



ELSEVIER

Contents lists available at ScienceDirect

## Pattern Recognition

journal homepage: [www.elsevier.com/locate/pr](http://www.elsevier.com/locate/pr)

# An analysis of the transition proportion for binarization in handwritten historical documents



Marte A. Ramírez-Ortegón<sup>a,b,\*</sup>, Lilia L. Ramírez-Ramírez<sup>e</sup>, Volker Märgner<sup>b</sup>,  
Ines Ben Messaoud<sup>c,b</sup>, Erik Cuevas<sup>f</sup>, Raúl Rojas<sup>d</sup>

<sup>a</sup> División Académica de Ciencias Básicas, UJAT, Apartado Postal 24, 86690 Cunduacán, Mexico

<sup>b</sup> Institut für Nachrichtentechnik, Technische Universität Braunschweig, Schleinitzstraße 22, 38106 Braunschweig, Germany

<sup>c</sup> Laboratoire des Systèmes et Traitement de Signal, LTS Ecole Nationale d'Ingénieurs de Tunis, ENIT Tunis, Tunisia

<sup>d</sup> Institut für Informatik, Freie Universität Berlin, Takustr. 7, 14195 Berlin, Germany

<sup>e</sup> Instituto Tecnológico Autónomo de México (ITAM), Rio Hondo No 1. Col. Progreso Tizapán, D.F., Mexico

<sup>f</sup> Departamento de Ciencias Computacionales, Universidad de Guadalajara, Av. Revolución, 1500 Guadalajara, Mexico

## ARTICLE INFO

### Article history:

Received 13 September 2012

Received in revised form

27 January 2014

Accepted 7 February 2014

Available online 18 February 2014

### Keywords:

Historical documents

Threshold

Thresholding

Binarization

Transition method

Transition pixel

Normal

Lognormal

Minimum error rate

Otsu

Bayes' theory

## ABSTRACT

In this paper, we will present a mathematical analysis of the transition proportion for the *normal threshold* (NorT) based on the transition method. The transition proportion is a parameter of NorT which plays an important role in the theoretical development of NorT. We will study the mathematical forms of the quadratic equation from which NorT is computed. Through this analysis, we will describe how the transition proportion affects NorT. Then, we will prove that NorT is robust to inaccurate estimations of the transition proportion. Furthermore, our analysis extends to thresholding methods that rely on Bayes rule, and it also gives the mathematical bases for potential applications of the transition proportion as a feature to estimate stroke width and detect regions of interest. In the majority of our experiments, we used a database composed of small images that were extracted from DIBCO 2009 and H-DIBCO 2010 benchmarks. However, we also report evaluations using the original (H-)DIBCO's benchmarks.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

Binarization is the process of identifying pixels that contain relevant information for the specific application. This set of pixels is named foreground and its complement set is the background. For the purposes of this paper, foreground pixels are those pixels that constitute ink strokes.

Applications like optical character recognition [1,2], thinning [3,4], document layout analysis [5,6], text segmentation [7,8], and writer identification [9,10], rely on features that are extracted either from foreground pixels at a binary level, or from information (gray

intensities, color, textures) subjected to foreground pixels. Hence, noisy binary inputs tend to systematically propagate noise throughout the whole system and, consequently, binarization is a crucial task for document analysis.

The binarization of historical documents is a difficult task since such documents frequently have complex structures and heavy degradation. The former refers to the nature of the text and background themselves, such as faded strokes, stroke width variability, different types and sizes of fonts, irregular angle text orientations, and background printing patterns. The latter refers to interfering patterns that are a product of artifact or document aging, such as irregular illumination, blurring, bleed-through, ink stains, smudged characters, and paper fold outlines.

Due to their complexity, many research groups currently work on binarization methods for historical documents [11–15]. Some of these methods are listed in Table 1 along with the types of features that the methods exploit. Our primary sources to construct this table were Elsevier and Springer digital libraries,

\* Corresponding author at: División Académica de Ciencias Básicas, UJAT, Apartado Postal 24, 86690 Cunduacán, Mexico.

E-mail addresses: [mars.sasha@gmail.com](mailto:mars.sasha@gmail.com) (M.A. Ramírez-Ortegón), [lilialeticia.ramirez@itam.mx](mailto:lilialeticia.ramirez@itam.mx) (L.L. Ramírez-Ramírez), [maergner@ifn.ing.tu-bs.de](mailto:maergner@ifn.ing.tu-bs.de) (V. Märgner), [ibmnoussa@gmail.com](mailto:ibmnoussa@gmail.com) (I. Ben Messaoud), [erik.cuevas@cucei.udg.mx](mailto:erik.cuevas@cucei.udg.mx) (E. Cuevas), [rojas@inf.fu-berlin.de](mailto:rojas@inf.fu-berlin.de) (R. Rojas).

**Table 1**  
Binarization methods that specialize in, or that have been applied to historical documents. The column experimental refers to methods whose authors reported experimental studies where the method's performance was evaluated by varying certain parameter(s). The columns for parameters refer to (a) Manual: The parameters are set as constants, or by training data. (b) Auto.: The parameters are in function of a variable, or are estimated from the data. The columns for features refer to: (a) Stat.: mean, variance, minimum, etc. (b) Dist.: analysis of the distribution of features. (c) Struc.: information such as stroke width, line height, and connected component size. (d) Surf.: estimation of the fore or background surface. (e) Cont.: measures based on contrast. (f) Edge: features subjected to, or based on edge pixels.

Method	Year	Expe.	Parameters			Features					
			Manual	Auto.	total	Stat.	Dist.	Struc.	Surf.	Cont.	Edge
Lelore and Bouchara [16] <sup>a</sup>	2013	•	9	9	18	•	•	•			•
Valizadeh and Kabir [17] <sup>a</sup>	2013		2	6	8	•				•	•
Howe [18]	2012	•	7	10	17	•	•	•		•	•
Valizadeh and Kabir [19]	2012	•	3	2	5		•	•	•		
Rivest-Hénault et al. [20]	2012		18	9	27			•		•	
Shi et al. [21] <sup>a</sup>	2012		3	1	4		•			•	
Moghaddam and Cheriet [22] <sup>a</sup>	2012		2	8	10	•	•	•		•	
Hedjam et al. [23]	2011		7	2	9	•		•			
Bataineh et al. [24] <sup>a</sup>	2011		8	2	10	•					
Ben Messaoud et al. [25] <sup>a</sup>	2011		3	0	3	•	•		•		
Su et al. [26] <sup>a</sup>	2010		0	4	4	•		•	•		•
Lu et al. [27] <sup>a</sup>	2010	•	7	3	10	•		•	•	•	•
Ramírez-Ortegón et al. [28]	2010		6	10	16	•	•			•	•
Moghaddam and Cheriet [29] <sup>a</sup>	2010		3	6	9	•	•	•			
Ntirogiannis et al. [30] <sup>a</sup>	2009	•	6	1	7	•		•		•	•
Gatos et al. [31]	2008		5	2	7			•			•
Mello and Schuler [32]	2008		15	0	15	•	•				
Mello et al. [33]	2008		4	0	4		•				
Gupta et al. (MROtsu) [34]	2007		2	2	4		•	•			
Gupta et al. (MarginED) [34]	2007		3	2	5	•					
Gatos et al. [35]	2006		9	6	15	•			•	•	
Kavallieratou and Stathis [36]	2006	•	4	0	4	•					

<sup>a</sup> Indicates that some features and parameters of auxiliary methods are not considered in the table.

although selected papers from other libraries and conference proceedings were also included.

From Table 1, we can see that these binarization methods involve a great deal of parameters. Readers may notice that some authors of these methods claim that their methods have none or few parameters. However, these authors neglect values that are taken as constants, or that have been computed in an unsupervised manner. In [22], for instance, the authors proposed a threshold that is a generalization of Otsu's method. They claim that it is a parameterless threshold, but their threshold equation involves a parameter, namely  $k_{\sigma}$ , which was empirically set to 1.6 in all their experiments.

According to Table 1, few authors report formal experimental analyses on how the method's parameters affect its performance, and none of them report formal mathematical analyses of the method's parameters. The lack of parameter analyses is a serious problem not only to select suitable parameters' values for different types of images, but also to ensure the system's robustness. For example, Sauvola and Pietikäinen [37] proposed a threshold for a variety of degraded documents (none of them historical). They empirically found that the optimal value of one of its crucial parameters, namely  $k$ , was 0.5. On the other hand, the same threshold is computed as a step in Gatos et al.'s method [35], but Gatos et al. claimed that  $k=0.2$  was a more suitable value for historical documents after some tests. In both papers, the optimal  $k$ 's values were empirically selected by a human expert and, as a consequence, we cannot infer if the optimal  $k$ 's value depends on the type of image (historical or non-historical for instance), or on the neighborhood size (since Sauvola's threshold is locally computed), or on some other features.

The importance of the parameter selection motivates us to present a parameter analysis of the transition method for binarization [38] (reviewed in Section 3), which is part of our ongoing

project of digitization of historical documents. The novelty of our contribution is the manner we tackled the analysis of a single parameter, namely transition proportion. For this aim, we examine the transition proportion to calculate the parameter intrinsic bias (Section 4.1), formulate analytical inequalities (Section 4.2), observe empirical distributions (Section 4.3), determine some value's convergences (Table 3), and compute the bias under theoretical distributions (Section 4.4). We believe that this paper is a reference for parameter analyses in binarization methods even when our analysis is focused on our method's parameter.

The transition proportion is deeply involved in the theoretical formulation of thresholding methods based on the transition method and the minimum error rate (Bayes' rule), as shown in Section 3.3. In the Bayesian approach, the product of the a priori class probability and the class-conditional probability determines the pixel classification in favor of the likeliest class; for binarization this is either foreground or background. Following the Bayesian approach, the transition method estimates the a priori class probability with the transition proportion. Hence, our mathematical analysis extends to other methods that also rely on Bayes' rule. For instance, Hedjam's method [23] assumes that the a priori probabilities of observing both fore and background pixels are equal within the pixel neighborhoods. This assumption is not fulfilled in most of the pixels' neighborhoods but, in spite of this simplification, the performance of Hedjam's method is comparable to state-of-the-art binarization methods; see Table 7. Our current analysis partially justifies why a simplification such as Hedjam's method does not crucially affect the method's performance.

In Section 4.2, we will describe some mathematical connections between the transition proportion and its performance. We will show that the a priori class probability has a small impact on the threshold selection as long as the gray intensities' means and variances are accurately estimated.

Later on, we will describe the statistical distribution of the transition proportion (Section 4.3). Subsequently, we will report two objective performance evaluations of the transition method (Section 5): the former evaluation shows that the transition sets within the region of interest can be accurately estimated by standard transition operators. The latter evaluation compares different variants of the transition method with the state-of-the-art methods.

In this paper, the method's implementation is an improved version of our method in [28] and it is the version submitted to DIBCO 2013 [15]. Our method estimates the transition proportion automatically for each pixel; previously, it was a constant for all pixels. Furthermore, it detects regions of interest by considering a lower bound of the transition proportion; previously it was a predefined number of pixels. We also added pre- and post-processes to reduce the pixel misclassification due to bleed-through and bi-level backgrounds. However, the details of the extra pre- and post-process are given in the paper's supplementary material since they are not the main goal of this paper and some of the processes are in preparation for publication.

## 2. Image dataset

Throughout this paper, we will support our theory with empirical data computed from a database composed of 245 images. We extracted these images from 15 handwritten images: 5 handwritten test images plus 2 handwritten examples of DIBCO 2009, and 8 of 10 images of H-DIBCO 2010. Two images of H-DIBCO 2010 (H05 and H07) were excluded from the analysis since both images have overwhelming artifacts due to compression artifact rather than to aging or printing degradation. These extracted images are small regions of interest usually containing a single word. Examples of these images are shown in Fig. 1.

From now on, we will refer to these images as our database, or our dataset. However, in Section 5.2, we used the original (H-)DIBCO's benchmark from 2009 to 2013 [12–15,39].

## 3. Overview of the transition method

Since our parameter analysis is derived from the transition method, we will devote this section to review the method's main steps. In particular, we will review the thresholds based on transition sets and define our parameter of interest, namely the transition proportion.

### 3.1. Notation

We denote pixels in bold while we denote the gray intensity of a pixel  $\mathbf{p}$  as  $I(\mathbf{p}) \in \mathbb{N}$ , where black is set to zero, and the color white is set to  $g$ . The image of gray intensities is denoted by  $I$ .

We denote the fore and background sets by  $\mathcal{F}$  and  $\mathcal{B}$ , respectively, such that  $\mathcal{P} = \mathcal{F} \cup \mathcal{B}$ . The neighborhood of a pixel  $\mathbf{p}$ , denoted by  $\mathcal{P}_r(\mathbf{p})$ , are those pixels within a square centered at the pixel  $\mathbf{p}$ , of sides with length  $2r+1$ . Moreover, given a set of pixels  $\mathcal{A}$ , we will write  $\mathcal{A}_r(\mathbf{p})$  as shorthand for  $\mathcal{A} \cap \mathcal{P}_r(\mathbf{p})$ . For instance,  $\mathcal{B}_r(\mathbf{p}) = \mathcal{B} \cap \mathcal{P}_r(\mathbf{p})$  denotes the pixels in the background within a neighborhood of radius  $r$ , around the pixel  $\mathbf{p}$ .

In the following sections, we will adopt the thresholding approach as

$$B(\mathbf{p}) = \begin{cases} 1 \text{ (foreground)} & \text{if } I(\mathbf{p}) \leq T(\mathbf{p}), \\ 0 \text{ (background)} & \text{otherwise,} \end{cases} \quad (1)$$

where  $B$  denotes the binary image of  $I$ , and  $T(\mathbf{p})$  is the threshold calculated for  $\mathbf{p}$ .

### 3.2. The transition method

The transition method was proposed in [38] and subsequently improved in [28,40]. It assumes that the distribution of both fore and background gray intensities can be approximated by the distribution of gray intensities particularly of those pixels within their corresponding extended contour, namely  $t$ -transition pixels; see Fig. 2.

A  $t$ -transition pixel is a pixel that contains both fore and background pixels in its neighborhood of radius  $t$ . The set of  $t$ -transition pixels  ${}_t\mathcal{P}$  can be divided into two disjoint subsets:  ${}_t\mathcal{F} = {}_t\mathcal{P} \cap \mathcal{F}$  and  ${}_t\mathcal{B} = {}_t\mathcal{P} \cap \mathcal{B}$ .

In the transition method, only information from  $t$ -transition pixels is used to compute the threshold surface such that  ${}_t\mathcal{F}$  and  ${}_t\mathcal{B}$  are considered as representative samples of  $\mathcal{F}$  and  $\mathcal{B}$ , respectively. This method is divided into six steps:

1. Compute the transition values for each pixel by

$$V(\mathbf{p}) = \max_{\mathbf{q} \in {}_t\mathcal{P}(\mathbf{p})} \{I(\mathbf{q})\} + \min_{\mathbf{q} \in {}_t\mathcal{P}(\mathbf{p})} \{I(\mathbf{q})\} - 2I(\mathbf{p}), \quad (2)$$

where  $t$  is small. In our experiments,  $t = 2$ .

2. We expect that (2) attains high values for positive transition pixels, high negative values for negative transition pixels, and values close to zero for non-transition pixels. Hence, the transition sets approximations are defined by

$${}_t\hat{\mathcal{F}} = \{\mathbf{p} | V(\mathbf{p}) \geq t_+\} \quad \text{and} \quad {}_t\hat{\mathcal{B}} = \{\mathbf{p} | V(\mathbf{p}) \leq -t_-\}, \quad (3)$$

where  $t_+$  and  $t_-$  are computed by the double linear threshold [28] which is a thresholding for unimodal histograms.

3. Enhance  ${}_t\hat{\mathcal{F}}$  and  ${}_t\hat{\mathcal{B}}$ ; see [28].
4. Calculate the region of interest:

$$\hat{\mathcal{R}} = \{\mathbf{p} \text{ such that } |{}_t\hat{\mathcal{F}}_r(\mathbf{p})| \geq n_+ \text{ and } |{}_t\hat{\mathcal{B}}_r(\mathbf{p})| \geq n_-\}, \quad (4)$$

where  $n_+$  and  $n_-$  are two positive integers:  ${}_t\hat{\mathcal{F}}_r(\mathbf{p}) = {}_t\hat{\mathcal{F}} \cap \mathcal{P}_r(\mathbf{p})$  and  ${}_t\hat{\mathcal{B}}_r(\mathbf{p}) = {}_t\hat{\mathcal{B}} \cap \mathcal{P}_r(\mathbf{p})$ .

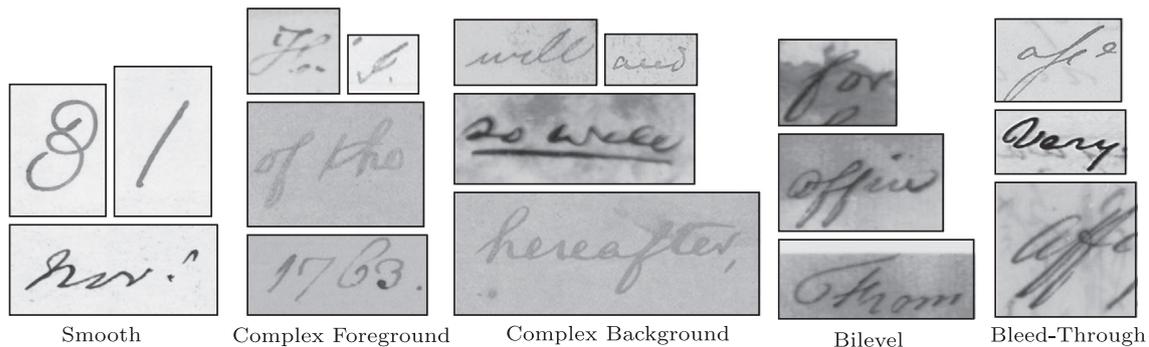
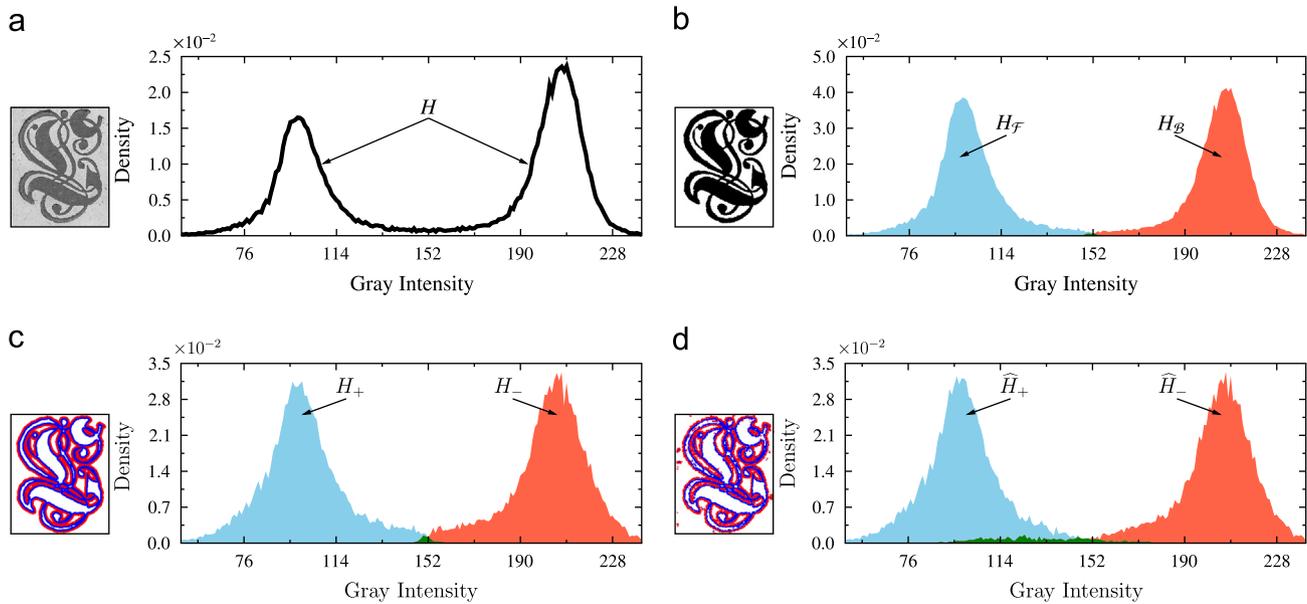


Fig. 1. Examples of our database.



**Fig. 2.** (a) Original image and its histogram of gray intensities. (b) Binary groundtruth and its histograms of both fore and background gray intensities. (c) Positive and negative transition sets and their histograms of gray intensities. (d) Approximation of the positive and negative transition sets and their histograms of gray intensities. Transition pixels are represented in two colors: in blue, positive transition pixels and in red, negative transition pixels. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

5. Label  $\mathbf{p}$  as background if  $\mathbf{p} \notin \hat{\mathcal{R}}$ . Otherwise, compute a local threshold  $T(\mathbf{p})$  from the information from the transition sets; see Section 3.3.
6. Restore  $\hat{\mathcal{F}}$  and  $\hat{\mathcal{B}}$  with standard algorithms. For example, removing small connected components.

### 3.3. Binarization by transition sets

In the following subsections, we assume that the thresholds are computed on regions of interest where both fore and background are not empty.

#### 3.3.1. Minimum-error-rate transition

In Bayesian decision theory, the probability of a pixel misclassification is minimized by Bayes' decision rule: classify  $\mathbf{p}$  as foreground if

$$\Pr(\mathbf{p} \in \mathcal{F}_r(\mathbf{p}) | I(\mathbf{p}) = i) \geq \Pr(\mathbf{p} \in \mathcal{B}_r(\mathbf{p}) | I(\mathbf{p}) = i). \quad (5)$$

Otherwise, classify  $\mathbf{p}$  as background.

Under Bayes' criteria, the probability of error in  $\mathcal{P}_r(\mathbf{p})$  is given by

$$Er_{min} \approx \frac{1}{|\mathcal{P}_r(\mathbf{p})|} \sum_{i=0}^g \min\{H_f(i), H_b(i)\}, \quad (6)$$

where  $H_f(i)$  and  $H_b(i)$  denote the occurrence of fore and background pixels with gray intensity  $i$  that are within  $\mathcal{F}_r(\mathbf{p})$  and  $\mathcal{B}_r(\mathbf{p})$ , respectively.

If the class-conditional densities of  $\mathcal{B}_r(\mathbf{p})$  and  $\mathcal{F}_r(\mathbf{p})$  are unknown,  $\hat{t}_{opt}$  cannot be computed. However, the transition method assumes that the intersection of such densities is approximately the same intersection between the gray-intensity densities of the positive and negative transition sets.

The minimum-error-rate threshold based on transition sets, named MER transition, is defined as

$$\hat{t}_{opt} \approx \arg \min_{t \in [0, g]} \{Er^*(t)\} \quad \text{such that} \quad (7)$$

$$Er^*(t) = [1 - w_+] \sum_{i=0}^t \frac{H_-(i)}{|{}_t\mathcal{B}_r(\mathbf{p})|} + w_+ \sum_{i=t+1}^g \frac{H_+(i)}{|{}_t\mathcal{F}_r(\mathbf{p})|} \quad (8)$$

where  $H_+(i)$  and  $H_-(i)$  denote the occurrence of positive and negative transition pixels with gray intensity  $i$  that are within  ${}_t\mathcal{F}_r(\mathbf{p})$  and  ${}_t\mathcal{B}_r(\mathbf{p})$ , respectively, and

$$w_+ = \frac{|{}_t\mathcal{F}_r(\mathbf{p})|}{|{}_t\mathcal{P}_r(\mathbf{p})|} \quad (9)$$

is the transition proportion in  ${}_t\mathcal{P}_r(\mathbf{p})$ .

#### 3.3.2. Normal transition

The normal threshold based on transition sets (normal transition) assumes that the transition pixels' gray intensities are approximately normally distributed within a neighborhood of radius  $r$ . Thus, the optimal threshold is the intersection between gray-intensity distributions of both the positive and negative transition sets.

Generally, the threshold is the root  $\mu_+ < \hat{t} < \mu_-$  of a quadratic equation with coefficients  $a \neq 0$ ,  $b$ , and  $c$  given by

$$a = \frac{1}{\sigma_+^2} - \frac{1}{\sigma_-^2}, \quad (10)$$

$$b = \frac{2\mu_-}{\sigma_-^2} - \frac{2\mu_+}{\sigma_+^2}, \quad \text{and} \quad (11)$$

$$c = \frac{\mu_+^2}{\sigma_+^2} - \frac{\mu_-^2}{\sigma_-^2} - 2 \ln \left( \frac{\sigma_- \cdot w_+}{\sigma_+ \cdot w_-} \right), \quad (12)$$

where  $\mu_+$  and  $\sigma_+^2$  denote the mean and variance of gray intensities in  ${}_t\mathcal{F}_r(\mathbf{p})$  and  $\mu_-$  and  $\sigma_-^2$  denote the mean and variance of gray intensities in  ${}_t\mathcal{B}_r(\mathbf{p})$ , and  $w_- = 1 - w_+$ .

In [38] a special case was omitted when  $\sigma_+ = \sigma_- = \sigma > 0$ , which implies that  $a=0$ . Thus, the threshold is given by

$$\hat{t} = \frac{\mu_+ + \mu_-}{2} - \frac{\sigma^2 \cdot \ln \left( \frac{w_-}{w_+} \right)}{\mu_- - \mu_+}. \quad (13)$$

To avoid numerical problems with the term  $a$ , we also computed (13) if  $|\sigma_+ - \sigma_-| < 1$ . In this manner, we also avoided numerical problems that arose when  $\sigma_+ \approx \sigma_-$ .

Another problem is caused when some estimated terms of (10) are considerably inaccurate. Namely when no root of (10) may be between  $\mu_+$  and  $\mu_-$ . In such cases, we computed the autolinear threshold [38] defined as

$$\hat{t} = \mu_+ + \frac{\sigma_+}{\sigma_+ + \sigma_-}(\mu_- - \mu_+) \tag{14}$$

ensuring a threshold between  $\mu_+$  and  $\mu_-$ .

### 3.3.3. Lognormal transition

The lognormal threshold based on transition sets (lognormal transition) is derived in a similar manner to the normal transition. It is computed by  $\exp(\hat{t})$ , where  $\hat{t}$  is the root of the quadratic equation with coefficients given by (10), but replaces  $\mu_+$  and  $\sigma_+^2$  with  $\tilde{\mu}_+$  and  $\tilde{\sigma}_+^2$ , where

$$\tilde{\sigma}_+^2 = \ln\left(1 + \frac{\sigma_+^2}{\mu_+^2}\right) \quad \text{and} \tag{15}$$

$$\tilde{\mu}_+ = \ln(\mu_+) - \frac{1}{2}\tilde{\sigma}_+^2. \tag{16}$$

Note that  $\tilde{\sigma}_+$  is used to compute  $\tilde{\mu}_+$ . Likewise,  $\tilde{\mu}_-$  and  $\tilde{\sigma}_-$  are estimated.

## 4. Transition proportion

As we mentioned in Section 1, our main goal with this paper is to analyze the transition proportion's role on binarization thresholds. In [28,38,41], the transition proportion was assumed to be a constant, namely 0.5. Without formal study, however, the positive outcomes in their evaluations could be attributed to the particular properties of the tested images. Hence, this section shows that the normal threshold based on transition sets is indeed robust to rough estimates of the transition proportion in the context of handwritten historical documents. In addition, we will give some suitable values for  $w_+$  depending on the pixel's neighborhood size and its type of strokes.

### 4.1. Analysis of bias

Before we study the effects of the transition proportion on the transition method, we will first study the transition method's bias when the distributions of transition sets are known. Then, we can determine how the transition proportion impacts the method's performance.

We conducted an experiment to study its bias: for the  $i$ th image and  $j$ th thresholding method, the threshold bias is defined as

$$\text{Bias}_{ij} = \widehat{Er}_{ij} - \text{MER}_i \tag{17}$$

where  $\text{MER}_i$  is the minimum error rate in the  $i$ th (globally computed) and  $\widehat{Er}_{ij}$  is the error rate of the  $j$ th thresholding method for the  $i$ th image.

Our results, summarized in Fig. 3 and Table 2, confirm that the bias of thresholds based on transition sets is marginal compared with the minimum error rate (lower than 0.2%). Therefore, this experiment ensures that any bias observed in the following sections is a result of an inaccurate transition proportion and of the thresholds' intrinsic bias.

### 4.2. Mathematical form of the transition proportion

The transition method can attain different estimated transition sets in such a manner that the means and variances of both approximations marginally differ from each other, but their positive transition proportions are significantly different. For example, Fig. 4 shows two approximations of the transition set with different transition proportions (Fig. 4(c) and (d)); nevertheless both approximations lead to approximately the same means.

To understand the transition proportion's influence in the normal transition, we must observe that the transition proportion only appears in the constant term of (10) and, as a consequence, the transition proportion's role in the threshold can be analyzed within the symmetry of the quadratic equation

$$F(x; w) = a \cdot x^2 + b \cdot x + \underbrace{\frac{\mu_+^2}{\sigma_+^2} - \frac{\mu_-^2}{\sigma_-^2}}_h - 2 \ln\left(\frac{\sigma_-}{\sigma_+}\right) - 2 \ln\left(\frac{w}{1-w}\right). \tag{18}$$

Let us assume that  $a < 0$ , which implies  $\sigma_+ > \sigma_-$  and  $b > 0$ , then  $F(x; w)$  has a vertical symmetry axis in  $-b/2a$  for all  $w$  as Fig. 5(a) shows. Therefore, the difference  $d$  between the threshold when  $w = w_+$  and the threshold when  $w = 0.5$  is

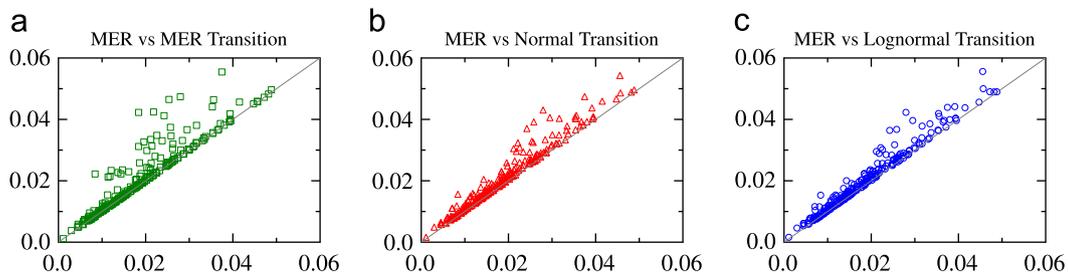
$$d = \sqrt{b^2 - 4ah - 4ak} - \sqrt{b^2 - 4ah}. \tag{19}$$

The value  $d$  is upper bounded by

$$d \leq 2\sqrt{-a}\sqrt{k} \leq 2\sqrt{\frac{\sigma_+^2 - \sigma_-^2}{\sigma_+^2 \cdot \sigma_-^2}}\sqrt{k} < \frac{2}{\sigma_+ \cdot \sigma_-}\sqrt{k} \tag{20}$$

**Table 2**  
Bias of thresholds based on transition sets from 245 images.

Quartile	Transition			
	MER	MER	Normal	Lognormal
MER approx.				
<b>First</b> $\times 10^{-2}$	0	0.52	3.82	4.64
<b>Second</b> $\times 10^{-2}$	1.02	3.49	9.14	9.66
<b>Third</b> $\times 10^{-2}$	4.08	12.51	19.74	19.04
<b>Maximum</b> $\times 10^{-2}$	21.43	238.90	147.97	143.29
Mean by quartiles				
<b>First</b> $\times 10^{-2}$	0.00	0.06	1.42	1.86
<b>Second</b> $\times 10^{-2}$	0.29	1.78	6.19	6.85
<b>Third</b> $\times 10^{-2}$	2.38	6.40	13.46	14.13
<b>Fourth</b> $\times 10^{-2}$	8.36	69.41	49.72	43.84
<b>Overall mean</b> $\times 10^{-2}$	2.77	19.58	17.78	16.73



**Fig. 3.** Pairwise comparison of the minimum error rate and the error of thresholds based on transition sets.

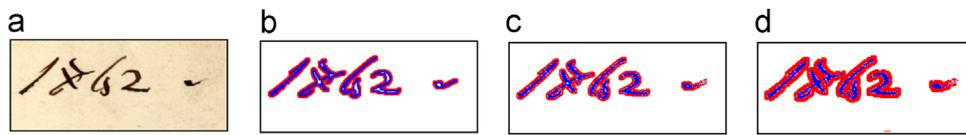


Fig. 4. (a) Original image, (b) transition set, and (c and d) two different approximations of the transition set with two different transition proportions: 0.343 and 0.238, respectively.

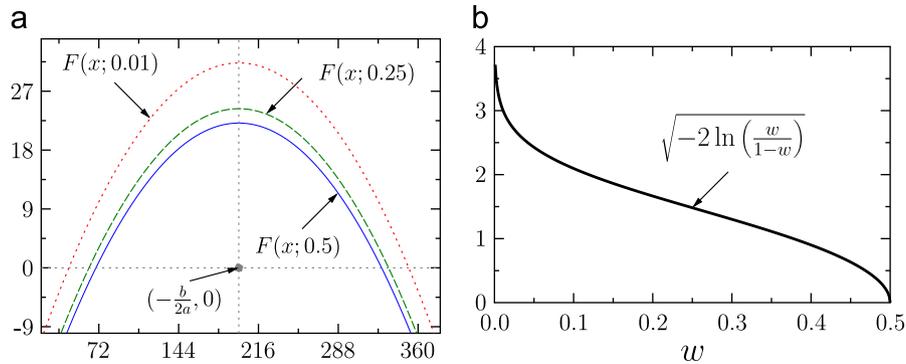


Fig. 5. (a) Three graphs of  $F(x; w)$ . (b) Graph of  $k$ 's values.

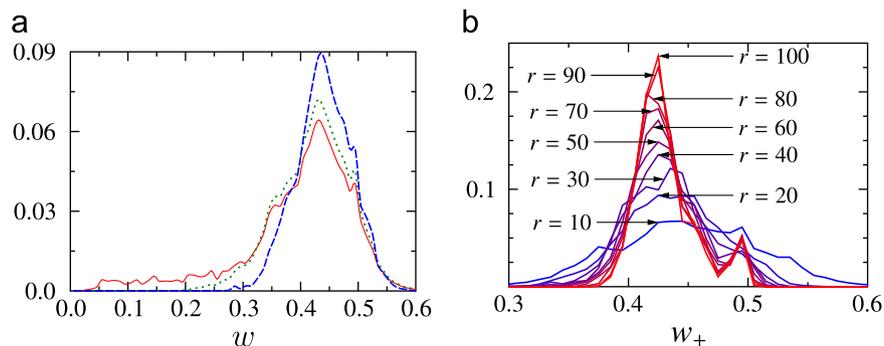


Fig. 6. (a) Different shapes of the transition-proportion histogram are computed from neighborhood with radius  $r=15$ . The solid line depicts the histogram of those pixels whose neighborhoods contain at least one positive and one negative transition pixel. The dotted line depicts the histogram of those pixels whose neighborhoods contain at least  $n_+$  positive and  $n_-$  negative transition pixels, where  $n_+ = n_- = r+1 = 16$ . The dashed line depicts the histogram of those pixels within the transition set. (b) Histograms of transition proportion are computed from transition pixels and with different neighborhood sizes.

since  $\sqrt{u+v} \leq \sqrt{u} + \sqrt{v}$  for  $u, v > 0$ . Furthermore, we know experimentally that  $\sigma_+, \sigma_- \geq 1$  when the gray intensities range between 0 and 255 (artifacts may produce outliers). Therefore,  $d < 2\sqrt{k}$ .

The magnitudes of  $\sqrt{k}$  are computed by sweeping the domain of  $w$  as Fig. 5(b) shows, where  $0 \leq \sqrt{k} < 3.04$  for  $w \in [0.01, 0.5]$ . Then,  $d < 6.08$ , which represents less than 2.4% of the number of gray levels ( $g=255$ ).

Notice that the upper bound for  $d$  does not depend on the specific definition of the transition method, in fact, it depends only on the a priori probability of the pixel class. Therefore, this upper bound is fulfilled for those methods that rely on Bayes' rule using the mixture of two normal distributions.

### 4.3. Distribution and convergence of the transition proportion

In the previous section, we determined the upper bound  $d$  of the difference between thresholds with different transition proportions; see (20). To refine this upper bound, we studied the theoretical and empirical distribution of  $w$  in this section.

The shape of the transition-proportion histogram depends on two factors:

(1) The proximity of background pixels to foreground pixels: if a background pixel is far from the foreground pixels, then the

positive transition set tends to be under sampled since the pixel neighborhood barely overlaps the foreground. As a result, the transition proportion tends to be small. To avoid this problem, the transition method selects those pixels whose neighborhoods have at least  $n_+$  positive and  $n_-$  negative transition pixels; see Fig. 6(a).

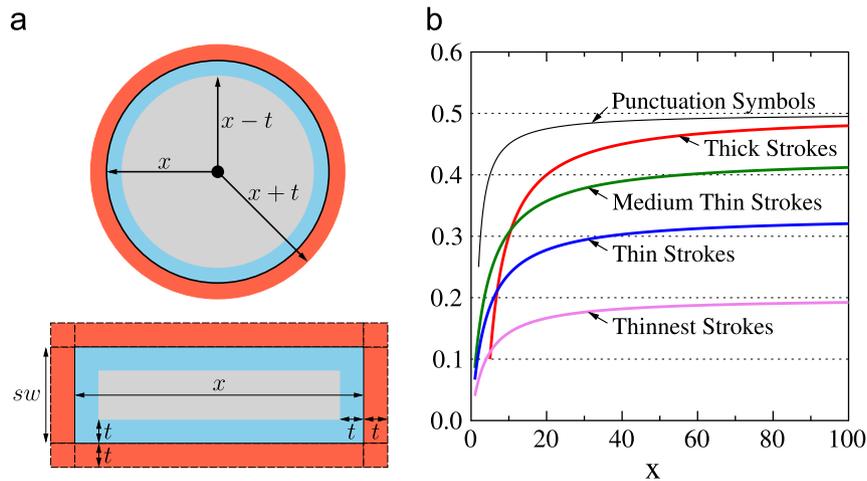
(2) The pixel neighborhoods' size from which the transition proportion is locally computed: the histogram converges to  $|_t\mathcal{F}|/|_t\mathcal{P}|$  as the pixel neighborhood size converges to the image size; see Fig. 6(b).

To explain the shapes of the transition-proportion histograms, we studied the theoretical limits of circles and lines. The circles represent punctuation symbols while the lines represent ink strokes.

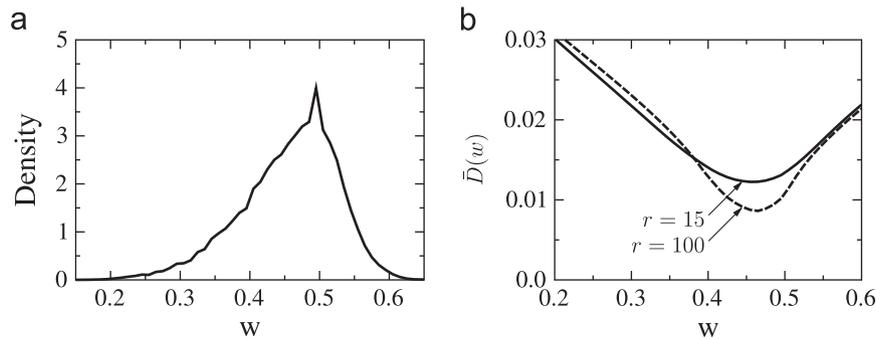
The equations in Table 3, derived from Fig. 7(a), establish that the convergence value of the transition proportion depends on the stroke width, denoted by  $sw$ , the pixel neighborhood size ( $x=2r+1$ ), and the size  $t$  of the transition neighborhood. Fig. 8(a) shows that the average distribution of 17 handwritten images ranges between 0.10 and 0.65 and clusters between 0.41 and 0.5, for  $t=2$  and  $x \geq 31$ . This suggests that  $sw \geq 3$  for the majority of the strokes as can be visually verified. From now on, we will assume that  $t=2$  for the rest of this paper.

**Table 3**  
Transition proportion of geometrical shapes.

	Punctuations circles ( $x \geq t$ )	Thick strokes line ( $x \geq sw \geq 2t$ )	Medium strokes line ( $x \geq sw = 3$ )	Thin strokes line ( $x \geq sw = 2$ )	Thinnest strokes line ( $x \geq sw = 1$ )
For small $t$ 's					
<b>Function</b>	$\frac{1}{2} - \frac{t}{4x}$	$\frac{1}{2 + \frac{4t}{(x+sw-2t)}}$	$\frac{3}{3+2t + \frac{6t+4t^2}{x}}$	$\frac{2}{2+2t + \frac{4t+4t^2}{x}}$	$\frac{1}{1+2t + \frac{2t+4t^2}{x}}$
<b>Converge to</b>	$\frac{1}{2}$ if $x \rightarrow \infty$	$\frac{1}{2}$ if $x+sw \rightarrow \infty$	$\frac{3}{3+2t}$ if $x \rightarrow \infty$	$\frac{1}{1+t}$ if $x \rightarrow \infty$	$\frac{1}{1+2t}$ if $x \rightarrow \infty$
For $t=2$					
<b>Function</b>	$\frac{1}{2} - \frac{1}{2x}$	$\frac{1}{2 + \frac{8}{x+sw-4}}$	$\frac{3}{7 + \frac{28}{x}}$	$\frac{1}{3 + \frac{12}{x}}$	$\frac{1}{5 + \frac{20}{x}}$
<b>Converge to</b>	$\frac{1}{2}$ if $x \rightarrow \infty$	$\frac{1}{2}$ if $x+sw \rightarrow \infty$	$\frac{3}{7}$ if $x \rightarrow \infty$	$\frac{1}{3}$ if $x \rightarrow \infty$	$\frac{1}{5}$ if $x \rightarrow \infty$



**Fig. 7.** (a) Geometrical shapes and their transition sets are shown: a circle of radius  $x$  (top) and a rectangle of sides  $x \times sw$  (bottom). The transition region has a neighborhood of radius  $t$ ; the positive transition set is shaded in blue while the negative transition set is shaded in red. (b) Functions of transition proportion for different geometrical shapes for  $t=2$  are shown. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)



**Fig. 8.** (a) Average distribution of the transition proportion is shown. This histogram is computed as the average of the histograms of 7 handwritten images from DIBCO 2010 and 10 images from H-DIBCO 2010. (b) Graphs of  $\bar{D}(w)$  assuming the empirical distributions of the transition proportion in (a) are depicted.

4.4. Bias of the normal transition

Since we have determined experimentally and theoretically that  $w_+ \in [0.1, 0.6]$  for  $r \geq 15$ , then we can determine by (20) that the difference  $d$  between the normal thresholds computed with  $w = w_+$  and the normal threshold computed with  $w = 0.5$  satisfies

$$d < 2\sqrt{-2 \ln \frac{0.1}{0.9}} < 4.20. \tag{21}$$

This represents less than 1.73% of the number of gray-intensity levels. Furthermore, the transition proportion within regions of

interest varies from 0.2 to 0.6 as Fig. 8(a) shows. Therefore,

$$d < 2\sqrt{-2 \ln \frac{0.2}{0.8}} < 3.34. \tag{22}$$

This represents less than 1.31% of the number of gray-intensity levels.

Notice that by choosing a different constant for  $w$ , we obtained a different bound. In fact, the maximum difference  $d$  between the normal thresholds computed with  $w_+$  and the normal threshold

computed with a value  $w$  is lower than

$$d < D(w_+, w) = \sqrt{\left| -2 \ln\left(\frac{w_+}{1-w_+}\right) + 2 \ln\left(\frac{w}{1-w}\right) \right|}. \quad (23)$$

Thus, assuming that  $w_+ \in [0.1, 0.6]$ , the inequality in (23) is minimized for  $w_* \approx 0.29$  such that  $D(w_*, w) = 3.23$  for  $w \in [0.1, 0.6]$ . Table 4 reports several values of  $D(w_*, w)$ .

The election of a constant for  $w$  is not straightforward because it depends on the distribution of  $w_+$  rather than on the maximum difference ( $D(w_+, w)$ ). Taking into account the distribution of  $w_+$ , we analyzed the average difference between the actual normal threshold and the estimated threshold and computed the optimal constant  $w_*$  that minimizes such an average. In the following two subsections, we will analyze the transition proportion assuming for discrete and continuous distributions.

#### 4.4.1. Discrete distribution

Let us assume that  $w_+$  follows a discrete distribution. So, the average difference for a constant  $w$  is bounded by

$$\bar{D}(w) < 2 \sum_{i=1}^n \Pr(W_+ = w_i) D(w_i, w), \quad (24)$$

where  $\Pr(W_+ = w_i)$  denotes the probability of the transition proportion of being equal to  $w_i$ .

Although we do not know the distribution of the stroke widths ( $sw$ ) in DIBCO 2009 and H-DIBCO 2010, we know, by visual inspection, that the majority of strokes fulfill  $sw \geq 3$ . In fact, the dominant  $sw$  appears to be four or more pixels. Nevertheless, let us assume a discrete uniform distribution for  $sw = \{ 3, 4+ \}$ , where

$4+$  denotes four or more pixels. Hence,  $\bar{D}(w)$  is minimized at  $w_* = 3/7$  and at  $w_* = 1/2$ .

When the pixel neighborhood is not large enough, as in our dataset, the transition proportion have to be computed according to Table 3. In particular, for our dataset, we can underestimate the transition proportion by assuming that the characters are longer than 40 pixels. With this restriction, two minimums are reached at  $w_* = 41/105$  and at  $w_* = 41/90$ ; see Table 4.

#### 4.4.2. Continuous distribution

Similar to (24), we can derive  $w_*$  when  $w_+$  follows a continuous distribution in  $[w_l, w_u]$ . In such a case, the average difference is bounded by

$$\bar{D}(w) < 2 \int_{w_l}^{w_u} \Pr(W_+ = x) D(x, w) dx. \quad (25)$$

For example, Fig. 8(a) shows the empirical distribution of  $w_+$  for DIBCO 2009 and H-DIBCO 2010 images in neighborhoods of radius  $r = 15$ . Following this distribution, we found that  $\bar{D}(w)$  is minimized at  $w_* = 41/90$ . Furthermore, if the radius changes to  $r = 100$ , then  $w_* = 0.465$ ; see Fig. 8(b).

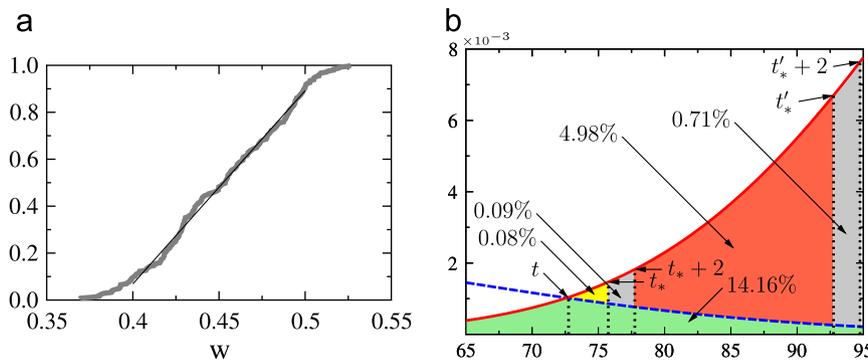
For our database of 245 small images, Fig. 9(a) shows that the empirical cumulative distribution of  $w_+$  is approximately uniformly distributed between 0.4 and 0.5, with 87.35% of the population. Therefore, we can assume that  $w_+ \sim U(0.4, 0.5)$  and, consequently,  $w_* = 0.4502$ ; see Table 4.

#### 4.4.3. Standardized error rate

At this point, we have shown that replacing  $w_+$  for  $w \in [0.1, 0.6]$  yields approximately the same thresholds ( $\pm 4$ ). Moreover, the difference between the actual optimal threshold

**Table 4**  
Summary of values for  $w$ . The column of empirical distribution refers to the empirical distribution in Fig. 8(a). DU stands for discrete uniform distribution.

$w$	$\bar{E}r(w) \times 10^{-2}$	$D(w_+, w)$ $w_+ \in [0.1, 0.6]$	$\bar{D}(w)$		$w_+ \sim U(0.4, 0.5)$	Empirical distribution
			$sw \sim DU\{3, 4+\}$			
			$r \rightarrow \infty$	$r = 20$		
0.5	2.5538	4.1926	0.7589	1.5408	0.1197	0.0135
41/90	2.3285	4.0190	1.0656	0.7448	0.0853	0.0123
0.4502	2.2497	3.9974	1.0520	0.9086	0.0849	0.0123
3/7	2.2365	3.9085	0.7589	1.0293	0.0910	0.0129
0.4	2.2129	3.7861	1.3853	0.9564	0.1207	0.0141
41/105	2.2156	3.7437	1.5044	0.7448	0.1350	0.0144
0.2899	2.8135	3.2266	2.4412	2.1473	0.2351	0.0226
$w_+$	2.4165	-	-	-	-	-



**Fig. 9.** (a) Empirical distributions of the transition proportion for our database of 245 images are shown. (b) Error rates for different thresholds are shown. In green, the minimum error rate that corresponds to  $t$  is shown. In yellow the error increments from  $t_{mer}$  to  $t_*$ . In gray (left), the error increments from  $t_*$  to  $t_* + 2$ . In red, the error increments from  $t_* + 2$  to  $t'_*$  and in gray (right), the error increment from  $t'_*$  to  $t'_* + 2$ . (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

and the threshold computed with a constant transition proportion is less than 0.03 gray levels in the average. See the empirical distribution column of Table 4. However, we should remark that the difference between the error rates of these thresholds also depends on  $\mu_+$ ,  $\mu_-$ ,  $\sigma_+^2$ , and  $\sigma_-^2$ .

To exemplify our above argument, observe Fig. 9(b), where the minimum error rate, achieved at  $t$ , is around 14.17%. For  $t_*$  and  $t_*+2$ , the error rate increases 0.08% and 0.17%, respectively. However, for  $t'_*$  and  $t'_*+2$ , the difference between their corresponding error rates is 0.72%: nine-fold the increase from 0.08% to 0.17%.

Fig. 9(b) also shows that the increase of the error rate from  $t_*$  to  $t_*+2$  depends on the difference between  $\mu_+$  and  $\mu_-$  and the magnitudes of  $\sigma_+^2$  and  $\sigma_-^2$ . Then, we empirically determined how the transition proportion affects the error rate when  $\mu_+$ ,  $\mu_-$ ,  $\sigma_+^2$ , and  $\sigma_-^2$  are accurately estimated.

In order to analyze the error rate, let us define the standardized error rate, which maps the error rate as

$$\bar{Er}(i, w) = \frac{Er(i, w) - Er_{mer}(i)}{Er_{white}(i) - Er_{mer}(i)} \quad (26)$$

where  $Er_{mer}(i)$  is the minimum error rate of the  $i$ th image,  $Er_{white}(i)$  is the error rate that corresponds to a white image (all pixels are classified as background) in the  $i$ th image, and  $Er(i, w)$  is the error rate that corresponds to the normal threshold such that the transition proportion is replaced by  $w$ .

Fluctuations of  $\bar{Er}(i, w)$  are related to the method's performance rather than related to fluctuations of the minimum error rates in the dataset. Hence, the standardized error rate is a more suitable measure than the error rate measure for thresholding methods since its mean is adjusted to the fluctuations of the minimum error rate. Because of this, we defined our error function as

$$\bar{Er}(w) = \frac{1}{n} \sum \bar{Er}(i, w), \quad (27)$$

where  $n$  is the number of images in our database.

The measure  $\bar{Er}(w)$  has mixed results; see Fig. 10 and Table 4. On one hand,  $\bar{Er}(0.4) = 2.21\%$  is smaller than  $\bar{Er}(w_+) = 2.42\%$  and smaller than  $\bar{Er}(0.5) = 2.55\%$ ; see Fig. 10(a). On the other hand, the conditional pairwise comparison indicates a triangle ranking, where  $w = w_+$  outperforms  $w = 0.5$ ,  $w = 0.5$  outperforms  $w = 0.4$ , and  $w = 0.4$  outperforms  $w = w_+$ ; see Fig. 10(b).

The mixed results in Fig. 10 suggest that the election of a constant for  $w$  between 0.3 and 0.5 has little influence on the error rate. The means are marginally different and their conditional pairwise comparisons are around 50%, which indicates similar performance.

### 5. Evaluations

As a complement to this research, we conducted two evaluations. The first evaluation assesses the transition method's efficiency

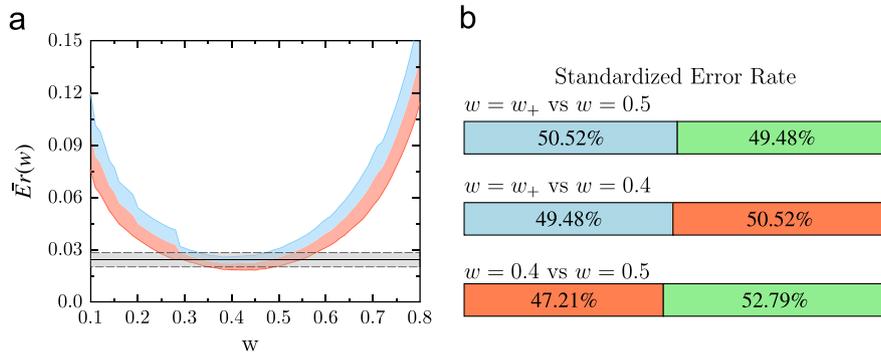


Fig. 10. (a) In the thick solid line, the graph of the standardized error rate is depicted. Its upper and lower bounds of the interval of confidence (95%) are shaded in blue and red, respectively. The gray rectangle is the interval of confidence (95%) for  $w = w_+$ . (b) Conditional pairwise comparison of standardized error rate are shown. The percent in the bar entitled  $w = w_+$  vs  $w = 0.5$  refers to  $\Pr(\bar{Er}(w_+) > \bar{Er}(0.5) | \bar{Er}(w_+) \neq \bar{Er}(0.5))$ ; while, the right percent refers to  $\Pr(\bar{Er}(w_+) < \bar{Er}(0.5) | \bar{Er}(w_+) \neq \bar{Er}(0.5))$ . The other bars are defined in a similar manner. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

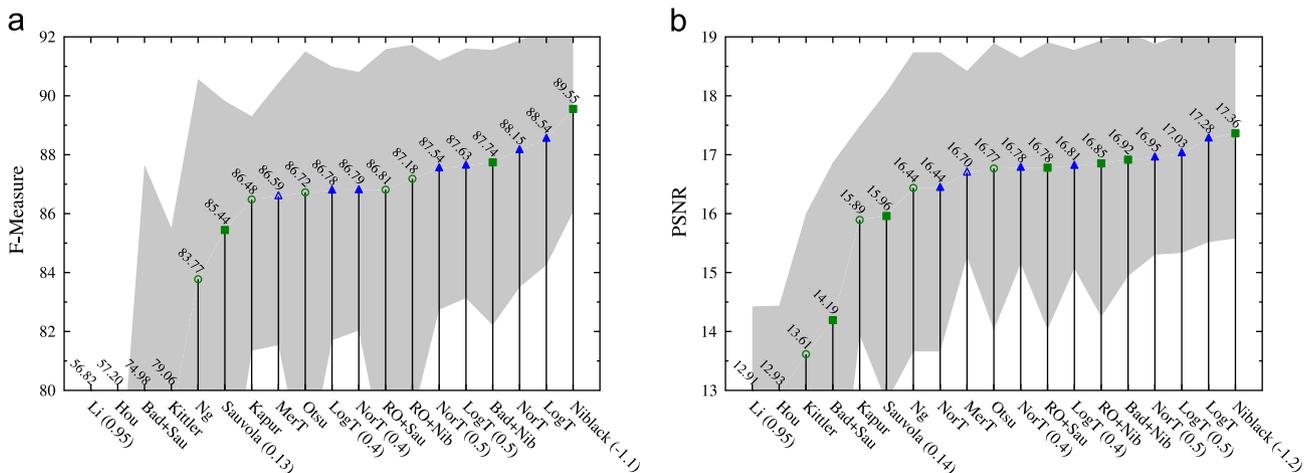


Fig. 11. Means of measurements (markers) are shown. In the shadows, the lower and upper means are depicted. Methods whose parameters are manually selected have filled markers. In blue, binarization algorithms based on the transition method are shown. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

in regions of interest. The second evaluation compares the transition method's performance with the state-of-the-art methods.

5.1. Performance measures

To evaluate the binarization method's performance, we computed the *F*-measure (FM), recall (RE), and precision (PR):

$$FM = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad \text{where} \quad (28)$$

$$RE = \frac{|F \cap \hat{F}|}{|\hat{F}|} \quad \text{and} \quad PR = \frac{|F \cap \hat{F}|}{|F|}, \quad (29)$$

where *F* and *B* are the fore and background, respectively; and their corresponding estimates are  $\hat{F}$  and  $\hat{B}$ . We also computed the peak signal-to-noise ratio (PSNR) which is a transformed measure

of the error rate (ER):

$$PSNR = -10 \log_{10}(ER) \quad \text{where} \quad (30)$$

$$ER = \frac{|F \cap \hat{B}(t)| + |B \cap \hat{F}(t)|}{|P|}. \quad (31)$$

We quantified the dispersion of a measure by the lower bound and upper bound mean defined, respectively, as

$$\mu_u = \sum_{x_i \geq \mu} (x_i - \mu) \quad \text{and} \quad (32)$$

$$\mu_l = \sum_{x_i \leq \mu} (\mu - x_i) \quad (33)$$

where  $x_i$  is the measurement of interest and  $\mu$  is the mean of the measurements.

**Table 5**  
Details of the measurements for our methods in (H-)DIBCO 2009–2011. RE, PR, and FM are given in percentages.

Image	LogT				ROI+LogT				ROI+LogT+BR				ROI+NorT+BR			
	RE	PR	FM	PSNR	RE	PR	FM	PSNR	RE	PR	FM	PSNR	RE	PR	FM	PSNR
<b>DIBCO 2009</b>																
H01	85.53	97.36	91.06	19.49	85.26	97.18	90.83	19.39	94.60	92.09	93.33	20.43	94.65	92.05	93.33	20.43
H02	92.85	33.73	49.48	13.87	92.86	90.21	91.51	24.29	94.03	88.77	91.32	24.13	94.89	86.48	90.49	23.65
H03	86.61	83.38	84.96	15.26	86.42	89.54	87.95	16.39	96.42	85.57	90.67	17.16	96.90	83.94	89.95	16.78
H04	85.73	80.33	82.94	15.87	86.79	74.99	80.46	15.10	93.78	89.20	91.44	18.90	93.98	88.21	91.00	18.66
H05	83.16	72.43	77.43	17.33	80.17	68.18	73.69	16.61	87.73	87.90	87.81	20.32	89.39	86.56	87.95	20.30
HW1 <sup>a</sup>	85.75	84.71	85.23	22.88	85.28	96.19	90.41	25.04	88.55	95.22	91.76	25.60	88.18	95.41	91.65	25.55
HW2 <sup>a</sup>	82.55	98.31	89.74	19.68	82.49	98.31	89.71	19.67	86.36	96.89	91.32	20.29	86.14	96.94	91.22	20.24
P01	88.98	90.62	89.79	16.13	88.94	93.43	91.13	16.80	94.01	90.91	92.44	17.31	94.24	90.31	92.23	17.18
P02	95.24	93.55	94.39	16.29	95.39	97.29	96.33	18.22	97.88	95.62	96.74	18.63	97.97	95.50	96.72	18.61
P03	93.93	97.06	95.47	18.17	33.22	98.49	49.68	9.39	96.66	98.72	97.68	21.05	96.58	98.78	97.66	21.03
P04	90.36	94.05	92.17	17.94	89.63	93.97	91.75	17.73	94.15	92.45	93.29	18.49	94.11	92.37	93.23	18.45
P05	79.40	88.93	83.89	13.51	80.59	90.94	85.46	13.97	86.65	89.68	88.14	14.67	87.59	88.46	88.02	14.57
PR1 <sup>a</sup>	92.33	74.21	82.29	18.25	92.37	98.15	95.17	24.53	94.67	97.50	96.06	25.34	94.54	97.55	96.02	25.31
PR2 <sup>a</sup>	94.94	98.14	96.51	18.79	95.01	98.80	96.87	19.27	98.81	96.11	97.44	20.00	98.78	96.14	97.44	20.00
Mean H <sup>b</sup>	86.78	73.45	77.18	16.37	86.30	84.02	84.89	18.36	93.31	88.70	90.91	20.19	93.96	87.45	90.54	19.96
Mean P <sup>b</sup>	89.58	92.84	91.14	16.41	77.55	94.82	82.87	15.22	93.87	93.48	93.66	18.03	94.10	93.08	93.57	17.97
<b>H-DIBCO 2010</b>																
H01	87.35	94.89	90.97	17.32	87.32	96.83	91.83	17.80	96.14	91.90	93.97	18.80	96.20	91.93	94.02	18.83
H02	92.85	93.90	93.37	22.26	91.96	95.69	93.79	22.60	96.72	90.26	93.38	22.09	96.67	90.34	93.40	22.10
H03	79.96	97.28	87.78	18.02	79.88	98.05	88.04	18.13	86.10	96.57	91.03	19.20	85.77	96.76	90.93	19.17
H04	83.55	95.16	88.98	17.64	83.48	95.06	88.90	17.60	90.36	89.81	90.08	17.81	90.28	89.87	90.08	17.81
H05	91.72	68.00	78.10	15.27	91.40	94.78	93.06	21.04	96.26	91.69	93.92	21.43	96.36	91.47	93.85	21.37
H06	73.57	94.96	82.90	17.16	72.99	94.81	82.48	17.07	82.41	92.00	86.94	18.05	82.25	92.17	86.93	18.05
H07	87.29	76.02	81.26	15.49	87.18	93.65	90.30	18.81	94.34	89.77	91.99	19.38	94.78	89.35	91.98	19.36
H08	75.62	95.48	84.40	16.56	75.37	95.69	84.32	16.55	87.38	91.30	89.30	17.81	87.43	91.15	89.25	17.79
H09	79.81	94.93	86.71	19.45	79.46	94.71	86.42	19.36	91.07	90.70	90.88	20.72	91.08	90.68	90.88	20.72
H10	76.56	96.23	85.28	17.96	74.00	96.37	83.72	17.59	79.01	91.20	84.67	17.61	79.32	90.99	84.76	17.62
Mean H	82.83	90.69	85.98	17.71	82.31	95.56	88.29	18.65	89.98	91.52	90.62	19.29	90.01	91.47	90.61	19.28
<b>DIBCO 2011</b>																
H01	96.07	63.42	76.40	11.24	83.87	94.17	88.72	15.68	98.21	62.69	76.53	11.17	98.19	63.17	76.88	11.26
H02	91.32	97.99	94.54	23.14	91.20	97.90	94.43	23.06	95.87	94.73	95.29	23.62	95.73	94.87	95.30	23.63
H03	77.72	98.78	86.99	17.56	76.19	98.95	86.09	17.31	85.54	96.21	90.56	18.71	85.35	96.24	90.47	18.68
H04	82.45	74.59	78.32	13.71	82.13	74.48	78.12	13.68	87.67	87.29	87.48	16.31	87.31	87.03	87.17	16.21
H05	91.34	88.33	89.81	16.33	91.25	88.22	89.71	16.29	94.93	85.17	89.79	16.15	94.86	84.68	89.48	16.01
H06	87.33	74.77	80.57	15.10	87.18	72.74	79.31	14.77	91.27	81.39	86.05	16.64	91.01	77.24	83.56	15.81
H07	86.04	60.85	71.28	15.45	83.86	91.60	87.56	20.09	86.48	90.99	88.68	20.42	85.83	92.12	88.87	20.53
H08	85.81	97.19	91.15	21.01	85.39	97.45	91.02	20.97	91.17	94.27	92.69	21.66	91.03	94.33	92.65	21.64
P01	85.07	74.25	79.29	11.26	85.32	98.45	91.42	15.69	92.40	97.70	94.98	17.84	92.51	97.54	94.96	17.82
P02	86.92	66.67	75.46	11.79	86.32	86.18	86.25	14.92	92.32	82.26	87.00	14.91	92.44	81.63	86.70	14.79
P03	93.03	90.17	91.58	15.01	93.27	92.95	93.11	15.94	95.07	91.34	93.17	15.90	95.06	91.23	93.11	15.86
P04	84.29	98.98	91.04	17.27	84.32	99.34	91.22	17.36	92.17	98.86	95.40	19.98	92.13	98.89	95.39	19.97
P05	83.22	77.91	80.48	12.54	83.23	90.78	86.84	14.58	90.93	92.36	91.64	16.40	91.21	91.41	91.31	16.21
P06	91.11	16.58	28.06	6.35	46.74	97.67	63.23	15.69	93.64	92.76	93.20	21.69	94.93	91.10	92.98	21.48
P07	92.62	05.22	09.89	3.80	92.42	75.59	83.16	20.34	95.30	86.27	90.56	23.09	95.36	86.08	90.48	23.05
P08	71.74	98.39	82.98	13.92	71.23	98.38	82.63	13.85	79.19	96.49	86.99	14.87	79.14	96.58	87.00	14.87
Mean H	87.26	81.99	83.63	16.69	85.13	89.44	86.87	17.73	91.39	86.59	88.38	18.08	91.17	86.21	88.05	17.97
Mean P	86.00	66.02	67.35	11.49	80.36	92.42	84.73	16.05	91.38	92.26	91.62	18.08	91.60	91.81	91.49	18.00

<sup>a</sup> Training images.

<sup>b</sup> These means do not consider the training images.

5.2. Evaluation I

Our first evaluation is designed to assess the performance of the algorithms in an atomic manner rather than in combination with other techniques. Hence, we compared our methods with several thresholding methods.

In this evaluation, all the methods were globally computed and no pre- or post-processing steps were done. In this manner, we assessed the quality of the threshold for different types of histograms.

We tested three variants of the transition method following the same steps but differing in gray-intensity threshold:

1. The minimum-error-rate transition, denoted by MerT.
2. The normal transition, denoted by NorT.
3. The lognormal transition, denoted by LogT.

When the transition proportion is replaced by a constant, it is indicated by ( $w = x$ ), where  $x$  is the constant. In these variants, the transition set was approximated as in Section 3.2. The implementation details can be found in this paper's supplementary material.

We compared the performance of our methods, with histogram-based methods that are related, recent, or top-ranked:

1. Otsu's threshold [42], top-ranked in [43,44].
2. Kapur's threshold [45], top-ranked in [46].

3. The minimum error thresholding (Kittler) [47], top-ranked in [48].
4. The minimum variance thresholding (Hou) [49], proposed in 2006.
5. The valley-emphasis method (Ng) [50], proposed in 2006.
6. Li's threshold [51], proposed in 2010.

We also selected statistical algorithms where only mean and standard deviation of gray intensities were computed:

1. Niblack's method [52], top ranked in [43].
2. Sauvola's method [37], top-ranked in [44,48,53].

Since both algorithms have parameters, we swept their parameter domain and reported the parameter sets that attained the highest performance for each measure. In addition to the manual parameter selection, we implemented two frameworks for automatic parameters selection: Badeskas and Papamarkos [54] and Ramírez-Ortegón et al. [44]. These frameworks were applied to Sauvola's and Niblack's and we reported the following combinations:

1. Badeskas and Papamarkos's framework applied to Niblack's threshold, denoted by BP+Nib.
2. Badeskas and Papamarkos's framework applied to Sauvola's threshold, denoted by BP+Sau.

**Table 6**  
Details of the measurements for our methods in (H-)DIBCO 2011–2013. RE, PR, and FM are given in percentages.

Image	LogT				ROI+LogT				ROI+LogT+BR				ROI+NorT+BR			
	RE	PR	FM	PSNR	RE	PR	FM	PSNR	RE	PR	FM	PSNR	RE	PR	FM	PSNR
H-DIBCO 2012																
H01	85.40	91.66	88.42	18.95	59.94	99.10	74.70	16.36	93.55	96.31	94.91	22.43	95.53	95.62	95.57	22.98
H02	78.63	93.92	85.60	16.49	53.64	97.09	69.10	13.91	86.56	93.19	89.75	17.76	87.01	93.12	89.96	17.84
H03	86.92	71.60	78.52	15.20	86.47	96.92	91.40	19.85	93.91	92.69	93.30	20.67	94.30	92.38	93.33	20.67
H04	86.04	92.99	89.38	20.04	85.87	96.73	90.98	20.83	90.29	95.06	92.61	21.56	90.80	94.25	92.49	21.46
H05	90.46	75.71	82.43	16.40	90.23	92.35	91.28	19.90	96.86	88.54	92.51	20.31	96.88	88.29	92.39	20.23
H06	92.28	70.90	80.19	15.22	68.16	96.55	79.91	16.46	96.30	89.66	92.86	20.11	96.36	89.19	92.64	19.96
H07	68.51	96.08	79.98	16.41	68.54	96.02	79.99	16.41	85.56	92.95	89.10	18.55	85.94	92.82	89.25	18.60
H08	91.27	96.53	93.83	20.71	91.33	96.46	93.82	20.70	93.92	93.13	93.52	20.36	93.96	92.30	93.12	20.08
H09	79.77	99.45	88.53	16.25	79.71	99.48	88.50	16.24	93.05	97.14	95.05	19.54	93.06	97.13	95.05	19.54
H10	87.91	92.48	90.14	17.17	87.42	95.60	91.33	17.81	95.72	95.62	95.67	20.64	95.80	95.59	95.69	20.66
H11	89.52	95.10	92.23	18.88	88.52	95.44	91.85	18.71	94.10	93.19	93.64	19.60	93.82	93.45	93.63	19.61
H12	84.53	96.61	90.17	19.73	84.42	96.53	90.07	19.69	92.69	93.28	92.99	20.93	92.88	93.22	93.05	20.97
H13	82.43	88.55	85.38	18.65	81.67	94.32	87.54	19.49	91.06	88.94	89.99	20.09	90.86	89.03	89.93	20.07
H14	88.08	93.53	90.72	20.83	87.94	97.39	92.43	21.80	92.55	96.11	94.30	22.89	92.45	96.38	94.37	22.96
Mean	85.13	89.65	86.82	17.92	79.56	96.43	86.63	18.44	92.58	93.27	92.87	20.39	92.83	93.05	92.89	20.40
DIBCO 2013																
H01	81.72	97.23	88.80	20.94	81.73	97.20	88.80	20.94	87.36	92.03	89.63	21.02	87.05	92.13	89.52	20.99
H02	88.32	84.89	86.57	17.36	88.13	96.51	92.13	19.96	91.55	95.05	93.27	20.52	91.33	95.26	93.25	20.52
H03	85.84	95.50	90.41	19.20	85.42	95.21	90.05	19.04	89.40	92.34	90.84	19.25	89.07	92.64	90.82	19.25
H04	96.02	53.20	68.47	13.75	95.84	96.48	96.16	24.38	97.78	95.52	96.64	24.89	97.86	95.28	96.55	24.78
H05	94.74	30.54	46.19	11.64	94.66	90.65	92.61	23.29	98.00	85.98	91.60	22.53	98.10	84.49	90.79	22.09
H06	90.45	94.19	92.28	19.95	90.32	95.01	92.61	20.16	96.04	92.92	94.46	21.23	96.09	92.85	94.44	21.21
H07	67.72	97.56	79.95	21.08	66.21	97.84	78.98	20.92	75.46	95.64	84.36	21.92	75.47	95.65	84.37	21.93
H08	91.15	85.86	88.43	18.42	91.25	84.98	88.00	18.24	94.29	93.60	93.94	21.35	94.21	93.83	94.02	21.41
P01	86.50	81.56	83.96	17.75	86.51	95.80	90.92	20.56	91.10	93.84	92.45	21.21	90.98	94.16	92.54	21.28
P02	91.99	95.34	93.63	19.52	92.25	97.81	94.95	20.57	95.66	95.58	95.62	21.06	95.63	95.67	95.65	21.09
P03	79.97	95.35	86.98	17.57	80.33	96.03	87.48	17.74	91.38	95.89	93.58	20.37	91.70	95.91	93.76	20.49
P04	91.22	92.74	91.98	17.04	74.11	97.95	84.38	14.68	78.72	95.38	86.25	15.07	78.83	95.17	86.23	15.05
P05	92.43	93.24	92.83	15.81	89.73	91.85	90.77	14.75	92.83	95.52	94.15	16.74	92.74	95.55	94.12	16.72
P06	92.45	65.24	76.50	11.91	93.12	66.68	77.71	12.17	97.07	59.62	73.87	11.08	97.31	58.84	73.33	10.95
P07	89.07	94.84	91.86	14.76	89.17	97.78	93.28	15.65	91.81	96.51	94.10	16.13	91.74	96.56	94.09	16.12
P08	87.54	67.87	76.46	13.35	87.61	71.36	78.65	13.89	90.01	75.51	82.12	14.73	89.90	76.09	82.42	14.82
Mean H	87.00	79.87	80.14	17.79	86.69	94.24	89.92	20.87	91.23	92.89	91.84	21.59	91.15	92.77	91.72	21.52
Mean P	88.90	85.77	86.78	15.96	86.61	89.41	87.27	16.25	91.07	88.48	89.02	17.05	91.10	88.49	89.02	17.07
Total from DIBCO 2009 to DIBCO 2013																
Mean H <sup>a</sup>	86.77	79.64	83.05	16.79	83.49	93.75	88.32	18.84	92.69	91.57	92.13	20.28	92.83	91.34	92.08	20.25
Mean P <sup>a</sup>	88.99	70.01	78.37	12.62	81.96	89.89	85.75	15.18	91.37	88.44	89.88	16.40	91.49	88.11	89.77	16.35

<sup>a</sup> These means do not consider the training images.

- Ramírez-Ortegón's framework applied to Niblack's threshold using the weighted variance measure (normal distribution), denoted by RO+Nib.
- Ramírez-Ortegón's framework applied to Sauvola's threshold using the weighted variance measure (normal distribution), denoted by RO+Sau.

From our results in Fig. 11, we have four comments:

- The transition method is moderately affected if the transition proportion is replaced by a constant. This supports our observations in Section 4.4.3, where we pointed out that the

accuracy of the means and variances of the transition sets may considerably affect the error rate.

- The crossed results for the normal transition methods point out that NorT has a more balanced error between false positive and false negatives. Observe that NorT performs better than NorT ( $w=0.5$ ) and NorT ( $w=0.4$ ) under FM while, at the same time, it performs the poorest under PSNR measure. In fact, NorT reports a better recall measurement than NorT ( $w=0.5$ ) and NorT ( $w=0.4$ ) which means that NorT ( $w=0.5$ ) and NorT ( $w=0.4$ ) tend to underestimate the foreground.
- The gray-intensity distributions of the positive and negative transitions sets may not be normally distributed since LogT outperforms NorT. A deep discussion of this is beyond the

**Table 7**  
Measurement averages of DIBCO 2009 for several state-of-the-art methods are given. We also listed the first top five methods in DIBCO 2009 competition sorted by the average of FM and PSNR. M stands for method.

Method	Handwritten documents						Printed documents					
	M	RE	PR	FM	M	PSNR	M	RE	PR	FM	M	PSNR
DIBCO 2009 [11]	26	–	–	88.65	26	19.42	14	–	–	94.09	14	17.90
	34a	–	–	86.16	14	18.57	26	–	–	93.81	26	17.89
	14	–	–	86.02	34a	18.32	24	–	–	93.18	24	17.44
	24	–	–	85.29	24	18.14	1	–	–	92.82	1	17.26
	6	–	–	84.75	6	17.82	19	–	–	92.76	33c	17.09
Bataineh [24]		88.00	83.50	85.08		18.06		88.72	93.43	90.93		16.37
Ben Messaoud [25]		89.08	90.11	89.56		19.77		91.28	95.59	93.37		17.55
Hedjam et al. [23]		89.57	88.61	88.93		19.51		92.63	94.90	93.74		17.89
Howe [18]		95.80	93.58	94.68		22.62		94.98	94.18	94.56		18.57
Lelore [16]		95.08	90.26	92.61		21.25		95.30	95.25	95.26		19.18
Lu [27]		90.67	86.72	88.55		19.42		94.14	93.48	93.72		17.89
Moghaddam [22]		90.64	91.12	90.84		20.31 <sup>a</sup>		96.04	90.79	93.29		17.40 <sup>a</sup>
Shi [21]		–	–	80.77		17.82		–	–	90.54		15.75
Su [26]		87.18	93.77	90.27		20.10		87.46	97.07	91.85		16.91
Su (DIBCO 2011) [13]		93.14	94.27	93.70		22.01		92.20	95.91	93.93		17.81
LogT		86.78	73.45	77.18		16.37		89.58	92.84	91.14		16.41
ROI+LogT		86.30	84.02	84.89		18.35		77.55	94.82	82.87		15.22
ROI+LogT+BR		93.31	88.70	90.91		20.19		93.87	93.48	93.66		18.03
ROI+NorT+BR		93.96	87.45	90.54		19.96		94.10	93.08	93.57		17.97
ROI+Niblack+BR		90.06	91.84	90.92		20.42		78.18	95.91	84.12		15.34
ROI+Otsu+BR		93.82	85.46	89.34		19.52		94.28	93.24	93.75		18.07

<sup>a</sup> Even though the method's publication does not report it, the value was estimated from the precision and recall measurements. The precision of some values is less than four decimals since the values were taken from the original publication rather than calculated from the binary images.

**Table 8**  
Measurements of H-DIBCO 2010 for several state-of-the-art methods are given. We also listed the first top five methods in the H-DIBCO 2010 competition sorted by the average of FM and PSNR. M stands for method. The precision of some values is less than four decimals since the values were taken from the original publication rather than calculated from the binary images.

Method	RE	PR	M	FM	M	PSNR
H-DIBCO 2010 [12]	–	–	3	91.78	1	19.78
	–	–	1	91.50	3	19.67
	–	–	14	89.73	2	19.15
	–	–	2	89.70	14	18.90
	–	–	10	87.98	10	18.26
Bataineh [24]	75.74	93.19		82.86		16.94
Ben Messaoud [25]	86.82	794.09		90.23		19.14
Howe [25]	92.40	95.34		93.81		21.12
Lelore [16]	93.67	94.42		93.99		21.20
Lu [27]	84.06	84.01		80.28		16.53
Su [26]	77.33	96.20		85.49		17.83
Su (DIBCO 2011) [13]	88.18	96.52		92.11		20.16
LogT	82.83	90.69		85.98		17.71
ROI+LogT	82.31	95.56		88.29		18.65
ROI+LogT+BR	89.98	91.52		90.62		19.29
ROI+NorT+BR	90.01	91.47		90.61		19.28
ROI+Niblack+BR	87.40	92.72		89.84		19.03
ROI+Otsu+BR	90.20	90.95		90.40		19.18

scope of this paper. For further discussion on such a topic, readers are referred to [55].

- Niblack's threshold slightly surpasses LogT. However, when Niblack's method is combined with a technique for an automatic parameter selection, like Badekas and Papamarkos' method, its performance is inferior to LogT; see BP+Nib.

### 5.3. Evaluation II

The goal of this evaluation is to show the potential of the transition method in historical documents. We tested our algorithms on (H-)DIBCO's benchmark 2009–2013 whereby we can compare our methods with the-state-of-the-art methods.

To deal with historical documents, we implemented the local version of our previous thresholds: ROI+LogT+BR and ROI+NorT+BR.

The inputs and outputs of these binarizations are pre-processed and post-processed for diverse techniques localizing the regions of interest, removing small connected components, restoring contours, and removing binary artifacts [56]. For further reference, details of these processes are given in the supplementary material.

We also implemented four other versions: LogT, ROI+LogT, ROI+Niblack+BR, and ROI+Otsu+BR. The former two methods are implemented in a similar manner to ROI+LogT+BR. However, no pre- or post-process was done for LogT, and no post-process was done for ROI+LogT. The latter two are identical to ROI+LogT+BR but replacing the threshold given by LogT with, respectively, Niblack's and Otsu's thresholds. In this manner, we can observe how pre- and post-processes enhance the final binary image.

Tables 5 and 6 report the measurements of each test image in detail. In addition, a comparison with the-state-of-art methods is given in Tables 7–11. For these tables, the reported measures were obtained in different manners:

- Computed from the binary images that were provided by the different authors of each method: methods of DIBCO 2011 [13], methods of H-DIBCO 2012 [14], methods of DIBCO 2013 [15], Bataineh's [24], Hedjam's [23], Lelore's [16], Lu's [27], and Su's [26] methods.
- Computed from binary images that were generated by an executable program of the different authors of each

**Table 9**

Measurements of DIBCO 2011 are given. We also listed the first top five methods on DIBCO 2011 competition sorted by the average of FM and PSNR. M stands for method.

Method	Handwritten documents					Printed documents				
	M	RE	PR	FM	PSNR	M	RE	PR	FM	PSNR
DIBCO 2011 [13]	10	90.63	94.38	92.38	19.93	10	91.82	85.97	87.87	17.08
	11	91.32	88.78	89.59	18.76	8	86.81	88.68	87.29	16.40
	8	85.57	93.54	88.74	18.62	11	88.37	87.13	87.20	16.26
	7	85.19	89.56	87.18	17.76	7	88.08	86.55	86.84	16.16
	5	82.96	89.17	85.71	17.68	6	84.41	90.44	86.54	16.10
Bataineh [24]		83.38	85.12	83.81	16.79		86.51	82.75	83.36	15.12
Ben Messaoud [25]		80.36	90.67	83.75	17.68		86.04	85.73	83.43	15.77
Howe [18]		85.25	96.10	89.24	20.08		93.02	89.08	90.33	18.01
Lelore [16]		94.58	94.07	94.31	21.03		91.48	91.20	90.89	17.86
Lu [27]		86.87	80.66	83.09	16.26		90.68	76.34	80.24	14.91
Su [26]		79.04	90.94	83.72	16.97		82.14	94.50	87.40	16.54
Su (DIBCO 2011) <sup>a</sup> [13]		86.38	93.46	89.23	18.86		88.65	83.31	81.66	15.59
LogT		87.26	81.99	83.63	16.69		86.00	66.02	67.35	11.49
ROI+LogT		85.13	89.44	86.87	17.73		80.36	92.42	84.73	16.05
ROI+LogT+BR		91.39	86.59	88.38	18.08		91.38	92.26	91.62	18.08
ROI+NorT+BR		91.17	86.21	88.05	17.97		92.83	93.05	92.89	20.40
ROI+Niblack+BR		87.48	93.65	90.38	19.10		87.40	95.74	91.25	17.90
ROI+Otsu+BR		90.53	86.34	87.72	17.85		92.22	91.63	91.76	18.16

<sup>a</sup> The binary executable program generates outputs whose measurements do not match with those measurements reported in [13].

**Table 10**

Measurements of H-DIBCO 2012 are given. We also listed the first top five methods in the H-DIBCO 2012 competition sorted by the average of FM and PSNR. M stands for method.

Method	M	RE	PR	FM	M	PSNR
DIBCO 2012 [14]	11	92.90	92.79	92.69	6	21.80
	7	92.02	91.67	91.55	11	20.57
	4a	86.19	97.61	91.34	4a	20.14
	18a	90.96	91.36	90.93	7	19.65
	13	85.39	96.00	90.23	8	19.44
Bataineh [24]		84.05	91.20	86.88		17.89
Ben Messaoud [25]		81.08	92.47	85.55		17.97
Howe [18]		94.98	95.66	95.29		22.37
Lelore [16]		93.89	94.51	94.12		21.44
Su (DIBCO 2011) [13]		86.92	93.55	88.87		19.63
LogT		85.13	89.65	86.82		17.92
ROI+LogT		79.56	96.43	86.63		18.44
ROI+LogT+BR		92.58	93.27	92.87		20.39
ROI+NorT+BR		92.83	93.05	92.89		20.40
ROI+Niblack+BR		86.58	94.55	90.15		19.28
ROI+Otsu+BR		93.08	92.13	92.54		20.15

**Table 11**  
Measurements of DIBCO 2013 are given. We also listed the first top five methods on DIBCO 2013 competition sorted by the average of FM and PSNR. M stands for method.

	Handwritten documents					Printed documents					
	M	RE	PR	FM	PSNR	M	RE	PR	FM	M	PSNR
DIBCO 2013 [15]	17 <sup>a</sup>	91.23	92.89	91.84	23.47	3	93.94	91.94	92.43	3	19.12
	3	87.56	96.99	90.30	22.55	13	96.86	89.40	92.39	15b	18.87
	15b	84.86	98.09	90.11	22.51	5	96.30	88.96	91.89	5	18.84
	13	87.48	94.93	90.09	22.48	15b	94.43	89.92	91.48	13	18.53
	5	88.41	93.58	89.30	21.59	17 <sup>a</sup>	10c	96.39	87.00	10c	17.90
Bataineh [24]		81.13	87.32	80.59	18.57		88.54	88.48	87.90		16.34
Ben Messaoud [25]		86.30	95.44	90.49	21.30		89.27	89.88	88.96		16.99
Howe [18]		87.10	97.29	89.54	23.87		93.65	89.45	90.65		18.23
Lelore [16]		88.63	94.85	90.89	22.66		94.54	89.76	91.45		18.52
Su (DIBCO 2011) [13]		86.69	94.24	89.92	20.87		86.61	89.41	87.27		16.25
LogT		87.00	79.87	80.14	17.79		88.90	85.77	86.78		15.96
ROI+LogT		86.69	94.24	89.92	20.87		86.61	89.41	87.27		16.25
ROI+LogT+BR		91.23	92.89	91.84	21.59		91.07	88.48	89.02		17.05
ROI+NorT+BR		91.15	92.77	91.72	21.52		91.10	88.49	89.02		17.07
ROI+Niblack+BR		79.05	82.32	80.45	20.72		86.87	92.44	88.91		17.07
ROI+Otsu+BR		90.70	91.90	90.96	21.14		90.46	88.31	88.42		16.91

<sup>a</sup> This method is the same method presented in this paper [13].

**Table 12**  
Measurements of all (H-)DIBCO's benchmarks (2009–2013) are given. For computing this table: the true positive, the true negative, the false positive, and the false negative pixels from all images were counted and, then, both FM and PSNR were computed.

Method	Handwritten documents				Printed documents			
	RE	PR	FM	PSNR	RE	PR	FM	PSNR
Bataineh [24]	84.30	87.84	86.03	17.91	89.22	87.13	88.16	15.74
Ben Messaoud [25]	84.18	93.44	88.57	18.91	88.14	89.33	88.73	16.04
Howe [18]	93.42	96.18	94.78	22.16	92.56	86.53	89.44	16.14
Lelore [16]	93.67	94.55	94.10	21.59	93.06	87.85	90.38	16.57
Su (DIBCO 2011) [13]	87.74	89.68	88.70	18.79	87.58	87.52	87.55	15.57
LogT	86.77	79.64	83.05	16.79	88.99	70.01	78.37	12.62
ROI+LogT	83.49	93.75	88.32	18.84	81.96	89.89	85.75	15.18
ROI+LogT+BR	92.69	91.57	92.13	20.28	91.37	88.44	89.88	16.40
ROI+NorT+BR	92.83	91.34	92.08	20.25	91.49	88.11	89.77	16.35
ROI+Niblack+BR	83.17	94.06	88.28	18.84	83.25	92.45	87.61	15.82
ROI+Otsu+BR	92.74	90.38	91.55	19.94	91.04	87.38	89.17	16.09



**Fig. 12.** (a) A sample of H06 of DIBCO 2010 is given. (b) Groundtruth. (c) Output of ROI+LogT+BR.

method: Ben Messaoud's [25], Howe's [18], and Su's (DIBCO 2011)<sup>1</sup> methods.

- Taken from the original papers: methods of DIBCO 2009 [11], methods of H-DIBCO 2010 [12], Moghaddam's [22] and Shi's [21] methods.

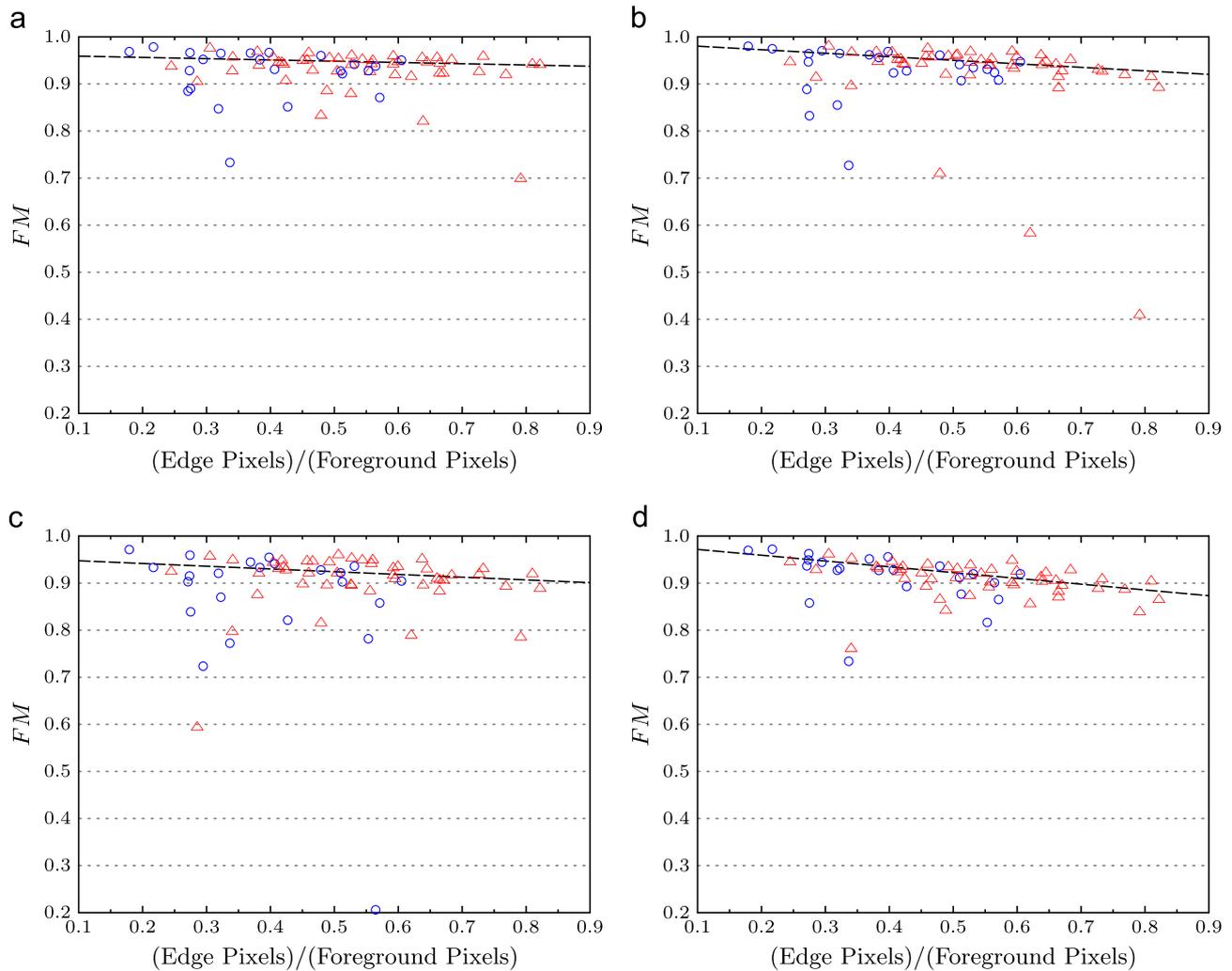
We summarized our results by computing the measures by the total of pixels from the images of all (H-)DIBCO's competitions: the true positive, the true negative, the false positive, and the false negative pixels from all images were counted and, then, both FM and PSNR were computed. Table 12 only reports those methods and their binary images of all the DIBCO's competitions.

In terms of measure means, Tables 5 and 6 show that ROI+LogT+BR is slightly better than ROI+NorT+BR. Even so, the pairwise comparisons between ROI+LogT+BR and ROI+NorT+BR show a ratio about 2:1 in favor of ROI+LogT+BR for both handwritten and printed documents.

We noticed that our method's performance in (H-)DIBCO benchmarks is correlated with the stroke width of the documents. This may be related to the fact that in thin strokes less transition pixels are available to compute the means and variances of gray intensities.

Another factor that acts upon the scores of our method is the definition of edge. We conjectured that the correlation between performance and stroke width happens due to our edge detectors: the edges in DIBCO's groundtruth images were constructed by a method based on Canny's edge detector [57] while our method

<sup>1</sup> <http://www.comp.nus.edu.sg/~subolan/bin/bin.html>



**Fig. 13.** Pairwise plot between the edge proportion and FM values for the images of (H-)DIBCO 2009–1013. The regression lines ( $y = mx + b$ ) were computed by the repeated median estimator. The handwritten documents are shown in red triangles while printed documents are shown in blue circles. (a) Lore's method ( $m = -0.0268$ ,  $b = 0.9617$ ). (b) Howe's method ( $m = -0.0749$ ,  $b = 0.9877$ ). (c) Su 2011's method ( $m = -0.0582$ ,  $b = 0.9534$ ). (d) ROI+LogT+BR method ( $m = -0.1230$ ,  $b = 0.9839$ ). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

relies on transition sets [28], and as a consequence, our contour definition differs from the contour definition in DIBCO's ground-truth images.

Note that slight contour differences may produce considerably different FM's scores in images where the ratio of edge pixels to foreground pixels is high, as in handwritten documents. For example, Fig. 12 shows one of the most representative images where edge pixels may be the cause of a low method's performance. Observe that our binarization is visually well-defined but, nevertheless, it scored a low FM (around 86%). In fact, the bigger the ratio between edge pixels to foreground pixels, the poorer our method's performance; Fig. 13 shows a correlation between the edge proportion and FM values for the images of (H-)DIBCO 2009–2013.

Among the other evaluated methods, Lore's, Howe's, and Su's (DIBCO 2011) methods have a remarkable performance. This is not a surprise since these methods have been top-ranked through out DIBCO's editions. In particular, Lore's and Howe's perform the best in handwritten and printed, respectively; see Table 12. All three methods detect edges by Canny's method, as the ground-truth generator, which is an advantage over our methods. Even so, observe in Table 12 that our method has comparable performance with Lore's and Howe's method for printing documents (less

than 0.1 of difference in FM and less than 0.2 in PSNR). Moreover, Su's (DIBCO 2011) method has a relatively low performance according to Table 12 but, at the same time, Tables 7 and 8 report a high performance for Su's (DIBCO 2011) method. It happens because such tables report FM averages at the image-level rather than at the pixel-level and, as a result, dreadful performance in certain images is covered.

## 6. Conclusions

In this paper, we have analyzed the transition proportion mathematically and experimentally for historical handwritten images. We concluded that the influence of the transition proportion on the normal transition is marginal. In concrete, we proved that inaccurate estimations of the transition proportion led to thresholds that differ less four or less gray levels of the actual optimal threshold (assuming 256 gray levels).

We determined the optimal values for the transition proportion when it was assumed to be a constant. These values depended on the distribution of the transition proportion and some assumptions on the gray-intensity variances.

For documents, we determined a criterion to detect regions of interest: pixels within regions of interest have a transition proportion greater than 0.1.

We also discovered that the convergence values of the transition proportion is a function of the stroke width and the radius of the pixel neighborhood. Hence, it can be used as a feature to determine dominant values of the stroke width within a document.

Our results not only have ensured the interaction of the transition method with the transition proportion, but also have established the mathematical bases for further research in stroke-width estimation and in detection of regions of interest. Furthermore, our analysis of the transition proportion can be extended to other methods based on the mixture of two normal distributions and Bayes rule.

## 7. Further work

To improve our method, we will explore better fitting models for the mixture of distributions. In particular, we will continue the work in [55], where the authors proposed a model for the gray-intensity distribution based on the mixture of three distributions (two normal distributions and another that is not normal).

Furthermore, in [55] was introduced the concept of frontier pixels which are pixels that contain both fore and background regions. This concept is strongly related to edge pixels and, consequently, it can be used to formulate transition functions that identify transition sets more accurately. Note that, in this paper, a transition pixel is considered to be entirely contained within the foreground (positive transition) or within the background (negative transition). Hence, frontier pixels were not modeled properly.

## Conflict of interest

None declared.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.patcog.2014.02.003>.

## References

- [1] S. Kompalli, S. Setlur, V. Govindaraju, Devanagari OCR using a recognition driven segmentation framework and stochastic language models, *Int. J. Doc. Anal. Recognit.* 12 (2009) 123–138.
- [2] H. Lee, B. Verma, Binary segmentation algorithm for English cursive handwriting recognition, *Pattern Recognit.* 45 (4) (2012) 1306–1317.
- [3] Z. Su, Z. Cao, Y. Wang, Stroke extraction based on ambiguous zone detection: a preprocessing step to recover dynamic information from handwritten Chinese characters, *Int. J. Doc. Anal. Recognit.* 12 (2009) 109–121.
- [4] S. Bag, G. Harit, An improved contour-based thinning method for character images, *Pattern Recognit. Lett.* 32 (14) (2011) 1836–1842.
- [5] G.E. Louloudis, B.G. Gatos, I. Pratikakis, C. Halatsis, Text line detection in handwritten documents, *Pattern Recognit.* 41 (2008) 3758–3772.
- [6] Y. Xiao, H. Yan, Location of title and author regions in document images based on the Delaunay triangulation, *Image Vis. Comput.* 22 (4) (2004) 319–329.
- [7] X. Wang, L. Huang, C. Liu, A video text location method based on background classification, *Int. J. Doc. Anal. Recognit.* 13 (2010) 173–186.
- [8] X. Peng, S. Setlur, V. Govindaraju, S. Ramachandru, Using a boosted tree classifier for text segmentation in hand-annotated documents, *Pattern Recognit. Lett.* 33 (7) (2012) 943–950 (special Issue on Awards from ICPR 2010).
- [9] H. Lv, W. Wang, C. Wang, Q. Zhuo, Off-line Chinese signature verification based on support vector machines, *Pattern Recognit. Lett.* 26 (15) (2005) 2390–2399.
- [10] A. Brink, J. Smit, M. Bulacu, L. Schomaker, Writer identification using directional ink-trace width measurements, *Pattern Recognit.* 45 (1) (2012) 162–171.
- [11] B. Gatos, K. Ntirogiannis, I. Pratikakis, DIBCO 2009: document image binarization contest, *Int. J. Doc. Anal. Recognit.* 14 (2011) 35–44.
- [12] I. Pratikakis, B. Gatos, K. Ntirogiannis, H-DIBCO 2010—handwritten document image binarization competition, in: International Conference on Frontiers in Handwriting Recognition, IEEE Computer Society, Los Alamitos, CA, USA, 2010, pp. 727–732.
- [13] I. Pratikakis, B. Gatos, K. Ntirogiannis, ICDAR 2011 document image binarization contest (DIBCO 2011), in: 2011 International Conference on Document Analysis and Recognition, IEEE, 2011, pp. 1506–1510.
- [14] I. Pratikakis, B. Gatos, K. Ntirogiannis, ICFHR 2012 competition on handwritten document image binarization (H-DIBCO 2012), in: 2012 International Conference on Frontiers in Handwriting Recognition (ICFHR), 2012, pp. 817–822.
- [15] I. Pratikakis, B. Gatos, K. Ntirogiannis, ICDAR 2013 document image binarization contest (DIBCO 2013), in: 12th International Conference on Document Analysis and Recognition, IEEE Computer Society, Los Alamitos, CA, USA, 2013, pp. 1471–1476.
- [16] T. Lelore, F. Bouchara, FAIR: a fast algorithm for document image restoration, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (8) (2013) 2039–2048.
- [17] M. Valizadeh, E. Kabir, An adaptive water flow model for binarization of degraded document images, *Int. J. Doc. Anal. Recognit.* 16 (2) (2013) 165–176.
- [18] N.R. Howe, Document binarization with automatic parameter tuning, *Int. J. Doc. Anal. Recognit.* 16 (3) (2013) 247–258.
- [19] M. Valizadeh, E. Kabir, Binarization of degraded document image based on feature space partitioning and classification, *Int. J. Doc. Anal. Recognit.* 15 (1) (2012) 57–69.
- [20] D. Rivest-Hénault, R. Farrahi Moghaddam, M. Cheriet, A local linear level set method for the binarization of degraded historical document images, *Int. J. Doc. Anal. Recognit.* 15 (2012) 101–124.
- [21] J. Shi, N. Ray, H. Zhang, Shape based local thresholding for binarization of document images, *Pattern Recognit. Lett.* 33 (1) (2012) 24–32.
- [22] R.F. Moghaddam, M. Cheriet, AdOtsu: an adaptive and parameterless generalization of Otsu's method for document image binarization, *Pattern Recognit.* 46 (6) (2012) 2419–2431.
- [23] R. Hedjam, R.F. Moghaddam, M. Cheriet, A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images, *Pattern Recognit.* 44 (9) (2011) 2184–2196.
- [24] B. Bataineh, S.N.H.S. Abdullah, K. Omar, An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows, *Pattern Recognit. Lett.* 32 (14) (2011) 1805–1813.
- [25] I. Ben Messaoud, H. El Abed, H. Amiri, V. Märgner, New method for the selection of binarization parameters based on noise features of historical documents, in: Proceedings of the 2011 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data, ACM, New York, NY, USA, 2011, pp. 1:1–1:8.
- [26] B. Su, S. Lu, C.L. Tan, Binarization of historical document images using the local maximum and minimum, in: Proceedings of the 9th IAPR International Workshop on Document Analysis Systems, ACM, 2010, pp. 159–166.
- [27] S. Lu, B. Su, C.L. Tan, Document image binarization using background estimation and stroke edges, *Int. J. Doc. Anal. Recognit.* 13 (2010) 303–314.
- [28] M.A. Ramírez-Ortegón, E. Tapia, R. Rojas, E. Cuevas, Transition thresholds and transition operators for binarization and edge detection, *Pattern Recognit.* 43 (10) (2010) 3243–3254.
- [29] R.F. Moghaddam, M. Cheriet, A multi-scale framework for adaptive binarization of degraded document images, *Pattern Recognit.* 43 (6) (2010) 2186–2198.
- [30] K. Ntirogiannis, B.G. Gatos, I. Pratikakis, A modified adaptive logical level binarization technique for historical document images, in: 10th International Conference on Document Analysis and Recognition, 2009, pp. 1171–1175.
- [31] B. Gatos, I. Pratikakis, S.J. Perantonis, Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information, in: 19th International Conference on Pattern Recognition, ICPR, 2008, pp. 1–4.
- [32] C.A. Mello, L.A. Schuler, Thresholding images of historical documents using a Tsallis-entropy based algorithm, *J. Softw.* 3 (6) (2008) 36–39.
- [33] C. Mello, A. Sanchez, A. Oliveira, A. Lopes, An efficient gray-level thresholding algorithm for historic document images, *J. Cult. Herit.* 9 (2) (2008) 109–116.
- [34] M.R. Gupta, N.P. Jacobson, E.K. Garcia, OCR binarization and image pre-processing for searching historical documents, *Pattern Recognit.* 40 (2007) 389–397.
- [35] B. Gatos, I. Pratikakis, S. Perantonis, Adaptive degraded document image binarization, *Pattern Recognit.* 39 (3) (2006) 317–327.
- [36] E. Kavallieratou, S. Stathis, Adaptive binarization of historical document images, in: ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition, IEEE Computer Society, Washington, DC, USA, 2006, pp. 742–745.
- [37] J. Sauvola, M. Pietikäinen, Adaptive document image binarization, *Pattern Recognit.* 33 (2) (2000) 225–236.
- [38] M.A. Ramírez-Ortegón, E. Tapia, L.L. Ramírez-Ramírez, R. Rojas, E. Cuevas, Transition pixel: a concept for binarization based on edge detection and gray-intensity histograms, *Pattern Recognit.* 43 (2010) 1233–1243.
- [39] B. Gatos, K. Ntirogiannis, I. Pratikakis, ICDAR 2009 document image binarization contest (DIBCO 2009), in: Tenth International Conference on Document Analysis and Recognition, 2009, pp. 1375–1382.
- [40] M.A. Ramírez-Ortegón, R. Rojas, Transition thresholds for binarization of historical documents, in: 20th International Conference on Pattern Recognition, IEEE Computer Society, 2010, pp. 2362–2365.
- [41] M.A. Ramírez-Ortegón, R. Rojas, Unsupervised evaluation methods based on local gray-intensity variances for binarization of historical documents, in: 20th International Conference on Pattern Recognition, IEEE Computer Society, 2010, pp. 2029–2032.

- [42] N. Otsu, A threshold selection method from grey-level histograms, *IEEE Trans. Syst. Man Cybern.* 9 (1) (1979) 62–66.
- [43] Ø.D. Trier, A.K. Jain, Goal-directed evaluation of binarization methods, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (12) (1995) 1191–1201.
- [44] M.A. Ramírez-Ortegón, E.A. Duéñez-Guzmán, R. Rojas, E. Cuevas, Unsupervised measures for parameter selection of binarization algorithms, *Pattern Recognit.* 44 (3) (2011) 491–502.
- [45] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, *Comput. Vis. Graph. Image Process.* 29 (1985) 273–285.
- [46] P.K. Sahoo, S. Soltani, A.K. Wong, Y.C. Chen, A survey of thresholding techniques, *Comput. Vis. Graph. Image Process.* 41 (2) (1988) 233–260.
- [47] J. Kittler, J. Illingworth, Minimum error thresholding, *Pattern Recognit.* 19 (1) (1985) 41–47.
- [48] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, *J. Electron. Imaging* 13 (1) (2004) 146–168.
- [49] Z. Hou, Q. Hu, W.L. Nowinski, On minimum variance thresholding, *Pattern Recognit. Lett.* 27 (2006) 1732–1743.
- [50] H.-F. Ng, Automatic thresholding for defect detection, *Pattern Recognit. Lett.* 27 (2006) 1644–1649.
- [51] Z. Li, C. Liu, G. Liu, Y. Cheng, X. Yang, C. Zhao, A novel statistical image thresholding method, *AEU—Int. J. Electron. Commun.* 64 (12) (2010) 1137–1147.
- [52] W. Niblack, *An Introduction to Digital Image Processing*, Prentice Hall, Birkeroed, Denmark, 1985.
- [53] P. Stathis, E. Kavallieratou, N. Papamarkos, An evaluation technique for binarization algorithms, *J. Univers. Comput. Sci.* 14 (18) (2008) 3011–3030.
- [54] E. Badekas, N. Papamarkos, Estimation of appropriate parameter values for document binarization techniques, *Int. J. Robot. Autom.* 24 (1) (2009) 66–78.
- [55] M.A. Ramírez-Ortegón, L.L. Ramírez-Ramírez, I.B. Messaoud, V. Märgner, E. Cuevas, R. Rojas, A model for the gray-intensity distribution of historical handwritten documents and its application for binarization, *Int. J. Doc. Anal. Recognit. (IJ DAR)* (2013) 1–22.
- [56] M.A. Ramírez-Ortegón, V. Märgner, E. Cuevas, R. Rojas, An optimization for binarization methods by removing binary artifacts, *Pattern Recognit. Lett.* 34 (11) (2013) 1299–1306.
- [57] J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 8 (1) (1986) 679–698.

**Marte A. Ramírez-Ortegón** is currently a full professor at the University Juárez Autónoma de Tabasco. In 2002 he received a B.S. degree (Computer Science) from the University of Guanajuato (UG). He obtained the doctor degree in 2011 at the Freie Universität Berlin, Berlin, Germany. In 2012 he was a post-doctoral fellow at the Institute for Communications Technology, Technische Universität Braunschweig. His current research fields include pattern recognition, edge detection, binarization, thresholding, detection or regions of interest, document noise removal, and document feature analysis.

**Lilia L. Ramírez-Ramírez** a full professor at the Department of Actuarial Science and Insurance, Instituto Tecnológico Autónomo de México (ITAM). She received the Ph.D. degree in Statistics at the University of Waterloo, Canada, in 2008. Lilia holds a M.S. degrees from the University of Guanajuato (UG)/the Centre for Mathematical Research (CIMAT), Guanajuato, Mexico (1997–2000) and a B.S. degree in Actuarial Science from Autonomous National University of Mexico (UNAM), Mexico (1992–1996). Her current research fields include statistics and actuarial science stochastic ‘epidemic’ models, networks, inference for complex systems.

**Volker Märgner** received his diploma (Dipl.-Ing.) and doctorate (Dr.-Ing.) degrees in electrical engineering from the Technische Universität Carolo Wilhelmina zu Braunschweig (TUBS), Germany, in 1974 and 1983, respectively. Since 1983, he has been working at the TUBS. Currently he is a member of the research and teaching staff at the Institute for Communications Technology, in the position of academic director. His research interests include recognition of cursive handwriting Arabic documents, robust pre-processing and feature extraction for historical documents, and evaluation of document segmentation results. He organizes the biennial competition on Arabic handwriting recognition within the ICDAR conference since 2005. He published more than 100 papers including journal papers and book chapters and as editor the book “Guide to OCR for Arabic Scripts” in 2012.

**Ines Ben Messaoud Ben Arbia** Ines Ben Messaoud Ben Arbia is a system software developer at PHCom GmbH Germany. In 2006 she received her diploma in computer science engineering from the National Engineering School of Sfax (ENIS), Tunisia. In 2008 she received her master in automation and signal processing at the ENIT. In 2010 she got a scholarship from the German Academic Exchange Service (DAAD) for a two years research stay at the Institute for Communications Technology (IfN), Braunschweig, Germany, to accomplish her Ph.D. work. She received the doctorate in Electrical Engineering from National Engineering School of Tunis (ENIT), Tunisia in 2013. Her research is focused on document analysis and recognition and on automatic evaluation of document processing methods.

**Erik Cuevas** received the B.S. and M.S. degrees in Computer Engineering from the University of Guadalajara, Guadalajara, Mexico, in 1996 and 2000, respectively. He received the Ph.D. degree in Computer Engineering at the Freie Universität Berlin, Berlin, Germany, in 2006. Currently, he is a full professor at the Computer Science Department, University of Guadalajara. His research interests include computer vision, thresholding, circle detection, corner detection, learning systems, and soft computing.

**Raúl Rojas** received the B.S. and M.S. degrees in Mathematics from the National Polytechnic Institute (IPN), Mexico City, Mexico. He obtained the doctor degree at the Freie Universität Berlin, Berlin, Germany. Currently, Prof. Dr. Rojas is a full professor in Computer Science at the Freie Universität Berlin and leader of the Work Group on Artificial Intelligence at that University. His research interests include artificial intelligence, computer vision, and robotics.