surface coverage and the vertical axis the prediction accuracy in terms of average ΔE_{94} difference between measured and predicted spectra.

3.3.3 Resilience of the ink spreading curves to noise

The ink spreading curves on black do not improve the accuracy of the spectral prediction model. Let us further show that they are not only irrelevant, but also highly subject to noise. There are indeed several possible origins of noise in measured spectra. Noise comes from the printing device itself, from dirty patches, from variations in the intrinsic reflectance of the substrate, from variations in the printing conditions, etc. These different sources produce each a different kind of noise, making them difficult to separate and model. We therefore use Gaussian noise and analyze its impact on the calibration of the ink spreading curves.

In the following experiment, we observe how stable the calibration of the ink spreading curves is in the presence of Gaussian noise added to each wavelength of the measured reflectance with both average and standard deviation between 0 and 0.05. For each ink spreading curve and each noise average, we calibrate 500 times the effective coverage of the 50% patch as follows: We first choose randomly the standard deviation of the noise between 0 and 0.05. We generate the noise independently for each wavelength and add it to the measured spectrum of the 50% patch. The ink spreading curve is established using this measured noise augmented reflectance spectrum. For each ink spreading curve and each noise average, we have 500 different fitted effective coverages.

Figures 3.9 and 3.10 show the results for the PIXMA (ink jet) and CIEL (offset) measurement sets, respectively. Both figures present the same behavior: The effective surface coverages superposed with solid black are very sensitive to noise, i.e. the same noise induces very large effective surface coverage calibration variations. On the other hand, the other effective surface coverages are mainly subject to the mean value of the Gaussian noise. As expected, when adding noise with a positive mean, the measured spectra become more highlight, as if they were printed with less ink. The IS-YNSN model compensates for this positive difference by using lower effective surface coverages.

The problem caused by the ink spreading curves on solid black is underlined in Figures 3.11 and 3.12. In Figure 3.11, we focus on the ink spreading curve $f_{m/ck}$ of magenta over solid cyan and solid black for the CIEL measurement set. The reflection spectrum without magenta corresponds to the beginning of the ink spreading curve (0% coverage of magenta). The spectrum with solid magenta corresponds to the end of the ink spreading curve (100% coverage of magenta). We see that the difference between the two spectra is not significant. Therefore, the ink spreading curve of magenta over solid cyan and black has a negligible influence when predicting reflection spectra. Moreover, noise has a large impact on the calibration of the ink spreading

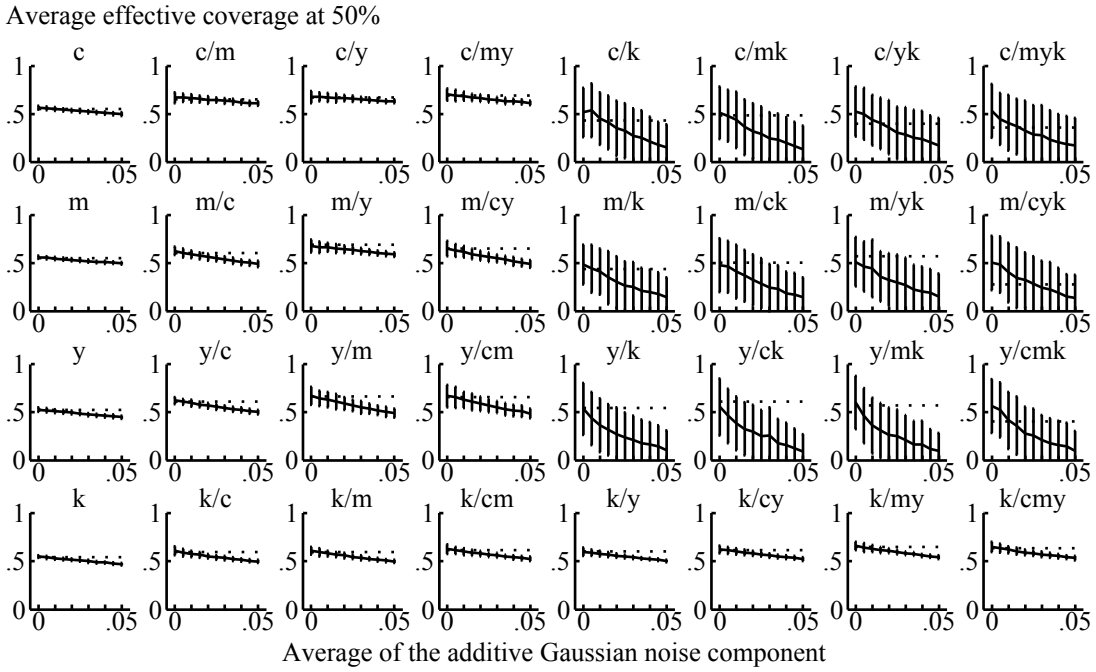


Figure 3.9: Impact of noise on the calibration of the ink spreading curves computed for the PIXMA measurement set. The x–axis shows the average of the additive Gaussian noise and y–axis the average and standard deviation of the effective coverage fitted at 50%.

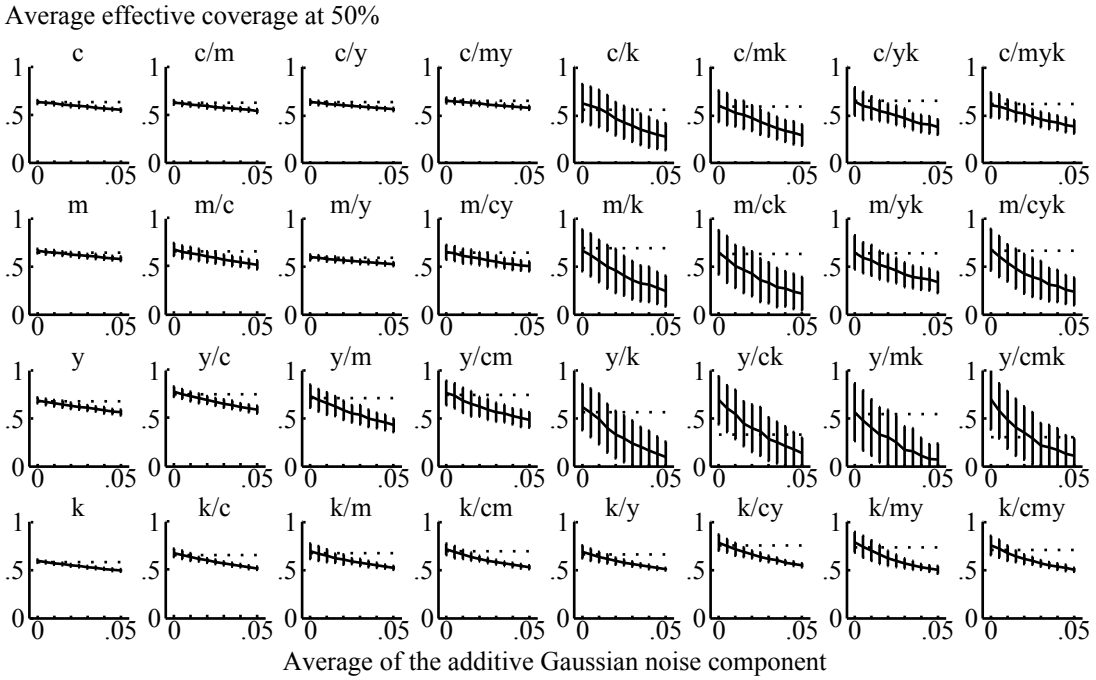


Figure 3.10: Impact of noise on the calibration of the ink spreading curves computed for the CIEL measurement set. The x–axis shows the average of the additive Gaussian noise and y–axis the average and standard deviation of the effective coverage fitted at 50%.

curve because small variations in the 50% magenta coverage spectrum must be compensated by large variations in the ink spreading curve.

The same reasoning applies to Figure 3.12 when considering the ink spreading curve $f_{m/k}$ of magenta of solid black (PIXMA set). The result of the calibration of the ink spreading curves for the PIXMA measurement set is shown in Figure 3.13. This figure shows the dot gain curves in function of nominal surface coverage. Dot gain is defined as the difference between effective and nominal surface coverage. All the dot gain curves except those on solid black show a consistent behavior. Conversely, we can see the effect of measurement noise on the dot gain curves superposed with solid black: The curves behave randomly.

3.3.4 Removing the ink spreading curves on solid black

In the previous two sections, we have shown that the ink spreading curves on solid black do not improve the accuracy of the spectral prediction model and are not resilient to noise. We therefore propose a new directive for the ink spreading model called *halftone black* that discards these 12 ink spreading curves, yielding the following set of equations:

$$
\begin{aligned}
c' &= (1-m')(1-y')f_c(c) \\
&+ \quad m' \quad (1-y')f_{c/m}(c) \\
&+ (1-m') \quad y' \quad f_{c/y}(c) \\
&+ \quad m' \quad\quad y' \quad f_{c/my}(c)
\end{aligned}
\qquad
\begin{aligned}
y' &= (1-c')(1-m')f_y(y) \\
&+ \quad c' \quad (1-m')f_{y/c}(y) \\
&+ (1-c') \quad m' \quad f_{y/m}(y) \\
&+ \quad c' \quad\quad m' \quad f_{y/cm}(y)
\end{aligned}
$$

$$
\begin{aligned}
m' &= (1-c')\,(1-y')f_m(m) \\
&+ \quad c' \quad (1-y')f_{m/c}(m) \\
&+ (1-c') \quad y' \quad f_{m/y}(m) \\
&+ \quad c' \quad\quad y' \quad f_{m/cy}(m)
\end{aligned}
\qquad
\begin{aligned}
k' &= (1-c')(1-m')\,(1-y')\,f_k(k) \\
&+ \quad c' \quad (1-m')\,(1-y')\,f_{k/c}(k) \\
&+ (1-c') \quad m' \quad\;\; (1-y')\,f_{k/m}(k) \\
&+ \quad c' \quad\quad m' \quad\;\; (1-y')\,f_{k/cm}(k) \\
&+ (1-c')(1-m') \quad y' \quad f_{k/y}(k) \\
&+ \quad c' \quad (1-m') \quad y' \quad f_{k/cy}(k) \\
&+ (1-c') \quad m' \quad\quad y' \quad f_{k/my}(k) \\
&+ \quad c' \quad\quad m' \quad\quad y' \quad f_{k/cmy}(k)
\end{aligned}
\tag{3.1}
$$

Using the *halftone black* directive, it is possible to break the computation of the effective coverages into two steps: First, we compute the effective coverages for cyan, magenta, and yellow using the iterative method proposed for the *top or below* directive. Then, we directly compute the black effective coverage using the effective coverages of the other inks.

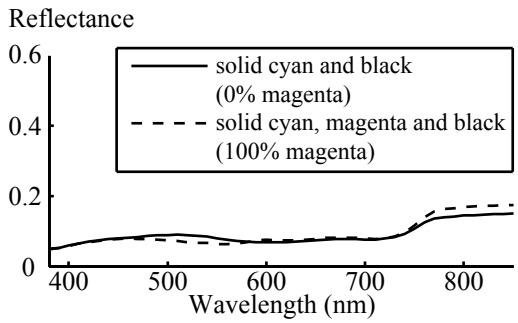


Figure 3.11: Spectra of solid cyan superposed with solid black (continuous line) and of solid cyan superposed with solid magenta and black (dashed line) from the CIEL measurement set.

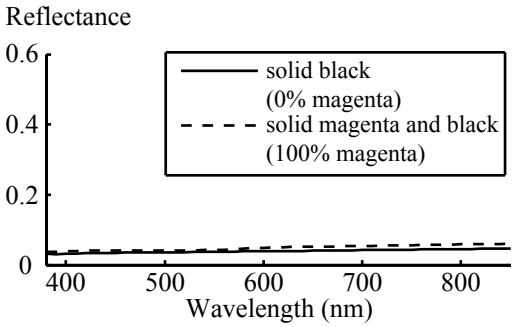


Figure 3.12: Spectra of solid black (continuous line) and of solid magenta superposed with solid black (dashed line) from the PIXMA measurement set.

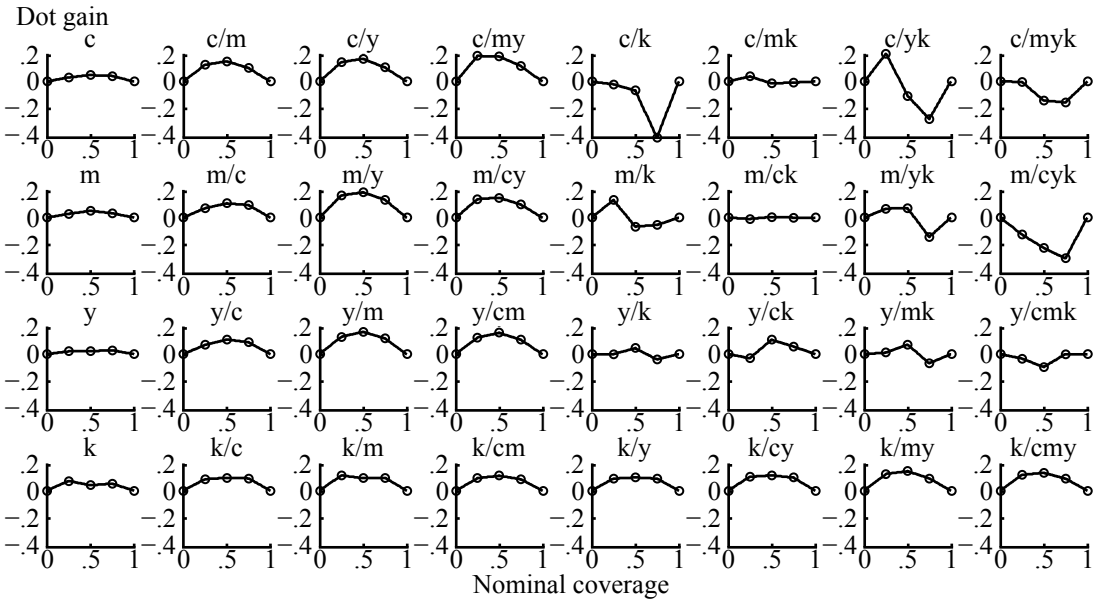


Figure 3.13: Dot gain curves, derived from the corresponding ink spreading curves, calibrated at 25%, 50%, and 75% nominal coverages on the PIXMA measurement set. The horizontal axis corresponds to the nominal surface coverage of the ink halftone and the vertical axis to the dot gain.

Table 3.5: Accuracy of the spectral prediction model on the PIXMA measurement set. For each of the 625 patches of the measurement set, we compute the ΔE_{94} between the predicted and measured spectra. The statistics show the average ΔE_{94}, the corresponding standard deviation, 95% percentile, and maximum ΔE_{94}. The Yule-Nielsen n-value is $n = 10$.

Directive	ΔE_{94}			
	Avg	Stdev	95%	Max
No ink spreading	4.29	2.70	1.08	11.90
Single	3.16	2.00	6.31	9.62
Top	2.28	1.69	5.50	8.32
Top or below	2.10	1.40	4.59	7.10
Halftone black	1.58	1.08	3.57	5.65

Table 3.6: Accuracy of the spectral prediction model on the CIEL measurement set. For each of the 750 patches of the measurement set, we compute the ΔE_{94} between the predicted and measured spectra. The statistics show the average ΔE_{94}, the corresponding standard deviation, 95% percentile, and maximum ΔE_{94}. The Yule-Nielsen n-value is $n = 7.8$.

Directive	ΔE_{94}			
	Avg	Stdev	95%	Max
No ink spreading	4.76	2.27	8.13	10.38
Single	2.20	1.11	4.03	5.27
Top	1.46	0.83	3.15	4.01
Top or below	1.47	0.82	3.03	3.99
Halftone black	1.44	0.83	3.08	3.98

The accuracy of the *halftone black* directive and the different ink spreading directives proposed in Section 2.3.2 are compared in Table 3.5 for the PIXMA measurement set and in Table 3.6 for the CIEL measurement set. On the PIXMA measurement set, the halftone black directive yields the best accuracy while the top or below directive suffers from the measurement noise as seen in Figure 3.13. Note that the *top* directive is shown even if it does not make sense for Inkjet printers since all the inks are printed at the same time. On the CIEL measurement set, both directives yield similar results. Considering that the patches are printed on newsprint paper, which is non-coated and less uniform than traditional paper, the achieved accuracy is remarkable. We still recommend using *halftone black*. It uses only 20 ink spreading curves instead of 32, reduces the calibration effort and decreases the time required to perform a spectral prediction.

3.3.5 Case study: silver ink

The *halftone black* directive is specific to CMYK printing devices. When other inks are used, each ink spreading curve must be analyzed to determine if it is relevant for the prediction accuracy and if it is resilient to measurement noise. The ink spreading equations must then be updated to include only the relevant ink spreading curves.

We apply this technique to the case of Offset printing with cyan, magenta, yellow, and silver inks (CMYS). This combination of inks is used in the *MetallicArt* technology to hide patterns in images [HCE03]. The silver ink behaves differently than the black ink. Because of its light gray appearance, overprinting other inks on solid silver creates significantly different shades. The ink spreading curves of inks on solid silver cannot be discarded as in the CMYK case. In order to determine the relevance of each ink spreading curve, we examine their surface coverage accuracy curves and their resilience to noise.

The surface coverage accuracy curves are shown in Figure 3.14. All the cyan, magenta, and yellow ink spreading curves behave similarly. They are optimally calibrated, i.e. the circles agree well with the crosses, and they are relevant to the prediction accuracy since the curvature of the parabolas is significant. Not all these ink spreading curves have the same importance, e.g. the $f_{y/ms}$ ink spreading curve is not perfectly calibrated and its surface coverage accuracy curve is flatter than the others. However, it remains much more important than the silver ink spreading curves. Except for the f_s ink spreading curve, all the silver ink spreading curves are nearly flat. They do not improve the prediction accuracy of the IS-YNSN model and are therefore not relevant.

The resilience to noise of these ink spreading curves is shown in Figure 3.15. We see that the silver ink spreading curves are not as resilient as the ink spreading curves of the other inks, again at the exception of the f_s ink spreading curve. The variance of silver ink spreading curves is twice as large as the other curves. Note the behavior of the $f_{s/cm}$ and $f_{s/cmy}$ ink spreading curves when adding a positive noise average. Due to the metallic particles composing the silver ink, a small part of the specular reflection is measured by spectrophotometers, resulting in the solid cyan and magenta spectrum being darker than the solid cyan, magenta, and silver spectrum. This is observed only for the two darkest colorants, i.e. CM and CMY, because the measured specular component is too small to impact the other colorants. The IS-YNSN model therefore interprets the positive noise average raising the spectrum as an increase of the silver effective coverage.

Since all the silver ink spreading curves at the exception of the f_s ink spreading curve are

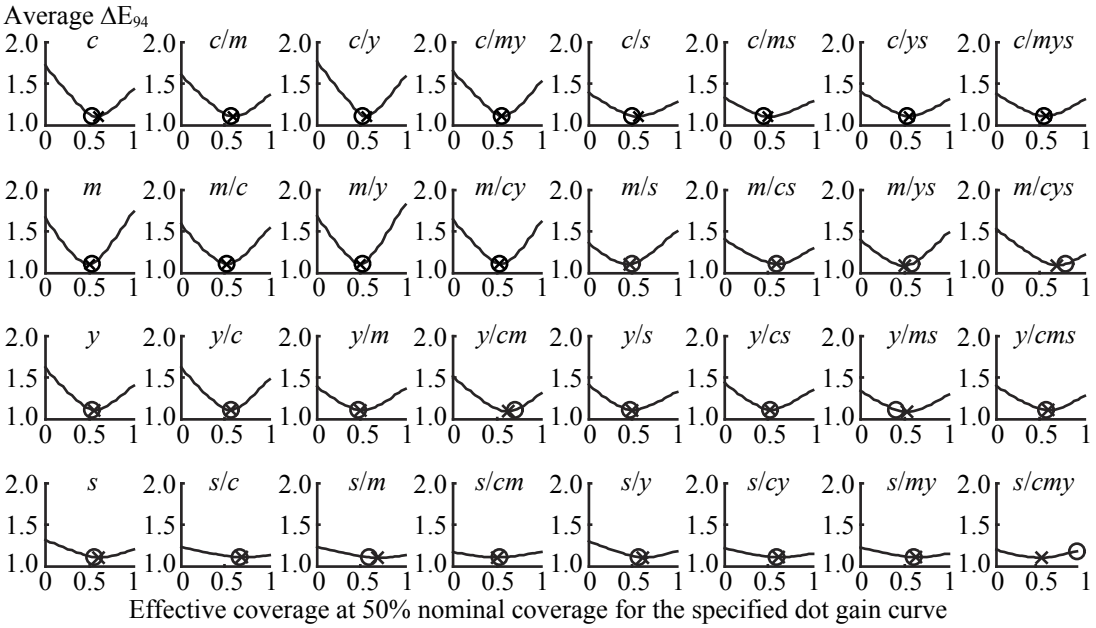


Figure 3.14: Surface coverage accuracy curves showing the impact of dot gain variations computed for the CMYS measurement set (offset, 150 lpi, Yule-Nielsen n-value $= 1.5$). The horizontal axis denotes the effective surface coverage at 50% nominal surface coverage and the vertical axis the prediction accuracy in terms of average ΔE_{94} between measured and predicted spectra.

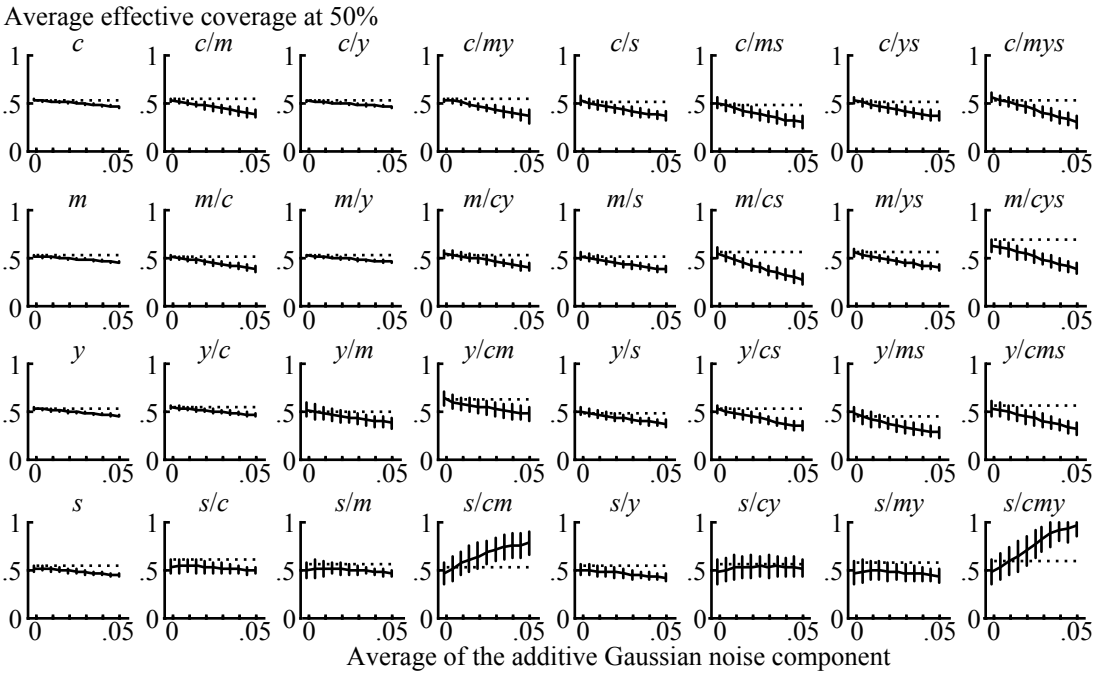


Figure 3.15: Impact of noise on the calibration of the ink spreading curves computed for the CMYS measurement set. The x–axis shows the average of the additive Gaussian noise and y–axis the average and standard deviation of the effective coverage fitted at 50%.

Table 3.7: Accuracy of the spectral prediction model on the CMYS measurement set. For each of the 750 patches of the measurement set, we compute the ΔE_{94} between the predicted and measured spectra. The statistics show the average ΔE_{94}, the corresponding standard deviation, 95% percentile, and maximum ΔE_{94}. The Yule-Nielsen n-value is $n = 1.5$.

Directive	ΔE_{94}			
	Avg	Stdev	95%	Max
No ink spreading	1.57	0.76	2.77	4.08
Single	1.20	0.68	2.53	3.92
Top	1.27	0.87	2.92	6.01
Top or below	1.11	0.71	2.59	3.85
Silver ink halftone	1.13	0.70	2.60	3.85

neither relevant for the prediction accuracy nor resilient to noise, we propose the following ink spreading directive, further referred to as the *silver ink halftone* directive, that discards these irrelevant ink spreading curves:

$$
\begin{aligned}
c' &= (1-m')(1-y')(1-s')\,f_c(c) \\
&+ \quad m' \quad (1-y')(1-s')\,f_{c/m}(c) \\
&+ (1-m') \quad y' \quad (1-s')\,f_{c/y}(c) \\
&+ \quad m' \quad\quad y' \quad (1-s')\,f_{c/my}(c) \\
&+ (1-m')(1-y') \quad s' \quad f_{c/s}(c) \\
&+ \quad m' \quad (1-y') \quad s' \quad f_{c/ms}(c) \\
&+ (1-m') \quad y' \quad\quad s' \quad f_{c/ys}(c) \\
&+ \quad m' \quad\quad y' \quad\quad s' \quad f_{c/mys}(c)
\end{aligned}
\qquad
\begin{aligned}
y' &= (1-c')(1-m')(1-s')\,f_y(y) \\
&+ \quad c' \quad (1-m')(1-s')\,f_{y/c}(y) \\
&+ (1-c') \quad m' \quad (1-s')\,f_{y/m}(y) \\
&+ \quad c' \quad\quad m' \quad (1-s')\,f_{y/cm}(y) \\
&+ (1-c')(1-m') \quad s' \quad f_{y/s}(y) \\
&+ \quad c' \quad (1-m') \quad s' \quad f_{y/cs}(y) \\
&+ (1-c') \quad m' \quad\quad s' \quad f_{y/ms}(y) \\
&+ \quad c' \quad\quad m' \quad\quad s' \quad f_{y/cms}(y)
\end{aligned}
\tag{3.2}
$$

$$
\begin{aligned}
m' &= (1-c')\,(1-y')(1-s')\,f_m(m) \\
&+ \quad c' \quad (1-y')(1-s')\,f_{m/c}(m) \\
&+ (1-c') \quad y' \quad (1-s')\,f_{m/y}(m) \\
&+ \quad c' \quad\quad y' \quad (1-s')\,f_{m/cy}(m) \\
&+ (1-c')\,(1-y') \quad s' \quad f_{m/s}(m) \\
&+ \quad c' \quad (1-y') \quad s' \quad f_{m/cs}(m) \\
&+ (1-c') \quad y' \quad\quad s' \quad f_{m/ys}(m) \\
&+ \quad c' \quad\quad y' \quad\quad s' \quad f_{m/cys}(m)
\end{aligned}
\qquad
s' = f_s(s)
$$

The prediction accuracy of the *silver ink halftone* directive is shown in Table 3.7. The *silver ink halftone* directive has the same excellent accuracy as the *top or below* directive not only for

the average ΔE_{94} error, but also for the standard deviation, 95 percentile, and maximum ΔE_{94} error. This confirms the validity of the proposed approach to determine which ink spreading curves are relevant.

3.3.6 Summary

We perform experiments to deduce the ink spreading curves that have a high impact on the accuracy of the spectral prediction model and that are resilient to measurement noise. Based on this in-depth analysis, we show that the ink spreading curves of ink halftones superposed with solid black are neither relevant for the accuracy of the spectral prediction model nor resilient to measurement noise. We therefore propose an improved ink spreading directive called *halftone black* relying on a reduced set of ink spreading equations.

This new directive is compared with the other directives for two different measurement sets, one from a consumer-grade inkjet printer and the other from a newspaper web-offset press. In both cases the *halftone black* directive performs best on average. Depending on the printing technology, the other directives may sometimes show similar performances, but since *halftone black* requires only 20 ink spreading curves instead of 32, we recommend it for CMYK spectral prediction models.

A similar analysis is conducted for the CMYS case. The silver ink spreading curves superposed with one or more solid inks are shown to have a small impact on the accuracy of the spectral prediction model. A new ink spreading directive called *silver ink halftone* is proposed, requiring only 25 ink spreading curves instead of the initial 32. The *silver ink halftone* directive is shown to have the same prediction accuracy as the *top or below* directive, validating the proposed approach of determining the relevant ink spreading curves.

In the remaining of this thesis, we always use the *halftone black* directive when performing predictions of CMYK measurement sets. However, when working with CMY data, we use the *top or below* directive.

Chapter 4

Calibrating the ink spreading model using color image tiles

4.1 Introduction

The IS-YNSN model presented in Chapter 2 and refined in Chapter 3 must be calibrated by measuring the reflection spectra of specially conceived color-constant calibration patches [1] and by deducing the ink spreading curves mapping nominal ink surface coverages to effective ink surface coverages. These special calibration patches comprise all solid colorants (paper, solid inks and solid ink superpositions) and single ink halftones superposed with all combinations of solid ink superpositions. However, these specially conceived calibration patches take valuable space on the printed product and must be removed, or are simply not available. For example, on offset presses, the special calibration patches are placed in the margins of the printed pages and are cut off before assembling the final print product.

In this chapter, we focus on the calibration of the ink spreading model. Instead of calibrating the ink spreading curves using specially conceived color-constant calibration patches, we propose a method based on image tiles extracted from printed images. Section 4.2 describes the calibration method and performs experiments based on the data used by Hersch et al. [HBBH09]. In Section 4.3, we focus on the algorithm extracting image tiles from printed color images and show that selecting image tiles specifically for the proposed calibration algorithm improves the accuracy of the ink spreading curves and requires less image tiles.

1. We use the term *color-constant* to describe patches composed of a single unique color in order to differentiate them from the *uniform* tiles extracted from images. This term is not a reference to the field of color constancy.

In this chapter, the colorants (Neugebauer primaries) must still be measured and the Yule-Nielsen n-value known. The calibration of entire IS-YNSN model is discussed in Chapter 5.

4.2 Constrained Acquisition of Ink Spreading Curves from Printed Color Images

Spectral reflection prediction models are able to accurately predict the reflection spectrum of any color-constant patch provided that the ink spreading curves are correctly calibrated. Such calibrations usually require printing specific color-constant calibration patches in addition to the printed content, patches that must be cut out from the final document.

In this section, we show how to calibrated the ink spreading curves using the measurements of image tiles extracted from printed color images. By deducing the ink spreading curves from image tiles, we eliminate the need of printing special color-constant halftone calibration patches. However, the extracted image tiles may not always provide sufficient information to accurately determine each ink spreading curve. To prevent the calibration procedure from setting these ink spreading curves to artificially large values, we impose additional constraints on the calibration procedure which rely on a relevance metric. We compare the predictions of the new model calibrations to the predictions of the classical calibrations. Thanks to the new calibration, the model predicts not only image tiles similar to the calibration tiles, but also color-constant patches spanning the entire printing gamut. Moreover, the new predictions are in some cases more accurate than classical predictions, suggesting that the patches used for classical calibrations may not be the most relevant patches for model calibration.

All the experiments have been performed on an inkjet printer, i.e. a device that provides a stable reproduction of colors across the entire page as well as from page to page. These assumptions are not valid for all reproduction devices. For example, in offset printing, the printable area is divided into several inking zones perpendicular to the printing direction. The ink flow is specific to each zone and depends on the amount of ink used within that zone. Ink spreading may therefore vary from zone to zone. The study of zone-dependent reproduction systems is out of the scope of the present dissertation.

4.2.1 Parabolic Ink Spreading Curves

The ink spreading model proposed in Section 2.3 relies on linearly interpolated ink spreading curves, which is standard in the literature and works well in practice, but they are not

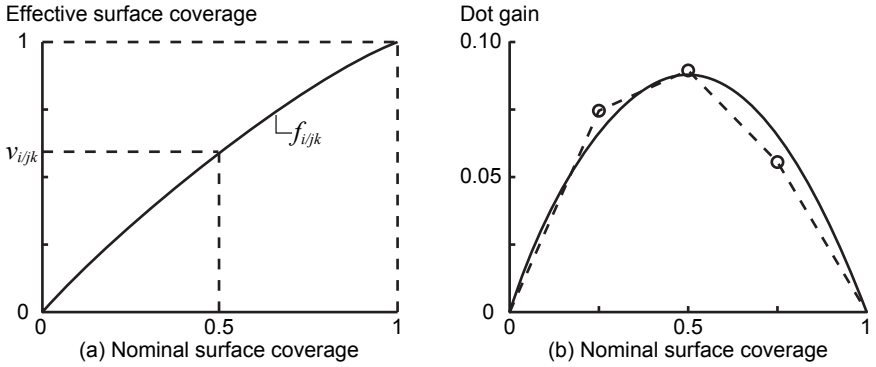


Figure 4.1: (a) Example of a parabolic ink spreading curve $f_{i/jk}$ characterized by its mid-point $v_{i/jk}$. (b) Corresponding dot gain curve, defined as the difference between effective and nominal surface coverages, with the fitted dot gains of three calibration patches (circles), the linearly interpolated dot gain curve (dashed line), and the corresponding least squares approximated parabolic dot gain curve (solid line).

smooth at the calibration points. As seen in Figure 2.1b, the different linear sections of the ink spreading curve do not smoothly connect, i.e. they are continuous, but not differentiable. Moreover, due to printing artifacts or measurement noise, the calibration points may not always define a monotonic increasing function, which is counterintuitive. Indeed, we can assume that increasing the nominal surface coverage of an ink increases its effective surface coverage as well.

Instead of linearly interpolated ink spreading curves, we propose to rely on parabolic ink spreading curves. As shown in Figure 4.1a, parabolic ink spreading curve $f_{i/jk}$ of halftone ink i superposed with solid inks j and k, is specified by a single parameter $v_{i/jk}$. This parameter, further referred to as *the mid-point of the ink spreading curve*, corresponds to the effective coverage of the halftone at 50% nominal ink surface coverage. Based on $v_{i/jk}$, ink spreading curve $f_{i/jk}$ is analytically described as follows:

$$\begin{aligned} f_{i/jk}(u_i, v_{i/jk}) &= u_i + (4v_{i/jk} - 2)(1 - u_i)u_i \\ &= u_i^2(2 - 4v_{i/jk}) + u_i(4v_{i/jk} - 1) \end{aligned} \tag{4.1}$$

where u_i is the nominal surface coverage of ink i. An ink spreading curve can be calibrated by measuring the halftone at 50% nominal surface coverage and deducing its effective coverage using Eq. (2.4), but as shown in Figure 4.1b, we prefer 3 calibration patches per ink spreading curve. The parabola is then fitted by minimizing the least-squares difference between the fitted and predicted dot gains.

Compared to linearly interpolated ink spreading curves, parabolic ink spreading curves are

differentiable, which is extremely important for optimization procedures. Moreover, they are monotonic increasing if $v_{i/jk}$ is in the interval $[0.25 \;\; 0.75]$, a simple constraint to enforce. Although slightly less accurate than their linearly interpolated counterparts, they play an significant role in the success of the proposed method to calibrate the ink spreading model based on image tiles extracted from printed color images.

4.2.2 Calibration of the ink spreading model using reflection spectra of image tiles extracted from color images

We now present a method to calibrate the parabolic ink spreading curves used by the IS-YNSN spectral prediction model, i.e. to set the mid-points of the 20 ink spreading curves for CMYK prints using a set of reflection spectra of image tiles extracted from printed color images, further referred to as *image tiles*, instead of the reflection spectra of specific color-constant calibration patches. The set of image tiles used for calibration is further referred to as the *image calibration set*, or *image calibration tiles*. The reflection spectra of specially conceived color-constant calibration patches composed of one ink halftone superposed with paper or superposed with other solid inks is further referred to as the *classical calibration set*, or *classical calibration patches*. The reflection spectra of the solid colorants (Neugebauer primaries) and the Yule-Nielsen n-value are assumed to be known for the target printing process.

Definition of the ink spreading weights

The problem with image tiles is that not all superpositions of inks and halftones are necessarily present. For example, if an image does not contain any cyan, it is not possible to calibrate the ink spreading curves of the cyan ink. When the contribution of an ink to a given image calibration tile is weak, the minimization algorithm tends to induce large variations of the values of the corresponding ink spreading curves in order to influence the difference metric. The consequence is that the ink spreading curves are set to artificially large values that do not reflect the actual dot gain. Such calibrations may be able to accurately predict image tiles similar to the calibration set, but fail to correctly predict image tiles of other colors [HBBH09].

To avoid this problem, we introduce weights, one for each ink spreading curve. A weight evaluates whether its ink spreading curve is relevant for predicting the image tiles of the calibration set or not. If modifying a given ink spreading curve changes significantly some of the predicted reflection spectra, its associated weight should be close to 1. If on the contrary the predicted spectra do not significantly change, the weight should be close to 0. The weights are

used to define upper and lower bounds for the mid-points of the ink spreading curves. As seen in Section 4.2.1, the mid-points must be in the interval $[0.25\ \ 0.75]$ in order to obtain monotonically increasing ink spreading curves. We modify the definition of the bounds for a given ink spreading curve as follows:

$$0.5 - 0.25 \cdot w_{i/jk} \leq v_{i/jk} \leq 0.5 + 0.25 \cdot w_{i/jk} \tag{4.2}$$

where $w_{i/jk}$ represents the weight of ink halftone i superposed with solid inks j and k. When the weight is equal to 1, the interval for the mid-point is not reduced. When the weight is equal to 0, i.e. when modifying the ink spreading curve does not modify the predicted spectra, the interval is reduced to a single mid-point, i.e. 0.5. When an ink spreading curve contributes only slightly to the calibration set, its weight is low and the interval for its mid-point is reduced. This prevents the fitting procedure from inducing large mid-point variations when minimizing the metric expressing the distance between predicted and measured reflection spectra.

Calculation of the ink spreading weights

Given a tile p of the image calibration set, let the nominal and effective coverages of ink i within tile p be respectively $u_{i,p}$ and $u'_{i,p}(v_i, v_{i/j}, v_{i/k}, v_{i/jk})$. The weight associated to ink spreading curve $f_{i/jk}$ of halftone ink i superposed with solid inks j and k is defined as the following gradient:

$$w_{i/jk,p} = \frac{\partial u'_{i,p}}{\partial v_{i/jk}} \tag{4.3}$$

Since ink spreading curve $f_{i/jk}$ is fully determined by its mid-point $v_{i/jk}$, the gradient in Eq. (4.3) expresses the influence of ink spreading curve $f_{i/jk}$ on the resulting effective surface coverage $u'_{i,p}$ of ink i within tile p.

Since a calibration set is composed of several tiles, we define the weight associated to ink spreading curve $f_{i/jk}$ for the full image calibration set as the maximum of the weights of all the tiles:

$$w_{i/jk} = \max_p [w_{i/jk,p}] = \max_p \frac{\partial u'_{i,p}}{\partial v_{i/jk}} \tag{4.4}$$

We take the maximum of the derivatives among all the tiles because the tiles with a high weight have the largest influence on the metric minimized by the ink spreading curve mid-point fitting algorithm, i.e. when there is at least one high weight tile and low weight tiles, the ink spreading curve mid-point is mainly fitted by the high weight tile.

Interval of the ink spreading weights

Let us show that the weights defined according to Eq. (4.4) are comprised between 0 and 1. We calculate the derivative of an ink spreading curve with respect to its mid-point using Eq. (4.1):

$$\frac{\partial f_{i/jk}}{\partial v_{i/jk}} = \frac{\partial}{\partial v_{i/jk}}[u_i + (4v_{i/jk} - 2)(1 - u_i)u_i] = 4(1 - u_i)u_i \tag{4.5}$$

Eq. (4.5) depends neither on the superposition condition nor on the mid-point value, but only on the nominal surface coverage of the considered ink. Since u_i is a nominal surface coverage comprised between 0 and 1, the derivative reaches a maximum of 1 when $u_i = 0.5$ and a minimum of 0 when $u_i = 0$ or $u_i = 1$.

Let us now calculate the derivative of the effective surface coverages with respect to the mid-point of a given ink spreading curve. This derivative is 0 for all effective coverages except for the effective coverage affected by the chosen ink spreading curve. Let us for example consider the cyan ink. The effective coverage of cyan c' is influenced by four ink spreading curves: f_c, $f_{c/m}$, $f_{c/y}$, and $f_{c/my}$. With Eqs. (3.1) and (4.5), the derivatives of the cyan effective coverage with respect to v_c, $v_{c/m}$, $v_{c/y}$, and $v_{c/my}$ are:

$$\frac{\partial c'}{\partial v_c} = (1 - m')(1 - y')\frac{\partial f_c(c)}{\partial v_c} = (1 - m')(1 - y') \cdot 4c(1 - c) \tag{4.6}$$

$$\frac{\partial c'}{\partial v_{c/m}} = m'(1 - y')\frac{\partial f_{c/m}(c)}{\partial v_{c/m}} = m'(1 - y') \cdot 4c(1 - c) \tag{4.7}$$

$$\frac{\partial c'}{\partial v_{c/y}} = (1 - m')y'\frac{\partial f_{c/y}(c)}{\partial v_{c/y}} = (1 - m')y' \cdot 4c(1 - c) \tag{4.8}$$

$$\frac{\partial c'}{\partial v_{c/my}} = m'y'\frac{\partial f_{c/my}(c)}{\partial v_{c/my}} = m'y' \cdot 4c(1 - c) \tag{4.9}$$

These derivatives represent the weights w_c, $w_{c/m}$, $w_{c/y}$, and $w_{c/my}$ for tile p given by its nominal surface coverages c, m, y, k. The nominal coverages are known from prepress image data and the effective coverages are obtained with the YNSN model, i.e. the effective coverages are found by minimizing the sum of square differences between the measured reflection spectra components of the considered image tile and the ones predicted by the YNSN model [Eq. (2.4)]. Eqs. (4.6)–(4.9) are therefore fully determined. Moreover, all the variables are either nominal or effective surface coverages which are comprised between 0 and 1. As shown above, $4c(1-c)$ is also comprised between 0 and 1. Eqs. (4.6)–(4.9) are therefore also comprised between 0

and 1. This demonstrates that the weights defined in Eq. (4.4) are indeed comprised between 0 and 1.

It is also interesting to observe the nominal surface coverages when a weight reaches its maximum, i.e. when it is equal to 1. If we consider the weight $\partial c'/\partial v_c$ associated with the ink spreading curve $f_c(c)$ of cyan halftones printed on paper, expressed by Eq. (4.6), it reaches its maximum when $m' = y' = 0$ and when $c = 0.5$. As $m' = 0$ implies $m = 0$, a patch maximizes the weight of the cyan $f_c(c)$ ink spreading curve only when its nominal surface coverages are $c = 0.5$, $m = 0$ and $y = 0$. These surface coverages correspond to the nominal surface coverages of the patch used to calibrate the cyan ink spreading curve in the classical calibration set. In the case of cyan superposed with solid magenta, computed by Eq. (4.7), the maximal weight is obtained with $c = 0.5$, $m = 1$ and $y = 0$. Similar statements are valid for the other two superposition conditions. Therefore, the classical calibration set maximizes the weights of all the ink spreading curves. Hence, when using the classical calibration set, the calibration of the ink spreading model is not constrained, i.e. the weights are always 1.

Least-squares calibration procedure

Let us now present the complete calibration procedure. We propose a least-squares approach to calibrate the ink spreading curves given the reflection spectra of an image calibration set. First, the effective surface coverages (c', m', y', k') of each image calibration tile are fitted by using the YNSN model [Eq. (2.4)]. Then, the weights associated to the ink spreading curves are computed using Eq. (4.4), with examples of the weights of a single tile given in Eqs. (4.6)–(4.9). Each weight $w_{i/jk}$ defines the bounds for its associated ink spreading curve mid-point $v_{i/jk}$. Finally, the mid-points of the ink spreading curves are determined using a constrained least-squares procedure where the constraints are the bounds of the different mid-points. The constrained least-squares equations are derived from Eq. (3.1). There is one least-squares equation per effective ink surface coverage. For example, let us take the equation computing the cyan effective surface coverage. Substituting Eq. (4.1) in Eq. (3.1) yields:

$$\begin{aligned} c' = &(1-m')(1-y')(c + (4v_c - 2)(1-c)c) \\ &+ \quad m' \quad (1-y')(c + (4v_{c/m} - 2)(1-c)c) \\ &+ (1-m') \quad y' \quad (c + (4v_{c/y} - 2)(1-c)c) \\ &+ \quad m' \quad\quad y' \quad (c + (4v_{c/my} - 2)(1-c)c) \end{aligned} \tag{4.10}$$

Let $\xi = [(1-m')(1-y') \quad m'(1-y') \quad (1-m')y' \quad m'y']$. Eq. (4.10) is rewritten in matrix

form:

$$c' + \xi \begin{bmatrix} 2(1-c)c - c \\ 2(1-c)c - c \\ 2(1-c)c - c \\ 2(1-c)c - c \end{bmatrix} = 4(1-c)c\ \xi \begin{bmatrix} v_c \\ v_{c/m} \\ v_{c/y} \\ v_{c/my} \end{bmatrix} \tag{4.11}$$

Since the sum of all terms of ξ is one, Eq. (4.11) becomes:

$$c' + 2(1-c)c - c = 4(1-c)c\ \xi \begin{bmatrix} v_c & v_{c/m} & v_{c/y} & v_{c/my} \end{bmatrix}^T \tag{4.12}$$

Eq. (4.12) has the form $b_p = A_p x$ where A_p is the row vector $4(1-c)c\xi$ and x the vector with the unknown ink spreading mid-point values. Knowing both nominal and effective surface coverages, each tile of the calibration set forms one line of matrix A and one element of vector b. Considering all the calibration tiles, we obtain the following constrained least-squares problem:

$$x_{\min} = \arg\min_x(\|b - Ax\|) \text{ such that } \begin{cases} v_{i/jk} \leq 0.5 + 0.25 \cdot w_{i/jk} \\ v_{i/jk} \geq 0.5 - 0.25 \cdot w_{i/jk} \end{cases} \tag{4.13}$$

where x_{min} is the vector containing the calibrated mid-points of the ink spreading curves. Such problems are readily solved using existing algorithms [CL96, GMW81], possibly implemented by computer programs such as Matlab. Note that it is also possible to perform the above calibration procedure using a calibration set composed of a mixture of classical calibration patches and image tiles, or even composed of classical calibration patches only.

4.2.3 Selection of image calibration and test tiles

Before verifying the accuracy of the proposed calibration on image tiles, we need to select tiles as uniform as possible from the available images, well distributed across the color space. For each image, we form two sets of tiles. The first set contains calibration tiles and the second set test tiles.

The first step consists in scanning each image horizontally and vertically in 2 mm steps. Successive 5 mm large square tiles form the candidate tiles from which calibration and test tiles are selected. Each tile is associated a uniformity value. We define the uniformity of a given color area as the RMS of the standard deviation of the CIELAB pixel values located within the considered image tile:

$$s = \sqrt{\left(\sigma(L*)^2 + \sigma(a*)^2 + \sigma(b*)^2\right)/3} \tag{4.14}$$

The calibration set associated with a given image contains approximately 50 of the most uniform tiles, i.e. those with the smallest s value, with the additional condition that each tile has no other tile within $\Delta E_{ab} < 2$ and no more than one other tile within $\Delta E_{ab} < 6$. The test set associated with that image contains the most uniform 30 tiles that are not part of the calibration set, and with the additional condition that each tile has no other tile within $\Delta E_{ab} < 6$.

We use the same images and the same image tile selections presented in [HBBH09] and shown in Appendix B, where the ink spreading curves were fitted without considering their relevance for a given image calibration set, i.e. without introducing constraints on the ink spreading curve mid-points.

4.2.4 Setup of the experiments

The experiments are performed on a Canon PIXMA Pro 9500 printer, with color images printed on Canon MP-101 matte paper with classical clustered dot halftoning, at a resolution of 600 dpi and a screen frequency of 100 lpi. The black ink is pigment based and absorbs in the near infrared wavelength range (cartridge PGI-9MBK). Reflectance spectra are measured with a Datacolor MF-45 spectrophotometer (geometry: 45°d:0°, i.e. 45° directed incident light and capture at 0°) at 10 nm intervals between 380 nm and 850 nm.

The reflectances of the printed colorants (Neugebauer primaries) have been measured. The Yule-Nielsen n-value is set to a suitable value according to the classical calibration, 7.8 in the current case. Note that both the reflectances of the colorants and the Yule-Nielsen n-value can be tabulated for each combination of inks and paper.

The *fruits* and *textile* images shown in Appendix B.1 are printed using 3 inks only (CMY) and using 4 inks (CMYK). In each case, 4 different sets of tiles are measured: the calibration and test sets of the *fruits* image (Appendix B.2), and the calibration and test sets of the *textile* image (Appendix B.3). Since the selection algorithm is performed in the CIELAB color space, the selected tiles are the same for CMY and CMYK.

4.2.5 Calibration of the ink spreading curves with spectra measured on the selected image tiles

The algorithm presented in Section 4.2.2 aims at calibrating the ink spreading model, i.e. at setting the ink spreading curves as accurately as possible, without printing specially conceived halftone patches. According to Eq. (3.1), there are 12 ink spreading curves to fit for CMY prints, and 20 ink spreading curves for CMYK prints. To each ink spreading curve $f_{i/jk}$ map-

ping nominal to effective surface coverages in a given superposition condition, we associate a dot gain curve $g_{i/jk}$ with the dot gain being defined as the effective surface coverage minus the nominal surface coverage:

$$g_{i/jk}(u_i) = f_{i/jk}(u_i, v_{i/jk}) - u_i = (4v_{i/jk} - 2)(1 - u_i)u_i \tag{4.15}$$

In the same way, the bounds associated with dot gain curve $g_{i/jk}$ are defined as follows based on Eqs. (4.2) and (4.15):

$$-0.25 \cdot w_{i/jk} \leq g_{i/jk}(0.5) = v_{i/jk} - 0.5 \leq 0.25 \cdot w_{i/jk} \tag{4.16}$$

In both the CMY and CMYK cases, we consider five different calibrations. The first case, where all the ink spreading curves are arbitrarily set to 10% dot gain, is referred to as the *no calibration* case. The second calibration is called the *minimal classical calibration* and is calibrated using a single color-constant patch per ink spreading curve. This single patch yields the effective surface coverage at 50% nominal surface coverage and is referred to as the *50% patch*. For example, the 50% patch of the $f_{c/m}$ ink spreading curve is 50% cyan superposed with solid magenta. Therefore, the minimal classical calibration requires 12 calibration patches for CMY prints, and 20 calibration patches for CMYK prints. The third calibration, referred to as the *extended classical calibration*, includes the 25%, 50% and 75% color-constant patches for the calibration of each ink spreading curve. Such a calibration requires 36 patches for CMY prints and 60 patches for CMYK prints. The fourth and fifth calibrations use the *fruits* image calibration set and the *textile* image calibration set, respectively. These last two calibrations are also referred to as the *image calibrations*.

Figure 4.2 compares no calibration, minimal classical calibration and extended classical calibration. The cyan and magenta dot gain curves are similar for the minimal and extended classical calibrations. However, the yellow and black dot gain curves show significant variations. This indicates that the 50% patches are enough to characterize the cyan and magenta dot gain curves, but not the yellow and black dot gain curves.

Figure 4.3 compares the extended classical, the fruits and the textile calibrations in the CMYK case. The figure shows that half the dot gain curves of both image calibrations agree well with the extended classical calibration. There are also a few cases where the dot gain curves of the image calibrations reach their respective bounds computed using Eq. (4.16), especially the black ink spreading curves. However, this occurs only when the bounds are tight.

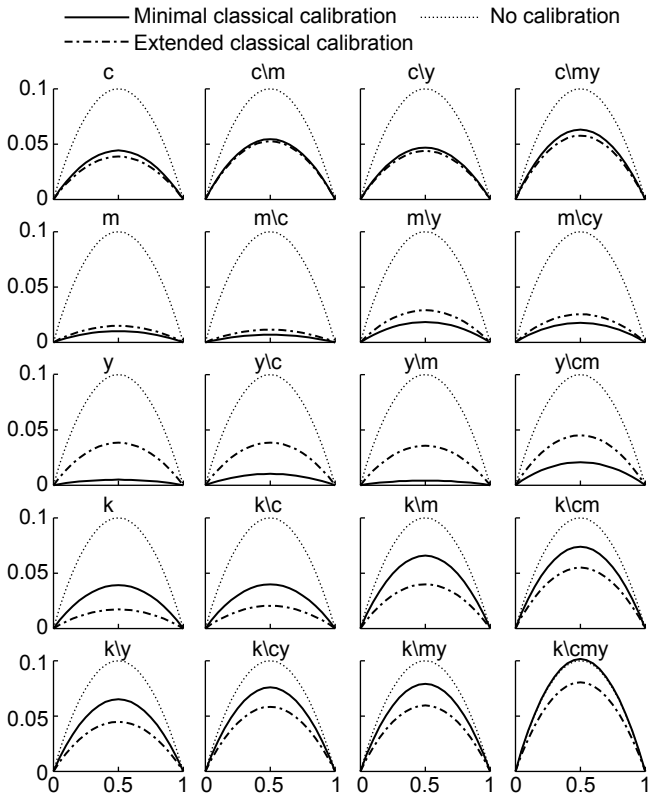


Figure 4.2: CMYK dot gain curves of the minimal classical calibration, extended classical calibration and no calibration.

When the bounds are loose, the fitted ink spreading curves are located close to the extended classical calibration and away from the bounds, which is the expected behavior and confirms that the bounds of a given ink spreading curve are loose only when the calibration set contains at least one tile highly relevant for that ink spreading curve.

The effect of the constraints can be seen in Figure 4.4. When the constraints are not used, five black and one yellow dot gain curves are set outside the defined bounds, with three of them even reaching the original bound of $\pm 25\%$. These dot gain curves are also the ones with the strictest bounds, meaning that the calibration set does not contain tiles that are relevant enough to calibrate these dot gain curves. The unconstrained calibration procedure sets the ink spreading curves to extreme values in order to have a small gain of accuracy. This gain improves only the accuracy of the image calibration set. When predicting another set of tiles, all gains are lost and the predictions are poor.

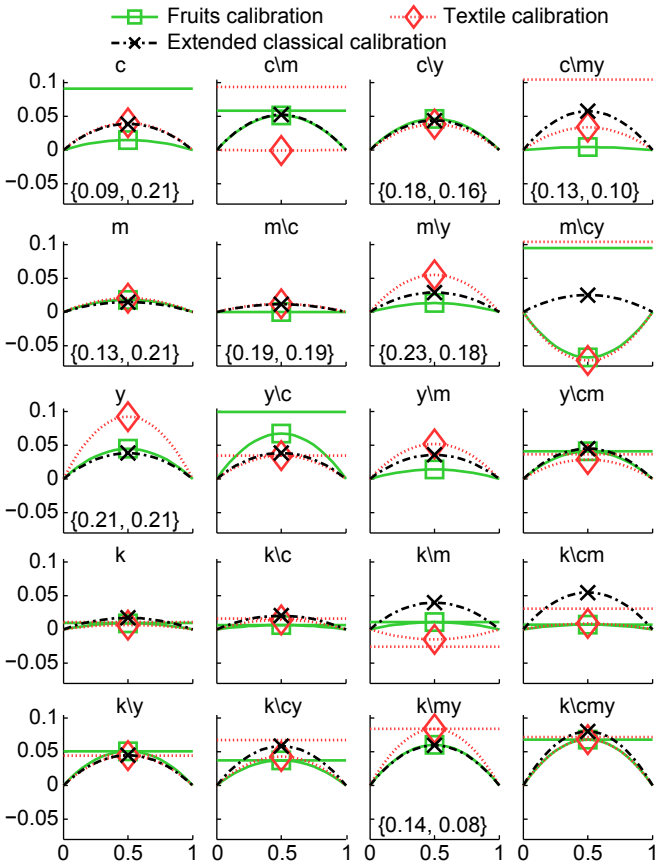


Figure 4.3: CMYK dot gain curves of the fruits, textile and extended classical calibrations. The upper bounds of the fruits (F_B) and textile (T_B) dot gain curves are indicated as horizontal lines or with the values $\{F_B, T_B\}$ respectively.

4.2.6 Prediction accuracy

The different calibrations performed in Section 4.2.5 are tested by comparing the difference between predicted and measured reflectances for different test sets. In the CMY case, we predict the CMY test sets of the *fruits* and *textile* images and a set composed of 125 CMY color-constant patches whose nominal surface coverages are all the CMY combinations at 0%, 25%, 50%, 75%, or 100%. In the CMYK case, we predict the CMYK test sets of the *fruits* and *textile* images, a set composed of 625 CMYK patches whose nominal surface coverages are all the CMYK combinations at 0%, 25%, 50%, 75%, or 100%, and a fourth set, further referred to as the 125 UCR set, composed of 125 CMYK patches corresponding to the 125 CMY set, but with nominal surface coverages obtained by converting the original CMY coverages to CMYK using a standard under color removal algorithm.

The reflection spectrum of each image tile or color-constant patch of the selected sets is

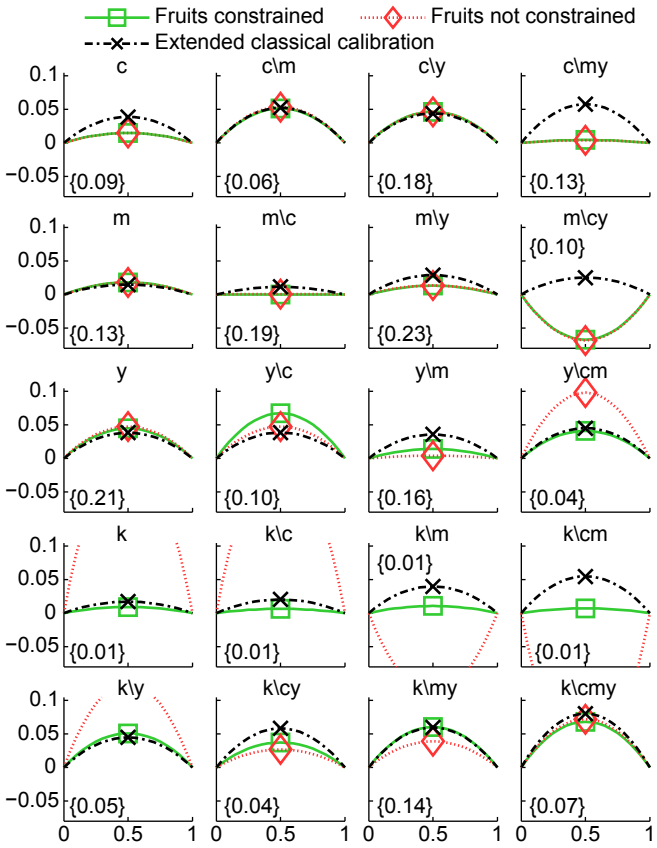


Figure 4.4: CMYK dot gain curves of the constrained fruits calibration, unconstrained fruits calibration, and extended classical calibration. The bounds of the constrained fruits calibration are given as $\{F_B\}$.

predicted and then compared to the corresponding measured reflection spectrum using the ΔE_{94} metric. For each measurement set, we compute the following statistics: average, 95 percentile and maximum ΔE_{94} between measured and predicted spectra. The results are shown in Table 4.1 for the CMY tests and Table 4.2 for the CMYK tests. In Figures 4.5a–4.5c, the prediction accuracy is defined as the inverse of the average ΔE_{94}.

Figure 4.5a shows the accuracy of "no calibration", minimal classical calibration and extended classical calibration for one CMY and two CMYK test sets. As expected the "no calibration" case yields the worst results and the extended classical calibration yields significantly better results than the minimal classical calibration, especially for the CMYK test sets, outlining the fact that, for the considered PIXMA Pro 9500 ink jet printer, the 25% and 75% surface coverage patches are important for classical calibrations.

Figures 4.5b–4.5c show the accuracy difference between the extended classical, fruits, and textile calibrations, as well as the fruits calibrations without constraints. These calibrations are compared on the CMY fruits, CMY textile, and 125 CMY test sets in Figure 4.5b and

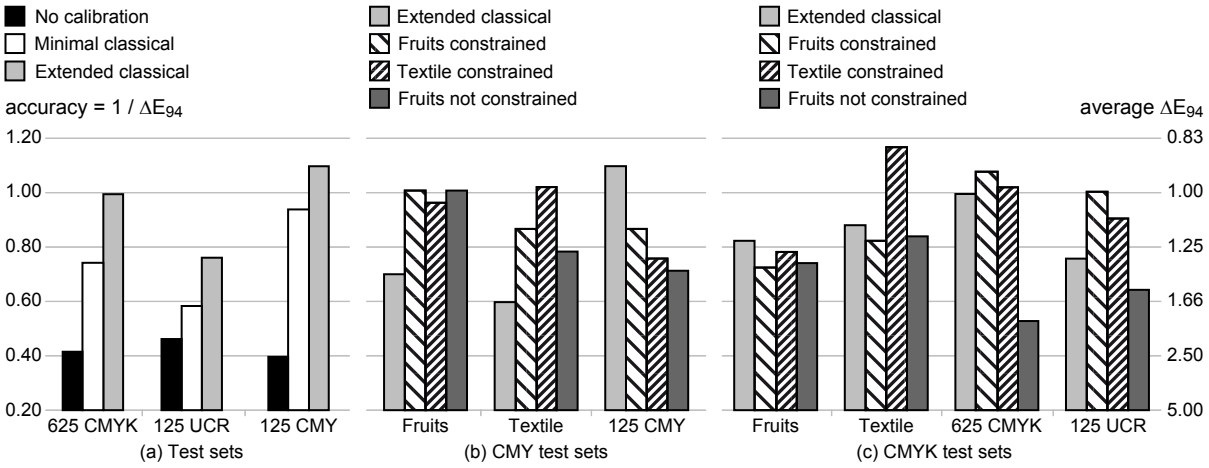


Figure 4.5: Comparison of the accuracy of (a) "no calibration", minimal classical calibration and extended classical calibration on the 625 CMYK, 125 UCR and 125 CMY test sets (b) extended classical, constrained textile, constrained fruits, and unconstrained fruits calibrations on the CMY fruits and textile test tiles and on the 125 CMY test set (c) extended classical, constrained textile, constrained fruits, and unconstrained fruits calibrations on the CMYK fruits and textile test tiles, and on the 625 CMYK and 125 UCR test sets.

Table 4.1: Prediction accuracy for different calibrations and CMY test sets. The statistics show the average, 95 percentile and maximum ΔE_{94}.

CMY test sets	ΔE_{94}		
Calibrations	avg	95%	max
CMY test set fruits			
No calibration	3.16	4.49	4.92
Minimal calibration	1.57	2.72	2.98
Extended calibration	1.32	2.51	2.81
Fruits with constraints	1.02	1.84	1.88
Textile with constraints	1.12	2.40	2.62
Fruits no constraints	1.02	1.74	1.87
Textile no constraints	1.14	2.61	2.95
CMY test set textile			
No calibration	3.34	5.11	5.29
Minimal calibration	1.98	3.41	3.49
Extended calibration	1.68	2.94	3.28
Fruits with constraints	1.15	2.01	2.02
Textile with constraints	0.98	1.82	2.24
Fruits no constraints	1.28	3.26	3.40
Textile no constraints	0.90	1.75	1.90
125 CMY test set			
No calibration	2.51	4.31	4.48
Minimal calibration	1.07	2.79	3.43
Extended calibration	0.91	1.81	2.36
Fruits with constraints	1.16	2.76	3.95
Textile with constraints	1.33	3.18	4.23
Fruits no constraints	1.41	3.85	4.51
Textile no constraints	1.71	4.42	5.62

Table 4.2: Prediction accuracy for different calibrations and CMYK test sets. The statistics show the average, 95 percentile and maximum ΔE_{94}.

CMYK test sets	ΔE_{94}		
Calibrations	avg	95%	max
CMYK test set fruits			
No calibration	1.66	3.53	4.12
Minimal calibration	1.61	2.85	3.26
Extended calibration	1.22	2.40	3.34
Fruits with constraints	1.38	2.55	4.67
Textile with constraints	1.27	2.62	4.26
Fruits no constraints	1.35	2.54	4.59
Textile no constraints	1.32	2.93	4.18
CMYK test set textile			
No calibration	1.56	3.31	3.37
Minimal calibration	1.50	2.50	2.95
Extended calibration	1.14	1.72	2.55
Fruits with constraints	1.21	2.11	2.17
Textile with constraints	0.86	1.38	1.57
Fruits no constraints	1.19	2.07	2.24
Textile no constraints	0.81	1.46	1.49
625 CMYK test set			
No calibration	2.41	3.48	3.94
Minimal calibration	1.35	2.50	3.21
Extended calibration	1.01	1.87	2.43
Fruits with constraints	0.93	1.99	3.62
Textile with constraints	0.98	2.23	3.56
Fruits no constraints	1.91	6.07	10.37
Textile no constraints	1.99	5.69	8.34
125 UCR test set			
No calibration	2.16	3.63	4.20
Minimal calibration	1.73	3.17	3.69
Extended calibration	1.31	2.30	2.79
Fruits with constraints	0.99	1.83	2.56
Textile with constraints	1.10	2.11	2.40
Fruits no constraints	1.56	4.06	6.03
Textile no constraints	2.03	5.24	6.99

on the CMYK fruits, CMYK textile, 625 CMYK, and 125 UCR test sets in Figure 4.5c. In Figure 4.5b, the results are as expected: The fruits calibrations have the best accuracy on the fruits test set, the textile calibration on the textile test set, and the extended classical calibration on the 125 CMY test. Moreover, the constrained fruits calibration has a significantly higher prediction accuracy on the 125 CMY test set than the unconstrained fruits calibration.

In Figure 4.5c, the extended classical and images calibrations, both with and without constraints, accurately predict both the CMYK textile and fruits test sets. The textile calibration is particularly accurate on the textile test set. The fact that the fruits calibration does not display such an accuracy on the fruits test set is due to the fact that the fruits test set is more different from the fruits calibration set (average distance between each test tile and its nearest calibration tile is $\Delta E_{94} = 2.94$) than the textile test set from the textile calibration set (average distance $\Delta E_{94} = 1.60$). The constraints are extremely important when predicting both the 625 CMYK set and the 125 UCR set. Removing the constraints leads to a sharp drop of accuracy. This indicates that the constraints considerably improve the accuracy of a calibration when predicting the reflection spectra of a set of patches or tiles significantly different from the calibration tiles.

Surprisingly, the image calibrations with constraints have predictions as accurate as the extended classical calibration or even better predictions in the case of the 625 CMYK or 125 UCR test sets. This indicates that calibrating the ink spreading model based on reflection spectra of image halftones is a valid approach. Whereas the classical calibration uses artificially crafted patches—there are at the macroscopic level few to no occurrences of superpositions of one halftone and solid inks in a conventional image—relying on tiles directly taken from real images can achieve higher accuracy because these image tiles better represent the normal operational conditions of the reproduction device.

4.2.7 Summary

When correctly calibrated, today's spectral reflection prediction models are able to accurately predict the reflection spectra of printed color image tiles. However, the required calibration procedures remain difficult and cumbersome because they require printing specially conceived color-constant calibration patches which use space and need to be removed from the final product.

We have shown how to calibrate the ink spreading model using spectral data of image tiles extracted from printed color images. Image tiles as uniform as possible are selected within printed images and the ink spreading model is calibrated with the measurements of the selected

calibration image tiles.

Depending on the available calibration image tiles, certain ink spreading curves cannot be reliably calibrated because these ink spreading curves are not relevant in the prediction of the calibration tiles. We therefore establish a simple relevance metric which relies on the nominal surface coverages of the considered halftone and on the effective surface coverages of the superposed other inks. Relying on this metric, we create bounds for the mid-points of the fitted ink spreading curves. When the relevance is low, the effective surface coverages remain close to the nominal surface coverages. This metric reaches its maximum relevance at 50% halftone surface coverage on top of solid inks, which corresponds to the classical calibration patches. Therefore, the classical calibration is a subset of the proposed constrained image tile calibration approach.

The performed experiments demonstrate that the new constraint-based calibration procedure is reliable when predicting patches significantly different from the calibration patches. Moreover, the predictions of patches similar to the calibration patches remain as accurate with the added constraints as without. Finally, the image calibration can be more accurate than a classical calibration when predicting test sets composed of color-constant patches distributed over the entire printing gamut of the reproduction device, suggesting that the classical calibration patches, despite incorporating the highest relevant patches formed by halftones superposed with solid inks, are not necessarily the most adequate patches to calibrate the ink spreading model.

4.3 Optimized selection of image tiles for ink spreading calibration

Section 4.2 describes a method to calibrate the ink spreading model using image tiles extracted from printed color images, but does not discuss the impact of the method selecting image tiles on the prediction accuracy.

The method to select image tiles is based on preliminary work by Hersch et al. [HBBH09]. However, the existing tile selection algorithm has a number of limitations. The first limitation is that it is based on the CIELAB color space and therefore requires the conversion of the images into CIELAB. The conversion is time-consuming, depends on a color profile and cannot distinguish between two tiles of the same CIELAB color, but printed using different amounts of cyan, magenta, yellow, and black inks. The second limitation is that it is not possible to select

tiles according to a given goal in an optimal manner. For example, if the image tiles are used to calibrate a cyan dot gain curve, we would like to choose halftone tiles with a predominance of cyan.

In the present section, we first present an algorithm to automatically select image tiles from color images based uniquely on the CMY or CMYK pixel values of these color images and show that the reflectances of these image tiles can be accurately predicted by the IS-YNSN model if they are uniform enough. This selection algorithm incorporates additional constraints expressing the impact of a halftone on a given ink spreading curve. By relying on 6 different color images from which we separately extract the calibration tiles, we show the calibration performances of both a standard and an optimized calibration tile selection method according to the number of image tiles used for the calibration.

4.3.1 Gradient-descent based calibration of the ink spreading model

The method to calibrate the ink spreading model based on image calibration tiles is described in Section 4.2.2. However, since this method relies on solving a least-squares problem, it requires a relatively large number of image calibration tiles to yield stable results and does not work with a limited number of calibration tiles. When too little information about ink spreading is available in the image calibration set, the least-squares problem becomes ill-posed and the noise contained in the measurements of the image calibration tiles significantly impacts the resulting calibrated model.

We therefore propose a variation based on a gradient-descent algorithm. The new method is slower, but employs a starting point for the optimization corresponding to the IS-YNSN model with all ink spreading curves set to 0.5 at 50% nominal surface coverage (no physical dot gain). If the image calibration set contains very little information about the ink spreading of an ink halftone, i.e. the calibration set contains few or no image tiles with the corresponding colorants, the optimization does not significantly deviate from the starting point.

The aim of the calibration procedure is to find the best values for the mid-points of the ink spreading curves. We group all the mid-points into a single vector $\vec{v}$ and use as objective function the sum of square differences between the predicted and measured reflection spectra of all the image calibration tiles for given ink spreading curve mid-points $\vec{v}$:

$$f(\vec{v}) = \sum_{p} \sum_{\lambda} \left[predSpectrum(\lambda, covs_p, \vec{v}) - measSpectrum(\lambda, p) \right]^2 \tag{4.17}$$

where λ is the wavelength, p an image calibration tile, $covs_p$ its nominal ink surface coverages, and $\vec{v}$ a vector whose elements are the mid-points of the ink spreading curves. The function $measSpectrum$ returns the measured spectrum of image calibration tile p and the function $predSpectrum$ uses the IS-YNSN model to predict the spectrum of the image calibration tile p.

The calibration procedure using Eq. (4.17) can be summarized as follows:

$$\vec{v}_{opt} = \arg\min_{\vec{v}}[f(\vec{v})] \text{ such that } \begin{cases} v_{i/jk} \leq 0.5 + 0.25 \cdot w_{i/jk} \\ v_{i/jk} \geq 0.5 - 0.25 \cdot w_{i/jk} \end{cases} \tag{4.18}$$

where $\vec{v}_{opt}$ is the vector containing the calibrated mid-points, i.e. the mid-points of the ink spreading curves minimizing the sum of square differences between the measured and predicted spectra, and $w_{i/jk}$ is the weight, comprised between 0 and 1 and defined in Section 4.2.2, used to constrain $v_{i/jk}$.

4.3.2 Selecting image tiles based on the CMYK color space only

In order to calibrate the ink spreading model with image tiles extracted from the color images to be printed, we must first select such tiles. Our main assumption, verified in Section 4.3.5, is that the best image tiles are the most uniform tiles. The non-uniformity value of a tile located within a small sub-domain of the considered image is computed using its c, m, y, and k pixel values as follows:

$$u = \sqrt{\left(\sigma(c)^2 + \sigma(m)^2 + \sigma(y)^2 + \sigma(k)^2\right)/4} \tag{4.19}$$

where u is the non-uniformity value; c, m, y, and k are the cyan, magenta, yellow, and black pixel values of the considered image tile ranging from 0 to 1; and the σ function is the standard deviation. If the values of the pixels of a given image tile are close to the average pixel value, the standard deviations are low and u is low, i.e. the patch is considered uniform. However, if many pixels have values distant from the average, the standard deviations are large and u is large, i.e. the tile is considered non-uniform.

The aim of the selection algorithm is to select the most uniform tiles. However, in some images, many uniform tiles are composed of similar amounts of inks. To prevent having two tiles too close to each other, we use the Euclidian distance (norm) in the CMYK color space to determine when a candidate tile is too close to currently selected tiles. The CMYK norm indicates the proximity between two CMYK values. Two tiles can indeed have the same color,

but be composed of radically different amounts of inks. As we are interested in selecting tiles representing as many different reproduction situations as possible, the CMYK norm is an appropriate metric to discriminate between uniform tiles.

The first step of the algorithm consists in scanning each image horizontally and vertically in 2 mm steps. Successive 5 mm large square tiles form the candidate tiles from which the tiles are selected. Each tile is associated with a non-uniformity value according to Eq. (4.19). Assuming c, m, y, and k values between 0 and 1, tiles with a non-uniformity value above 0.1 are discarded.

Given a CMY image, the selection algorithm creates a set of 20 tiles composed of the most uniform tiles, i.e. those with the smallest u values, with the additional condition that each tile has no other tile within a CMY norm of 0.5 and no more than one other tile within a CMY norm of 1.5. Given a CMYK image, the selection algorithm creates a set of the 30 most uniform tiles so that each tile has no other tile within a CMYK norm of 0.3 and no more than one other tile within a CMYK norm of 0.9. Such image tile sets are further referred to as *standard image calibration tile sets*.

Each image tile has its associated nominal ink coverages. The nominal ink coverages of an image tile are computed as the average of the CMY or CMYK values of the pixels located within a radius of 1.5 mm from the center of the tile.

4.3.3 Adding constraints to the tile selection algorithm

The standard selection algorithm proposed in Section 4.3.2 selects the most uniform image tiles under the constraint that these tiles are not too close to each other in the CMYK color space. However, the best set of image tiles to calibrate the ink spreading model according to the procedure detailed in Section 4.3.1 are the tiles maximizing the weights $w_{i/jk}$ associated to the ink spreading curves in Eq. (4.4). Since the standard selection algorithm does not take this aspect into account, we propose an optimized selection algorithm which also maximizes these weights. This algorithm selects uniform enough image tiles maximizing the weights $w_{i/jk}$. Such a selection includes tiles highly relevant for the calibration of the ink spreading curves as well as the most uniform tiles.

Since there are 12 or 20 ink spreading weights $w_{i/jk}$ in the CMY or CMYK case, respectively, the strategy of the optimized algorithm is to maximize the sum of these weights by selecting the image tiles yielding the highest sum. However, the image tiles must still be uniform enough to be pertinent. In order to provide a flexible threshold ensuring a high enough unifor-

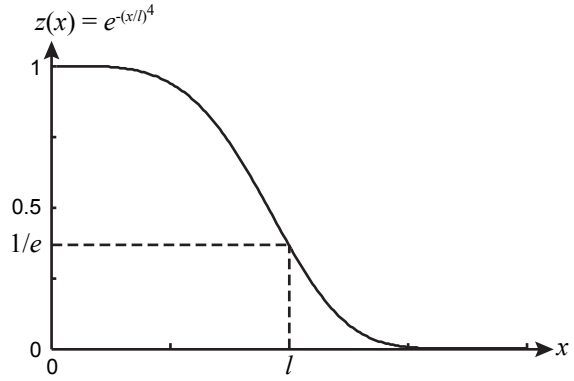


Figure 4.6: Adjustment function $z(x)$ inspired by the Weibull distribution.

mity, we use the following function inspired by the Weibull probability distribution [Wei]:

$$z(x) = e^{-(x/l)^4} \tag{4.20}$$

Function $z(x)$ is shown in Figure 4.6 and is used to adjust the ink spreading weights $w_{i/jk}$ as follows:

$$\tilde{w}_{i/jk,p} = w_{i/jk,p} z(u_p) \tag{4.21}$$

where $\tilde{w}_{i/jk,p}$ is the uniformity-adjusted weight of tile p, $w_{i/jk,p}$ is the weight of tile p computed in Eq. (4.2), and u_p is the non-uniformity value of tile p computed in Eq. (4.19). If the given tile is uniform, i.e. its non-uniformity value u_p is low, the adjustment function is close to unity and the weight is not modified. If the tile is non-uniform, its non-uniformity value is high and the adjustment function is close to 0. In that case, the tile is discarded. There is a sharp transition in which the less uniform the tile, the more the criterion function is decreased, creating a "smooth threshold".

The function $z(x)$ can be tailored to accommodate the available image tiles. The l value positions the transition zone and the exponent, set to 4 in Eq. (4.20), can be increased or decreased to shrink or enlarge the transition zone. If the exponent is set to infinity, the distribution becomes a step function and acts as a real threshold, discarding tiles with non-uniformity values higher than l. We used a value of $l = 0.10$ in Section 4.3 and $l = 0.05$ in Chapter 5.

Given ink spreading curve $f_{i/jk}$ and candidate image tile p, uniformity-adjusted weight $\tilde{w}_{i/jk,p}$ indicates if tile p is suitable to calibrate ink spreading curve $f_{i/jk}$. By taking for each ink spreading curve the image tile with the maximum uniformity-adjusted weight $\tilde{w}_{i/jk,p}$, we create an image calibration set containing the most suitable tiles to calibrate all the ink spreading

curves. This set is further referred to as the *adjusted calibration tile set*. Since a given tile can be the most suitable tile for more than one ink spreading curve, the adjusted calibration set contains at most as many calibration tiles as ink spreading curves. This simple selection algorithm creates therefore calibration sets with a fixed number of tiles.

If we want to create calibration sets containing fewer image tiles, we cannot select for each ink spreading curve its most suitable image calibration tile. A second criterion must be used to perform the selection. This second criterion is the sum s of the ink spreading weights $w_{i/jk}$ computed as follows:

$$s = \sum_{w_{i/jk}} w_{i/jk} = \sum_{w_{i/jk}} \max_{p} [w_{i/jk,p}] \tag{4.22}$$

where the weights $w_{i/jk}$ are computed according to Eq. (4.4). Since ink spreading weight $w_{i/jk}$ indicates the amount of information the entire calibration set provides about ink spreading curve $f_{i/jk}$, the sum s of all ink spreading weights indicates the amount of information available to calibrate all the ink spreading curves. In order to maximize s while selecting only suitable image tiles, we first create for each ink spreading curve a list of the candidate image tiles ordered from the most suitable tile to the least suitable tile. The ordering criterion is given by $\tilde{w}_{i/jk,p}$ according to Eq. (4.21). The heads of the lists therefore correspond to the most suitable tiles. The calibration set is then created by selecting one tile after another, each time selecting the tile among the heads of the lists that yields the maximum sum s of the ink spreading weights across all selected tiles. That tile is then removed from the lists. Note that as in Section 4.3.2, for an image tile to be selected, its non-uniformity must be below 0.1 and it must not be too close to the already selected tiles according to the CMYK norm. When no more tile increases the sum s of the weights, we revert to the standard algorithm and complete the set with the most uniform tiles that satisfy the CMYK norm constraint (Section 4.3.2). Image tile sets selected using this algorithm are further referred to as *optimized image calibration tile sets*.

Figure 4.7 shows the difference in the sum of the 20 CMYK ink spreading weights between the standard and optimized image tile sets. For all 6 test images shown in Appendix C.1 and described in Section 4.3.4, the sum of the weights is larger using the optimized selection algorithm than using the standard selection algorithm for any number of selected tiles. Moreover, when the optimized image calibration tile sets are composed of the same number of tiles as the adjusted calibration tile sets, the sum of the weights of the optimized sets are always equal to or slightly higher than their corresponding adjusted calibration tile set.

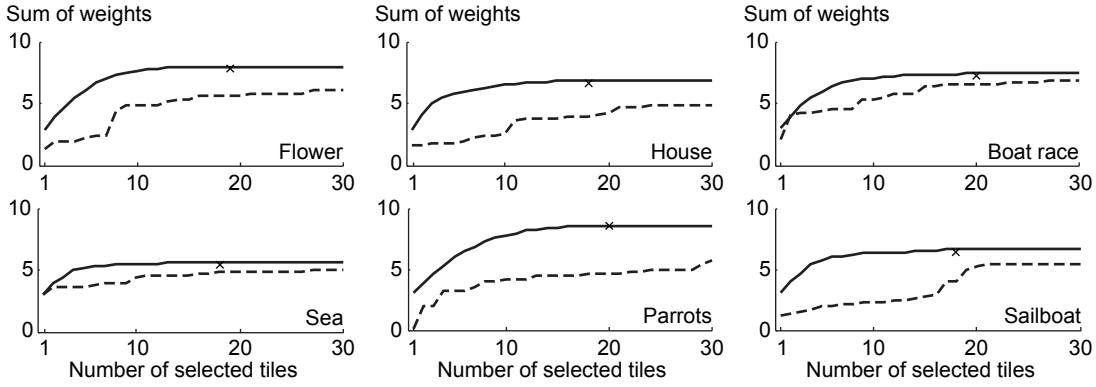


Figure 4.7: Sum of the 20 CMYK ink spreading curve weights for the six test images according to the number of selected tiles using either the standard (dashed line) or optimized (solid lines) tile selection algorithm. The crosses indicate the sum of the ink spreading weights of the adjusted calibration tile sets.

4.3.4 Setup of the experiments

In order to verify the accuracy of the spectral reflection predictions performed by the IS-YNSN model, the image tiles need to be measured. The size and shapes of the image tiles must be adapted according to the measurement geometry and aperture of the measurement device. For the Datacolor MF45IR spectrophotometer, the tiles are conceived as disks with a diameter of 3 mm. Since this device illuminates the sample using a directed source (45°d:0° geometry), we measure each sample four different times, each time rotating it by 90° and taking the average. This reduces errors due to the positioning of the device and irregularities of the illumination geometry. The non-uniformity value of a given patch is computed on an area slightly larger than the actual tile, i.e. a 5 mm disk instead of a 3 mm disk. Ensuring that the area surrounding the actual tile is also uniform further reduces positioning errors.

In order to analyze how the different selection algorithms perform, we use the 6 different sRGB images shown in Appendix C.1. The first image comes from a digital camera and the other five from different sources providing standard test images [Fra, Web, Lev]. The sizes of the images have not been modified, but they have been resampled to 600 dpi. Moreover, they have been converted to two different color spaces: CMY and CMYK. Both conversions have been performed using Photoshop and the U.S Web coated (SWOP) v2 CMYK profile. The conversion to CMYK has been performed with medium GCR (gray component replacement).

To verify the impact of the uniformity of the image tiles on the accuracy of the IS-YNSN model when predicting these tiles, we print eight sets of 50 image tiles further referred to as the uniformity test sets and shown in Appendix C.4. Each set contains tiles randomly selected among the candidate tiles of the 6 images having a non-uniformity value u in a given range.

Each set is referred to by the range of non-uniformity values of the tiles it contains, e.g. the 0.12–0.16 group contains 50 tiles with non-uniformity values between 0.12 and 0.16. Note that these 8 sets were only created for the CMYK images.

The image calibration tiles selected using the algorithms detailed in Section 4.3.2 (standard image calibration tile sets) and Section 4.3.3 (optimized image calibration tile sets) are shown in Appendix C.2 for the CMY images and in Appendix C.3 for the CMYK images. They are printed on a Canon PIXMA Pro 9500 inkjet printer at 600 dpi and 150 lpi using classical rotated halftone screens. For CMY, the halftone screen angles are 75°, 15°, and 45°, respectively. For CMYK, the halftone screen angles are 75°, 15°, 0° and 45°, respectively.

To compare the prediction accuracy of calibrations performed using different methods, we use two test sets composed of color-constant patches. The first test set, further referred to as the 125 CMY set, is composed of 125 color-constant patches with all possible combinations of cyan, magenta, and yellow at 0%, 25%, 50%, 75%, and 100%. The second set, further referred to as the 125 UCR set, is composed of 125 CMYK color-constant patches with nominal surface coverages obtained by converting the 125 CMY set to CMYK using the standard under color removal algorithm described above.

4.3.5 Prediction accuracy of the classical calibration relying on color-constant patches

All spectral predictions are performed using the IS-YNSN model calibrated using the classical uniform calibration patches, i.e. the 16 solid colorants (Neugebauer primaries) and $20 \cdot 3 = 60$ color-constant patches composed of one halftone ink at 25%, 50%, or 75% superposed with zero, one, or more solid inks. In both the CMY and CMYK cases, the Yule-Nielsen n-value is set to 10. The first part of Tables 4.3, 4.5, and 4.6 shows the reference prediction accuracy of the IS-YNSN model, i.e. the accuracy when predicting the color-constant patches of the 125 CMY set (Table 4.5) or the 125 UCR set (Tables 4.3 and 4.6).

Table 4.3 shows the accuracy when predicting the uniformity test sets. We can see that the average ΔE_{94} prediction error, derived from the measured and predicted reflectances of the different sets, consistently increases when the non-uniformity values of the image tiles increase. This correlation also holds for the 95 percentile and maximum ΔE_{94}.

An additional experiment shows that the information contained in non-uniform tiles cannot be used to calibrate the ink spreading curves. With the classically calibrated IS-YNSN model, we can either predict the effective coverages of a tile using the ink spreading model or de-

duce these effective coverages by directly fitting them with the YNSN model using Eq. (2.4). Table 4.4 shows that the classically calibrated IS-YNSN model is able to accurately predict non-uniform image tiles using the deduced effective coverages. However, the effective coverages predicted by the classically calibrated ink spreading model do not correspond to the deduced effective surface coverages, i.e. the difference between predicted and deduced coverages increases when the non-uniformity of the tiles increases.

When calibrating the ink spreading curves using image tiles, the deduced effective ink surface coverages of the calibration tiles are implicitly computed in Eq. (4.17), i.e. the calibration algorithm tries to set the ink spreading curves so that the predicted and deduced effective coverages agree. However, when non-uniform image tiles are used, the calibration algorithm sets the ink spreading curves to incorrect values to force the agreement between predicted and deduced coverages and therefore reduces the prediction accuracy. This demonstrates the importance of using image tiles as uniform as possible.

The second part of Tables 4.5 and 4.6 show the prediction accuracy of the classically calibrated IS-YNSN model when predicting the CMY and CMYK image calibration tile sets, respectively. We can see that the IS-YNSN model is also able to accurately predict image tiles, i.e. the average ΔE_{94} is either below or close to 2 and the 95 percentile is either below or close to 3, which confirms our choice of limiting the non-uniformity of the selected image tiles to 0.1. Note that there is no difference in the prediction accuracy of the standard and optimized image tile sets since they are all composed of tiles whose non-uniformity is below 0.1.

4.3.6 Accuracy of the IS-YNSN image calibrations

In the previous section, we have demonstrated that image tiles can be accurately predicted by an IS-YNSN model calibrated using the classical uniform calibration patches. We now compare predictions performed on models calibrated according to the different methods. The first method is the classical calibration detailed in Section 4.3.5. The other methods are based on the optimized and standard image calibration tile sets selected from one of the six test images. For each optimized or standard image calibration tile set, we perform a series of ink spreading curve calibrations using the algorithm detailed in Section 4.3.1. The first calibration of the series is performed using only the first image tile of the set. The second calibration is performed using the first two image tiles of the set, and so on. The last calibration is therefore performed using all the image tiles of the set. In the CMY case, there are 20 tiles per calibration set, an optimized and a standard calibration set for each image, and 6 separate images on which cali-

Table 4.3: Prediction accuracy of the classically calibrated IS-YNSN model when predicting the color-constant patches of the 125 UCR set and sets of tiles grouped according to their non-uniformity values. The statistics show the average, 95 percentile, and maximum ΔE_{94}.

	ΔE_{94}		
	avg	95%	max
Uniform patches			
125 UCR set	1.16	1.94	2.54
Non-uniformity interval			
0.00 - 0.04	1.46	2.75	3.77
0.04 - 0.08	1.69	2.95	4.04
0.08 - 0.12	2.12	3.70	4.88
0.12 - 0.16	3.04	5.26	5.65
0.16 - 0.20	4.11	7.53	8.47
0.20 - 0.24	6.31	10.25	11.46
0.24 - 0.28	10.37	15.82	16.94
0.28 - 0.32	16.38	21.35	24.94

Table 4.4: Average accuracy of predicting sets of tiles grouped according to their non-uniformity values using effective coverages either predicted using the classically calibrated ink spreading model (*Predicted* column) or deduced by direct fitting of surface coverages according to Eq. (2.4) (*Deduced* column). ΔCovs indicates the average CMYK norm between predicted and deduced effective coverages.

	Average ΔE_{94}		ΔCovs
	Predicted	Deduced	
Non-uniformity interval			
0.00 - 0.04	1.46	0.29	0.0008
0.04 - 0.08	1.69	0.23	0.0005
0.08 - 0.12	2.12	0.21	0.0022
0.12 - 0.16	3.04	0.25	0.0021
0.16 - 0.20	4.16	0.24	0.0028
0.20 - 0.24	6.31	0.27	0.0045
0.24 - 0.28	10.37	0.33	0.0064
0.28 - 0.32	16.38	0.67	0.0123

Table 4.5: Prediction accuracy of the classically calibrated IS-YNSN model when predicting the 125 CMY set and the different CMY image calibration tile sets. The statistics show the average, 95 percentile, and maximum ΔE_{94}.

	ΔE_{94}		
	avg	95%	max
Uniform patches			
125 CMY set	0.89	1.63	2.53
CMY standard image tiles			
Flower	1.64	2.50	2.53
House	1.67	2.57	3.33
Boat race	1.89	2.83	3.40
Sea	1.74	2.63	2.76
Parrots	1.49	2.29	2.35
Sailboat	1.98	2.75	2.91
CMY optimized image tiles			
Flower	1.66	2.38	2.47
House	1.61	2.50	2.58
Boat race	1.98	2.87	2.97
Sea	1.66	2.40	2.44
Parrots	1.35	2.07	2.09
Sailboat	1.91	2.58	2.69

Table 4.6: Prediction accuracy of the classically calibrated IS-YNSN model when predicting the 125 UCR set, and the different CMYK image calibration tile sets. The statistics show the average, 95 percentile, and maximum ΔE_{94}.

	ΔE_{94}		
	avg	95%	max
Uniform patches			
125 UCR set	1.16	1.94	2.54
CMYK standard image tiles			
Flower	1.49	3.05	3.08
House	1.41	2.52	2.77
Boat race	1.63	3.08	3.12
Sea	1.59	2.60	2.64
Parrots	1.96	3.11	3.21
Sailboat	2.05	2.81	3.13
CMYK optimized image tiles			
Flower	1.64	3.01	3.03
House	1.78	3.00	3.08
Boat race	1.79	3.19	3.40
Sea	1.70	2.80	3.48
Parrots	2.07	3.04	3.10
Sailboat	2.00	3.24	3.32

brations are performed. In the CMYK case, there are 30 tiles per calibration set, an optimized and a standard calibration set for each image, and 6 separate images on which calibrations are performed. The classical calibration and all the image tile based calibrations are compared against the reference YNSN model, i.e. the Yule-Nielsen model without ink spreading curves. The comparison is performed by predicting the color-constant patches of the 125 CMY set, respectively the 125 UCR set. The lower the average ΔE_{94}, the higher the accuracy.

Figures 4.8 and 4.9 show the prediction results for the CMY and CMYK images, respectively. In both cases, the IS-YNSN model with calibrated ink spreading curves is much more accurate than the reference model without calibrated ink spreading curves as shown by the gap between the two dotted lines. This difference outlines the importance of accounting for superposition dependent physical dot gain. Moreover, all the different image calibrations have a better accuracy than the reference model, indicating that even a single image tile provides useful information about physical dot gain to the calibration procedure.

A deeper analysis of the standard image calibrations reveals that adding more image tiles to the calibration set improves the prediction accuracy. A given additional image tile may slightly degrade the prediction accuracy, but additional tiles will cancel the degradation. Degradations may happen when the new added tile is too similar to previous tiles. For example, the 3rd tile of the CMY Flower standard calibration set is a red tile similar to the 2nd tile of the set (Appendix C.2). The 9th tile of the CMYK Parrots standard set is the 3rd dark brown tile of the set (Appendix C.3). In such cases, the calibration procedure yields a model specialized in predicting tiles occurring repeatedly in the calibration set and fails to accurately predict rarely occurring tiles. Adding calibration tiles outside the specialized area quickly improves the prediction accuracy. For example, the first 15 tiles of the CMYK Flower standard set are red or dark brown, except the 9th and 13th tiles that are green (Appendix C.3). When the 1st and 2nd green tiles are added in the set, we observe a sharp decrease of the prediction error to 2.5 and 1.9, respectively. With enough image tiles in the calibration set, the prediction accuracy stabilizes close to the prediction accuracy of the classical calibration. It is however difficult to predict how many tiles are required to stabilize the standard image calibration. For the CMYK Boat race image, less than 5 tiles are required. Conversely, for the CMY Flower image or the CMYK Sailboat image, more than 20 image tiles are necessary.

In contrast, optimized image calibration tile sets offer a high prediction accuracy even with sets composed of as few as 5 to 10 image tiles. Their accuracy is therefore always higher or, when both calibrations are stable, equal to the accuracy of standard image calibrations. When there is a large difference in the sum of the weights between the standard and optimized image

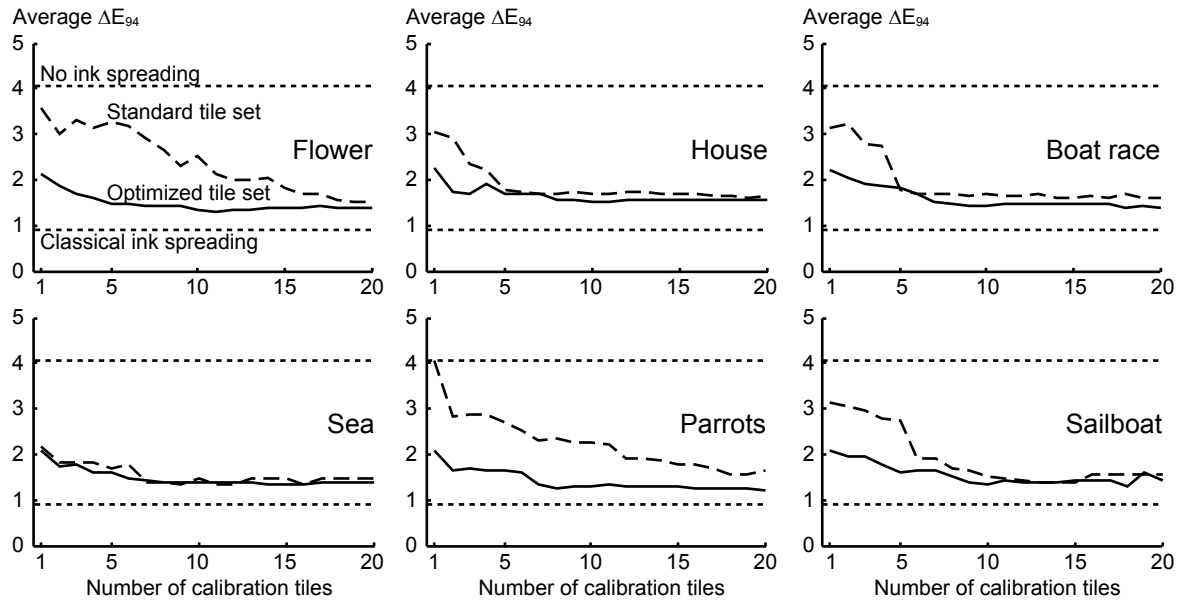


Figure 4.8: Accuracy of different calibrated IS-YNSN models when predicting the color-constant patches of the 125 CMY set. The following calibrations for six different CMY images are shown: no ink spreading (upper dotted lines), classical ink spreading calibration (lower dotted line), ink spreading calibrations using standard image tile sets (dashed line), and ink spreading calibrations using optimized image tile sets.

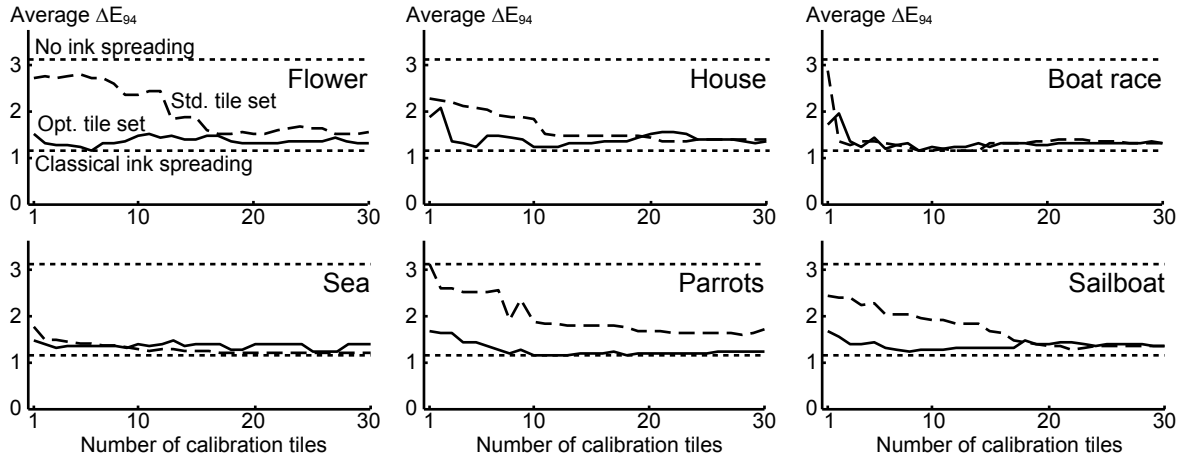


Figure 4.9: Accuracy of different calibrated IS-YNSN models when predicting the CMYK color-constant patches of the 125 UCR set. The following calibrations for six different CMYK images are shown: no ink spreading (upper dotted lines), classical ink spreading calibration (lower dotted line), ink spreading calibrations using standard image tile sets (dashed line), and ink spreading calibrations using optimized image tile sets.

calibration sets (Figure 4.7, Flower image), there is also a large difference in the prediction accuracy when the model is calibrated with a limited number of patches (Figures 4.8 and 4.9). When the difference in the sum is small (Figure 4.7, Sea or Boat race images), the difference in prediction accuracy is also small, i.e. the two curves in Figures 4.8 and 4.9 remain close. Optimized image tile sets therefore ensure that the model achieves good prediction accuracy with the fewest possible number of tiles. Moreover, the number of necessary tiles is lower than the minimum required by a classical calibration. An optimized image tile based calibration requires 5 to 10 image tiles in both CMY and CMYK cases, whereas a classical calibration requires 12 and 20 patches in the CMY and CMYK cases, respectively.

The Parrot image is particularly interesting. It is a colorful image containing many uniform image tiles. Because the standard selection algorithm selects tiles based on uniformity only and although the tiles are selected not too close to each other, the first selected tiles of the standard set are all different shades of green (Appendices C.2 and C.3). Yellow and red tiles appear only later in the set and no blue tile is selected. Since many selected tiles have similar shades, the sum of the weights only slowly increases each time a tile is added. This contrasts with the optimized calibration set whose first selected tiles are well distributed in the color space because the selection algorithm selects tiles maximizing the ink spreading weights. The sum of the weights therefore quickly increases (Figure 4.7). As a consequence, the prediction accuracy achieved by the standard algorithm only slowly improves compared to the accuracy achieved by the optimized algorithm (Figures 4.8 and 4.9). In images with less color diversity or less uniform tiles, e.g. the Sailboat image, the standard algorithm still requires more tiles, but is eventually able to achieve the same accuracy as the optimized algorithm since the number of candidate tiles not too close to each other is reduced and the selected tiles are forced to be well distributed. In some cases such as the Sea or Boat race images, both algorithms have the same performances.

4.3.7 Summary

To facilitate the calibration of the ink spreading curves, we propose a method based on tiles extracted from color images. The detailed selection algorithm automatically selects tiles in an image based on the CMYK pixel values of that image. Moreover, it ensures that the selected image tiles are uniform and not too close to each other in the CMYK domain. We then propose an optimized tile selection algorithm that selects the image tiles maximizing the ink spreading weights, i.e. the tiles having the largest impact on the calibration of the ink spreading curves.

The performed experiments show that an image tile can be accurately predicted only if its non-uniformity is low. Image tiles selected by the optimized selection algorithm largely improve the sum of the ink spreading weights compared to the tiles selected by the standard selection algorithm. Since large ink spreading weights indicate a strong influence of the tiles on the ink spreading curves, the ink spreading calibration is improved. An optimized image calibration set composed of 5 to 10 image tiles is indeed sufficient to reliably calibrate the ink spreading curves whereas it is not possible to provide an upper bound of image tiles for the standard image calibration set. Finally, the number of required image tiles is lower than the number of color-constant patches required for a classical calibration.

Chapter 5

Recovering Neugebauer colorant reflectances and ink spreading curves from printed color images

5.1 Introduction

The YNSN model incorporates parameters which must be learned for each combination of paper, inks, and printing device. This step, known as the calibration of the model, consists in determining the best values for the different model parameters. The calibration of the YNSN model requires the spectral measurements of predetermined color patches representing the Neugebauer primaries, i.e. the reflectances of paper, of the solid inks, and of superpositions of the solid inks, as well as a few halftones to determine the Yule-Nielsen n-value. The YNSN model actually performs a linear interpolation between the extremities of the space formed by the reflectances of the Neugebauer primaries raised to the power of $1/n$. In order to further improve its accuracy, the YNSN model has been enhanced with an ink spreading model accounting for physical dot gain (IS-YNSN, Chapter 2). The calibration of the IS-YNSN model requires the spectral measurements of a set of predetermined halftone patches.

Using the improvements proposed in Chapter 3, 36 spectral measurements of predetermined color-constant halftone and solid patches are required. Since these color-constant patches are not present within normal printed images, they need to be added, for example in the margins of printed pages. Once measured, they must be removed. Instead of calibrating the IS-YNSN model with predetermined color-constant patches, we propose a calibration procedure relying

on the measurements of image tiles extracted from printed color images. In Chapter 4, the content of printed color images has already been used for the calibration of ink spreading, but not for the YNSN model itself, i.e. the Neugebauer primaries still had to be measured and the Yule-Nielsen n-value determined. Balasubramanian proposed to recalibrate the Neugebauer primaries based on a training set of measured samples, but only as an optimization step based on initially measured primary reflectances [Bal96].

In this chapter, we show how to recover the Neugebauer primaries from printed color images. We propose a two-step procedure to calibrate both the YNSN model and the ink spreading model using tiles extracted from printed color images without relying on prior knowledge. In the first step, constraints are imposed both on the ink spreading model and on the Neugebauer primaries to prevent large variations when the information contained in the image calibration tiles is not sufficient. In the second step, the Neugebauer primaries are refined to improve the calibration accuracy. We then show how to determine the Yule-Nielsen n-value.

5.2 Unconstrained calibration of the IS-YNSN model

The IS-YNSN model must be calibrated for each combination of printer, halftone screens, inks, and paper. Such a calibration consists in recovering the reflection spectra of the colorants (Neugebauer primaries), fitting the mid-points of the ink spreading curves, and setting the Yule-Nielsen n-value to a suitable scalar value. In a classical calibration, the colorant reflectances are directly measured and the effective surface coverages used for creating the ink spreading curves are fitted using a predetermined set of color-constant halftone calibration patches.

Instead of using predetermined color-constant patches, we show how the IS-YNSN model can be calibrated using image calibration tiles extracted from color images and grouped in an image calibration set. The selection of these tiles is described in Section 5.5. In this section, we propose an unconstrained procedure to calibrate the IS-YNSN model using an image calibration set. This approach relies on a gradient-descent algorithm to find the optimal colorant reflectances and the mid-points of the ink spreading curves. In the present setting, we assume that the Yule-Nielsen n-value is known. Its calibration is discussed in Section 5.9.

The gradient-descent algorithm uses as objective function the sum of square differences between the measured and predicted reflectance components of the image calibration tiles:

$$f(C, \vec{v}) = \sum_{p} \sum_{k} \left[predSpectrum(\lambda_k, p, C, \vec{v}) - R_p(\lambda_k) \right]^2 \tag{5.1}$$

where λ_k is the wavelength, p an image calibration tile, C the 48x16 matrix whose columns correspond to the 16 CMYK colorant reflectances sampled each 10 nm between 380 nm and 850 nm, $\vec{v}$ a vector containing the mid-points of the 20 CMYK ink spreading curves (see Section 4.2.1), and R_p the measured reflection spectrum of image calibration tile p. Function $predSpectrum$ uses the IS-YNSN model to predict the spectrum of image calibration tile p.

The gradient-descent algorithm performs the following minimization using Eq. (5.1):

$$\{C_{opt} \quad \vec{v}_{opt}\} = \arg\min_{C,\vec{v}}[f(C,\vec{v})] \text{ such that } \begin{cases} R_{\min} \leq R_i \leq R_w \\ 0.25 \leq v_{i/jk} \leq 0.75 \end{cases} \tag{5.2}$$

where matrix C_{opt} and vector $\vec{v}_{opt}$ contain respectively the colorant reflectances and the mid-points of the ink spreading curves minimizing the sum of square differences between the measured and predicted reflectance spectra of the image calibration tiles. The only constraints imposed on the colorant reflectances and the mid-points of the ink spreading curves are their intrinsic bounds. In respect to the ink spreading curve mid-points, 0.25 and 0.75 are the bounds beyond which the quadratic ink spreading curves are not monotonic increasing functions anymore [Eq. (4.1)]. In respect to the colorants reflectances, R_{min} and R_w correspond respectively to the minimal value allowed for a reflectance and the maximal value expressed by the measured paper reflectance. The lower bound for reflectances is actually 0, but R_{min} is set to 0.001 to avoid searching for negative values. Moreover, on real prints, the darkest black has a reflectance larger than 0.001. At the start of the gradient-descent algorithm, the mid-points of the ink spreading curves in $\vec{v}$ are set to 0.5 and the colorant reflectances in C to $(R_{min} + R_{max})/2$.

For several reasons, calibrations of the IS-YNSN model using Eq. (5.2) are not reliable. Because the colorant reflectances (Neugebauer primaries) are not constrained, they are set to spectra that are not smooth and therefore do not look like real CMYK colorant reflectances. In addition, the relationship between the reflectances of the inks and the reflectances of their superpositions is not preserved. Finally, the image calibration tiles do not always provide enough information to accurately calibrate the colorant reflectances. For example, if the image calibration tiles are primarily composed of cyan ink, it is not possible to calibrate the yellow colorant. In such cases, unconstrained procedures such as Eq. (5.2) tend to induce large variations in some spectral components in order to influence the difference metric of Eq. (5.1). As a consequence, the fitted reflectances do not reflect real measurements. Examples of such fitted reflectances are given in Section 5.7.

5.3 Constrained calibration of the IS-YNSN model

In order to prevent the calibration procedure from setting the colorant reflectances and mid-points of the ink spreading curves to extreme values, we impose constraints on these parameters. For the ink spreading model, we use the constraints proposed in Section 4.2.2 that were successful in calibrating the ink spreading model by relying on tiles extracted from color images. For the colorants, we propose a set of constraints that forces the calibration procedure to set their reflectances to spectra that are similar to real reflection spectra.

5.3.1 Ink spreading constraints

The problem faced by the calibration procedure when fitting the mid-points of the ink spreading curves is that the image calibration tiles do not always provide sufficient information about each ink spreading curve in order to set it with high confidence to a sensible value. The goal of the constraints imposed on the ink spreading curves is therefore to restrict the range of values that can be assigned to the mid-points when there is not enough information. Since we use the same constraints as in Section 4.2.2, i.e. Eqs. 5.3, 5.4, 5.5, and 5.6 are the same as Eqs. 4.2, 4.3, 4.4, and 4.7, we only give a short summary of these constraints and discuss them in the context of calibrating the entire IS-YNSN model.

Ink spreading weight $w_{i/jk}$, associated to ink spreading curve $f_{i/jk}$ of halftone ink i superposed with solid inks j and k, indicates if there is enough information to fit the mid-point $v_{i/jk}$ of ink spreading curve $f_{i/jk}$. A weight equal to 0 indicates that no information is available. In this case, the effective surface coverage is made equal to the nominal surface coverage and the mid-point is set to 0.5, i.e. we assume that there is no ink spreading. A weight equal to 1 indicates that the mid-point $v_{i/jk}$ can be freely calibrated between 0.25 and 0.75, the bounds beyond which the ink spreading curves are not monotonic increasing functions anymore. For a weight $w_{i/jk}$ between 0 and 1, the mid-point $v_{i/jk}$ is constrained as follows:

$$0.5 - 0.25 \cdot w_{i/jk} \leq v_{i/jk} \leq 0.5 + 0.25 \cdot w_{i/jk} \tag{5.3}$$

The ink spreading weights $w_{i/jk}$ are computed from the nominal ink coverages of the image tiles as follows. Given a tile p of the image calibration set, let $u_{i,p}$ be the nominal surface coverage and $u'_{i,p}(v_i, v_{i/j}, v_{i/k}, v_{i/jk})$ be the effective surface coverage of ink i within tile p computed using the ink spreading model. The weight associated to ink spreading curve $f_{i/jk}$ of

halftone ink i superposed with solid inks j and k is defined as the following gradient:

$$w_{i/jk,p} = \frac{\partial u'_{i,p}}{\partial v_{i/jk}} \tag{5.4}$$

Since ink spreading curve $f_{i/jk}$ is fully determined by its mid-point $v_{i/jk}$, the gradient $w_{i/jk,p}$ in Eq. (5.4) expresses the influence of ink spreading curve $f_{i/jk}$ on the resulting effective surface coverage $u'_{i,p}$ of ink i within tile p.

Since a calibration set is composed of several image tiles, we define the weight associated to ink spreading curve $f_{i/jk}$ for the entire image calibration set as the maximum of the weights of all the tiles:

$$w_{i/jk} = \max_p[w_{i/jk,p}] = \max_p \frac{\partial u'_{i,p}}{\partial v_{i/jk}} \tag{5.5}$$

We take the maximum of the derivatives across all the tiles because tiles with a high weight have the largest influence on the metric minimized by the ink spreading curve mid-point fitting algorithm Eq. (5.1), i.e. when there is at least one high weight tile and low weight tiles, the ink spreading curve mid-point is mainly fitted by the high weight tile.

By inserting Eq. (4.1) into Eq. (3.1), we can compute the gradients of Eq. (5.4) for each ink spreading curve. For example, the weight $w_{c/m}$ of the ink spreading curve $f_{c/m}$ of cyan halftones printed on solid magenta is:

$$w_{c/m} = \frac{\partial c'}{\partial v_{c/m}} = m'(1-y')\frac{\partial f_{c/m}(c)}{\partial v_{c/m}} = m'(1-y') \cdot 4c(1-c) \tag{5.6}$$

where c', m' and y' are the effective surface coverages of the cyan, magenta, and yellow halftones, respectively, and c is the nominal surface coverage of the cyan halftone. Since neither the colorant reflectances nor the ink spreading curves are calibrated yet, the effective coverages m' and y' required in Eq. (5.6) cannot be computed. We use the nominal coverages m and y instead. This modification influences the weights, but since the difference between effective and nominal coverages is usually small, the weights are not largely affected and the ink spreading mid-point values can be accurately calibrated. This assertion is further discussed in Section 5.10.4.

5.3.2 Constraints in respect to colorant reflectances

The constraints imposed on the ink spreading curves rely on the existence of a default value for the mid-points $v_{i/jk} = 0.5$ that can be used when no information is available. There is no

such default value for the colorant reflectances. Instead, we make the distinction between the following three groups of colorants: the paper, the inks, and the ink superpositions. We assume that the paper reflectance representing the white colorant can be measured. The white colorant is therefore known. We then impose different constraints on the ink reflectances and on the ink superposition reflectances.

Ink reflectance constraints

The reflectance of an ink depends on the substrate it is printed on. The first constraint imposed on ink reflectance R_i expresses this dependence by assuming that light reflected by the print is attenuated twice by the ink layer and once by being reflected by the paper:

$$R_i = R_w T_i^2 \tag{5.7}$$

where R_i is the reflectance of ink i on paper, R_w the paper reflectance, and T_i the transmittance of ink i.

Using Eq. (5.7), ink reflectances are expressed by the corresponding ink transmittances. We can therefore impose constraints on these transmittances. The spectrum of a transmittance should be smooth and have a general shape given by the color of the ink. Using principal component analysis (PCA [FB04, TB05]), we can constrain ink transmittances by limiting the number of principal components expressing these transmittances while retaining both smoothness and general shape of their spectra. For each ink color, we assemble a list of ink transmittances based on measurements coming from offset prints, ink jet prints, thermal transfer prints, etc. We then extract the average transmittance as well as the first three principal vectors from these ink transmittances. The principal vectors for ink i are stored in a matrix W_i. W_i has therefore three columns. The transmittance of an ink is then:

$$T_i = W_i \begin{bmatrix} c_{i,1} & c_{i,2} & c_{i,3} \end{bmatrix}^T + \overline{T_i} \tag{5.8}$$

where T_i is the transmittance of ink i; W_i is the matrix holding the principal vectors for the transmittances of ink i; $\overline{T_i}$ is the average transmittance of ink i; and $c_{i,1}$ to $c_{i,3}$ are three scalars, further referred to as the principal components of ink i. Using Eqs. (5.7) and (5.8), the reflectance R_i of ink i can be expressed using only three principal components, i.e. $c_{i,1}$ to $c_{i,3}$. The 12 principal components required for CMYK are listed in Table 5.1. Appendix A provides more details about the accuracy of the implemented PCA.

Table 5.1: List of the 12 principal components and 28 ink thickness variation factors required for CMYK prints. All these parameters are scalar values.

Ink	Principal Components			Ink thickness variation factors						
Cyan	$c_{c,1}$	$c_{c,2}$	$c_{c,3}$	d_{Cm}	d_{Cy}	d_{Ck}	d_{Cmy}	d_{Cmk}	d_{Cyk}	d_{Cmyk}
Magenta	$c_{m,1}$	$c_{m,2}$	$c_{m,3}$	d_{cM}	d_{My}	d_{Mk}	d_{cMy}	d_{cMk}	d_{Myk}	d_{cMyk}
Yellow	$c_{y,1}$	$c_{y,2}$	$c_{y,3}$	d_{cY}	d_{mY}	d_{Yk}	d_{cmY}	d_{cYk}	d_{mYk}	d_{cmYk}
Black	$c_{k,1}$	$c_{k,2}$	$c_{k,3}$	d_{cK}	d_{mK}	d_{yK}	d_{cmK}	d_{cyK}	d_{myK}	d_{cmyK}

Ink superposition reflectance constraints

The reflectance of superposed inks depends on the substrate on which they are printed as well as on the individual ink transmittances. However, a simple extension of Eq. (5.7) to multiple inks is not appropriate. Hersch et al. have shown that the relative thickness of a printed solid ink layer depends on the presence of other superposed inks [HBB+09]. Beer's law enables expressing these ink thickness variations. We adapt Eq. (5.7) to account for modified relative ink thicknesses:

$$R_{ij} = R_w T_i^{2d_{Ij}} T_j^{2d_{iJ}} \tag{5.9}$$

$$R_{ijk} = R_w T_i^{2d_{Ijk}} T_j^{2d_{iJk}} T_k^{2d_{ijK}} \tag{5.10}$$

$$R_{ijkl} = R_w T_i^{2d_{Ijkl}} T_j^{2d_{iJkl}} T_k^{2d_{ijKl}} T_l^{2d_{ijkL}} \tag{5.11}$$

where R_{ij}, R_{ijk}, and R_{ijkl} are the reflectances of the superposition of 2, 3, and 4 solid inks, respectively; R_w is the paper reflectance; T_i, T_j, T_k, and T_l are the transmittances of inks i, j, k, and l computed by Eq. (5.8); and d_{Ij} and d_{iJ} are the scalar ink thickness variation factors of ink i superposed with ink j and of ink j superposed with ink i. For four inks, we have 6 colorants that are composed of two inks, 4 colorants composed of 3 inks, and a single colorant composed of all 4 inks. Based on Eqs. (5.9) to (5.11), only the 28 scalar ink thickness variation factors listed in Table 5.1 are required to model the 11 colorant reflectances composed of 2 or more inks.

5.3.3 Constrained gradient-descent calibration of the IS-YNSN model

Using the constrained colorant reflectances expressed by Eqs. (5.7) to (5.11), we can modify the objective function of Eq. (5.1) as follows:

$$f(\vec{v},\vec{c},\vec{d}) = \sum_{p}\sum_{k}\left[predSpectrum(\lambda_k,p,\vec{v},\vec{c},\vec{d}) - R_p(\lambda_k)\right]^2 \tag{5.12}$$

where we have added $\vec{c}$ the vector containing the principal components of the inks, and $\vec{d}$ the vector containing the ink thickness variation factors of the ink superpositions.

Finally, using Eq. (5.2) and the ink spreading constraints of Eq. (5.3), the calibration procedure becomes:

$$\{\vec{v}_{opt},\vec{c}_{opt},\vec{d}_{opt}\} = \underset{\vec{v},\vec{c},\vec{d}}{\arg\min}\left[f(\vec{v},\vec{c},\vec{d})\right] \text{ such that } \begin{cases} v_{i/jk} \leq 0.5 + 0.25\cdot w_{i/jk} \\ v_{i/jk} \geq 0.5 - 0.25\cdot w_{i/jk} \\ R_{min} \leq R_i \leq R_w \end{cases} \tag{5.13}$$

where $\vec{v}_{opt}$, $\vec{c}_{opt}$, and $\vec{d}_{opt}$ are respectively the optimal mid-points of the ink spreading curves, optimal principal components of the ink transmittances, and optimal ink thickness variation factors. These optimal values minimize the sum of square differences between the measured and predicted spectra. Note that we maintain the intrinsic bounds for the colorants expressed by R_{min} and R_w, i.e. the minimal value allowed for a reflectance and the maximal value expressed by the paper white. At the start of the gradient-descent algorithm, the mid-points of the ink spreading curves in $\vec{v}$ are set to 0.5, the principal components of the ink transmittances in $\vec{c}$ to 0, and the ink thickness variation factors in $\vec{d}$ to 1.

The colorant reflectances recovered from the optimal principal components and optimal ink thickness variation factors fitted in Eq. (5.13) are close to the measured colorant reflectances if there is enough information to fit them. However, these recovered colorant reflectances are subject to strong constraints and cannot exactly match the measured colorant reflectances, even when there would be enough information in the image calibration set to achieve such a match. We therefore propose a second calibration step to improve the recovered colorant reflectances.

5.4 Surface coverage based optimization of the colorant reflectances

The constraints imposed on the colorant reflectances in Eqs. (5.7) to (5.11) enable the first calibration step in Eq. (5.13) to recover colorant reflectances similar to real reflectances even when there is not enough information in the image calibration set. However, because of the

PCA, the recovered colorant reflectances are restricted to a limited set of reflection spectra. If the real reflectances lay outside of this set, it is not possible to optimally fit the colorant reflectances to match these real reflectances and the IS-YNSN model will not be able to offer the most accurate predictions. To improve the match between the recovered and real colorant reflectances, we perform a second calibration step in which only the colorant reflectances are modified. The mid-points of the ink spreading curves are considered to be accurately calibrated by the minimization performed by Eq. (5.13) and are not modified in this second step.

5.4.1 Surface coverage based colorant reflectance constraints

Thanks to the first calibration step performed in Eq. (5.13), an approximation of the colorant reflectances is available and we can use a similar approach as in Section 5.3.1. If no information is available in the image calibration set to fit reflectance R_i of colorant i, the colorant reflectance recovered in the first calibration step R_i^{old} is not modified. The more information is available, the more freedom is given to fit colorant reflectance R_i. *Colorant weight* ω_i, associated to colorant i, indicates the amount of information available to fit colorant reflectance R_i. A weight equal to 0 indicates that no information is available. In this case, the colorant reflectance is set to R_i^{old}, the spectrum recovered in Eq. (5.13). A weight equal to 1 indicates that the colorant reflectance can be freely calibrated between the minimal value allowed for a reflectance and the paper reflectance. For colorant weight ω_i between 0 and 1, colorant reflectance R_i is constrained as follows:

$$\begin{cases} R_i(\lambda) \leq (1-\omega_i)R_i^{old}(\lambda) + \omega_i \cdot R_w(\lambda) \\ R_i(\lambda) \geq (1-\omega_i)R_i^{old}(\lambda) + \omega_i \cdot R_{\min} \end{cases} \tag{5.14}$$

where R_i is the considered colorant reflectance, R_i^{old} the corresponding colorant reflectance recovered in Eq. (5.13), R_w the measured paper reflectance, and R_{min} the minimal value allowed for a reflectance.

The colorant weights ω_i are computed from the nominal ink coverages of the image tiles as follows. For each tile p of the image calibration set, we compute the effective ink coverages using Eq. (3.1). We then input these effective ink coverages to the Demichel equations to obtain the effective colorant coverages [Eq. (2.2)]. An effective colorant coverage $a_{i,p}$ indicates the area effectively covered by colorant i of reflectance R_i in tile p. The larger the area covered by colorant i, the more information about that colorant is available. We therefore consider the effective colorant coverage $a_{i,p}$ as a metric indicating the amount of information about colorant

i contained in image tile p. We also define the weight ω_i associated to colorant i for the entire image calibration set as the maximum effective colorant coverage among the tiles:

$$\omega_i = \max_p [a_{i,p}] \tag{5.15}$$

Note that in contrast to the ink spreading weights, the colorant weights are computed from effective coverages instead of nominal coverages. This is possible because the ink spreading curves are calibrated in the first calibration step performed in Eq. (3.1). Moreover, a weight $\omega_i = 1$ indicates that within the measured reflectances of the calibration set, there is also the measured colorant reflectance R_i. This is always true for the white colorant since the paper reflectance is measured. When a measured colorant reflectance is present, that colorant is removed from the set of unknown colorant reflectances to be optimized.

The colorant weights ω_i define the bounds within which the colorant reflectances can be set. However, because each colorant reflectance is a spectrum composed of many elements, usually 36 or 48, the optimization procedure has a tendency to let not only high weight tiles influence the fitted colorant reflectances, but also middle weight tiles. Since information contained in middle weight tiles is not as relevant as information contained in high weight tiles, the influence of middle weight tiles must be reduced. We achieve this by defining weights ω_p associated to the image tiles. An image tile is considered relevant for the calibration if at least one of its effective colorant coverages is large. Image tile weight ω_p associated to image tile p is set to the maximum effective colorant surface coverage present:

$$\omega_p = \max_{\{i|\omega_i \neq 1\}} [a_{i,p}] \tag{5.16}$$

Note that the maximum only considers the effective coverages of colorants that must be optimized. Since the sum of the effective colorant coverages of a tile is equal to 1, a tile cannot have both middle and high colorant coverages. Either it has a single large colorant coverage and the remaining colorant coverages are small, or it has two or three middle colorant coverages and the remaining colorant coverages are also small. In the first case, its associated image tile weight ω_p is large and the influence of this tile within the optimization procedure is not reduced. In the second case, its associated image tile weight ω_p is small and its influence on the optimization procedure is reduced.

Using both image tile weights ω_p and colorant weights ω_i, we show in Section 5.4.2 how to design an optimization procedure ensuring respectively that (1) an image tile contributes to

Table 5.2: List of variables used to express matrix and vector sizes.

Variable	Description	Typical value
n_c	Number of colorants	16 in the case of CMYK
n_i	Number of colorants satisfying $\omega_i \neq 1$	15 if only paper is measured
n_p	Number of calibration patches	50 in Sections 5.6 through 5.10
n_k	Number of wavelength components composing reflectances	48 when sampled every 10 nm between 380 nm and 850 nm

the computation of the colorant reflectances only if it provides valuable information, and (2) a given colorant reflectance is modified only when valuable information is available.

5.4.2 Least-squares method to optimize the colorant reflectances

The optimization procedure used for the second calibration step is based on a weighted and constrained least-squares approach derived from Eq. (2.3). The matrix and vector sizes have the values shown in Table 5.2. Let us first rewrite Eq. (2.3) in a linear form [UG06]:

$$Q_p(\lambda_k) = R_p(\lambda_k)^{1/n} = \sum_i a_{i,p} R_i(\lambda_k)^{1/n} = \sum_i a_{i,p} Q_i(\lambda_k) \tag{5.17}$$

where $Q = R^{1/n}$ is a n_k x 1 reflectance raised to the power $1/n$, further referred to as an n–modified reflectance. Since each wavelength λ_k is independent, we can rewrite Eq. (5.17) as follows:

$$q_{p,k} = \sum_i a_{i,p} q_{i,k} = \vec{a}_p^T \vec{\theta}_k \tag{5.18}$$

where $q_{p,k}$ is the k^{th} component of the n–modified reflectance of image tile p, $a_{i,p}$ the effective coverage of colorant i for image tile p, $q_{i,k}$ the k^{th} component of the n–modified reflectance of colorant i, $\vec{a}_p$ the n_c x 1 vector containing the effective coverages of all the colorants for image tile p, and θ_k the n_c x 1 vector containing the k^{th} components of all the colorant reflectances. If some colorant reflectances are known, i.e. $\omega_i = 1$, we adjust Eq. (5.18) by determining the reflectance part $\tilde{q}_{p,k}$ of the unknown colorant reflectances:

$$\tilde{q}_{p,k} = q_{p,k} - \sum_{\{i|\omega_i=1\}} a_{i,p} q_{i,k} = \sum_{\{i|\omega_i\neq 1\}} a_{i,p} q_{i,k} = \vec{a}_p'^T \vec{\theta}_k' \tag{5.19}$$

where $\vec{a}'_p$ and $\vec{\theta}'_k$ are both n_i x 1 vectors. Eq. (5.19) can be written for each of the p image calibration tiles. By combining the equations of all the image calibration tiles, we obtain the following equation:

$$\vec{q}_k = \begin{bmatrix} \tilde{q}_{1,k} \\ \vdots \\ \tilde{q}_{n_p,k} \end{bmatrix} = \begin{bmatrix} \vec{a}'^T_1 \\ \vdots \\ \vec{a}'^T_{n_p} \end{bmatrix} \vec{\theta}'_k = A\,\vec{\theta}'_k \tag{5.20}$$

where A is a n_p x n_i matrix whose elements correspond to the effective colorant coverages $a_{i,p}$. Since the effective colorant coverages $a_{i,p}$ can be computed using the ink spreading curves and since the reflectances $R_p(\lambda_k)$ of the image calibration tiles are measured, colorant coverage matrix A and vector $\vec{q}_k$ of size n_p x 1 are known. Based on Eq. (5.20), we formulate the following typical least-squares problem:

$$\vec{\theta}'_k = \arg\min_{\vec{\theta}} \left[(\vec{q}_k - A\vec{\theta})^T (\vec{q}_k - A\vec{\theta}) \right] = LSQ(\vec{q}_k, A) \tag{5.21}$$

Let us now introduce the image tile weights ω_p to weight the least-squares residuals and reduce the influence of calibration tiles that do not provide valuable information. We first construct a diagonal matrix W of size n_p x n_p with the following diagonal elements:

$$W_{pp} = \omega_p \tag{5.22}$$

Using matrix W, we formulate the following weighted least-squares problem [Bjö96]:

$$\begin{aligned} \vec{\theta}'_k &= WLSQ(\vec{q}_k, A, W) \\ &= \arg\min_{\vec{\theta}} \left[(\vec{q}_k - A\vec{\theta})^T W (\vec{q}_k - A\vec{\theta}) \right] \\ &= \arg\min_{\vec{\theta}} \left[(W^{1/2}\vec{q}_k - W^{1/2}A\vec{\theta})^T (W^{1/2}\vec{q}_k - W^{1/2}A\vec{\theta}) \right] \\ &= LSQ(W^{1/2}\vec{q}_k, W^{1/2}A) \end{aligned} \tag{5.23}$$

The final least-squares problem becomes:

$$\vec{\theta}'_k = LSQ(W^{1/2}\vec{q}_k, W^{1/2}A) \quad \text{such that} \begin{cases} \theta_{i,k} = Q_i(\lambda_k) \leq \left((1-\omega_i)R_i^{old}(\lambda_k) + \omega_i \cdot R_w(\lambda_k)\right)^{1/n} \\ \theta_{i,k} = Q_i(\lambda_k) \geq \left((1-\omega_i)R_i^{old}(\lambda_k) + \omega_i \cdot R_{\min}\right)^{1/n} \end{cases} \tag{5.24}$$

where $Q_i(\lambda_k)$ is the k^{th} component of the reflectance of the i^{th} colorant raised to the power $1/n$, ω_i the colorant weight associated to colorant reflectance R_i, $R_i^{old}(\lambda)$ the reflection spectrum of colorant i recovered in Eq. (5.13), $R_w(\lambda)$ the reflection spectrum of the paper substrate, and R_{min} the minimum value allowed for a reflectance. This least-squares problem is solved for each wavelength λ_k to obtain the entire optimized colorant reflectances.

5.5 Selection of image calibration tiles

As seen in Chapter 4 as well as in [HBBH09], we should use uniform image tiles to calibrate the IS-YNSN model. The algorithm used in this chapter to extract such uniform tiles from color images is similar to the optimized selection algorithm described in Section 4.3.3 and relies on the same non-uniformity metric and adjustment function, i.e. Eqs. (5.25) and (5.26) correspond to Eqs. (4.19) and (4.20). However, the algorithm creates two different sets of image tiles. The first set, further referred to as the *image calibration set*, is used to calibrate the IS-YNSN model and is composed of 50 image tiles. The second set, further referred to as the *image test set*, is used to test the accuracy of the calibrated model and is composed of 30 image tiles.

Two concepts, non-uniformity and proximity in the CMYK color space, are used to select tiles. The non-uniformity value of a tile located within a small sub-domain of the considered color image is computed using its c, m, y, and k pixel values as follows:

$$u = \sqrt{\left(\sigma(c)^2 + \sigma(m)^2 + \sigma(y)^2 + \sigma(k)^2\right)/4} \tag{5.25}$$

where u is the non-uniformity value; c, m, y, and k are the cyan, magenta, yellow, and black pixel values of the considered image tile ranging from 0 to 1; and the σ function is the standard deviation. In addition to selecting uniform tiles, the algorithm ensures that the tiles are not too close to other already selected tiles. We define *not too close* using the Euclidian distance (norm) in the CMY or CMYK color space. For a CMY tile to be selected, it must not have any tile within a CMY norm of 0.5 and no more than one other tile within a CMY norm of 1.5. For a CMYK tile to be selected, it must not have any tile within a CMYK norm of 0.3 and no more than one other tile within a CMYK norm of 0.9. When a tile respects these constraints, it is said to be not too close to the already selected tiles.

The first phase of the algorithm consists in listing the candidate tiles from which calibration and test tiles are selected. Each image is scanned horizontally and vertically in 2 mm steps. Successive 5 mm large square tiles form the candidate tiles. Each tile is associated with a

non-uniformity value according to Eq. (5.25). Tiles with a non-uniformity value above 0.1 are discarded.

The second phase consists in the actual selection of tiles. The selection algorithm not only selects uniform tiles, but also tiles best suited to calibrate the model. Because it is more important to accurately calibrate the colorant reflectances than the ink spreading curves, we define the best tiles as the tiles maximizing the colorant weights ω_i computed in Eq. (5.25). This is different from Section 4.3.3 in which the best tiles where defined as the tiles maximizing the ink spreading weights $w_{i/jk}$. To maximize the colorant weights ω_i, the selection algorithm needs to find among the candidate tiles the tiles that both have large effective colorant coverages a_i and are uniform. To account for both aspects, the effective colorant coverages are adjusted according to their non-uniformity value u as follows:

$$\tilde{a}_{i,p} = a_{i,p} \cdot e^{-(u_p/0.05)^4} \tag{5.26}$$

where $\tilde{a}_{i,p}$ is the uniformity-adjusted effective colorant coverage of tile p, $a_{i,p}$ is the effective colorant coverage of tile p, and u_p is the non-uniformity value of tile p computed in Eq. (5.25). If the given tile is uniform, i.e. its non-uniformity value u_p is low, the adjustment function is close to unity and the effective colorant coverage is not modified. If the tile is non-uniform, its non-uniformity value is high and the adjustment function is close to 0. In that case, the tile is discarded. For each colorant, the selection algorithm chooses the tiles (two for CMY images and one for CMYK images) having the maximum uniformity-adjusted effective colorant coverage $\tilde{a}_{i,p}$ and that are not too close to already selected tiles according to the CMY or CMYK norm. Since image tiles having a large maximum uniformity-adjusted effective colorant coverage $\tilde{a}_{i,p}$ also have a large colorant weight ω_i, the proposed selection algorithm ensures that the colorant weights ω_i computed in Eq. (5.15) are optimal. Once the 16 tiles maximizing the colorant weights ω_i are selected, the algorithm completes the image calibration set by selecting the 34 most uniform image calibration tiles that are not too close to already selected tiles. The final calibration set is therefore composed of 50 image tiles. In Section 5.10, we examine if smaller sets can be used.

The image test set created by the selection algorithm is composed of the 30 most uniform tiles not part of the image calibration set and not too close to already selected test tiles according to the CMY or CMYK norm.

5.6 Setup of the experiments

In order to analyze how the different calibration procedures perform, we use the 21 different sRGB images shown in Appendix D.1. The first three images come from a digital camera and the other eighteen from different sources providing standard test images [Fra, Web, Lev]. The sizes of the images have not been modified, but they have been resampled to 600 dpi. Moreover, they have been converted to two different color spaces: CMY and CMYK. Both conversions have been performed using Photoshop and the U.S. Web coated (SWOP) v2 CMYK profile. The conversions to CMY and CMYK have been performed without GCR (Gray Component Replacement) and with medium GCR, respectively.

The image calibration and test tiles selected using the algorithm described in Section 5.5 are shown in Appendix D.2 for the CMY images and in Appendix D.3 for the CMYK images. They are printed on a Canon PixmaPro 9500 inkjet printer at 600 dpi using classical rotated halftone screens. The CMY images are printed at 150 lpi with CMY halftone screen angles at 75°, 15°, and 45°, respectively. The CMYK images are printed at 100 lpi with CMYK halftone screen angles at 75°, 15°, 0° and 45°, respectively. The size and shape of the image tiles are adapted to the Datacolor MF45IR spectrophotometer. The tiles are conceived as disks with a diameter of 3 mm. Since this device illuminates the sample using a directed source (45°d:0° geometry), we measure each sample four different times, each time rotating it by 90° and taking the average. This reduces errors due to the positioning of the device and irregularities of the illumination geometry. The non-uniformity value of a given tile is computed on an area slightly larger than the actual tile, i.e. a 5 mm disk instead of a 3 mm disk. Ensuring that the area surrounding the actual tile is also uniform further reduces positioning errors.

To compare the prediction accuracy of the different calibrations on color-constant patches well distributed across the printable gamut, we use four test sets composed of color-constant patches. The first test set, further referred to as the *117 CMY set*, is composed of 125 color-constant CMY patches with all possible combinations of cyan, magenta, and yellow at 0%, 25%, 50%, 75%, and 100% surface coverages from which we exclude the 8 CMY colorants. The second set, further referred to as the *64 CMY halftone set*, is composed of 64 color-constant CMY patches with all possible combinations of cyan, magenta, and yellow at 20%, 40%, 60%, and 80% surface coverages. The third and fourth sets, further referred to as the *117 UCR set* and *64 UCR halftone set*, respectively, are composed of 117 and 64 color-constant CMYK patches with nominal surface coverages obtained by converting the 117 CMY set and 64 CMY set to CMYK using the standard under color removal algorithm described above.

5.7 Accuracy of the recovered colorant reflectances

For each of the 21 image calibration sets extracted from the 21 sRGB images, each composed of 50 image tiles, we perform the following three calibrations: the *unconstrained calibration* described in Section 5.2, the *constrained calibration* described in Section 5.3, and the surface coverage based *optimized image calibration* described in Section 5.4. We then compare the recovered colorant reflectances with the measured reflectances. Figures 5.1 to 5.5 focus on five different cases: the yellow colorant in the *Hats* image (Figure 5.1), the red colorant in the *Yacht* image (Figure 5.2), the blue colorant in the *Hats* image (Figure 5.3), the green colorant in the *Lighthouse II* image (Figure 5.4), and the black colorant in the *Handa Island* image (Figure 5.5). In all the figures, the light gray region indicates the unconstrained colorant bounds introduced in Eq. (5.2), i.e. the area between R_{min} and R_w. The dark gray region indicates the optimized colorant bounds determined by the colorant weights ω_i in Eq. (5.14). The colorant reflectances recovered using the unconstrained (thin solid line) and constrained (dotted line) image calibrations are restricted to the light gray region, whereas the colorant reflectances recovered using the optimized image calibration (dashed line) are restricted to the dark gray region. The measured colorant reflectances (thick solid line) are also shown.

As seen in Figures 5.1 to 5.5, the recovered unconstrained reflectances (thin solid lines) do not match the measured reflectances (thick solid lines) and show a CIELAB ΔE_{94} error close to or above 20. Moreover, the unconstrained red colorant in Figure 5.2 reaches the lower bound of the light gray region, which creates acute angles in the spectrum, angles that are not found in real reflectances. Such reflectances are not suitable for predictions and show the limits of the unconstrained calibration procedure.

The recovered constrained reflectances (dotted lines) fit much better the measured reflectances, but the shapes of the reflectances are not always accurate. For example, in Figure 5.1, the transition at 500 nm is not as sharp as the transition of the measured reflectance. In Figure 5.2, the bump at 500 nm is not present. Conversely, in Figure 5.3, a hollow is present at 570 nm in the constrained reflectance and not in the measured reflectance. Although not always perfectly aligned with the corresponding measured reflectances, the constrained reflectances have shapes similar to the measured reflectances and display an average ΔE_{94} error of 5.79 as shown in Table 5.3.

The largest ΔE_{94} errors occur for the recovered reflectances of colorants incorporating the black ink, e.g. the colorant shown in Figure 5.5. Because the GCR algorithm uses the black ink only to replace large percentages of cyan, magenta, and yellow, it is not possible for an image

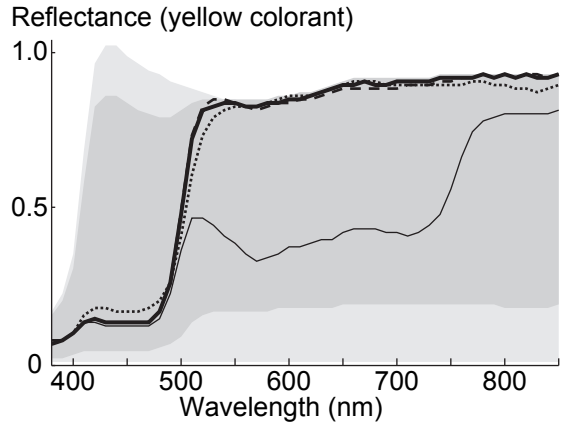


Figure 5.1: Measured reflectance (thick solid line), unconstrained (thin solid line), constrained (dotted line), and optimized (dashed line) recovered reflectances for the yellow colorant of the *Hats* image (colorant weight $\omega_{yellow} = 0.80$). The ΔE_{94} error between the measured reflectance and the unconstrained, constrained, and optimized fitted reflectances are respectively 26.52, 2.66, and 0.56.

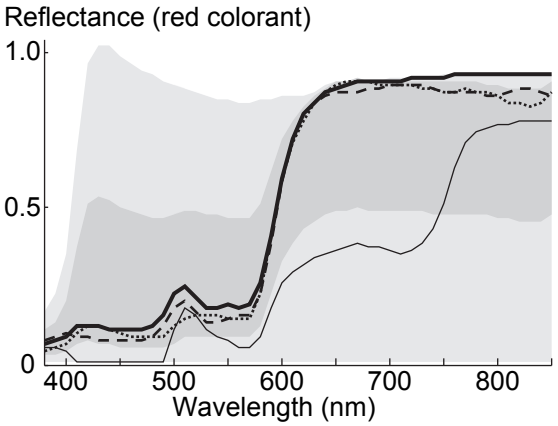


Figure 5.2: Measured reflectance (thick solid line), unconstrained (thin solid line), constrained (dotted line), and optimized (dashed line) recovered reflectances for the red colorant of the *Yacht* image (colorant weight $\omega_{red} = 0.46$). The ΔE_{94} error between the measured reflectance and the unconstrained, constrained, and optimized fitted reflectances are respectively 25.42, 4.72, and 3.74.

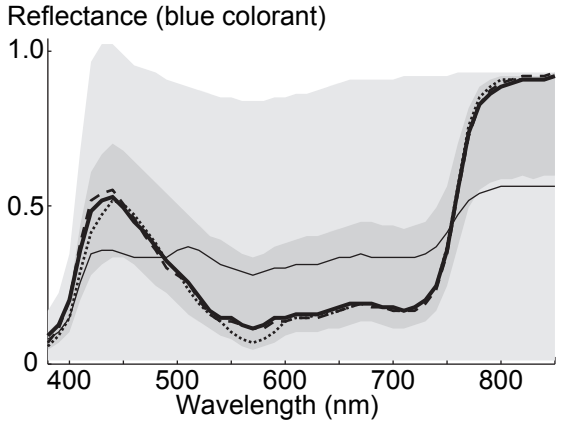


Figure 5.3: Measured reflectance (thick solid line), unconstrained (thin solid line), constrained (dotted line), and optimized (dashed line) recovered reflectances for the blue colorant of the *Hats* image (colorant weight $\omega_{blue} = 0.35$). The ΔE_{94} error between the measured reflectance and the unconstrained, constrained, and optimized fitted reflectances are respectively 34.82, 4.08, and 1.98.

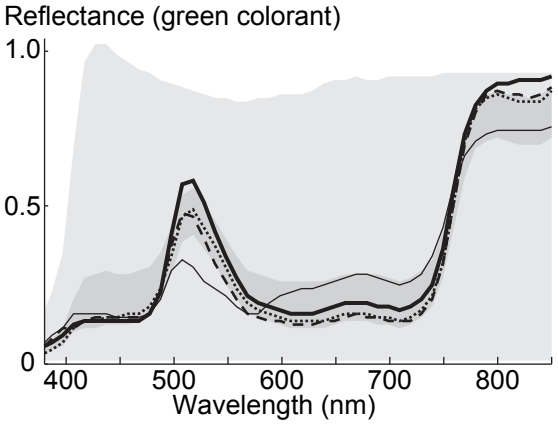


Figure 5.4: Measured reflectance (thick solid line), unconstrained (thin solid line), constrained (dotted line), and optimized (dashed line) recovered reflectances for the green colorant of the *Lighthouse II* image (colorant weight $\omega_{green} = 0.17$). The ΔE_{94} error between the measured reflectance and the unconstrained, constrained, and optimized fitted reflectances are respectively 19.95, 5.48, and 8.25.

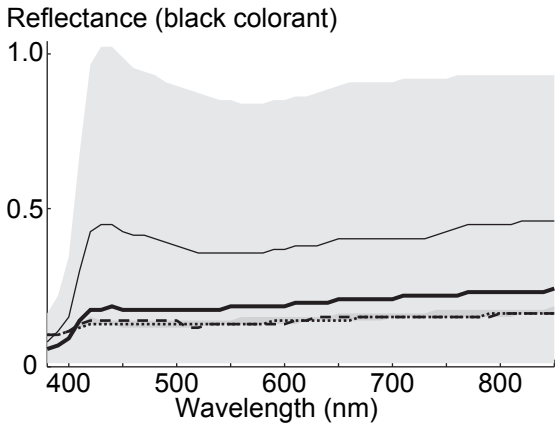


Figure 5.5: Measured reflectance (thick solid line), unconstrained (thin solid line), constrained (dotted line), and optimized (dashed line) recovered reflectances for the black colorant (black ink) of the *Handa Island* image (colorant weight $\omega_{black} = 0.02$). The ΔE_{94} error between the measured reflectance and the unconstrained, constrained, and optimized fitted reflectances are respectively 19.06, 7.23, and 8.23.

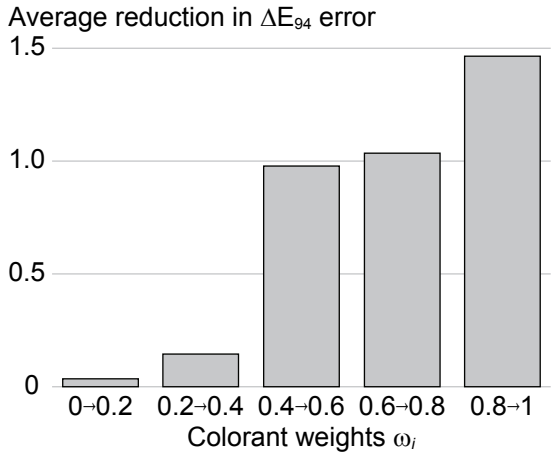


Figure 5.6: Average reduction in ΔE_{94} error between recovered optimized and constrained colorant reflectances according to colorant weights ω_i. A reduction of 1 ΔE_{94} indicates that the optimized reflectances are on average more accurate than the constrained reflectances by 1 ΔE_{94}.

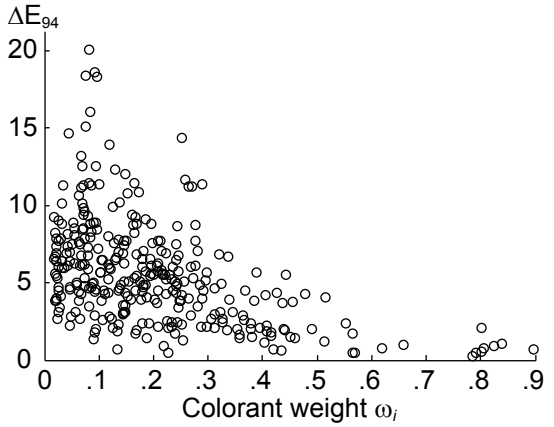


Figure 5.7: ΔE_{94} error between measured and recovered optimized colorant reflectances according to colorant weights ω_i. The figure shows that the higher colorant weight ω_i, the more accurate the recovered optimized colorant reflectance.

Table 5.3: Average ΔE_{94} error between recovered and measured colorant reflectances according to different calibrations. For the constrained and optimized calibrations, the average is computed from the 21 different calibrations, each performed using the 50 calibration tiles extracted from one of the 21 images. The averages do not include the white colorant which is measured.

	Average ΔE_{94}		
	Constrained	Optimized	Optimized distributed 117 patches
CMY prints			
7 CMY colorants	4.16	3.63	1.34
CMYK prints			
15 CMYK colorants	5.79	5.61	3.83
7 CMY colorants	4.34	4.02	2.19

Table 5.4: Average ΔE_{94} error between the measured and 21 recovered colorant reflectances for the constrained and optimized image calibrations, as well as the average and standard deviation of the corresponding colorant weights.

	Colorants						
	C	M	CM	Y	CY	MY	CMY
Average ΔE_{94}							
Constrained	3.45	3.54	4.46	2.59	4.42	3.46	8.45
Optimized	4.72	3.10	4.99	1.62	4.12	2.44	7.14
Colorant weights							
Average	0.27	0.16	0.18	0.45	0.32	0.45	0.20
Std. dev.	0.10	0.12	0.10	0.19	0.11	0.26	0.05

tile to have a large weight for a colorant including black. Since there is not much information about these colorants, their fitted reflectances are less accurate than the reflectances of colorants that do not include black. If we consider only the 7 CMY colorants, the average ΔE_{94} error between the measured colorant reflectances and those recovered using the constrained calibration algorithm drops to 4.34. By comparison, the average ΔE_{94} error is 4.16 for CMY prints (Table 5.3). It is slightly better than the 7 CMY colorant average in CMYK prints because the GCR algorithm reduces the surface coverage of some colorants, especially the CMY colorant.

The detailed analysis of the 7 CMY colorants for CMYK prints is given in Table 5.4. It confirms that the recovered reflectances that best match the measured reflectances are the optimized reflectances. However, this holds only when the colorant weights ω_i are large enough. When the weight is low, as in Figures 5.4 and 5.5, the optimized colorant reflectances may show a larger ΔE_{94} error than the constrained colorant reflectances. In such cases, the dark gray regions may not always entirely cover the measured spectra and prevent the optimization algorithm from achieving better matches. However, the dark gray regions prevent the recovered optimized colorant reflectances from deviating too much from the constrained colorant reflectances due to lack of information. For example, in Figure 5.4, the dark gray region prevents a good match in the near-infrared region above 750 nm, but also prevents a larger deviation between 550 nm and 600 nm.

According to Table 5.3, the average ΔE_{94} error between measured and optimized colorant reflectances is 5.61, or respectively 4.02 if the colorants containing black are excluded. How-

ever, as shown in Figure 5.6, colorants whose weights ω_i are above 0.4 induce a large accuracy improvement. When the colorant weight is lower, the accuracy remains stable, confirming that the constraints imposed on the optimized colorant reflectances successfully prevent the optimized calibration procedure from setting the colorant reflectances to spectra unrelated to the measured spectra because of lack of information. Finally, Figure 5.7 shows that the higher the colorant weight ω_i, the lower the ΔE_{94} error between the measured and the recovered optimized colorant reflectances.

The last column of Table 5.3 shows the average ΔE_{94} error between the fitted and measured colorants for a calibration further referred to as the optimized distributed 117 patch calibration, i.e. the optimized calibration according to Section 5.4 performed using the 117 UCR set instead of an image calibration set. Such a calibration set is optimal since the tiles are evenly distributed across the entire printer gamut and do not contain any colorant. In this case, the average ΔE_{94} error is further reduced to 3.83 for CMYK prints, respectively to 2.19 when the colorants containing black are excluded, and 1.34 for CMY prints. These averages show the importance of measuring image tiles well distributed within the printer gamut.

5.8 Comparison of different model calibrations

Five different calibrations are compared. The first is the *unconstrained calibration* described in Section 5.2. The second is the *constrained calibration* described in Section 5.3. The third is the surface coverage based *optimized image calibration* described in Section 5.4. These first three calibrations are performed for each of the 21 image calibration sets. There are therefore 21 unconstrained calibrations, 21 constrained calibrations, and 21 optimized image calibrations. The fourth calibration is the *optimized distributed 117 patch calibration* introduced in Section 5.7. It is the optimized calibration described in Section 5.4, performed on the 117 CMY or respectively for CMYK on the 117 UCR set instead of an image calibration set. The fifth calibration, used as reference, is the classical calibration of the IS-YNSN model relying on the measurements of the colorant reflectances and of predetermined color-constant halftone patches.

Two test sets are used to compare the accuracy of the CMY calibrations: the 64 CMY halftone set and the test set merging all the 21 CMY image test sets. Since each image test set is composed of 30 image tiles, this merged test set is composed of 630 image tiles. The following CMYK test sets are used to compare the accuracy of the CMYK calibrations: the 64 UCR set and the test set merging the 21 CMYK image test sets.

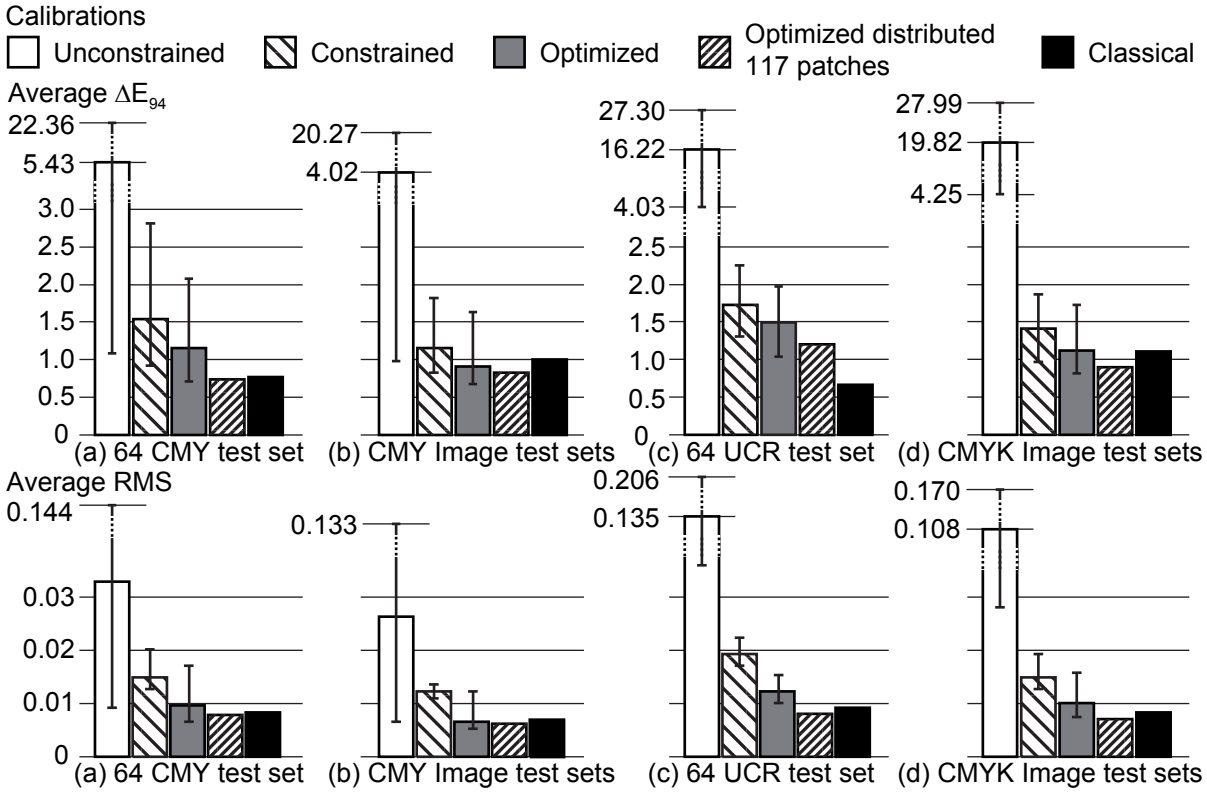


Figure 5.8: Comparison of the accuracy of 5 different types of calibrations. For the first three types of calibrations, the bar corresponds either to the average ΔE_{94} (above) or to the average root mean square error (RMSE, below) and the error bar to the best and worst among the 21 calibrations. The test sets predicted are (a) the 64 CMY test set, (b) the 21 merged CMY image test sets, (c) the 64 UCR test set, and (d) the 21 merged CMYK image test sets.

In Figure 5.8, each calibration is used to predict each test set. We use the average ΔE_{94} as a measure of accuracy. A lower average ΔE_{94} prediction error implies a more accurate calibration. Since there are 21 different calibrations for the first three types of calibrations, the first three bars represent the mean average ΔE_{94} and the vertical lines indicate the minimum and maximum average ΔE_{94}. We first observe that the unconstrained calibration is not able to accurately calibrate the IS-YNSN model. The average ΔE_{94} prediction error is 4 to 5 times larger in the CMY case (Figure 5.8a,b) and up to 20 times larger in the CMYK case (Figure 5.8c,d) compared with the classical calibration based on color-constant patches. When constraining the calibration, the average ΔE_{94} prediction error is on average reduced below 2. Optimizing the colorant reflectances further reduces the average ΔE_{94} to the point where even the worst optimized image calibration has an average ΔE_{94} of 2. The surface coverage based optimized image calibration is therefore the best method to calibrate the IS-YNSN model using tiles extracted from color images.

By comparing the performance of the optimized image calibrations to the performance of the optimized distributed 117 patch calibration, we see that the accuracy of the optimized distributed 117 calibration corresponds to the best optimized image calibration. It is therefore important for the image calibration tiles to be well distributed in the printer gamut to achieve the most accurate calibration of the model. Moreover, a classical calibration may perform better than an optimized image calibration when predicting color-constant patches (Figure 5.8a,c), but not when predicting image tiles (Figure 5.8b,d). Classical calibrations are the most accurate when predicting color-constant patches, but optimized image calibrations are on average more accurate when predicting tiles extracted from color images.

The average ΔE_{94} error achieved by the optimized image calibrations may seem surprisingly low when compared with the larger ΔE_{94} difference between the measured and fitted colorant reflectances shown in Tables 5.3 and 5.4 of Section 5.7. This difference in prediction accuracy is explained as follows. The IS-YNSN model performs predictions by interpolating between the colorants. The colorants are therefore the extreme points of the interpolation domain. In a classical calibration, these extreme points are measured, but are not the best references because they do not correspond to the normal halftoning operation of the printing device. Conversely, by recovering the reflectances of the colorants from tiles extracted from color images, we rely on information that better represents these normal operational conditions. The resulting IS-YNSN model may not be the most accurate when predicting reflectances close to the colorants, but is highly accurate when predicting reflectances of halftones and image tiles.

5.9 Calibration of the Yule-Nielsen n-value

In the preceding sections, the Yule-Nielsen n-value is set to a fixed scalar value. Let us investigate how to fit this n-value for both the constrained and the optimized image calibrations. In a classical calibration, the n-value is fitted as follows. A list of candidate n-values is established, for example between 1 and 10 in 0.5 steps. For each n, the IS-YNSN is calibrated and used to predict the calibration set. The average root mean square error (RMSE) between the predicted and measured calibration patch reflectances is recorded for each n. The chosen n-value is the one yielding the lowest average RMSE. We apply the same selection procedure for both the constrained and the optimized image calibrations.

The results are shown in Figure 5.9. The average RMSE of the classical calibration [line (1)] decreases smoothly when n increases. n is therefore set to 10, but the difference is not significant between $n = 7$ and higher values of n. When performing constrained image calibrations,

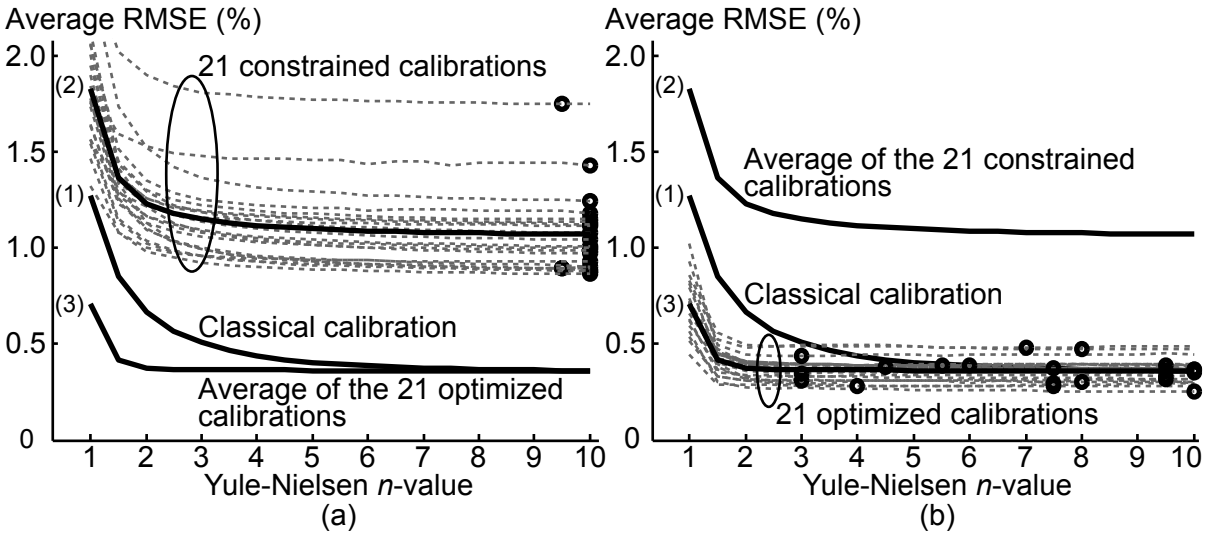


Figure 5.9: Average RMSE between the measured and predicted calibration patches, or image calibration tiles, according to different Yule-Nielsen n-values. Solid line (1) corresponds to the classical calibration, the dashed lines in (a) to each of the 21 constrained image calibrations, solid line (2) to the average of the dashed lines in (a), the dashed lines in (b) to each of the 21 optimized image calibrations, and solid line (3) to the average of the dashed lines in (b). The circles correspond to the n-value along each dotted line yielding the minimum average RMSE.

the n-value is usually set to 10, except for the *Child*, *Peppers*, and *Sailboat III* image calibrations where n is set to 9.5 (circles on the left). Generally, the average RMSE of such calibrations also decreases when n increases (dashed lines on the left), but some small local increases lead to a few n-values being set to values lower than 10. However, since the average RMSE of the 21 constrained image calibrations behaves on average in the same manner as the average RMSE of the classical calibration [line (2)], the selected n-values correspond to the n-value chosen for the classical calibration.

When optimized image calibrations are used to select n, the general shape of the average RMSE is a sharp drop between 1 and 2 followed by a straight line [line (3)], providing no useful information about the best n-value. As seen in Figure 5.9b, many different optimal n-values are selected. However, since the lines are nearly flat, any n-value of 2 or above could be chosen. The flexibility introduced by the optimal calibration seems to make the n-value a redundant parameter. Nevertheless, we rely on the constrained image calibrations to choose the best n-value since this n-value corresponds to the choice for the classical calibration. Once the best n-value is chosen, we can optimize the colorants of the corresponding constrained image calibration to obtain an IS-YNSN model fully calibrated from color printed images without relying on prior knowledge about colorant reflectances or ink spreading parameters.

5.10 Importance of accurately calibrating the colorant reflectances

In the previous sections, we have shown how to calibrate the IS-YNSN model using a set of 50 image calibration tiles maximizing the colorant weights ω_i defined in Eq. (5.15). In this section, we show the importance of maximizing the colorant weights ω_i. We also provide more insights about the required number of image calibration tiles and their most appropriate location in the CMYK color space. We finally show that using nominal ink surface coverages instead of effective coverages to compute the ink spreading weights $w_{i/jk}$ leads to more stable calibrations.

5.10.1 Iterative image calibration tile sets

In order to study the influence of the colorant weights ω_i, we perform the following experiment. We first merge the 1050 available calibration tiles selected in the 21 different test images (see Appendices D.2 and D.3). From this pool of image calibration tiles, we create calibration sets by randomly selecting tiles one after another and adding the chosen tile to the calibration set only if it increases the sum s of the colorant weights ω_i:

$$s = \sum_i \omega_i = \sum_i \max_p [a_{i,p}] \tag{5.27}$$

where the colorant weights ω_i are computed with the current calibration set according to Eq. (5.15). Therefore, a tile is selected only if at least one of its colorant coverages is larger than the corresponding colorant coverages of all the previously selected tiles. Since a colorant weight ω_i is computed as the maximum across all the p selected tiles, the addition of a new tile replaces one or more tiles for the computation of the increased colorant weights. Once the calibration set is composed of a sufficient number of tiles, some of these tiles do not contribute to the computation of the sum s of the colorant weights anymore. However, they remain in the calibration set. Such image calibration sets are further referred to as *iterative calibration sets*.

We create 100 different CMY iterative calibration sets containing each 25 image calibration tiles and 100 different CMYK iterative calibration sets containing each 40 image calibration tiles. Figure 5.10 shows how the sum s of the colorant weights ω_i increases with each additional calibration tile. The sum of the weights may slightly decrease when a tile is added because the colorant weights are computed based on the effective colorant coverages [Eq. (5.15)], relying

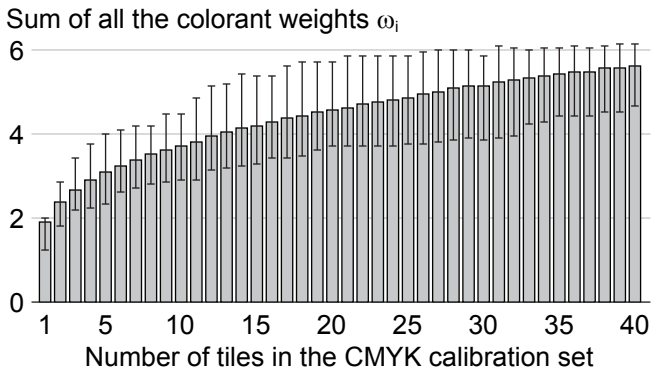


Figure 5.10: Sums of the CMYK colorant weights ω_i according to the number of image tiles in the iterative calibration sets. The bars represent the average sums over 100 different iterative calibration sets. The error bars represent the highest and lowest sums among the 100 iterative calibration sets.

themselves on ink spreading curves recalibrated each time a new calibration tile is added. As a reference, the CMYK image calibration sets created in Section 5.5 and shown in Appendix D.3 are composed each of 50 image tiles and have colorant weight sums between 3.14 (*Sea* image) and 5.37 (*Hats* image) with an average of 4.0, which corresponds to CMYK iterative calibration sets composed of 12 image tiles.

5.10.2 Prediction accuracy of calibrations based on iterative calibration sets

For each iterative calibration set, we perform as many optimized calibrations of IS-YNSN model as tiles in the calibration set, optimized calibrations performed according to the algorithm described in Section 5.4. The first calibration uses only the first tile of the set, the second calibration the first 2 tiles, the third calibration the first 3 tiles, etc. The last calibration therefore uses all the tiles of the calibration set.

The accuracy of the calibrations based on the iterative calibration sets is shown in Figure 5.11 when predicting the 64 CMY halftone set and the 21 merged CMY image test sets, and in Figure 5.12 when predicting the 64 UCR halftone set and the 21 merged CMYK image test sets. The image test sets are described in Section 5.5. The 64 CMY halftone set and 64 UCR halftone set are described in Section 5.6. We compare these predictions with the predictions performed in Section 5.8, i.e. with the average predictions of the 21 optimized calibrations (*) and with the predictions of the classical calibration (‡) . For both the CMY and CMYK test sets, the prediction error decreases with each additional calibration tile. This decrease correlates perfectly with the increase of the sum of the colorant weights shown in Figure 5.10. Sets composed of 10 CMY or 16 CMYK image calibration tiles are sufficient to achieve an

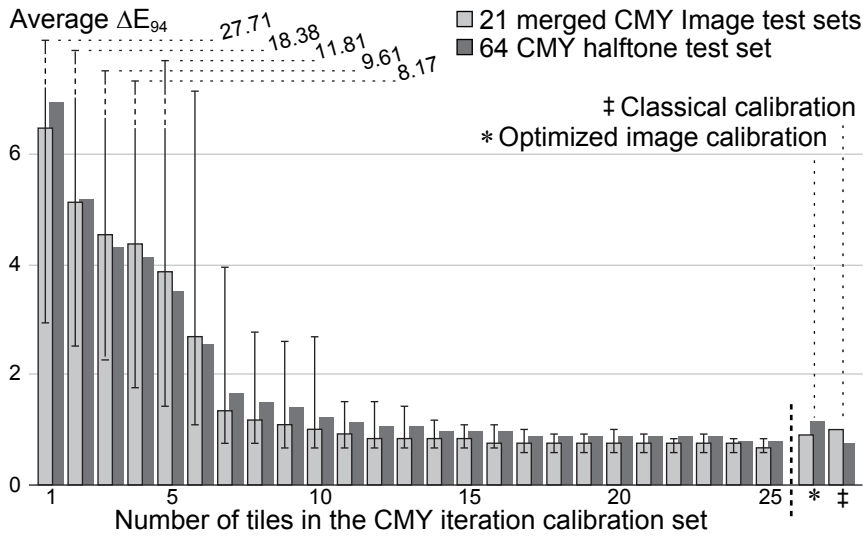


Figure 5.11: Average ΔE_{94} between the measured and predicted image tiles of the 21 merged CMY image test sets (light gray) and the CMY 64 halftone test set (dark gray) according to the number of tiles in the iterative calibration sets. The iterative image calibrations are compared with the optimized image calibration performed in Section 5.8 (*) and the classical calibration (‡). The bars indicate the mean average ΔE_{94} of all the 100 iterative calibrations. The error bars indicate the average ΔE_{94} of the best and worst calibrations among the 100 iterative calibrations performed. Since the error bars are similar for both test sets, they are shown only once.

accuracy similar to the calibration sets used in Section 5.8 (*). The prediction error keeps decreasing slowly with each additional tile, but not significantly once the image calibration set contains 20 CMY or 30 CMYK tiles. Calibrations using the iterative calibration sets therefore show significant improvement over the optimized calibrations performed in Section 5.8 (*).

In respect to the classical calibration (‡), calibrations using the iterative calibration sets are more accurate when predicting the image test tiles and almost as accurate when predicting the CMY color-constant halftone patches. Iterative optimized image calibrations are however not as accurate as a classical calibration when predicting the CMYK color-constant halftone patches. These results show the importance of maximizing the colorant weights ω_i. Since these weights indicate the amount of information available to calibrate the colorant reflectances, large colorant weights ω_i improve the accuracy of the recovered colorant reflectances, and therefore the accuracy of the resulting IS-YNSN model.

In order to ensure that the least-squares optimization of Section 5.4.2 is not rank deficient, the number of image calibration tiles should not be less than the number of colorant reflectances to calibrate. If the white colorant is the only measured colorant, there should be no less than 7 CMY or respectively 15 CMYK image calibration tiles. We can see in Figures 5.11 and 5.12 that calibrations using an insufficient number of calibration tiles are not stable and induce large

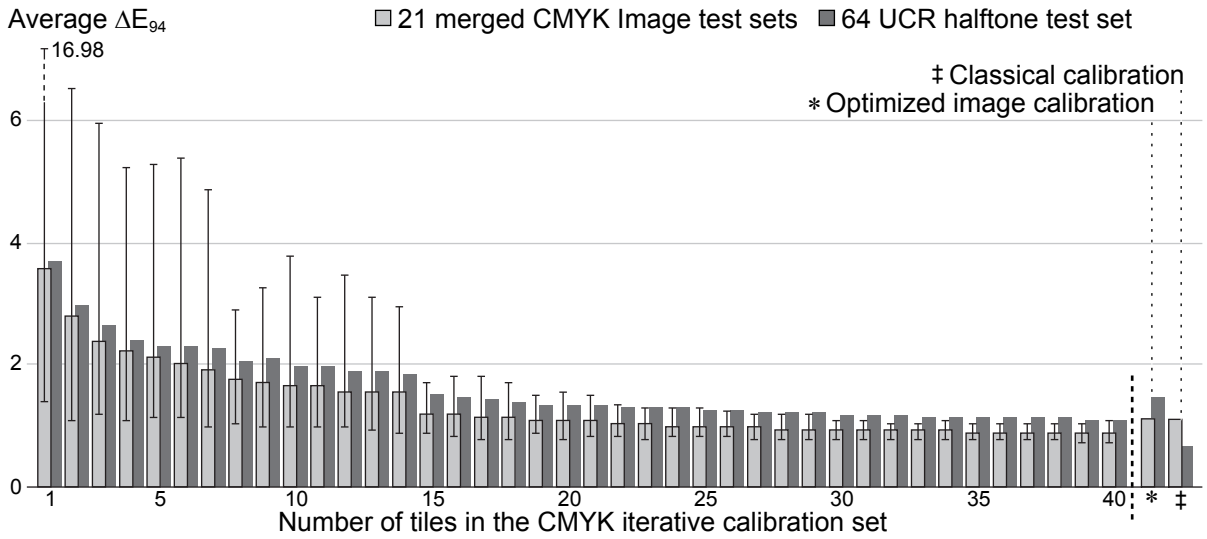


Figure 5.12: Average ΔE_{94} between the measured and predicted image tiles of the 21 merged CMYK image test sets (light gray) and the CMYK UCR 64 halftone test set (dark gray) according to the number of tiles in the iterative calibration sets. The iterative image calibrations are compared with the optimized image calibration performed in Section 5.8 (*) and the classical calibration (‡). The bars indicate the mean average ΔE_{94} of all the 100 iterative calibrations. The error bars indicate the average ΔE_{94} of the best and worst calibrations among the 100 iterative calibrations performed. Since the error bars are similar for both test sets, they are shown only once.

variations in accuracy. In such cases, we advise using the constrained calibration (Section 5.3) instead of the optimized calibration (Section 5.4).

5.10.3 Evaluating the importance of maximizing the colorant weights ω_i

In Section 5.10.1, iterative calibration sets are created by adding randomly chosen image tiles that keep increasing the sum s of the colorant weights ω_i. We have seen in Section 5.10.2 that these iterative calibration sets yield more accurate calibrations. We now show that a calibration set entirely composed of tiles chosen to maximize the sum s of the colorant weights does not yield the best calibration.

Instead of randomly choosing the image tiles, we create an iterative calibration set further referred to as the *maximum iterative set*. Like in Section 5.10.1, this set is created by selecting tiles one after another among the 1050 available calibration tiles. At each iteration, the selected tile is the tile that maximizes the sum s. When no more tile is able to increase s, which happens in our case after 18 image tiles are selected, s is reset to 0 and the process continues until the 40 image calibration tiles have been selected. The sum of the colorant weights ω_i achieved by the maximum iterative set is 6.51.

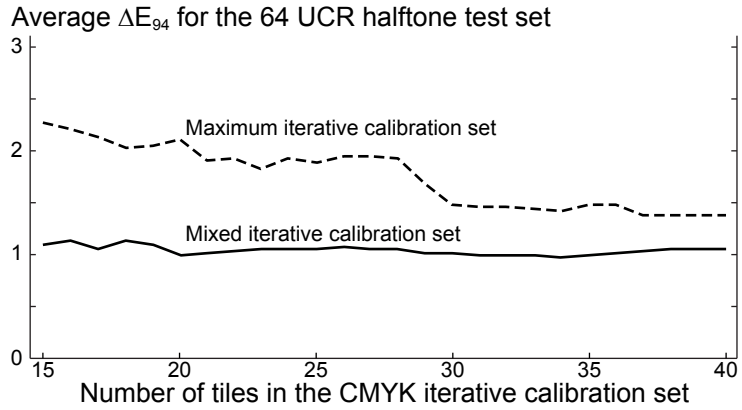


Figure 5.13: Accuracy of the optimized calibrations based on the maximum and mixed iterative calibration sets when predicting the 64 UCR halftone test set (Section 5.6) according to the number of calibration tiles. The dashed line is the accuracy of the maximum iterative set and the solid line the accuracy of the mixed iterative set.

The maximum iterative set is compared to the *mixed iterative set*. This set is created using two different strategies. The odd tiles are selected to maximize s like for the maximum iterative set. The even tiles are randomly chosen among the 50 image calibration tiles closest to the center of the CMYK color space. The sum of the colorant weights ω_i achieved by the mixed iterative set is the same as the sum achieved by the maximum iterative set, i.e. 6.51.

As shown in Figure 5.13, the optimized calibration based on the mixed iterative set is more accurate at predicting the 64 UCR halftone test set than the optimized calibration based on the maximum iterative set. It demonstrates that maximizing the colorant weights, which confines the tiles at the periphery of the CMYK color space, is not the only criterion that should be taken into considerations when selecting image calibration tiles. The selected tiles should also be well-distributed in the CMYK color space in order to provide information about the entire color space. In this way, the IS-YNSN model performs accurate predictions for all the color-constant patches and uniform image tiles and is not specialized in smaller areas of the CMYK color space.

5.10.4 Impact of using nominal ink coverages to compute the ink spreading weights $w_{i/jk}$

In Section 5.3.1, we mentioned that the ink spreading weights should theoretically be computed using effective ink surface coverages. However, since these effective coverages are not available when performing a constrained calibration, we have used nominal coverages instead.

In order to determine the impact of this modification, we perform the following experi-

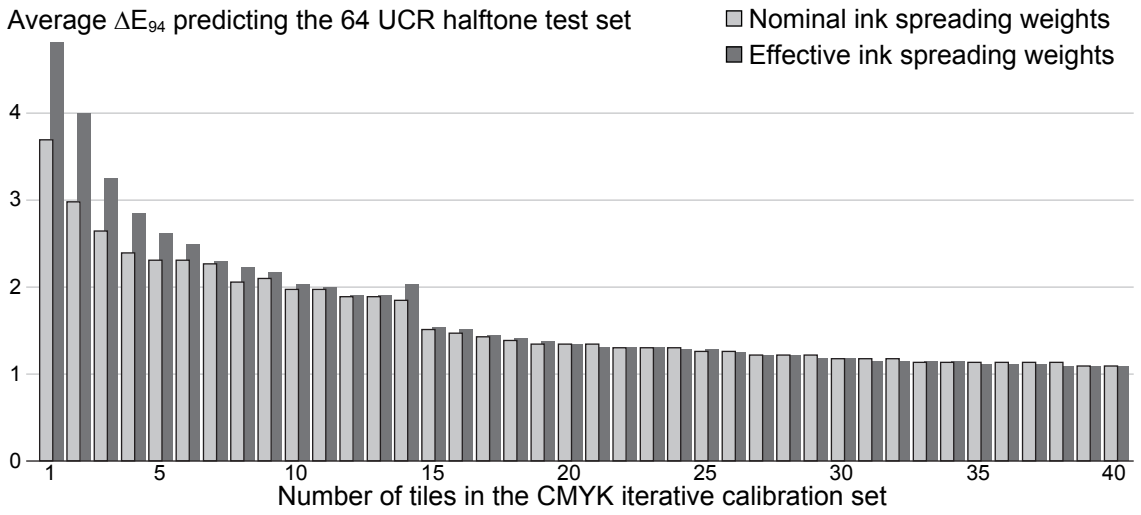


Figure 5.14: Average prediction accuracy of iterative calibrations according to two different methods to compute the ink spreading weights $w_{i/jk}$. The predicted test set is the CMYK UCR 64 halftone test set. The light gray bars show the accuracy when only nominal ink surface coverages are used to compute the ink spreading weights and the dark gray bars show the accuracy when effective coverages are used.

ment, based on the iterative image calibration sets created in Section 5.10.1. Instead of setting the starting point of the gradient-descent algorithm to the default values, i.e. setting the mid-points of the ink spreading curves to 0.5, the principal components of the ink transmittances to 0, and the ink thickness variation factors to 1, we use the values obtained by the previous calibration, i.e. the calibration performed with one calibration tile less. By using the values of the previous calibration, the ink spreading curves are already partly calibrated and the effective ink coverages can be computed to determine the ink spreading weights $w_{i/jk}$.

The results are shown in Figure 5.14. When 15 CMYK image tiles or more are used, i.e. when the least-squares problem has full rank, there is no difference between the two calibration methods. When the least-squares problem is rank deficient, using nominal coverages to compute the ink spreading weights $w_{i/jk}$ is preferable since it yields more stable and accurate calibrations. Using effective coverages to compute the ink spreading weights is more complicated and does not improve the calibration accuracy. It is therefore not necessary.

5.11 Summary

Image tiles have been used in Chapter 4 to calibrate the ink spreading curves, but not to recover the reflectances of the Neugebauer primaries. In this chapter, we propose a method to recover the colorant reflectances (Neugebauer primaries), the ink spreading curves, and the

Yule-Nielsen n-value using only tiles extracted from printed color images. There is no prior knowledge about the reproduction device. The first calibration step introduces constraints for colorant reflectances derived from the principal component analysis of a large number of ink transmittances and surface coverage related constraints for the ink spreading curves. Such a constrained calibration is performed for a list of candidate Yule-Nielsen n-values. The n-value yielding the most accurate calibration is selected. In a second step, in order to improve the accuracy of the calibration, the colorant reflectances are optimized by a weighted least-squares approach where the weights are set according to the colorant surface coverages.

The proposed optimized image calibration is almost as accurate as a classical calibration when predicting color-constant halftone patches and more accurate when predicting image tiles extracted from color images. Moreover, this high accuracy is achieved with sets composed of less than 25 well-chosen image calibration tiles. By entirely eliminating the need to print predetermined patches, these calibration procedures broaden the field of use of spectral reflection prediction models. Such models can now be used when it is impossible or too expensive to print extra patches or when the reproduction device is not available. It also enables the real-time monitoring of printing devices without modifying their print jobs.

Chapter 6

Conclusion

Spectral reflection prediction models are important tools for the characterization of color reproduction devices. Among such models, due to its simplicity and accuracy, the Yule-Nielsen model is the most widely used. However, many calibration patches are required for its calibration, patches that are usually not available within the printed document. In order to simplify the calibration of the prediction model and to prevent printing special halftone and solid, color-constant calibration patches, we use image tiles extracted from printed color images. The model used in this contribution, presented in Chapter 2, is the ink spreading enhanced Yule-Nielsen modified Spectral Neugebauer model (IS-YNSN model), a variant of the Yule-Nielsen model including an ink spreading model accounting for physical dot gain.

The IS-YNSN model is reviewed in the context of CMYK prints in Chapter 3. We show that the fourth black ink introduces two sources of ambiguity. The first follows from the two possibilities to create the color black, either using the black ink or superposing cyan, magenta, and yellow. When the IS-YNSN model is used to deduce the surface coverages of printed color halftone patches from their measured reflectance spectra, it cannot accurately distinguish between the two different blacks, introducing inaccuracies in the deduced coverages. The second ambiguity follows from the fact that printing solid black along with other inks yields black. There is no possibility to detect the presence of the other inks from measured spectra, generating instabilities when calibrating certain ink spreading curves. We therefore propose solutions to remove these sources of errors. This more robust version of the IS-YNSN model is used as basis for our revised calibration process.

The calibration of the IS-YNSN model consists in recovering the Neugebauer primaries and acquiring the ink spreading curves. In Chapter 4, we propose a method to acquire the

ink spreading curves from printed color images. A simple optimization like a gradient-descent algorithm is not sufficient to acquire the ink spreading curves because the information contained in the printed color images may be incomplete. We propose to rely on constraints based on a metric evaluating the relevance of each ink spreading curve to the set of image tiles extracted from printed color images. We optimize the algorithm automatically selecting the image tiles to be extracted from the printed color images and show that 5 to 10 well-chosen image tiles are sufficient to accurately acquire all the ink spreading curves.

The calibration of the entire IS-YNSN model from printed color images is demonstrated in Chapter 5. Again, a simple gradient-descent algorithm is not sufficient. Thanks to a set of constraints based on Principal Component Analysis (PCA) and the relationships between composed Neugebauer primaries and the ink transmittances, good approximations of the Neugebauer primaries are achieved. These approximations are then optimized, yielding an accurately calibrated IS-YNSN model comparable to one obtained by classical calibrations. A detailed analysis of these calibrations shows that 25 well-chosen CMYK image calibration tiles are sufficient to accurately recover both the Neugebauer primaries and the ink spreading curves.

This dissertation opens a whole new field in color reproduction. From online monitoring to self-calibrating devices via the extraction of device properties, many fields can benefit from the flexibility of the proposed calibration. Moreover, since no prior knowledge is required, the flexible calibration can be readily applied to new devices provided that the CMYK inks are not too different from classical process inks.

Future work

For three reasons, the proposed calibration has only been tested for CMY or CMYK inks. First, most color reproduction devices propose at least these four inks and can therefore benefit from this work. Second, the gamut generated by CMYK inks is one of the largest 4 ink gamuts. This increases the difficulty of calibrating the IS-YNSN with image tiles for such a large gamut. Third, color reproduction with more than 4 inks can often be represented by several 4 ink reproduction devices. The proposed calibration may therefore apply to a wide range of color reproduction devices. In addition, the model can be extended beyond 4 inks. In that case, the flexible calibration may further reduce the calibration effort.

Regarding image tiles, non-uniform tiles cannot be used because they are not accurately predicted by the IS-YNSN model. A deeper analysis of the prediction errors for non-uniform tiles may lead to a better understanding of the information contained in these image tiles, in-

formation that could then be used in the calibration process. Having a larger choice of image calibration tiles could further improve the accuracy of the calibrated IS-YNSN model.

When colorant reflectances are relaxed from their initial constraints and further optimized, the Yule-Nielsen n-value seems to become a redundant parameter. Further research is needed to establish the relationship between colorant reflectance optimization and n-value. The calibration of the n-value could then be performed at the same time as the acquisition of the ink spreading curves.

Finally, all the measurements have been performed spectrally. Instead of relying on spectral measurements, we could use sensor responses to perform the calibration. Compared to spectral acquisition devices, sensors are inexpensive and fast. However, the limited amount of information provided by these sensor responses poses new challenges to the proposed flexible calibration, especially in the last step when the colorant reflectances are optimized.

Appendix A

Accuracy of the reconstructed solid ink spectral reflectances using PCA

In Section 5.3.2, principal component analysis (PCA) is used to constrain the transmittances of the inks. Given a set of ink transmittances sampled at k different wavelengths, a PCA yields a list of k principal components that allow the reconstruction of the ink transmittances such that the first component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible.

By limiting the number of principal components to the first m components, only m scalars are required to reconstruct a transmittance. Moreover, as the first m principal components account for most of the variability of the data, we ensure that the reconstructed transmittances keep the intrinsic properties of real transmittances, i.e. their smoothness and characteristic shape. Note that one PCA is performed separately for each of the 4 basic c, m, y, and k inks.

The data used for the PCAs is composed of the spectral reflectance measurements of the cyan, magenta, yellow, and black colorants as well as the substrate on which these colorants are printed. We measured the ink reflectances of 25 CMY and 37 CMYK reproduction devices, comprising a wide range of technologies such as ink jet, thermal transfer, and electrophotographic printers; offset presses; and proofing devices. Note that the Canon PixmaPro 9500 printer used in the present experiments was not used to build the principal vectors of the PCA. Once the reflectance measurements are available, the transmittances of the inks are computed using Eq. (5.7) and a PCA is performed on these transmittances. For different values of the number of components m, we reconstruct the transmittances and compute the reconstructed

colorant reflectances using Eq. (5.7). We then calculate the following error metrics between the measured and reconstructed colorant reflectances: root mean square error (RMS), ΔE_{94} average, 95 percentile (95%), and maximum ΔE_{94}. We also indicate the score of each principal component, i.e. the percentage of variability explained by the considered principal component, as well as the total percentage of variability explained by the m selected principal components.

Tables A.1 to A.4 show the accuracy of the reconstructed transmittances according to the number m of retained principal components for the cyan, magenta, yellow, and black inks, respectively. Except for yellow, the first two principal components already account for more than 90% of the variance (score). Using three principal components, we are able to accurately reconstruct all the colorants with a maximum ΔE_{94} of 2.97 and a 95 percentile ΔE_{94} of 0.89. This accuracy is sufficient and provides us with a method to express the ink reflectances with only 3 scalars per ink.

Table A.1: Accuracy of the PCA for the **cyan** transmittances according to the number of principal components used (#PC).

#PC	RMS	ΔE_{94}			Score	
		Avg	95%	Max	Current	Total
1	0.012	0.59	1.08	3.78	69.32	69.32
2	0.007	0.53	0.96	2.96	24.31	93.63
3	0.005	0.49	0.89	2.97	3.99	97.62
4	0.003	0.42	0.83	2.87	1.02	98.64
5	0.003	0.42	0.85	2.75	0.75	99.39
6	0.001	0.28	0.65	1.75	0.43	99.81
7	0.001	0.04	0.09	0.11	0.14	99.95
8	0.001	0.04	0.08	0.10	0.02	99.97
9	0	0.02	0.05	0.08	0.01	99.98
10	0	0.02	0.05	0.06	0.01	99.99

Table A.2: Accuracy of the PCA for the **magenta** transmittances according to the number of principal components used (#PC).

#PC	RMS	ΔE_{94}			Score	
		Avg	95%	Max	Current	Total
1	0.010	1.08	2.19	4.60	88.13	88.13
2	0.006	1.03	2.09	4.96	6.13	94.26
3	0.004	0.37	0.91	1.86	3.84	98.09
4	0.003	0.25	0.54	0.64	1.27	99.36
5	0.002	0.11	0.24	0.49	0.36	99.72
6	0.001	0.09	0.23	0.45	0.16	99.88
7	0.001	0.04	0.09	0.13	0.09	99.97
8	0	0.03	0.07	0.08	0.01	99.98
9	0	0.02	0.04	0.08	0.01	99.99
10	0	0.01	0.02	0.02	0	99.99

Table A.3: Accuracy of the PCA for the **yellow** transmittances according to the number of principal components used (#PC).

#PC	RMS	ΔE_{94}			Score	
		Avg	95%	Max	Current	Total
1	0.014	0.83	2.34	4.54	61.94	61.94
2	0.008	0.69	2.11	4.14	24.92	86.86
3	0.004	0.32	0.84	1.71	9.18	96.05
4	0.003	0.28	0.80	1.60	2.08	98.13
5	0.002	0.23	0.56	1.13	1.23	99.36
6	0.001	0.07	0.18	0.28	0.44	99.80
7	0.001	0.04	0.09	0.11	0.11	99.90
8	0.001	0.02	0.05	0.07	0.04	99.95
9	0	0.02	0.05	0.07	0.03	99.97
10	0	0.01	0.03	0.04	0.01	99.98

Table A.4: Accuracy of the PCA for the **black** transmittances according to the number of principal components used (#PC).

#PC	RMS	ΔE_{94}			Score	
		Avg	95%	Max	Current	Total
1	0.008	1.11	2.78	17.98	74.63	74.63
2	0.001	0.36	0.79	1.65	24.74	99.37
3	0.001	0.34	0.73	1.40	0.59	99.96
4	0	0.13	0.28	0.89	0.03	99.99
5	0	0.08	0.21	0.28	0.01	100.00
6	0	0.07	0.12	0.29	0	100.00
7	0	0.04	0.09	0.16	0	100.00
8	0	0.02	0.06	0.07	0	100.00
9	0	0.02	0.05	0.07	0	100.00
10	0	0.01	0.02	0.03	0	100.00

Appendix B

Test images and selected tiles for the original tile selection algorithm

B.1 Test images

Fruits image

Textile image

B.2 Image tile sets of the *fruits* image

Image calibration tiles of the *fruits* image

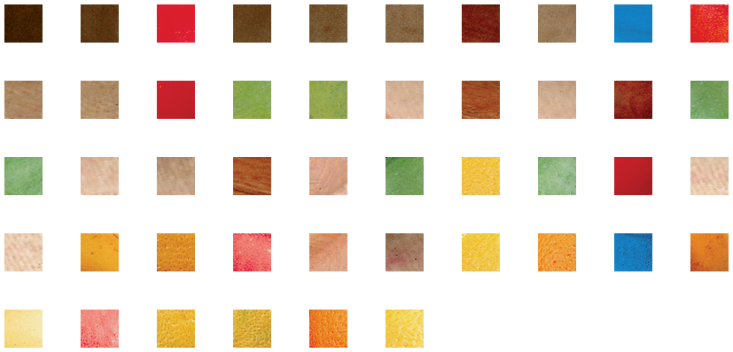


Image test tiles of the *fruits* image

B.3 Image tile sets of the *textile* image

Image calibration tiles of the *textile* image

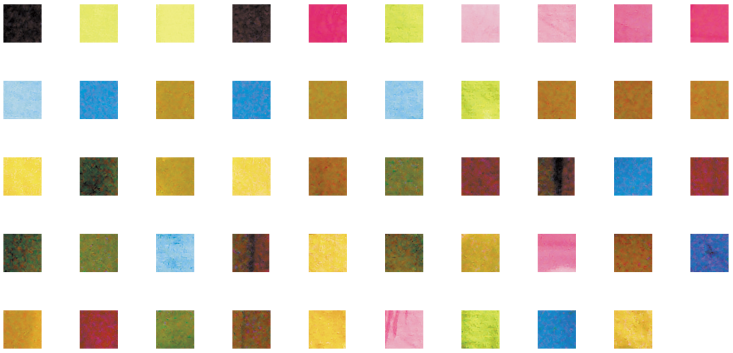
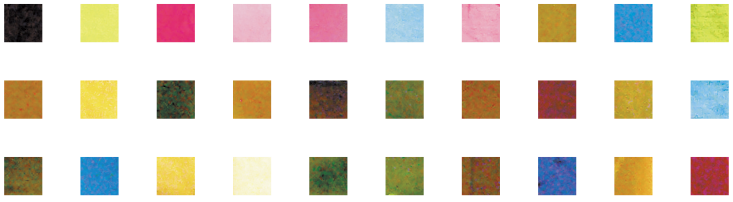


Image test tiles of the *textile* image



Appendix C

Test images and selected tiles for the optimized tile selection algorithm

C.1 Test images (scale 1:5)

N°1: Flower

N°3: Boat race

N°5: Parrots

N°2: House

N°4: Sea

N°6: Sailboat

C.2 CMY standard and optimized image calibration tile sets

N°1 Flower standard set

N°1 Flower optimized set

N°2 House standard set

N°2 House optimized set

N°3 Boat race standard set

N°3 Boat race optimized set

N°4 Sea standard set

N°4 Sea optimized set

N°5 Parrots standard set

N°5 Parrots optimized set

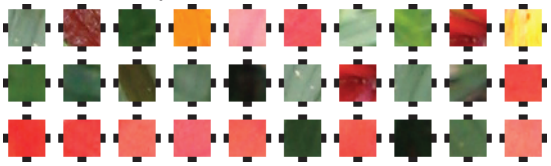
N°6 Sailboat standard set

N°6 Sailboat optimized set

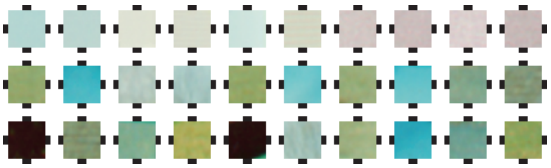
C.3 CMYK standard and optimized image calibration tile sets

N°1 Flower standard set

N°1 Flower optimized set



N°2 House standard set

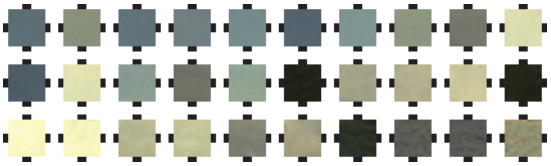


N°2 House optimized set

N°3 Boat race standard set

N°3 Boat race optimized set

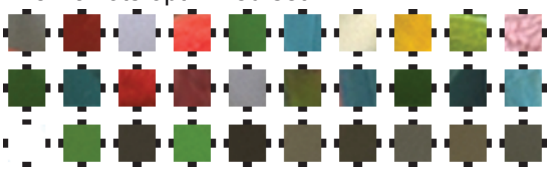
N°4 Sea standard set



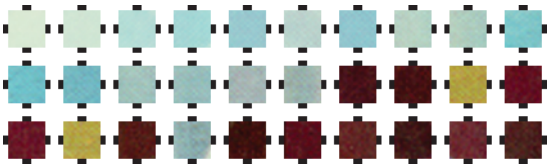
N°4 Sea optimized set

N°5 Parrots standard set

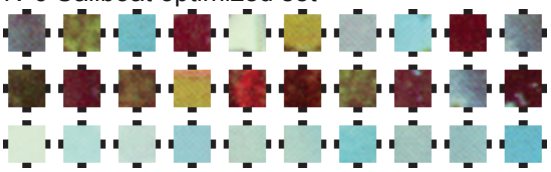
N°5 Parrots optimized set



N°6 Sailboat standard set

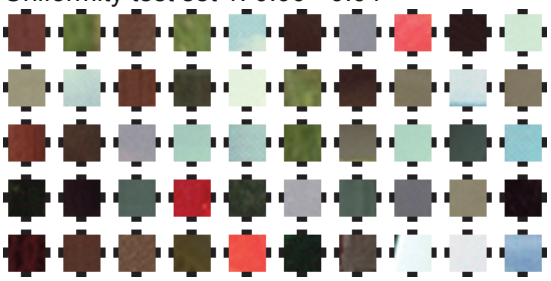


N°6 Sailboat optimized set

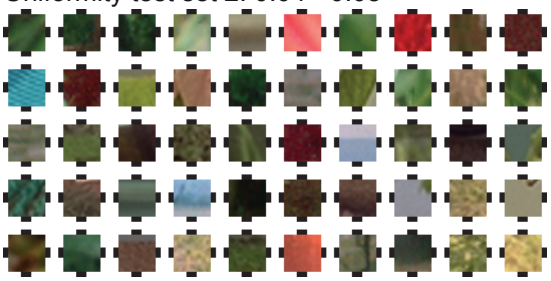


C.4 CMYK uniformity test sets relying on Eq. (4.19)

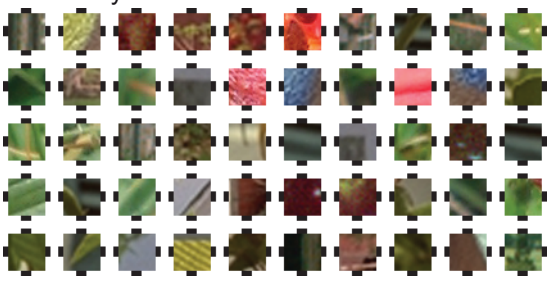
Uniformity test set 1: 0.00 - 0.04



Uniformity test set 2: 0.04 - 0.08



Uniformity test set 3: 0.08 - 0.12

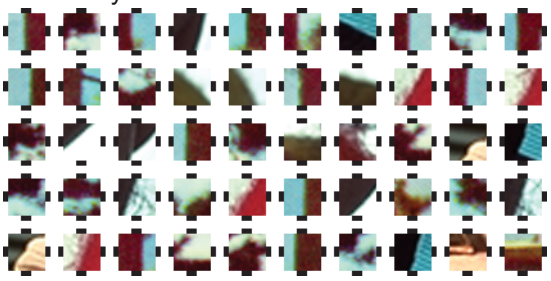


Uniformity test set 4: 0.12 - 0.16

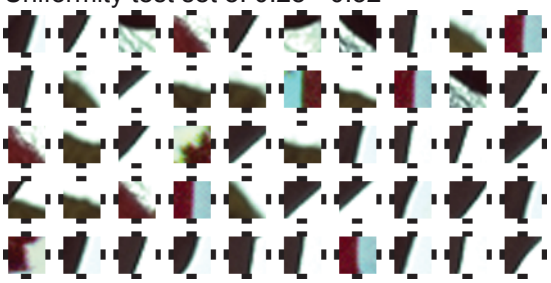
Uniformity test set 5: 0.16 - 0.20

Uniformity test set 6: 0.20 - 0.24

Uniformity test set 7: 0.24 - 0.28



Uniformity test set 8: 0.28 - 0.32



Appendix D

Test images and selected tiles for the recovery of the Neugebauer primaries

D.1 Test images (scale 1:5)

N°1: Handa Island

N°2: Beach

N°3: Flower I

N°4: Fruits

N°5: Goldhill


N°7: Hats


N°9: Sailboat I


N°10: Flower II


N°12: Sailboat II


N°13: Child


N°14: Sea


N°16: Lighthouse II


N°17: House II


N°18: Parrots


N°8: Woman


N°11: Boat race


N°15: Lighthouse I

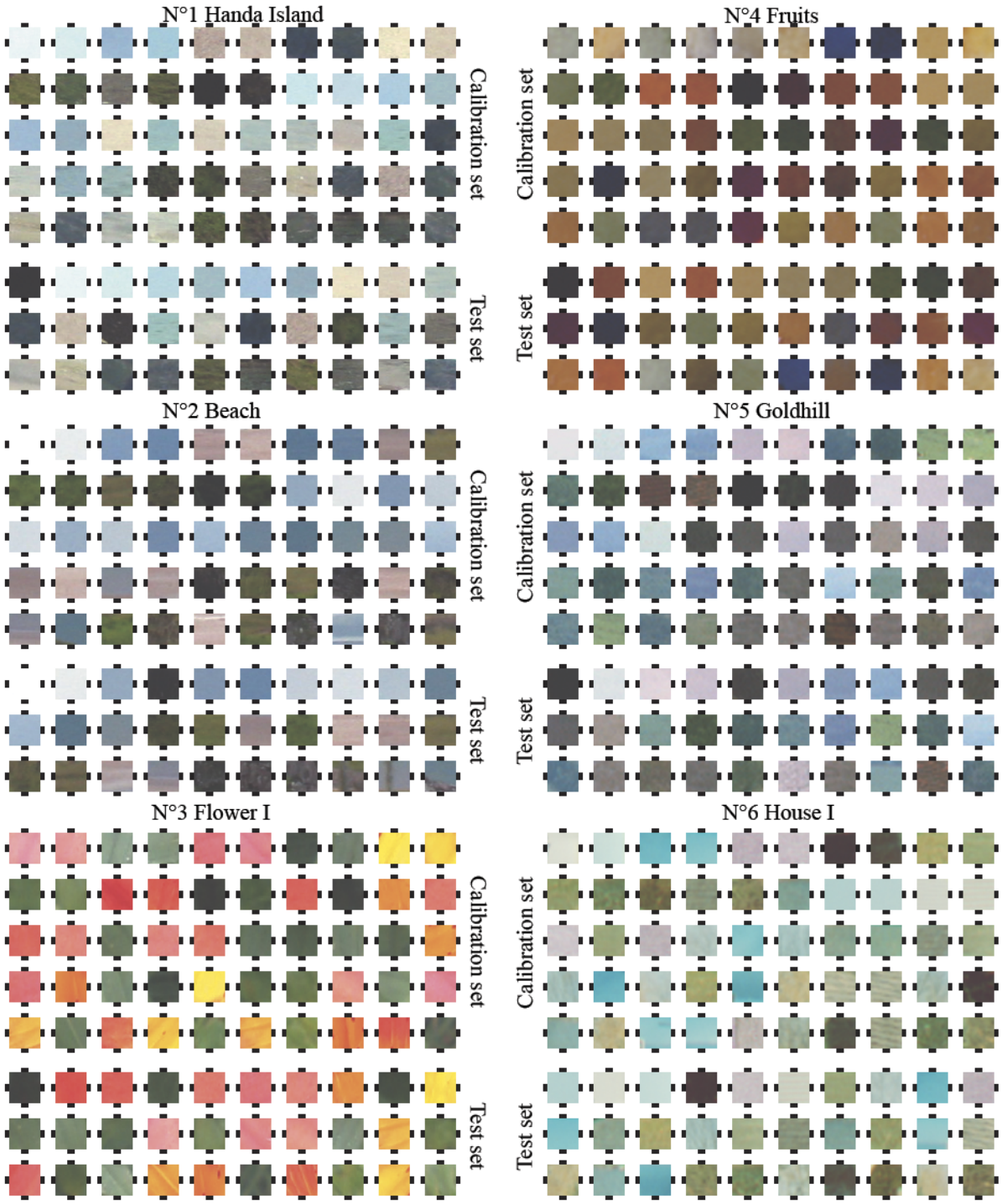

N°6: House I

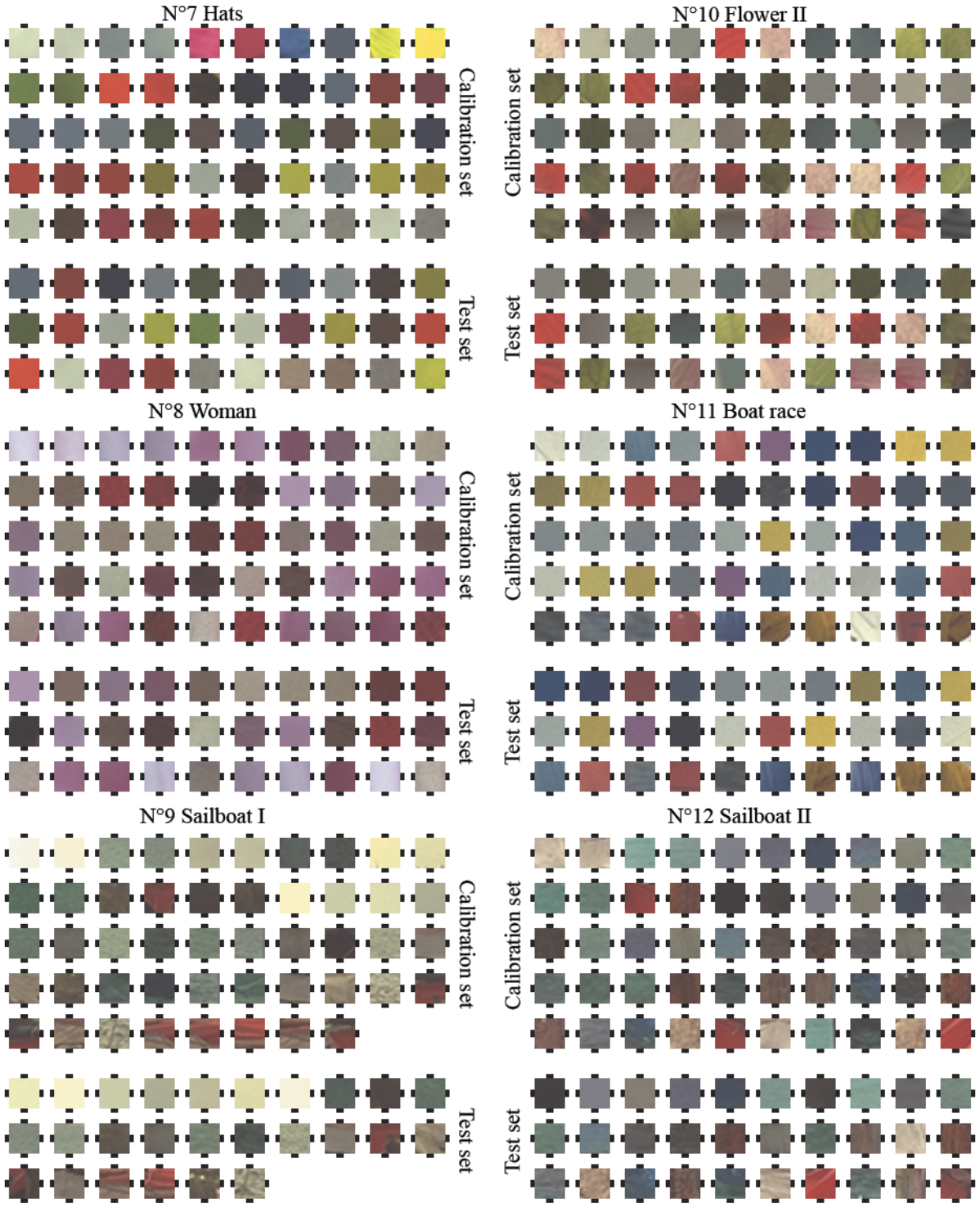

N°19 Peppers

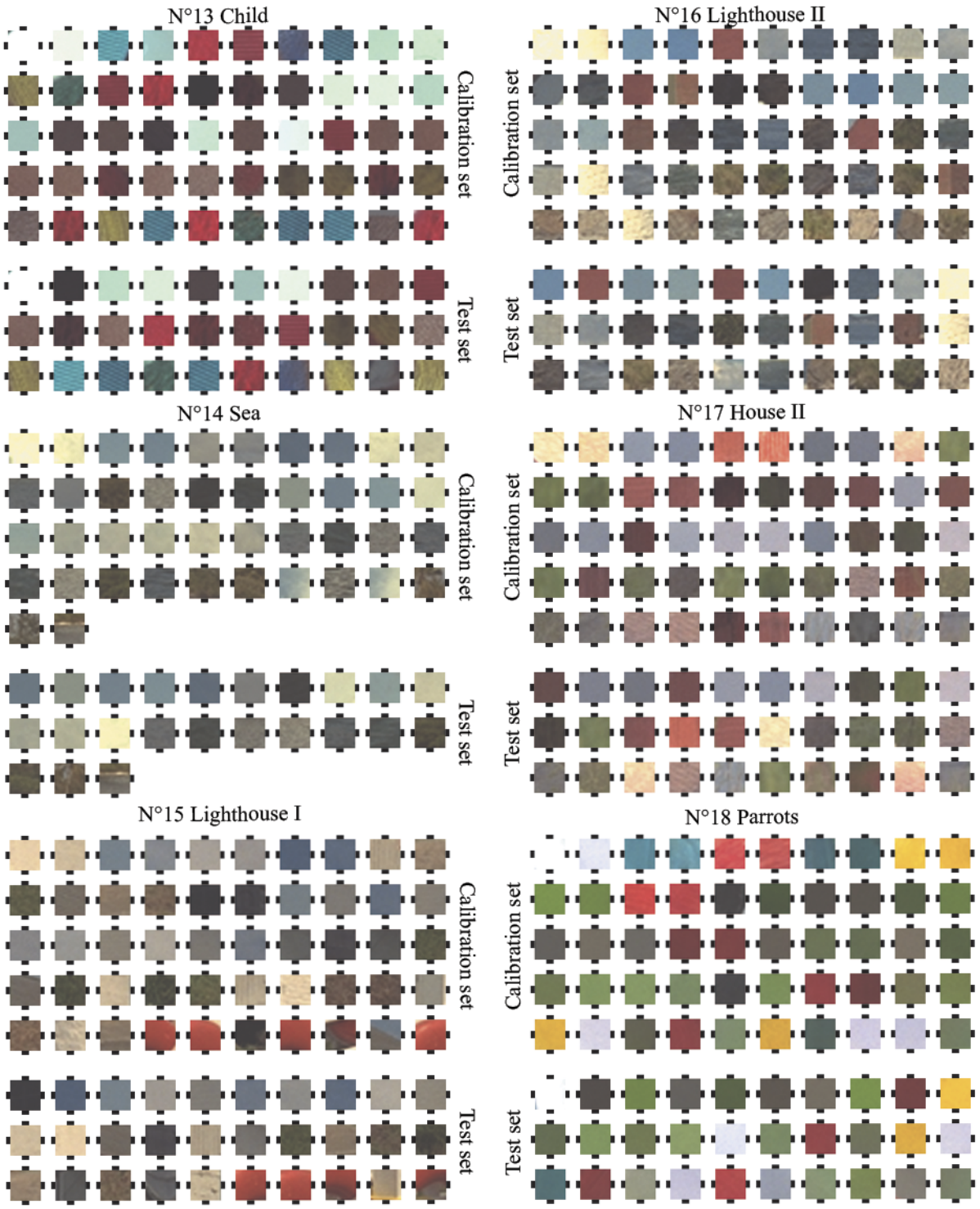

N°20 Sailboat III

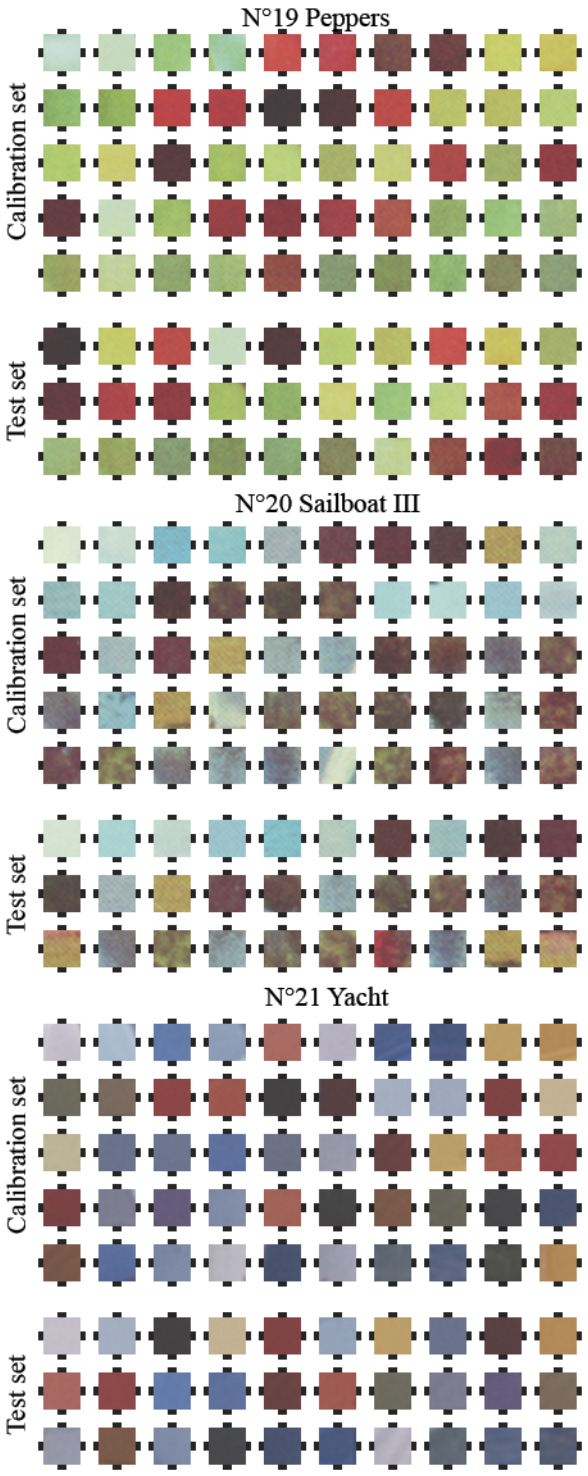

N°21: Yacht

D.2 CMY calibration and test image tiles

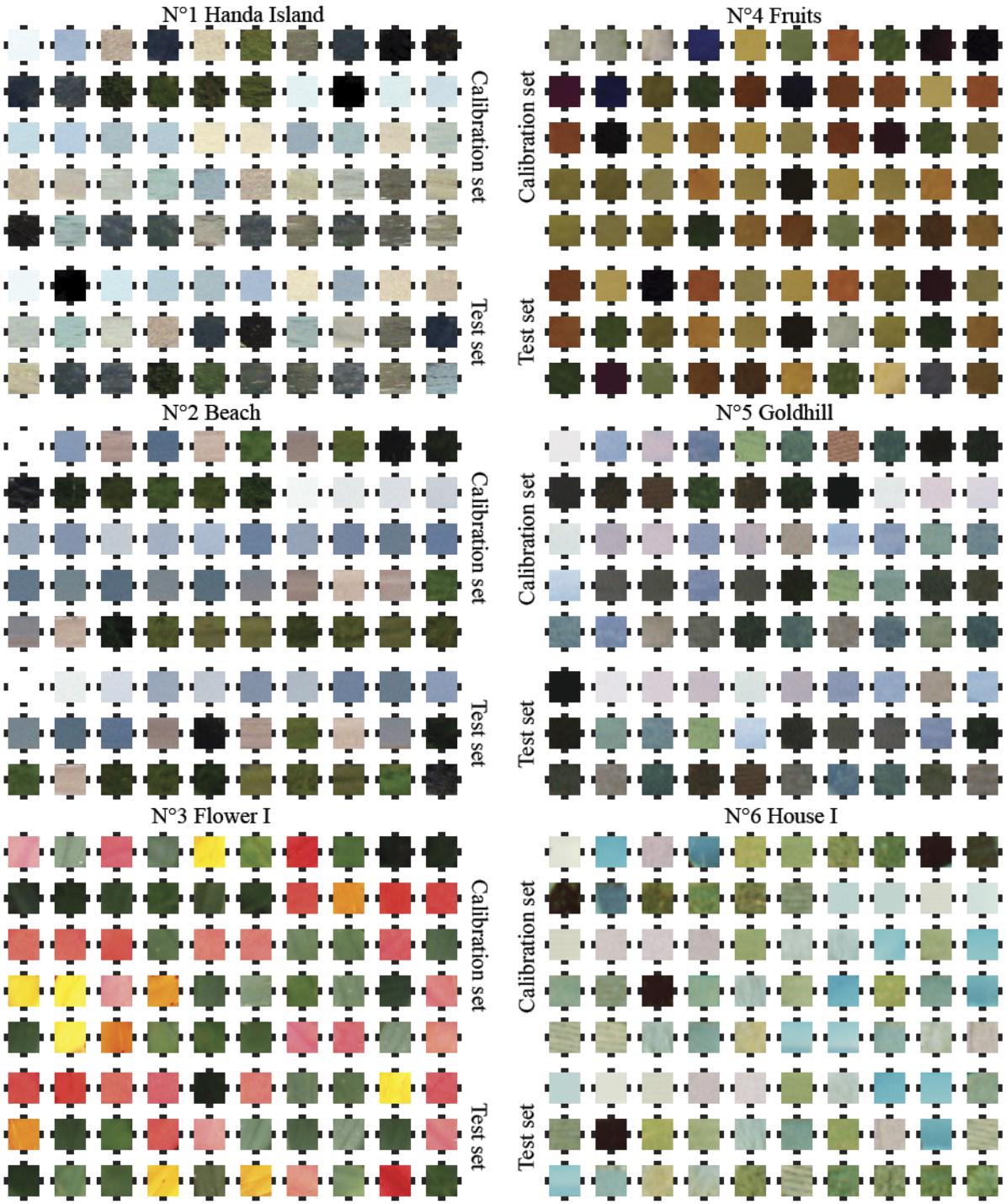


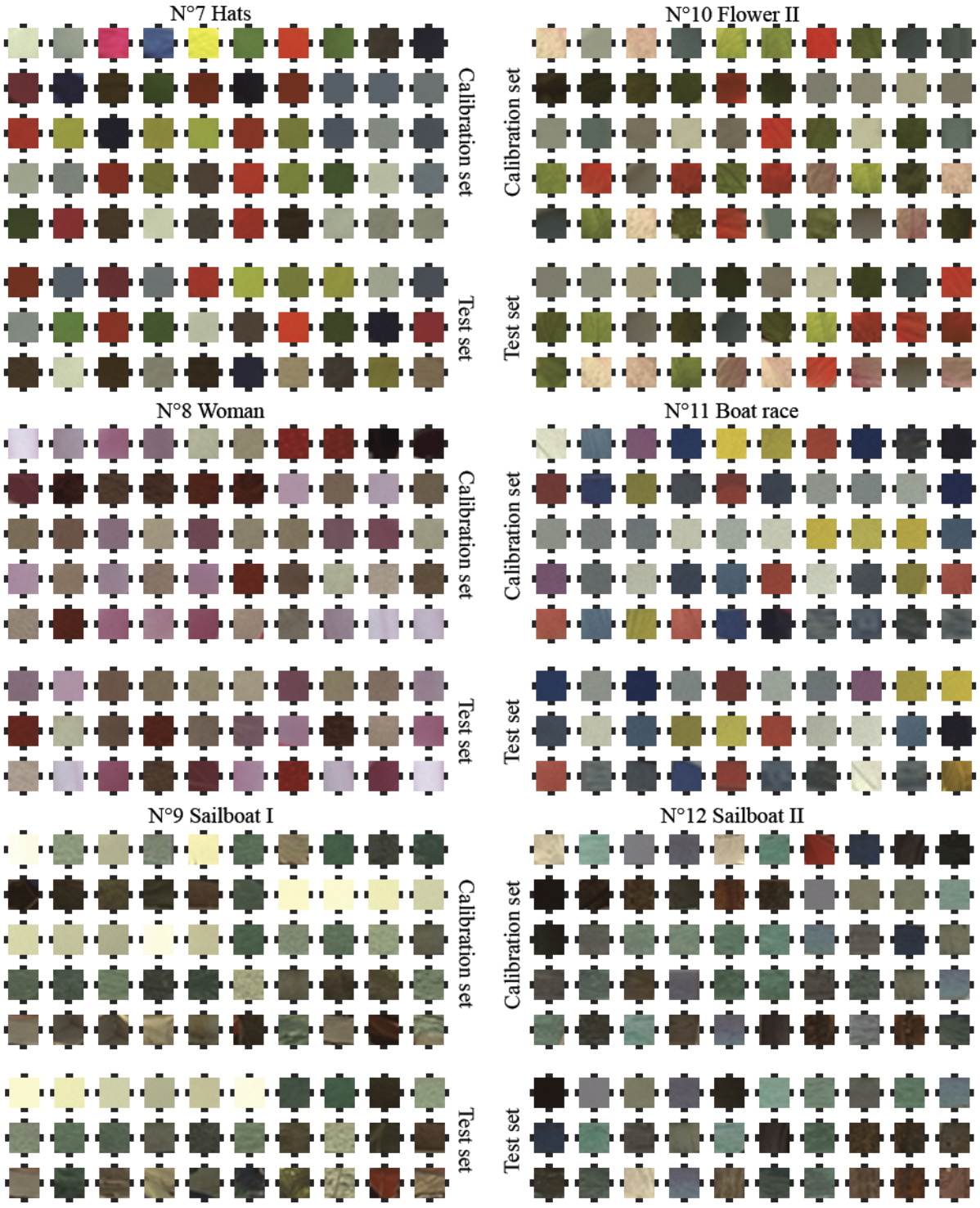

N°7 Hats
Calibration set
Test set
N°10 Flower II
Calibration set
Test set
N°8 Woman
Calibration set
Test set
N°11 Boat race
Calibration set
Test set
N°9 Sailboat I
Calibration set
Test set
N°12 Sailboat II
Calibration set
Test set

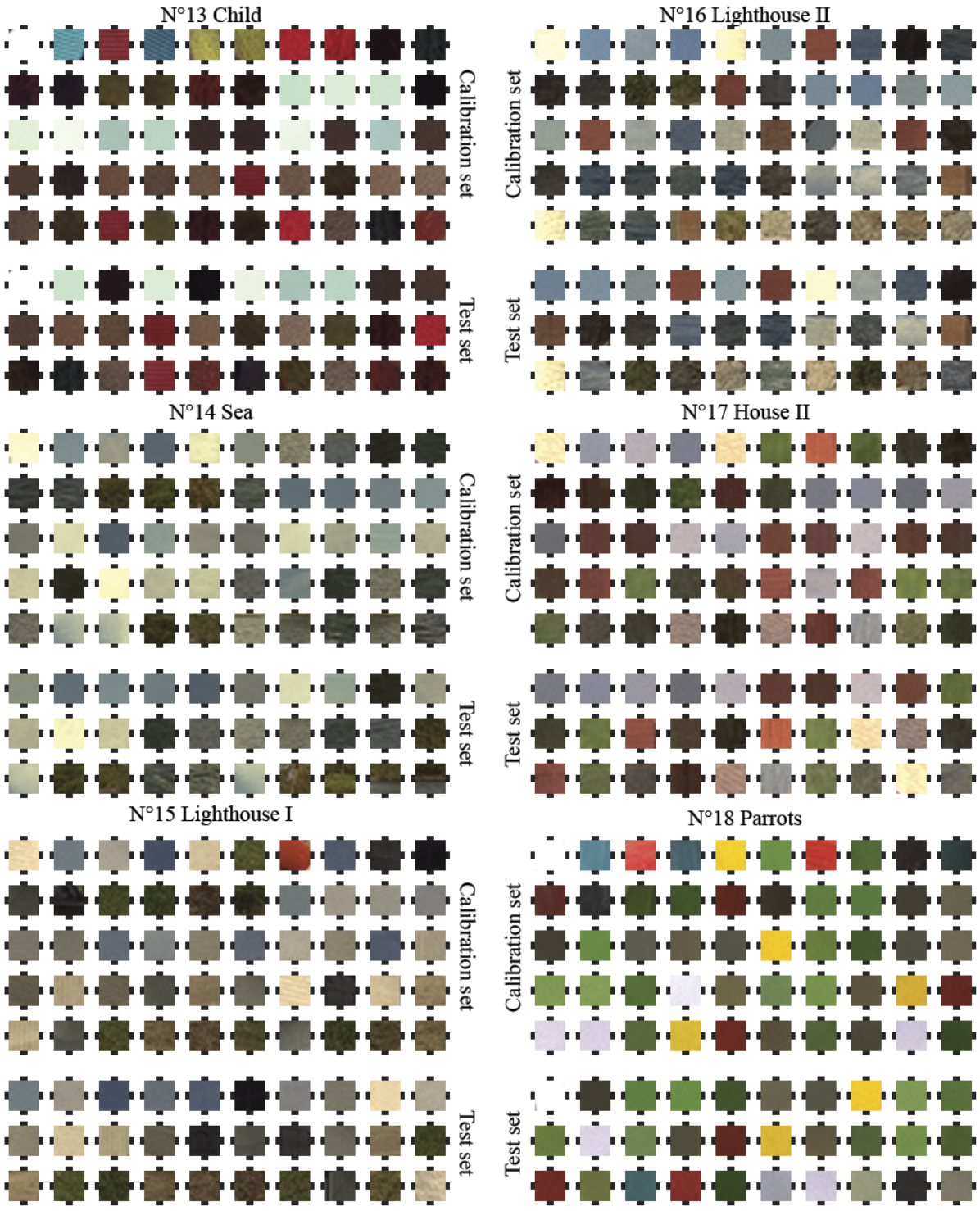

N°13 Child
Calibration set
Test set
N°16 Lighthouse II
Calibration set
Test set
N°14 Sea
Calibration set
Test set
N°17 House II
Calibration set
Test set
N°15 Lighthouse I
Calibration set
Test set
N°18 Parrots
Calibration set
Test set

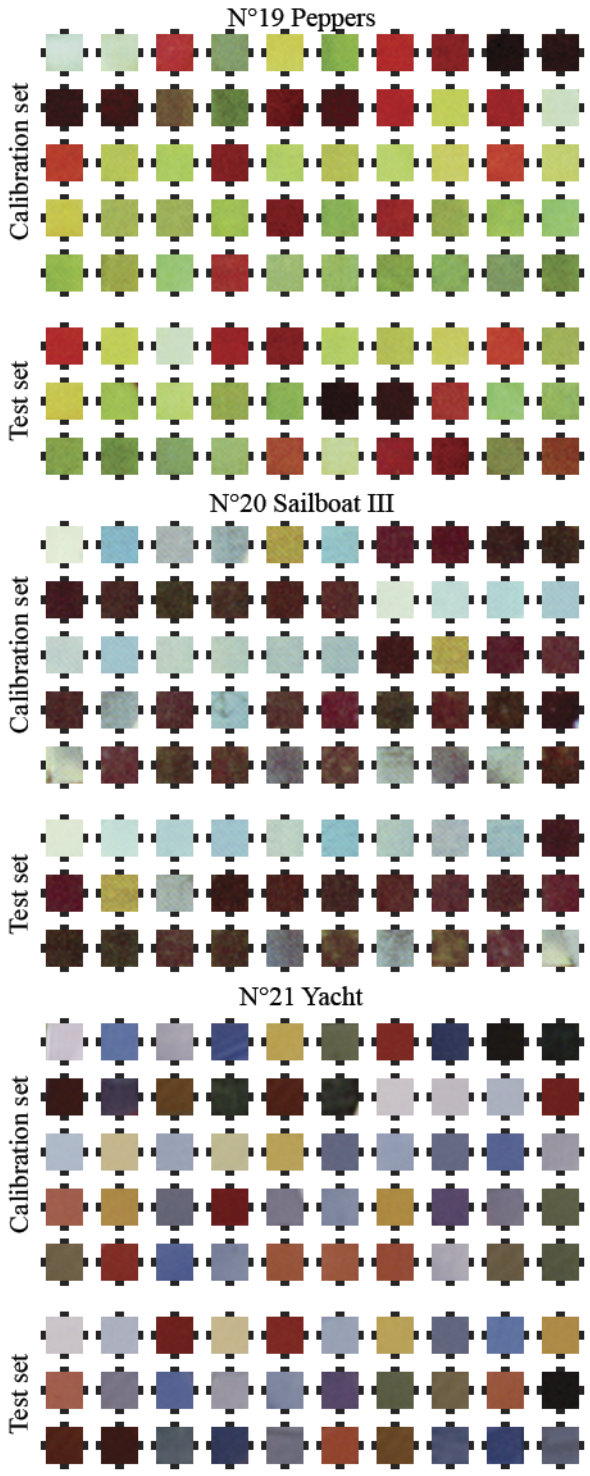

N°19 Peppers
Calibration set
Test set
N°20 Sailboat III
Calibration set
Test set
N°21 Yacht
Calibration set
Test set

D.3 CMYK calibration and test image tiles




N°7 Hats
Calibration set
Test set
N°10 Flower II
Calibration set
Test set
N°8 Woman
Calibration set
Test set
N°11 Boat race
Calibration set
Test set
N°9 Sailboat I
Calibration set
Test set
N°12 Sailboat II
Calibration set
Test set


N°13 Child
Calibration set
Test set
N°16 Lighthouse II
Calibration set
Test set
N°14 Sea
Calibration set
Test set
N°17 House II
Calibration set
Test set
N°15 Lighthouse I
Calibration set
Test set
N°18 Parrots
Calibration set
Test set


N°19 Peppers
Calibration set
Test set
N°20 Sailboat III
Calibration set
Test set
N°21 Yacht
Calibration set
Test set

Bibliography

[AA98] A.U. Agar and J.P. Allebach. An iterative cellular YNSN method for color printer calibration. In *Proceedings of the 6th IS&T/SID Color Imaging Conference, Scottsdale AZ*, pages 17–20, 1998.

[Bal95] R. Balasubramanian. Colorimetric modeling of binary color printers. In *International Conference on Image Processing*, volume 2, pages 327–330. IEEE, 1995.

[Bal96] R. Balasubramanian. The use of spectral regression in modeling halftone color printers. In *IS&T/OSA Optics & Imaging in the Information Age Proceedings*, pages 372–375, 1996.

[Bal99] R. Balasubramanian. Optimization of the spectral Neugebauer model for printer characterization. *Journal of Electronic Imaging*, 8(2):156–166, 1999.

[Bal03] R. Bala. Device Characterization. In S. Gaurav, editor, *Digital Color Imaging Handbook*, chapter 5, pages 269–379. CRC Press, 2003.

[BBH07] T. Bugnon, M. Brichon, and R.D. Hersch. Model-Based Deduction of CMYK Surface Coverages from Visible and Infrared Spectral Measurements of Halftone Prints. In *Color Imaging XII: Processing, Hardcopy, and Applications*. Proc. SPIE volume 6493, paper 649310, pages 1–10, 2007.

[BBH08] T. Bugnon, M. Brichon, and R.D. Hersch. Simplified ink spreading equations for CMYK halftone prints. In *Color Imaging XIII: Processing, Hardcopy, and Applications*. Proc. SPIE volume 6807, paper 680717, pages 1–12, 2008.

[BBT96] R.S. Berns, A. Bose, and D. Tzeng. The spectral modeling of large-format ink jet printers. *Research and Development Final Report (RIT Munsell Color Science Laboratory)*, 1996.

[BH11a] T. Bugnon and R.D. Hersch. Constrained Acquisition of Ink Spreading Curves from Printed Color Images. *IEEE Transactions on Image Processing*, 20(2):513–522, Feb 2011.

[BH11b] T. Bugnon and R.D. Hersch. Optimized Selection of Image Tiles for Ink Spreading Calibration. In *Color Imaging XVI: Processing, Hardcopy, and Applications*. Proc. SPIE volume 7866, paper 786612, pages 1–16, 2011.

[BHeda] T. Bugnon and R.D. Hersch. Calibrating ink spreading curves by optimal selection of tiles from printed color images. *Journal of Electronic Imaging*, to be published.

[BHedb] T. Bugnon and R.D. Hersch. Recovering Neugebauer colorant reflectances and ink spreading curves from printed color images. *Color Research & Application*, to be published.

[Bjö96] Å. Björk. *Numerical Methods for Least-Squares Problems*. SIAM press, 1996.

[BMHed] T. Bugnon, A. Maesani, and R.D. Hersch. Enhancing the specular effect of metallic color prints by reducing the use of yellow ink. *Journal of imaging science and technology*, to be published.

[BSMVdC05] R. Bala, G. Sharma, Vishal Monga, and J.-P. Van de Capelle. Two-dimensional transforms for device color correction and calibration. *IEEE Transactions on Image Processing*, 14(8):1172 –1186, Aug 2005.

[CBT04] Y. Chen, R.S. Berns, and L.A. Taplin. Six color printer characterization using an optimized cellular Yule-Nielsen spectral neugebauer model. *Journal of Imaging Science and Technology*, 48(6):519–528, 2004.

[CBTI03] Y. Chen, R.S. Berns, L.A. Taplin, and F.H. Imai. A multi-ink color-separation algorithm maximizing color constancy. In *Proceeding of IS&T/SID Eleventh Color Imaging Conference*, pages 277–281, 2003.

[CL96] T.F. Coleman and Y. Li. A Reflective Newton Method for Minimizing a Quadratic Function Subject to Bounds on Some of the Variables. *SIAM Journal on Optimization*, 6(4):1040–1058, 1996.

[CY53] F.R. Clapper and J.A.C. Yule. The effect of multiple internal reflections on the densities of half-tone prints on paper. *J. Opt. Soc. Am. A*, 43(7):600–603, 1953.

[Dem24] M.E. Demichel. Le procédé. *Vol*, 26:17–21, 1924.

[EBMZ09] R. Eschbach, R. Bala, M. Maltz, and I. Zhao. Creating variable data infrared signals for security applications. In *Color Imaging XIV: Displaying, Processing, Hardcopy, and Applications*. Proc. SPIE volume 7241, paper 72411E, pages 1–9, 2009.

[EH00] P. Emmel and R.D. Hersch. A unified model for color prediction of halftoned prints. *Journal of Imaging Science and Technology*, 44(4):351–359, 2000.

[FB04] H.S. Fairman and M.H. Brill. The principal compoenents of reflectances. *Color Research & Application*, 29(2):104–110, 2004.

[Fra] R. Franzen. Kodak Lossless True Color Image Suite. `http://r0k.us/graphics/kodak/`.

[Gau03] S. Gaurav. Color fundaments. In *Digital Color Imaging Handbook*, chapter 1, pages 35–36. CRC Press, 2003.

[GL96] P. Green and MacDonald L. *Colour Engineering: Achieving Device Independent Colour*. John Wiley & Sons, 1996.

[GMW81] P.E. Gill, W. Murray, and M.H. Wright. *Practical optimization*. Academic Press, London, UK, 1981.

[HBB+09] R.D. Hersch, M. Brichon, T. Bugnon, P. Amrhyn, F. Crété, S. Mourad, H. Janser, Y. Jiang, and M. Riepenhoff. Deducing ink thickness variations by a spectral prediction model. *Color Research & Application*, 34(6):432–442, 2009.

[HBBH09] R.D. Hersch, M. Brichon, T. Bugnon, and M. Hebert. Deducing Ink Spreading Curves from Reflection Spectra Acquired Within Printed Color Images. *Journal of Imaging Science and Technology*, 53(3):1–7, 2009.

[HC05] R.D. Hersch and F. Crété. Improving the Yule-Nielsen modified Neugebauer model by dot surface coverages depending on the ink superposition conditions. In *Color Imaging X: Processing, Hardcopy, and Applications*, volume 5667, pages 434–447. SPIE, 2005.

[HCE03] R.D. Hersch, F. Collaud, and P. Emmel. Reproducing color images with embedded metallic patterns. *ACM Transactions on Graphics (TOG)*, 22(3):427–434, 2003.

[HECC05] R.D. Hersch, P. Emmel, F. Collaud, and F. Crété. Spectral reflection and dot surface prediction models for color halftone prints. *Journal of Electronic Imaging*, 14(3):033001, 2005.

[HH06] M. Hébert and R.D. Hersch. Reflectance and transmittance model for recto-verso halftone prints. *J. Opt. Soc. Am. A*, 23(10):2415–2432, Oct 2006.

[HJP92] K.J. Heuberger, Z.M. Jing, and S. Persiev. Color transformations and lookup tables. In *TAGA*, pages 863–863, 1992.

[IB98a] K. Iino and R.S. Berns. Building color-management modules using linear optimization. I. Desktop color system. *The Journal of imaging science and technology*, 42(1):79–94, 1998.

[IB98b] K. Iino and R.S. Berns. Building color management modules using linear optimization. II. Prepress system for offset printing. *The Journal of imaging science and technology*, 42(2):99–114, 1998.

[Kan94] H.R. Kang. Applications of color mixing models to electronic printing. *Journal of Electronic Imaging*, 3(3):276–287, 1994.

[KM31] P. Kubelka and F. Munk. Ein Beitrag zur Optik der Farbanstrische. *Z. Tech. Phys.*, 12, 1931.

[Kub48] P. Kubelka. New Contributions to the Optics of Intensely Light-Scattering Materials. Part I. *J. Opt. Soc. Am. A*, 38(5):448–457, 1948.

[Lev] H. Levkin. Set of additional test images. `http://www.hlevkin.com/TestImages/additional.htm`.

[LLBD06] A. Lewandowski, M. Ludl, G. Byrne, and G. Dorffner. Applying the Yule-Nielsen equation with negative n. *J. Opt. Soc. Am. A*, 23(8):1827–1834, Aug 2006.

[MD96] M. Mahy and P. Delabastita. Inversion of the Neugebauer equations. *Color Research & Application*, 21(6):404–411, 1996.

[Neu37] H.E.J. Neugebauer. Die Theoretischen grundlagen Des Mehrfarbendruckes. *Zeitschrift für wissenschaftliche Photographie Photophysik und Photochemie*, 36(4), 1937.

[Nob85] J.H. Nobbs. Kubelka-Munk Theory and the Prediction of Reflectance. *Review of Progress in Coloration and Related Topics*, 15(1):66–75, 1985.

[PP72] I. Pobboravsky and M. Pearson. Computation of dot areas required to match a colorimetrically specified color using the modified Neugebauer equations. In *Proc. TAGA*, volume 32, pages 65–77, 1972.

[RB93] R. Rolleston and R. Balasubramanian. Accuracy of various types of Neugebauer models. In *Proc. IS&T/SID Color Imaging Conference: Transforms and Portability of Color*, volume 2, pages 32–37, 1993.

[RBH08] F. Rousselle, T. Bugnon, and R.D. Hersch. Spectral prediction model for variable dot-size printers. In *Proc. IS&T/SID 16th Color Imaging Conference*, pages 73–78, 2008.

[RBH10] R. Rossier, T. Bugnon, and R.D. Hersch. Introducing ink spreading within the cellular Yule-Nielsen modified Neugebauer model. In *Proc. IS&T/SID 18th Color Imaging Conference*, pages 295–300, 2010.

[TB00] D.-Y. Tzeng and R.S. Berns. Spectral-based six-color separation minimizing metamerism. In *Proc. IS&T/SID Eighth Color Imaging Conference*, pages 342–347, 2000.

[TB01] L.A. Taplin and R.S. Berns. Spectral color reproduction based on a six-color inkjet output system. In *Proc. IS&T/SID Ninth Color Imaging Conference*, pages 209–213, 2001.

[TB05] D.-Y. Tzeng and R.S. Berns. A review of principal component analysis and its applications to color technology. *Color Research & Application*, 30(2):84–98, 2005.

[UG06] P. Urban and R.-R. Grigat. Spectral-based color separation using linear regression iteration. *Color Research & Application*, 31(3):229–238, 2006.

[Urb09] P. Urban. Distributions of Paramers and Paramer Mismatch Gamuts. In *Proc. IS&T/SID Seventeenth Color Imaging Conference*, pages 143–148, 2009.

[Vig83] J.A.S. Viggiano. The GRL Dot Gain Model. In *Proc. TAGA*, pages 423–439, 1983.

[Vig90] J.A.S. Viggiano. Modeling the color of multi-colored halftones. In *Proc. TAGA*, volume 42, pages 44–62, 1990.

[Vig06] J.A.S. Viggiano. Physical significance of negative Yule-Nielsen n-value. In *ICIS '06: Proceedings of the International Congress of Imageing Science*, pages 607–610, 2006.

[Vig10] J.A.S. Viggiano. Ink Penetration, Isomorphic Colorant Mixing, and Negative Values of Yule-Nielsen n. In *Proc. IS&T/SID Eighteenth Color Imaging Conference*, pages 285–290, 2010.

[WB00] D.R. Wyble and R.S. Berns. A critical review of spectral models applied to binary color printing. *Color Research & Application*, 25(1):4–19, 2000.

[WC53] F.C. Williams and F.R. Clapper. Multiple internal reflections in photographic color prints. *J. Opt. Soc. Am. A*, 43(7):595–597, 1953.

[Web] A. Weber. The USC-SIPI Image Database Volume 3: Miscellaneous. `http://sipi.usc.edu/database/database.cgi?volume=misc`.

[Wei] Weibull distribution. `http://en.wikipedia.org/wiki/Weibull_distribution`.

[WK05] D. Wyble and A. Kraushaar. The theoretical basis of multicolor letterpress printing, Hans EJ Neugebauer. *Color Research & Application*, 30(5):322–331, 2005.

[XSSMT99] Minghui Xia, E. Saber, G. Sharma, and A. Murat Tekalp. End-to-end color printer calibration by total least squares regression. *Image Processing, IEEE Transactions on*, 8(5):700 –716, May 1999.

[YC51] J.A.C. Yule and R. Colt. Colorimetric investigation in multicolor printing. In *Proc. TAGA*, volume 3, pages 77–82, 1951.

[YK04] L. Yang and B. Kruse. Revised Kubelka-Munk theory. I. Theory and application. *J. Opt. Soc. Am. A*, 21(10):1933–1941, Oct 2004.

[YLK01] L. Yang, R. Lenz, and B. Kruse. Light scattering and ink penetration effects on tone reproduction. *J. Opt. Soc. Am. A*, 18(2):360–366, Feb 2001.

[YN51] J.A.C. Yule and W.J. Nielsen. The penetration of light into paper and its effect on halftone reproduction. In *Proc. TAGA*, volume 3, pages 65–76, 1951.

[ZS04] S. Zuffi and R. Schettini. Modeling dot gain and inks interaction. In *Proc. IS&T/SID 12th Color Imaging Conference*, pages 181–186, 2004.

[ZSM03] S. Zuffi, R. Schettini, and G. Mauri. Spectral-based characterization of inkjet printers by neural and genetic algorithms. In *IS&T'S PICS Conference*, pages 479–484. Society for Imaging Science & Technology, 2003.

[ZSS06] S. Zuffi, S. Santini, and R. Schettini. Accounting for Inks Interaction in the Yule-Nielsen Spectral Neugebauer Model. *Journal of Imaging Science & Technology*, 50(1):35–44, 2006.

List of Figures

List of Tables

Curriculum Vitæ

Thomas Bugnon was born in 1979 in Delémont. During his studies in Communication Systems at the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, he spent a year at DTU in Copenhagen, Denmark, as an Erasmus student. From March to October 2003, he worked as intern software developer at Whitestein Technologies AG, Zurich. He received his M.S degree in Communication Systems in 2005.

He joined the group of Prof. Roger D. Hersch at the Peripheral Systems Lab in 2005 as a research assistant. Responsible for the Microstructures and MetallicArt anti-counterfeiting technologies, he worked in collaboration with industrial partners to improve these technologies. Since 2007, he has been a PhD candidate focusing on spectral prediction models for color reproduction.

Personal bibliography

[BBH07] T. Bugnon, M. Brichon, and R.D. Hersch. Model-based Deduction of CMYK Surface Coverages from Visible and Infrared Spectral Measurements of Halftone Prints. In *Color Imaging XII: Processing, Hardcopy, and Applications*, Proc. SPIE volume 6493, paper 649310, pages 1–10, 2007.

[BBH08] T. Bugnon, M. Brichon, and R.D. Hersch. Simplified Ink Spreading Equations for CMYK Halftone Prints. In *Color Imaging XIII: Processing, Hardcopy, and Applications*, Proc. SPIE volume 6807, paper 680717, pages 1–12, 2008.

[BH11a] T. Bugnon and R.D. Hersch. Constrained Acquisition of Ink Spreading Curves from Printed Color Images. *IEEE Transactions on Image Processing*, 20(2), pages 513–522, Feb 2011.

[BH11b] T. Bugnon and R.D. Hersch. Optimized Selection of Image Tiles for Ink Spreading Calibration. In *Color Imaging XVI: Processing, Hardcopy, and Applications*, Proc. SPIE volume 7866, paper 786612. pages 1–16, 2011.

[BHedb] T. Bugnon and R.D. Hersch. Recovering Neugebauer colorant reflectances and ink spreading curves from printed color images. *Color Research & Application*, to be published.

[BHeda] T. Bugnon and R.D. Hersch. Calibrating ink spreading curves by optimal selection of tiles from printed color images. *Journal of Electronic Imaging*, to be published.

[BMHed] T. Bugnon, A. Maesani, and R.D. Hersch. Enhancing the specular effect of metallic color prints by reducing the use of yellow ink. *Journal of imaging science and technology*, to be published.

Personal contributions

[RBH08] F. Rousselle, T. Bugnon, and R.D. Hersch. Spectral Prediction Model for Variable Dot-Size Printers. In *Proc. IS&T/SID 16th Color Imaging Conference*, pages 73–78, 2008.

[HBBH09] R.D. Hersch, M. Brichon, T. Bugnon, and M. Hebert. Deducing Ink Spreading Curves from Reflection Spectra Acquired Within Printed Color Images. *Journal of Imaging Science and Technology*, 53(3), paper 030502, pages 1–7, 2009.

[HBB+09] R.D. Hersch, M. Brichon, T. Bugnon, P. Amrhyn, F. Crété, S. Mourad, H. Janser, Y. Jiang, and M. Riepenhoff. Deducing Ink Thickness Variations by a Spectral Prediction Model. *Color Research & Application*, 34(6):432–442, 2009.

[RBH10] R. Rossier, T. Bugnon, and R.D. Hersch. Introducing Ink Spreading Within the Cellular Yule-Nielsen modified Neugebauer Model. In *Proc. IS&T/SID 18th Color Imaging Conference*, pages 295–300, 2010.