

## Calibrating ink spreading curves by optimal selection of tiles from printed color images

**Thomas Bugnon\***

thomas.bugnon@a3.epfl.ch

**Roger David Hersch\***

rd.hersch@epfl.ch

Tel: +41 21 6934357

Fax: +41 21 6936680

EPFL-IC-LSP

Station 14

1015 Lausanne

Switzerland

\*School of Computer and Communication Sciences

École Polytechnique Fédérale de Lausanne (EPFL)

### ABSTRACT

The Yule-Nielsen modified spectral Neugebauer model (YNSN) enables predicting reflectance spectra from ink surface coverages of halftones. In order to provide an improved prediction accuracy, this model is enhanced with an ink spreading model accounting for ink spreading in all superposition conditions (IS-YNSN). As any other spectral reflection prediction model, the IS-YNSN model is conceived to predict the reflection spectra of color-constant patches. Instead of color-constant patches, we investigate if tiles located within color images can be accurately predicted and how they can be used to facilitate the calibration of the ink spreading model. In the present contribution, we detail an algorithm to automatically select image tiles as uniform as possible from color images by relying on their CMY or CMYK pixel values. The tile selection algorithm incorporates additional constraints relying on surface coverages of the inks. We demonstrate that an ink spreading model calibrated with as few as 5 to 10 optimally chosen image tiles allows the corresponding YNSN model to provide accurate spectral predictions.<sup>1</sup>

---

<sup>1</sup> This is an extended version of a paper previously published at the SPIE/IS&T 2011 Electronic Imaging conference: T. Bugnon, R.D.Hersch, "Optimized Selection of Image Tiles for Ink Spreading Calibration," *Color Imaging XVI: Processing, Hardcopy, and Applications*, Proc. SPIE, Vol. 7866, paper 786612, pp. 1-16, 2011.

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

## 1. INTRODUCTION

In color reproduction, the Yule-Nielsen modified spectral Neugebauer model (YNSN) is one of the most widely used models because of its simplicity and its relatively good accuracy<sup>1-4</sup>. In order to improve its accuracy, this model has been enhanced with an ink spreading model accounting for physical dot gain (IS-YNSN<sup>5-7</sup>). As any spectral reflection prediction model, the IS-YNSN model is designed to predict the reflection spectra of color-constant patches. Instead of color-constant patches, we investigate if tiles located within color images can be accurately predicted and how they can be used to facilitate the calibration of the IS-YNSN model. Facilitating the calibration of such models is important because a calibration should be performed each time the paper, the inks or the printing configuration is modified. Providing the means to perform this calibration in a more flexible manner by removing constraints such as printing specific color-constant patches leads to systems that can calibrate themselves automatically and are less expensive to operate. For example, in order to print with high color fidelity, one may not have to recalibrate an offset press or desktop printer each time a new type of paper is used.

Image tiles are formed by small square regions (5mm x 5 mm) of a color as uniform as possible. They are extracted from a color image and used to perform the model calibration, as mentioned in previous publications of the authors. These publications describe an algorithm to automatically select image tiles<sup>8</sup> and use them to calibrate the ink spreading model<sup>8-9</sup>. 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 contribution, 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.

This paper is organized as follows. The Yule-Nielsen modified Spectral Neugebauer model (YNSN) and the ink spreading enhanced YNSN model (IS-YNSN) are described in Sections 2 and 3, respectively. Section 4 shows how to calibrate the IS-YNSN model from image tiles. In Section 5, a non-uniformity metric based on the CMYK color values of the image pixels is detailed as well as an algorithm to automatically select image tiles based on this non-uniformity metric. Section 6 presents an optimized version of the selection algorithm maximizing the sum of the ink spreading weights, i.e. the overall impact of the halftones on the ink spreading curves. The experimental setup is presented in Section 7. Section 8 shows that uniform tiles can be accurately predicted with a classical calibration. In Section 9, we investigate the impact of the tile selection method and of the number of calibration tiles on the ink spreading calibration. A conclusion is drawn in Section 10.

## **2. THE YULE-NIELSEN MODIFIED SPECTRAL NEUGEBAUER MODEL (YNSN)**

One of the first color prediction models is the Neugebauer model<sup>10</sup>. In its original form, it predicts the RGB values of a color halftone patch. Yule and Nielsen modified this model to account for optical dot gain<sup>1</sup>, Yule and Colt applied it for CIE-XYZ tri-stimulus values<sup>11</sup>, and Viggiano extended it to the spectral domain<sup>2</sup>. This model, now known as the Yule-Nielsen modified Spectral Neugebauer model (YNSN), predicts the reflection spectra of color-constant patches whose ink coverages are given. The equation used to perform a prediction is the following:

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

where  $R(\lambda)$  is the predicted reflection spectrum,  $R_i(\lambda)$  the reflection spectra of the colorants (or Neugebauer primaries),  $a_i$  the relative area coverages of the colorants, and  $n$  the Yule-Nielsen factor.

The spectra of the colorants must be measured and the scalar  $n$  factor needs to be calibrated<sup>5</sup>. The relative area coverages of the colorants are computed from the relative ink surface coverages using the Demichel equations. For four inks, the Demichel equations are the following:

$$\begin{array}{ll}
a_w = (1-c) \cdot (1-m) \cdot (1-y) \cdot (1-k) & a_k = (1-c) \cdot (1-m) \cdot (1-y) \cdot k \\
a_c = c \cdot (1-m) \cdot (1-y) \cdot (1-k) & a_{ck} = c \cdot (1-m) \cdot (1-y) \cdot k \\
a_m = (1-c) \cdot m \cdot (1-y) \cdot (1-k) & a_{mk} = (1-c) \cdot m \cdot (1-y) \cdot k \\
a_{cm} = c \cdot m \cdot (1-y) \cdot (1-k) & a_{cmk} = c \cdot m \cdot (1-y) \cdot k \\
a_y = (1-c) \cdot (1-m) \cdot y \cdot (1-k) & a_{yk} = (1-c) \cdot (1-m) \cdot y \cdot k \\
a_{cy} = c \cdot (1-m) \cdot y \cdot (1-k) & a_{cyk} = c \cdot (1-m) \cdot y \cdot k \\
a_{my} = (1-c) \cdot m \cdot y \cdot (1-k) & a_{myk} = (1-c) \cdot m \cdot y \cdot k \\
a_{cmy} = c \cdot m \cdot y \cdot (1-k) & a_{cmyk} = c \cdot m \cdot y \cdot k
\end{array} \tag{2}$$

where  $c$ ,  $m$ ,  $y$  and  $k$  are the ink surface coverages.

When calibrated, the YNSN model can be seen as a function taking the ink surface coverages of a 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:

$$covs_{opt} = \arg \min_{covs} \left[ \sum_{\lambda} (predSpectrum(\lambda, covs) - R(\lambda))^2 \right] \tag{3}$$

where  $covs = \{c, m, y, k\}$  are the ink coverages,  $R(\lambda)$  the measured spectrum, and  $predSpectrum$  the YNSN model predicting function combining Eqs. (1) and (2).

Due to physical dot gain, ink coverages deduced using Eq. (3) are usually larger than the corresponding nominal ink coverages. As a consequence, spectral predictions made with nominal ink coverages are not accurate. In order to make accurate predictions, the YNSN model should be augmented with an ink spreading model<sup>6</sup>.

### 3. THE INK SPREADING MODEL (IS-YNSN)

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 coverages usually larger than the requested nominal ink coverages. This phenomenon is known as *physical dot gain* or *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<sup>6</sup>. The

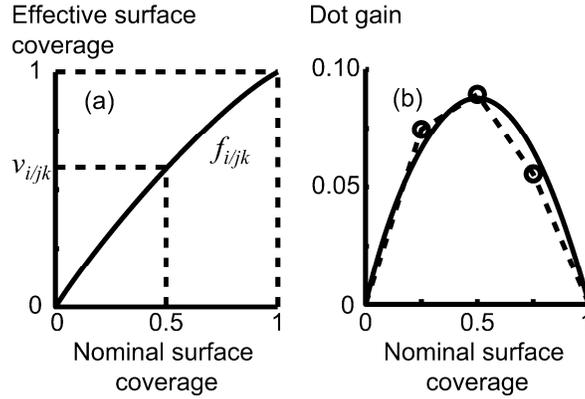


Fig. 1. (a) Example of an ink spreading curve characterized by its mid-point  $v$ . (b) Corresponding dot gain curve defined as the difference between effective and nominal surface coverages, with the fitted dot gains of three calibration patches (circles), the linearly interpolated dot gain curve (dashed line), and the corresponding least-squares approximated parabolic dot gain curve (solid line).

goal of the ink spreading model is therefore to deduce *effective ink coverages* from *nominal ink coverages*. The effective surface coverages are then used as input to the YNSN model.

### 3.1 Ink spreading curves

The ink spreading model relies on ink spreading curves (Fig. 1a). 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 once printed. We can approximate an ink spreading curve by a parabola characterized by its mid-point  $v$  as follows<sup>9</sup>:

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

where  $u_i$  is the nominal surface coverage of ink  $i$ ,  $f_{i/jk}$  is the ink spreading curve of ink  $i$  superposed with solid inks  $j$  and  $k$ , and  $v_{i/jk}$  is the effective surface coverage at 50% nominal surface coverage, also called mid-point of the ink spreading curve. Equation (4) represents a parabola passing through the three points  $(0,0)$ ,  $(0.5, v_{i/jk})$ , and  $(1,1)$ , as shown in Fig. 1a.

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,

Table I. List of the considered ink spreading curve indicia.

Cyan	Magenta	Yellow	Black	
$c$	$m$	$y$	$k$	$k/y$
$c/m$	$m/c$	$y/c$	$k/c$	$k/cy$
$c/y$	$m/y$	$y/m$	$k/m$	$k/my$
$c/my$	$m/cy$	$y/cm$	$k/cm$	$k/cmy$

yielding a total of 32 ink spreading curves. However, since any halftone superposed with solid black yields a reflection spectrum very close to the reflection spectrum of solid black, ink spreading curves where one ink halftone is superposed with solid black are discarded<sup>7</sup>. Table I lists the 20 ink spreading curves considered.

Fig. 1b shows the dot gain curve corresponding to the ink spreading curve of Fig. 1a, where dot gain is defined as the difference between effective and nominal coverages. Each ink spreading curve is calibrated using three calibration patches at 0.25, 0.5 and 0.75 nominal surface coverages. The nominal coverage of the ink halftone determines the position on the x-axis of the ink spreading curve. The corresponding effective surface coverage is fitted using Eq. (3) and determines the position on the y-axis. The parabola is then fitted by minimizing the least-squares difference between the fitted and predicted dot gains. An ink spreading curve can be entirely characterized by the halftone at 50% nominal surface coverage, but as shown in Fig. 1b, we prefer 3 calibration patches per ink spreading curve.

In order to avoid the ambiguity between chromatic black and pure black<sup>12</sup>, the spectral measurements span both the visible wavelength range (380-730 nm) and the near infrared (NIR) wavelength range (730-850 nm). In many printing systems, the NIR wavelength range helps to distinguish between black ink and chromatic black obtained by superpositions of solid cyan, magenta and yellow inks.

### 3.2 Ink Spreading Equations

In order to obtain the effective surface coverages of the ink dots forming a color halftone, we weight the contributions of the different ink spreading curves as follows. The effective coverage of halftone ink  $i$  superposed with inks  $j$  and  $k$  is the weighted average of the ink spreading functions  $f_i$ ,  $f_{i/j}$ ,  $f_{i/k}$ , and  $f_{i/jk}$ . The weights are computed according to the surface

coverages of the underlying colorants (Neugebauer primaries) formed by inks  $j$  and  $k$ . For CMYK, the ink spreading equations are:

$$\begin{aligned}
c' &= (1-m')(1-y') f_c(c) & m' &= (1-c')(1-y') f_m(m) \\
&+ m'(1-y') f_{c/m}(c) & &+ c'(1-y') f_{m/c}(m) \\
&+ (1-m') y' f_{c/y}(c) & &+ (1-c') y' f_{m/y}(m) \\
&+ m' y' f_{c/my}(c) & &+ c' y' f_{m/cy}(m) \\
y' &= (1-c')(1-m') f_y(y) & k' &= (1-c')(1-m')(1-y') f_k(k) \\
&+ c'(1-m') f_{y/c}(y) & &+ c'(1-m')(1-y') f_{k/c}(k) \\
&+ (1-c') m' f_{y/m}(y) & &+ (1-c') m'(1-y') f_{k/m}(k) \\
&+ c' m' f_{y/cm}(y) & &+ c' m'(1-y') f_{k/cm}(k) \\
& & &+ (1-c')(1-m') y' f_{k/y}(k) \\
& & &+ c'(1-m') y' f_{k/cy}(k) \\
& & &+ (1-c') m' y' f_{k/my}(k) \\
& & &+ c' m' y' f_{k/cmy}(k)
\end{aligned} \tag{5}$$

In Eqs. (5), e.g. when calculating the effective surface coverage  $c'$  of the cyan ink,  $(1-m')(1-y')$ ,  $m'(1-y')$ ,  $(1-m') y'$ , and  $m' y'$  express, respectively, the surface coverage of colorants white, magenta, yellow, and red superposed with the cyan halftone layer. Since a cyan, magenta, or yellow halftone over solid black is assumed to yield black, the superposition conditions corresponding to an ink halftone over solid black have been discarded. We have therefore 4 superposition conditions for the cyan, magenta and yellow ink halftones and 8 superposition conditions for the black ink halftone.

We solve Eqs. (5) iteratively, starting by assigning the nominal ink halftone coverages ( $c$ ,  $m$ ,  $y$ ,  $k$ ) to the effective ink halftone coverages ( $c'$ ,  $m'$ ,  $y'$ ,  $k'$ ). Four to five iterations ensure sufficient convergence to determine the effective ink halftone surface coverages<sup>7</sup>.

#### 4. CALIBRATION OF THE INK SPREADING CURVES USING IMAGE TILES

As explained in Section 3.1, an ink spreading curve is fitted using calibration patches with predetermined nominal coverages. Such calibration patches cannot be found in the content of a traditional printing job. Instead of adding calibration patches in the margins of the page, we calibrate the ink spreading curves using measurements of image tiles selected within the printed content.

A method developed to perform such a calibration was proposed by the current authors<sup>9</sup>. 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. This starting point is also always valid according to the constraints imposed on the mid-points of the ink spreading curves detailed in Eq. (7), i.e. the largest possible constraint forces a mid-point to be set to 0.5.

#### 4.1 Gradient-descent based calibration of the ink spreading model

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  $\vec{v}$ :

$$f(\vec{v}) = \sum_p \sum_{\lambda} [\text{predSpectrum}(\lambda, \text{covs}_p, \vec{v}) - \text{measSpectrum}(\lambda, p)]^2 \quad (6)$$

where  $\lambda$  is the wavelength,  $p$  an image calibration tile,  $\text{covs}_p$  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. (6) can be summarized as follows:

$$\vec{v}_{opt} = \arg \min_{\vec{v}} [f(\vec{v})] \text{ such that } \begin{cases} v_{i/jk} \leq 0.5 + 0.25 \cdot w_{i/jk} \\ v_{i/jk} \geq 0.5 - 0.25 \cdot w_{i/jk} \end{cases} \quad (7)$$

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, used to constrain  $v_{i/jk}$ .

To each ink spreading curve  $f_{i/jk}$  of ink  $i$  superposed with solid inks  $j$  and  $k$ , we associate a weight  $w_{i/jk}$  that prevents the minimization algorithm from setting the mid-points  $v_{i/jk}$  at an artificially large value when that ink spreading curve does not significantly influence any image tile. Without such constraints, the minimization algorithm tends to induce large variations of the values of the 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<sup>9</sup>.

## 4.2 Imposing constraints on the ink spreading mid-points

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 coverage and  $u_{i,p}'(v_i, v_{i/j}, v_{i/k}, v_{i/jk})$  be the effective 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:

$$w_{i/jk,p} = \frac{\partial u_{i,p}'}{\partial v_{i/jk}} \quad (8)$$

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. (8) 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 across all the considered calibration tiles:

$$w_{i/jk} = \max_p [w_{i/jk,p}] = \max_p \frac{\partial u_{i,p}'}{\partial v_{i/jk}} \quad (9)$$

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.

## 5. 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 8, 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:

$$\nu = \sqrt{(\sigma(c)^2 + \sigma(m)^2 + \sigma(y)^2 + \sigma(k)^2) / 4} \quad (10)$$

where  $\nu$  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  $\sigma$  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  $\nu$  is low, i.e. the tile is considered uniform. However, if many pixels have values distant from the average, the standard deviations are large and  $\nu$  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. (10). 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  $\nu$  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*. Within a standard image calibration tile set, the tiles are ordered according to their non-uniformity value.

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.

## 6. ADDING CONSTRAINTS TO THE TILE SELECTION ALGORITHM

The standard selection algorithm proposed in Section 5 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 are the tiles maximizing the weights  $w_{i/jk}$  associated to the ink spreading curves in Eq. (9). 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 uniformity, we use the following function inspired by the Weibull probability distribution<sup>13</sup>:

$$z(x) = e^{-(x/l)^4} \quad (11)$$

Function  $z(x)$  is shown in Fig. 2 and is used to adjust the weights  $w_{i/jk}$  as follows:

$$\tilde{w}_{i/jk,p} = w_{i/jk,p} z(\nu_p) \quad (12)$$

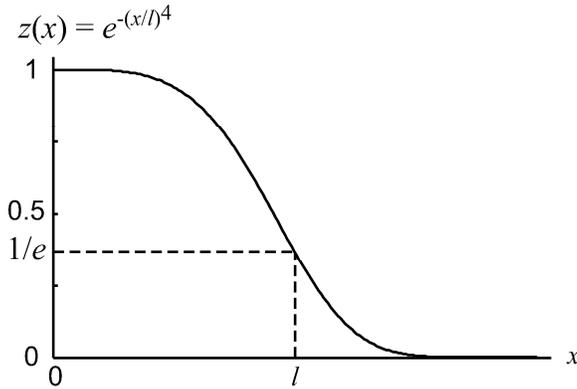


Fig. 2. Adjustment function  $z(x)$  inspired by the Weibull distribution.

where  $\tilde{w}_{ijk,p}$  is the uniformity-adjusted weight of tile  $p$ ,  $w_{ijk,p}$  is the weight of tile  $p$  computed in Eq. (8), and  $\nu_p$  is the non-uniformity value of tile  $p$  computed in Eq. (10). If the given tile is uniform, i.e. its non-uniformity value  $\nu_p$  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. (11), 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  $l$ .

Given ink spreading curve  $f_{ijk}$  and candidate image tile  $p$ , uniformity-adjusted weight  $\tilde{w}_{ijk,p}$  indicates if tile  $p$  is suitable to calibrate ink spreading curve  $f_{ijk}$ . By taking for each ink spreading curve the image tile with the maximum uniformity-adjusted weight  $\tilde{w}_{ijk,p}$ , 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

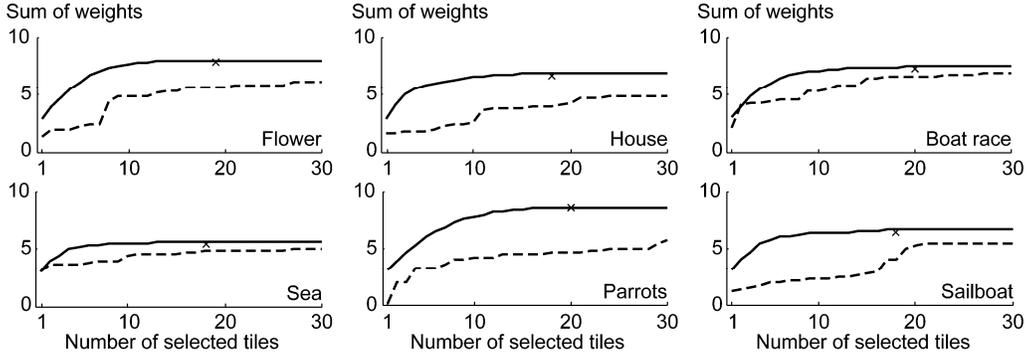


Fig. 3. Sum of the 20 CMYK ink spreading curve weights for the six test images according to the number of selected tiles using either the standard (dashed line) or optimized (solid lines) tile selection algorithm. The crosses indicate the sum of the ink spreading weights of the adjusted calibration tile sets.

perform the selection. This second criterion is the sum  $s$  of the ink spreading weights  $w_{i/jk}$  computed using all the combination of  $i, j$ , and  $k$  listed in Table I as follows:

$$s = \sum_{i,j,k} w_{i/jk} = \sum_{i,j,k} \max_P [w_{i/jk,p}] \quad (13)$$

where the weights  $w_{i/jk}$  are computed according to Eq. (9). 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. (12). 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 5, 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 5). Image tile sets selected using this algorithm are further referred to as *optimized image calibration tile sets*.

Fig. 3 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 Figure B1 in Appendix B and described in Section 7, 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.

## 7. 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 ( $45^\circ\text{d}:0^\circ$  geometry), we measure each sample four different times, each time rotating it by  $90^\circ$  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.

In order to analyze how the different selection algorithms perform, we use the 6 different sRGB images shown in Figure B1 in Appendix B. The first image comes from a digital camera and the other five from different sources providing standard test images<sup>14-16</sup>. 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 Figure B4 in Appendix B. Each set contains tiles randomly selected among the candidate tiles of the 6 images having a non-uniformity value  $\nu$

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 5 (standard image calibration tile sets) and Section 6 (optimized image calibration tile sets) are shown in Figure B2 for the CMY images and in Figure B3 for the CMYK images (Appendix B). They are printed on a Canon PixmaPro 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.

## **8. 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 color-constant 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 calibration yielded the same optimal Yule-Nielsen  $n$  factor of 10. The first part of Tables II, IV, and V 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 IV) or the 125 UCR set (Tables II and V).

Table II shows the accuracy when predicting the uniformity test sets. We can see that the average  $\Delta E_{94}$  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  $\Delta E_{94}$ .

Table II. Prediction accuracy of the classically calibrated IS-YNSN model when predicting the uniform 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  $\Delta E_{94}$ .

	$\Delta E_{94}$		
	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 IV. 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  $\Delta E_{94}$ .

	$\Delta E_{94}$		
	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 III. Average accuracy of predicting sets of tiles grouped according to their non-uniformity values using effective coverages predicted with the classically calibrated ink spreading model (*Predicted* column) and accuracy obtained by direct fitting of individual tile surface coverages according to Eq. (3) (*Deduced* column).  $\Delta Covs$  indicates the average CMYK norm between predicted and deduced effective coverages.

	Average $\Delta E_{94}$		$\Delta Covs$
	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	4.11	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 V. 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  $\Delta E_{94}$ .

	$\Delta E_{94}$		
	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

An additional experiment shows that the surface coverages deduced from 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 deduce these effective coverages by directly fitting them with the YNSN model using Eq. (3). Table III shows that 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. (6), 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 IV and V show the prediction accuracy of the classically calibrated IS-YNSN model when predicting the CMY and CMYK image calibration tile sets, respectively. The IS-YNSN model is also able to accurately predict image tiles, i.e. the average  $\Delta E_{94}$  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.

## 9. 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 color-constant 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 8. 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. The first calibration of the series is performed using only the first image tile of the set. The second calibration is performed

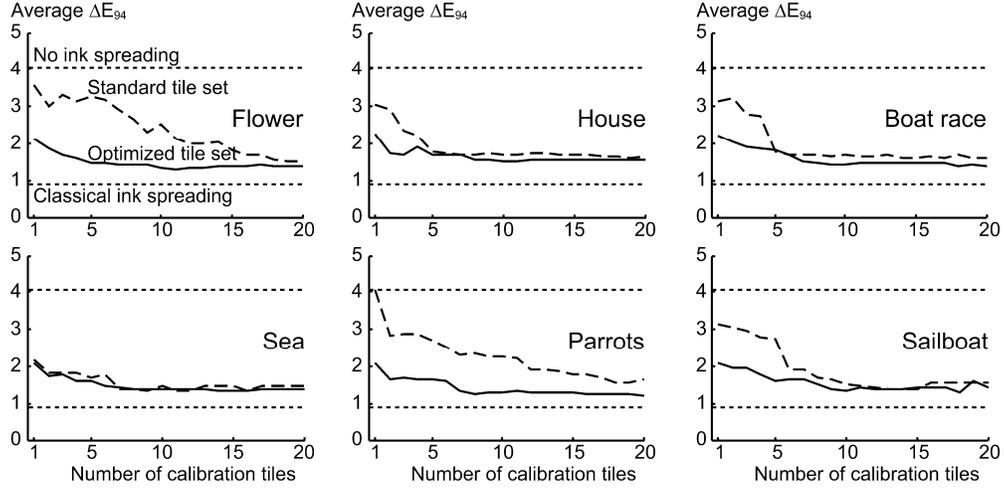


Fig. 4. Accuracy of different calibrated IS-YNSN models when predicting the uniform patches of the 125 CMY set. The following calibrations for six different CMY images are shown: no ink spreading (upper dotted lines), classical ink spreading calibration (lower dotted line), ink spreading calibrations using standard image tile sets (dashed line), and ink spreading calibrations using optimized image tile sets (solid line).

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 calibrations 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  $\Delta E_{94}$ , the higher the accuracy.

Figs. 4 and 5 show the prediction results for the CMY and CMYK images, respectively. We show the average  $\Delta E_{94}$  difference between predictions and measurements. We also show the corresponding 95 percentiles in Appendix A. The  $\Delta E_{94}$  prediction differences in term of 95 percentiles behave similarly to the average  $\Delta E_{94}$  difference. The larger variations are due to the fact that we minimize the average spectral difference in Eq. (6) and not the 95 percentile. In both cases, the IS-YNSN model with calibrated ink spreading curves is much more

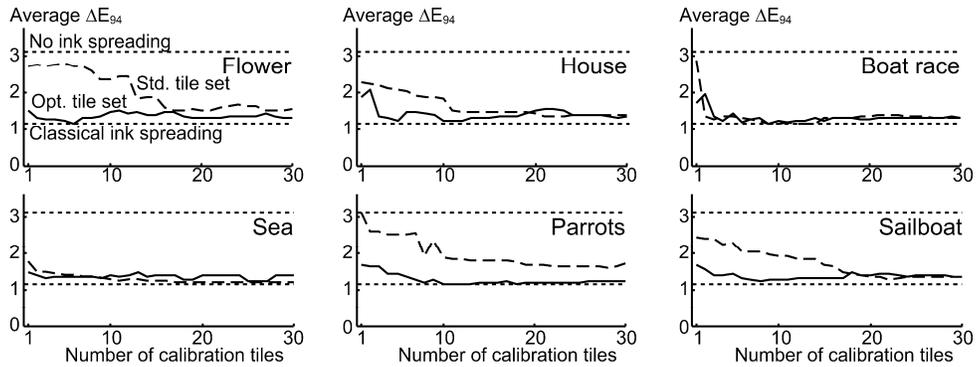


Fig. 5. Accuracy of different calibrated IS-YNSN models when predicting the CMYK uniform patches of the 125 UCR set. The following calibrations for six different CMYK images are shown: no ink spreading (upper dotted lines), classical ink spreading calibration (lower dotted line), ink spreading calibrations using standard image tile sets (dashed line), and ink spreading calibrations using optimized image tile sets (solid line).

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 3<sup>rd</sup> tile of the CMY Flower standard calibration set is a red tile similar to the 2<sup>nd</sup> tile of the set (Figure B2, Appendix B). The 9<sup>th</sup> tile of the CMYK Parrots standard set is the 3<sup>rd</sup> dark brown tile of the set (Figure B3, Appendix B). 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 9<sup>th</sup> and 13<sup>th</sup> tiles that are green (Figure B3, Appendix B). When the 1<sup>st</sup> and 2<sup>nd</sup> 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 calibration sets (Fig. 3, Flower image), there is also a large difference in the prediction accuracy when the model is calibrated with a limited number of tiles (Figs. 4 and 5). When the difference in the sum is small (Fig. 3, Sea or Boat race images), the difference in prediction accuracy is also small, i.e. the two curves in Figs. 4 and 5 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 (Figures B2 and B3, Appendix B). 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 (Fig. 3). As a consequence, the prediction accuracy achieved by the standard algorithm only slowly improves compared to the accuracy achieved by the optimized algorithm (Figs. 4 and 5). 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

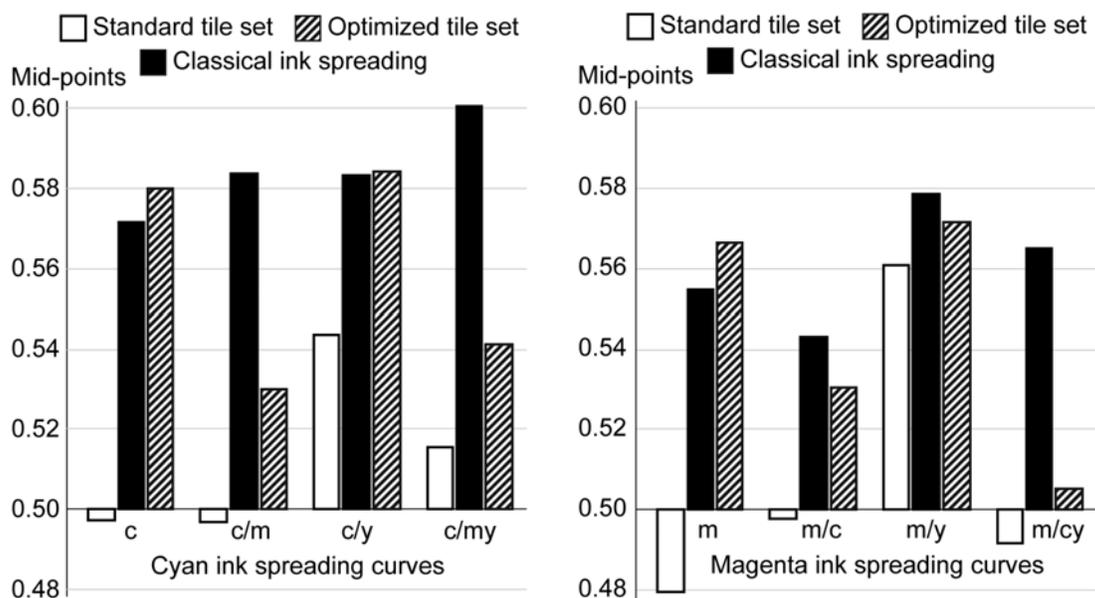


Fig. 6. Mid-points of the four cyan (left) and four magenta (right) ink spreading curves of the Flower image for the classical calibration (black bars), the ink spreading calibration using a standard tile set containing 10 image tiles (white bars), and the ink spreading calibration using an optimized tile set containing 10 image tiles (hatched bars).

are forced to be well distributed. In some cases such as the Sea or Boat race images, both algorithms have the same performances.

## 10. COMPARISON OF INK SPREADING CURVES BETWEEN CLASSICAL, STANDARD AND OPTIMIZED CALIBRATIONS

Optimized calibrations achieve high accuracy with only a limited number of image calibration tiles, an accuracy similar to the classical calibration used as a reference, whereas standard calibrations may require many more image calibration tiles. In this section, we compare the ink spreading curves calibrated by the classical, standard and optimized calibrations.

Fig. 6 shows the mid-points of the 4 cyan and 4 magenta ink spreading curves of the Flower image when calibrated using 10 image calibration tiles. The white bars correspond to the *standard mid-points*, i.e. the mid-points obtained using standard image calibration tiles. The hatched bars correspond to *optimized mid-points*, i.e. the mid-points obtained using optimized image calibration tiles. The black bars in-between correspond to the mid-points of the classical calibration and are used as *reference mid-points*.

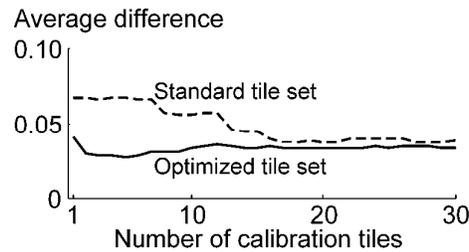


Fig. 7. Average mid-point surface coverage difference between the 20 reference mid-points of the classical calibration and either the 20 standard (dashed line) or optimized (solid line) mid-points. The differences are shown as a function of the number of tiles in the respective image tile sets. The image tiles are extracted from the Flower image.

The difference between  $c$  and  $m$  mid-points of the standard and reference ink spreading curve calibrations is large, whereas the difference between the corresponding optimized and reference mid-points is small. This can be explained by looking at the actual image tiles used for the calibrations (Figure B3, Appendix B). In the first ten tiles of the standard calibration set of the Flower image, there are no pure cyan or magenta tiles, whereas we have tiles 5 and 7 in the optimized calibration set. In the  $c/y$  and  $m/y$  cases, both the standard and optimized mid-points are well calibrated because there are both red and green tiles in the respective calibration sets. There is however no dark or blue tile in either calibration sets, therefore the  $c/m$ ,  $c/my$  and  $m/cy$  mid-points are not well calibrated. In general, thanks to the diversity of the optimized calibration tiles, the optimized mid-points are close to the reference mid-points. Moreover, the fact that they are always closer to the reference mid-points compared with the standard mid-points shows the benefit of the optimized tile selection strategy.

As mentioned in Section 9, when more calibration tiles are added, the accuracy of the standard calibrations improves and becomes comparable with the accuracy of the optimized calibrations. This is also true for the difference between the reference and standard mid-points and the difference between the reference and optimized mid-points.

As shown in Fig. 7 for the Flower image, the difference between the reference and optimized mid-points quickly decreases and then remains stable. However, the difference between the reference and standard mid-points remains large when using a small number of calibration tiles and slowly decreases to achieve a similar average difference for 17 tiles or more. In this regard, Fig. 7 is very similar to Fig. 5 and this similarity is valid for the other 5 test images as well. There is therefore a strong correlation between the prediction accuracy and the

difference between calibrated and reference mid-points. Fig. 7 also shows that the same prediction accuracy can be obtained with differently calibrated ink spreading curves. There is indeed an average difference of 0.04 remaining between the reference and either the optimized or the standard mid-points for 20 image calibration tiles or more.

## 11. CONCLUSION

Spectral reflection prediction models are able to accurately predict the reflection spectrum of any color-constant color 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.

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 considerably 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 well-chosen 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 calibration image tiles is lower than the number of color-constant patches required for a classical calibration.

The correlation between the sum of the ink spreading weights and the stabilization of the image calibration accuracy has only been superficially explored in this paper. Further exploiting this correlation may lead to an improved control over the calibration procedure. Future work should also consider the use of non-uniform image tiles for the calibration procedure.

## ACKNOWLEDGEMENT

We thank the Swiss National Science Foundation for their support, grant n° 200020-126757/1.

## REFERENCES

- [1] J.A.C. Yule, W.J. Nielsen, "The penetration of light into paper and its effect on halftone reproductions," Proc. TAGA, Vol. 3, 1951, pp. 65-76.
- [2] J.A.S. Viggiano, "Modeling the color of multi-colored halftones," Proc. TAGA, vol. 42, 1990, pp. 44-62.
- [3] R. Balasubramanian, "Optimization of the spectral Neugebauer model for printer characterization," Journal of Electronic Imaging, Vol. 8, No. 2, 1999, pp. 156-166.
- [4] A. U. Agar and J. P. Allebach, "An Iterative Cellular YNSN Method for Color Printer Calibration," Proc. of the 6th IS&T/SID Color Imaging Conference, Scottsdale AZ, 1998, pp. 197-200.
- [5] R. D. Hersch, F. Crété, "Improving the Yule-Nielsen modified spectral Neugebauer model by dot surface coverages depending on the ink superposition conditions," Color Imaging X: Processing, Hardcopy, and Applications, Proc. SPIE Vol. 5667, 2005, pp. 434-445.
- [6] R. D. Hersch, P. Emmel, F. Collaud, F. Crété, "Spectral reflection and dot surface prediction models for color halftone prints," Journal of Electronic Imaging, Vol. 14, No. 3, paper 033001, 2005, pp. 1-12.
- [7] Bugnon, Brichon, Hersch, "Simplified Ink Spreading Equations for CMYK Halftone Prints", Color Imaging XIII: Processing, Hardcopy, and Applications, SPIE Vol. 6807, 2008, paper 680717, pp. 1-12.
- [8] R.D. Hersch, M. Brichon, T. Bugnon, M. Hébert, "Deducing ink spreading curves from reflection spectra acquired within printed color images," Journal of Imaging Science and Technology, Vol. 53, No.3, 2009, paper 030502, pp. 1-7.
- [9] Th. Bugnon, R. D. Hersch, "Constrained Acquisition of Ink Spreading Curves from Printed Color Images", IEEE Trans. Image Process., Vol. 20, No. 2, 2011, pp 513-522.
- [10] H.E.J. Neugebauer, "Die theoretischen Grundlagen des Mehrfarbendrucks", Zeitschrift fuer wissenschaftliche Photographie, Vol. 36, 1937, pp. 36-73, translated by D. Wyble

and A. Kraushaar in “The theoretical basis of multicolor letterpress printing,” *Color Res. Appl.*, Vol. 30, 2005, pp. 323-331.

[11] J.A.C. Yule, R. Colt, “Colorimetric investigation in multicolor printing,” *Proc. TAGA*, Vol. 3, 1951, pp. 77-82.

[12] Th. Bugnon, M. Brichon and R.D. Hersch, “Model-Based Deduction of CMYK Surface Coverages from Visible and Infrared Spectral Measurements of Halftone Prints,” *Color Imaging XII: Processing, Hardcopy, and Applications*, Proc. SPIE Vol. 6493, 2007, paper 649310, pp. 1-10.

[13] [http://en.wikipedia.org/wiki/Weibull\\_distribution](http://en.wikipedia.org/wiki/Weibull_distribution)

[14] <http://r0k.us/graphics/kodak/>

[15] <http://sipi.usc.edu/database/database.cgi?volume=misc>

[16] <http://www.hlevkin.com/TestImages/additional.htm>

**Thomas Bugnon** received his M.S degree in Communication Systems from the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, in 2005. He joined the group of Prof. R. D. Hersch at the Peripheral Systems Lab in 2005 as a research assistant. After having worked on anti-counterfeiting techniques, he has been a PhD candidate since 2007. His research focuses on spectral prediction models for color reproduction.

**Roger David Hersch** received the engineering and PhD degrees from ETH Zurich in 1975 and from EPFL in 1985, respectively. He is a professor of computer science and head of the Peripheral Systems Laboratory at the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. He has published more than 100 scientific papers, is the editor of several books, and is inventor or coinventor in many patent applications. He is interested in novel imaging techniques related to color prediction, color reproduction, artistic imaging, and security printing. He is a fellow of IS&T, the Society for Imaging Science and Technology.

## APPENDIX A: 95 PERCENTILE ACCURACY

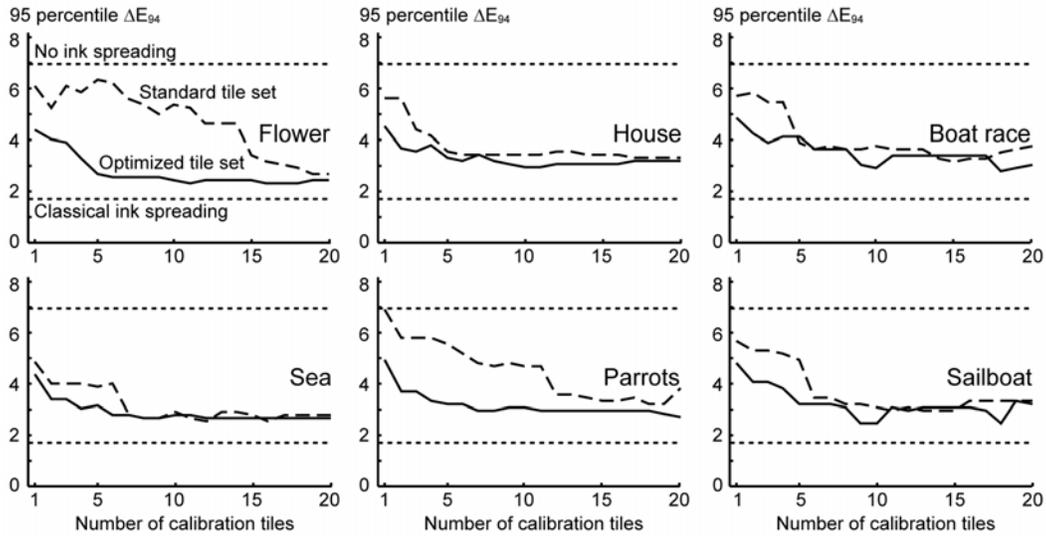


Fig. A1. The 95 percentile accuracy of differently calibrated IS-YNSN models when predicting the uniform patches of the 125 CMY set. The following calibrations for six different CMY images are shown: no ink spreading (upper dotted lines), classical ink spreading calibration (lower dotted line), ink spreading calibrations using standard image tile sets (dashed line), and ink spreading calibrations using optimized image tile sets (solid line).

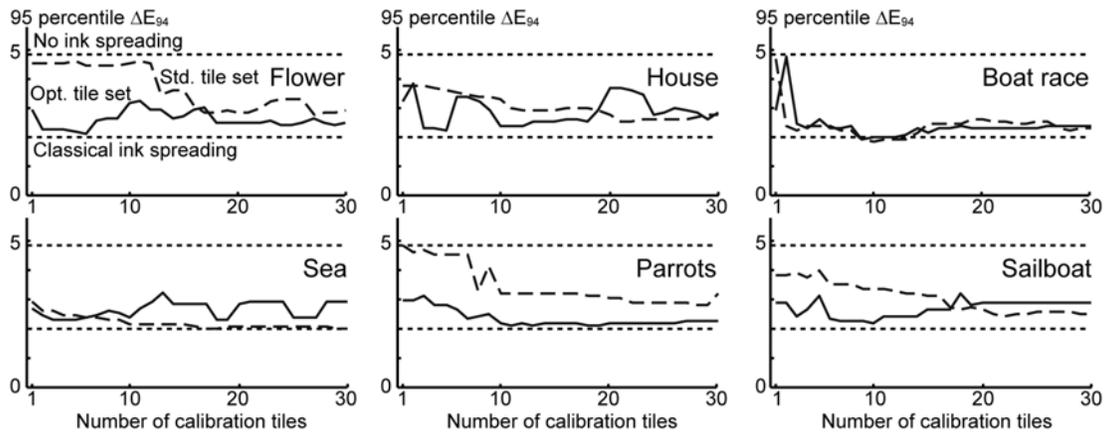


Fig. A2. The 95 percentile accuracy of differently calibrated IS-YNSN models when predicting the CMYK uniform patches of the 125 UCR set. The following calibrations for six different CMYK images are shown: no ink spreading (upper dotted lines), classical ink spreading calibration (lower dotted line), ink spreading calibrations using standard image tile sets (dashed line), and ink spreading calibrations using optimized image tile sets (solid line).

## APPENDIX B: TEST IMAGES AND IMAGE TILE SETS

N°1: Flower



N°3: Boat race



N°5: Parrots



N°2: House



N°4: Sea



N°6: Sailboat



Fig. B1. Test images (scale 1:5).

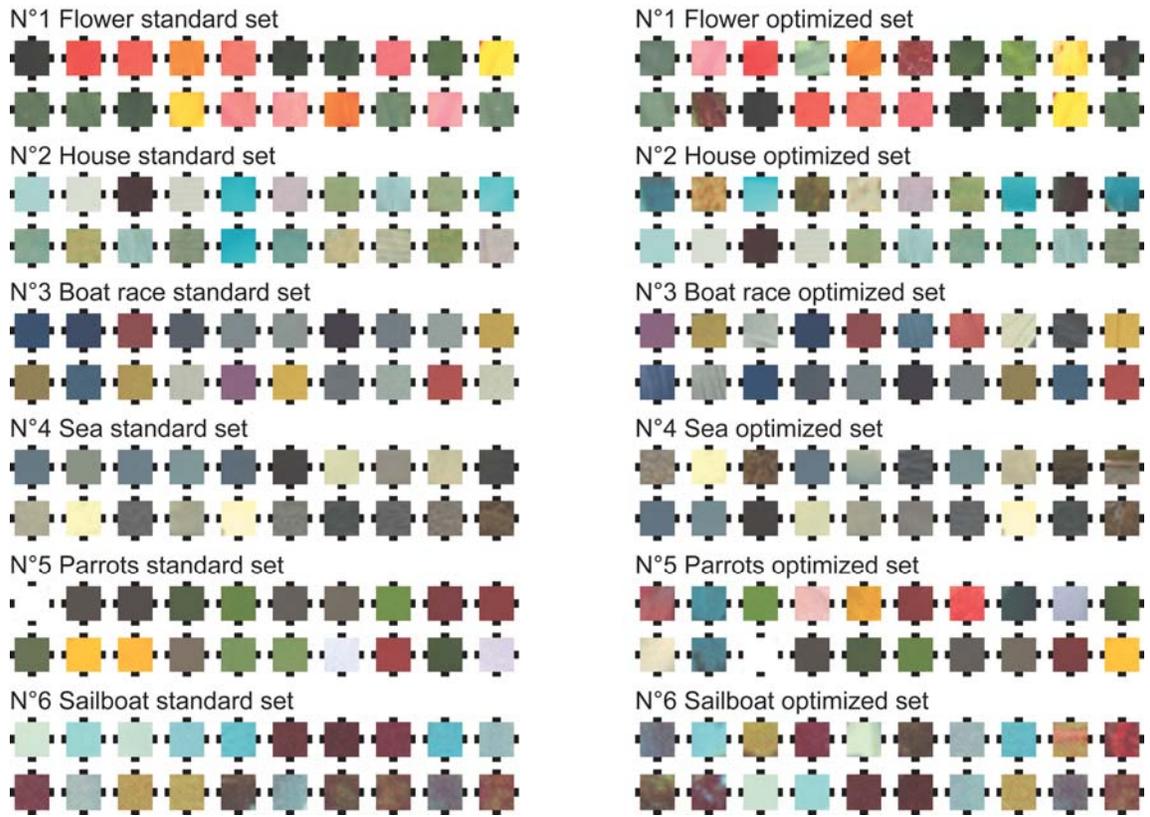


Fig. B2. CMY standard and optimized image calibration tile sets.

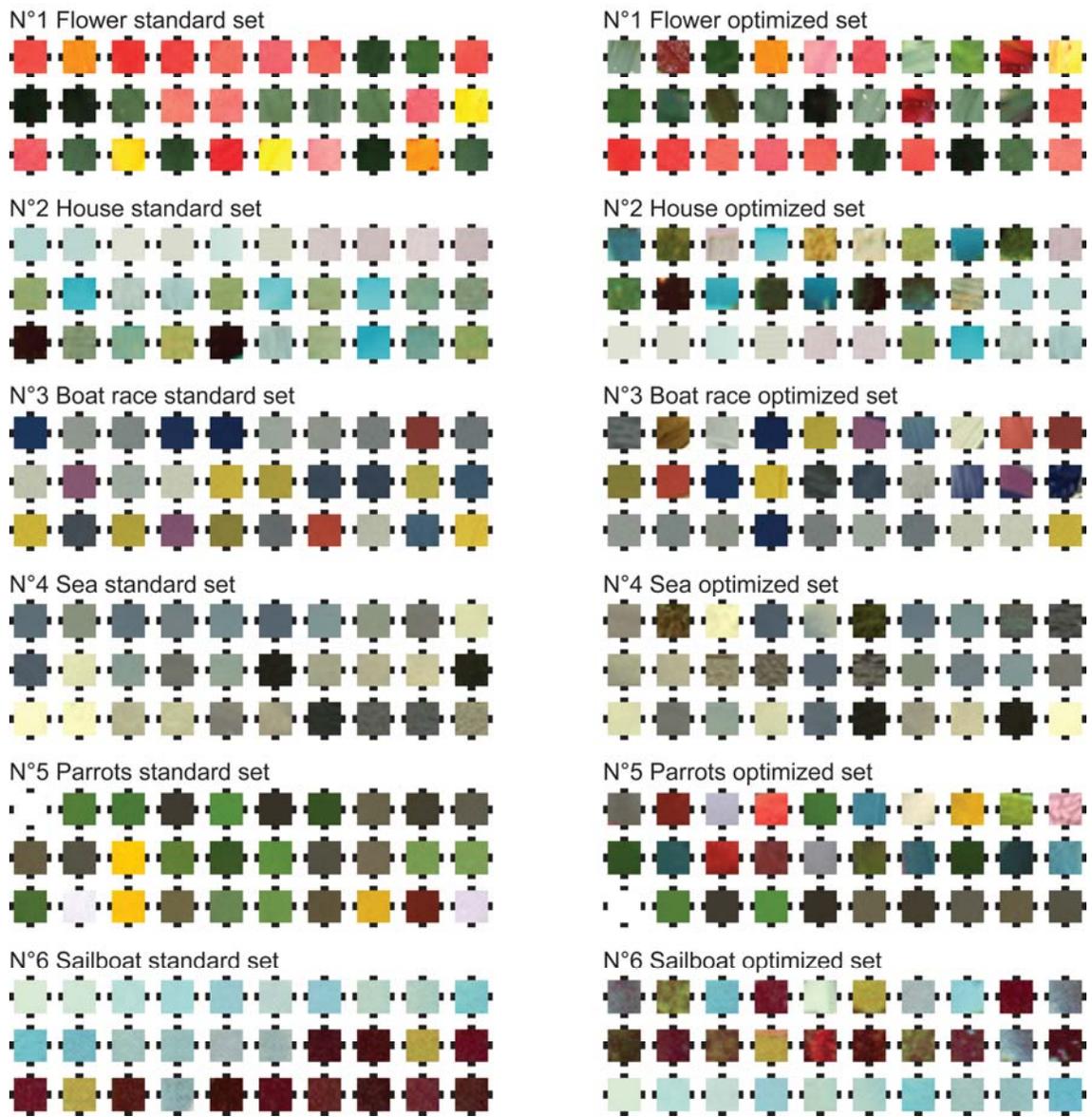


Fig. B3. CMYK standard and optimized image calibration tile sets.

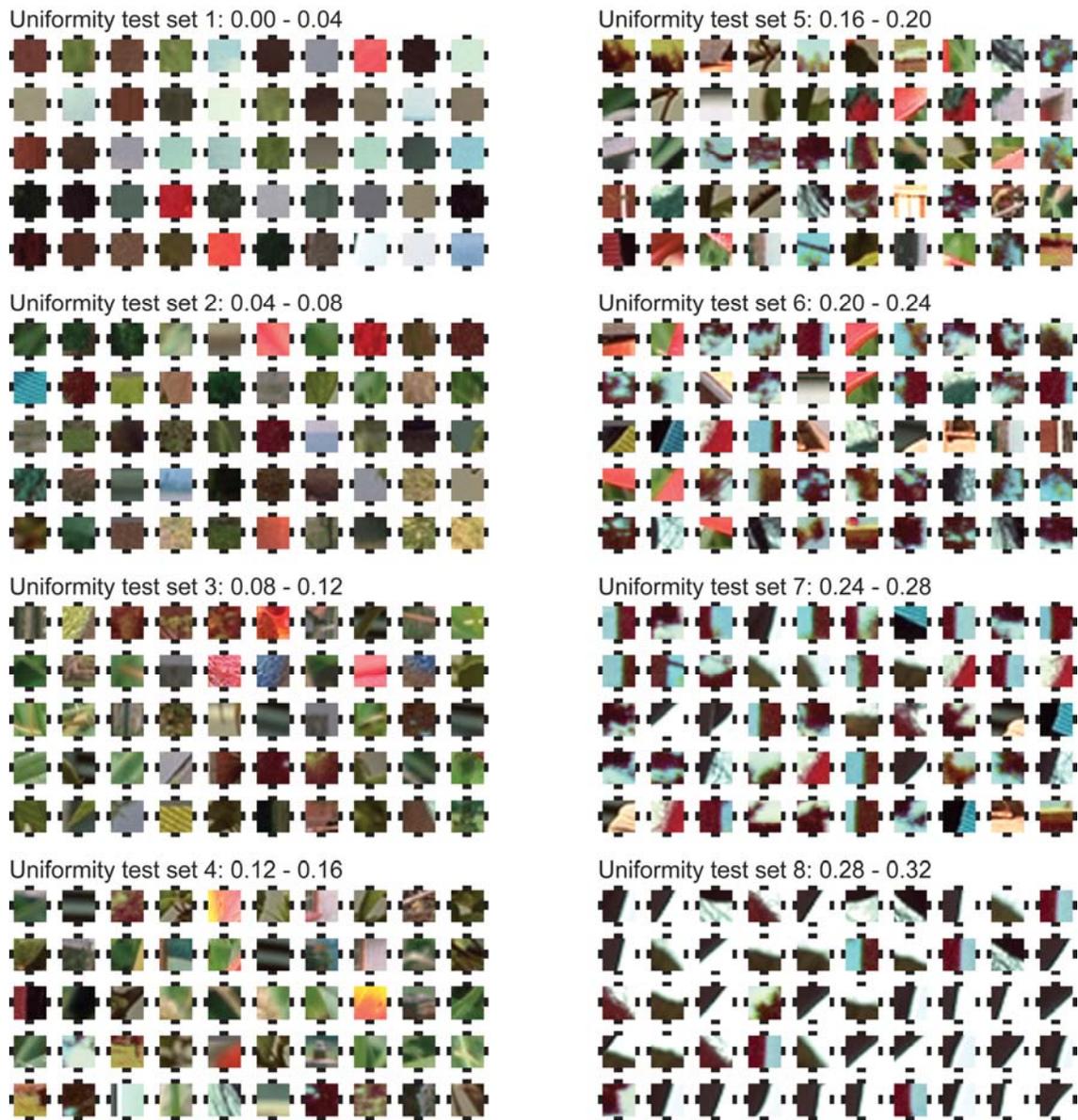


Fig. B4. CMYK standard and optimized image calibration tile sets.

## LIST OF FIGURES

Fig. 1. (a) Example of an ink spreading curve characterized by its mid-point  $v$ . (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).

Fig. 2. Adjustment function  $z(x)$  inspired by the Weibull distribution.

Fig. 3. 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.

Fig. 4. Accuracy of different calibrated IS-YNSN models when predicting the uniform 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 (solid line).

Fig. 5. Accuracy of different calibrated IS-YNSN models when predicting the CMYK uniform 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 (solid line).

Fig. 6. Mid-points of the four cyan (left) and four magenta (right) ink spreading curves of the Flower image for the classical calibration (black bars), the ink spreading calibration using a standard tile set containing 10 image tiles (white bars), and the ink spreading calibration using an optimized tile set containing 10 image tiles (hatched bars).

Fig. 7. Average mid-point surface coverage difference between the 20 reference mid-points of the classical calibration and either the 20 standard (dashed line) or optimized (solid line) mid-points. The differences are shown as a function of the number of tiles in the respective image tile sets. The image tiles are extracted from the Flower image.

Fig. A1. The 95 percentile accuracy of differently calibrated IS-YNSN models when predicting the uniform 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 (solid line).

Fig. A2. The 95 percentile accuracy of differently calibrated IS-YNSN models when predicting the CMYK uniform 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 (solid line).

Fig. B1. Test images (scale 1:5).

Fig. B2. CMY standard and optimized image calibration tile sets.

Fig. B3. CMYK standard and optimized image calibration tile sets.

Fig. B4. CMYK standard and optimized image calibration tile sets.

## LIST OF TABLES

Table I. List of the considered ink spreading curve indicia.

Table II. Prediction accuracy of the classically calibrated IS-YNSN model when predicting the uniform 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  $\Delta E_{94}$ .

Table III. Average accuracy of predicting sets of tiles grouped according to their non-uniformity values using effective coverages predicted with the classically calibrated ink spreading model (*Predicted* column) and accuracy obtained by direct fitting of individual tile surface coverages according to Eq. (3) (*Deduced* column).  $\Delta C_{ovs}$  indicates the average CMYK norm between predicted and deduced effective coverages.

Table IV. 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  $\Delta E_{94}$ .

Table V. 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  $\Delta E_{94}$ .