



Transition pixel: A concept for binarization based on edge detection and gray-intensity histograms

Marte A. Ramírez-Ortegón^{a,*}, Ernesto Tapia^a, Lilia L. Ramírez-Ramírez^b, Raúl Rojas^a, Erik Cuevas^c

^a Institut für Informatik, Freie Universität Berlin, Takustr. 9, 14195 Berlin, Germany

^b Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West Waterloo, Ontario, Canada N2L 3G1

^c Department of Computer Science, University of Guadalajara, Av. Revolución 1500, Guadalajara, Jalisco, México, Mexico

ARTICLE INFO

Article history:

Received 14 July 2009

Received in revised form

2 November 2009

Accepted 9 November 2009

Keywords:

Transition function

Thresholding

Binarization

Edge detection

Integral image

Image pre-processing

ABSTRACT

This paper introduces a novel binarization method based on the concept of *transition pixel*, a generalization of edge pixels. Such pixels are characterized by extreme transition values computed using pixel-intensity differences in a small neighborhood. We show how to adjust the threshold of several binary threshold methods which compute gray-intensity thresholds, using the gray-intensity mean and variance of the pixels in the transition set. Our experiments show that the new approach yields segmentation performance superior to several with current state-of-the-art binarization algorithms.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Binarization is a crucial pre-processing used in most image systems for document analysis and recognition. It consists in labeling each pixel in an image either as *foreground* or *background*. The former is a pixel subset \mathcal{F} that represents a region of interest containing forms and objects used for further analysis and recognition, while the latter \mathcal{B} is the complement of \mathcal{F} .

Few binarization researchers work with color images directly [1], the rest transforms an image from color to gray levels before applying the binarization algorithm [2,3]. We follow the latter approach taking a gray-level image I as input and returning a binary image B where pixels in white represent the background and pixels in black represent the foreground.

What constitute foreground depends on the objects to be recognized. While in document analysis one is interested in the location and extraction of ink with high contrast [4,5], in other contexts the information to be extracted can depend on the objects and their relationships. Fig. 1, for instance, shows (a) a triangle and (b) grid lines with similar gray intensity. Both images contain dark pixels which belong to the foreground for sure. However, in (a) we may keep the triangle in the binary image, in

(b) we could possibly remove the grid. This example shows that binarization is a complex problem if one considers only gray intensities in order to segment the foreground. Contextual information is needed in order to solve problems similar to those in Fig. 1.

We can identify three categories of binarization algorithms [6,7]. *Global algorithms* label a given pixel using information from the whole image, while *local algorithms* rely on information from the pixel neighborhood. *Hybrid algorithms* combine information from the whole image and from the neighborhood. Note that global algorithms can be converted into local versions by restricting the analysis to the pixel's neighborhood. All algorithms considered in this article are hybrid and local, even though some of them were originally formulated for global analysis. Local versions usually perform better than global ones [8].

Thresholding binarization algorithms label a pixel as foreground if its gray value is darker than a threshold. Sezgin and Sankur [9] present an exhaustive categorization of image thresholding methods. They affirm Kittler's *minimum error thresholding* [10] and Sauvola's *thresholding* [7] are the best-performing document binarization algorithms for images uniformly illuminated and degraded with noise and blur. In essence, Kittler's criterion maximizes the likelihood of the joint distribution of gray intensities assuming that foreground and background are normally distributed with different variances. The well-known Otsu's *algorithm* [11] also handles background and foreground as normally distributed. This method minimizes the gray-intensity variances of both foreground and background.

* Corresponding author.

E-mail addresses: mars_sasha@hotmail.com (M.A. Ramírez-Ortegón), tapia@inf.fu-berlin.de (E. Tapia), llramire@uwaterloo.ca (L.L. Ramírez-Ramírez), rojas@inf.fu-berlin.de (R. Rojas), erik.cuevas@cucei.udg.mx (E. Cuevas).

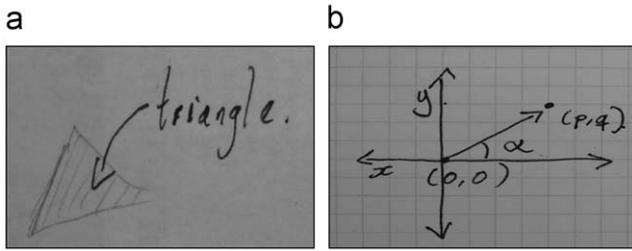


Fig. 1. The context changes the definition of foreground.

Sauvola's method is actually an improved version of *Niblack's algorithm* [12]. Both algorithms assume that the gray-intensity histogram of the background is approximately normally distributed and select a threshold as the lower limit of an interval centered in the mean of the gray-intensity histogram.

The key ideas in the methods mentioned above are the following: *Niblack* and *Sauvola* use statistics from a neighborhood of the pixel of interest, while *Otsu* and *Kittler* assume *a priori* statistical distribution for background and foreground. However, all four algorithms disregard edge information and this leads to small false foreground spots or diminishing foreground boundaries.

Li [13] and *Chen* [14], on the other hand, exploit edge information. *Li's algorithm* uses the *Laplace operator* and the gray-level covariance matrix to compute a gray-intensity threshold. *Chen's method* applies the *Canny edge detector* [15] for generating the edge image. Subsequently, several morphological operators help to generate an enhanced binary image. Both algorithms apply a criterion for selecting pixels with high information content.

A different approach to the binarization problem consist in postulating smooth gray-intensity surfaces for foreground and background in small neighborhoods. *Kavallieratou's algorithm* [16,17] uses an iterative variant of the *local white algorithm* [18]. *Kavallieratou* sets to white those pixel values above the local mean while the rest of pixels is normalized. The process is iterated until a stopping criterion is satisfied. Instead of using a gray-intensity average *Lu's method* [19] computes a polynomial gray-intensity surface for modeling shading fluctuations.

Our method assumes that both background and foreground vary smoothly, exhibiting high contrast at the boundary. We propose a criterion to select pixels by modeling statistically the gray-intensity histograms of those pixels. We also introduce the term *ideal image* based on smooth surfaces and contrast. Later on, the concepts of the *transition pixel* and the *transition set* are used as an extension of edge pixels and the edge set, respectively. Transition pixel's properties are analyzed in an ideal image providing the mathematical foundations for deriving discriminant functions which we named *transition functions*. Each pixel is thus associated with a *transition value* (varying from negative to positive) computed by a corresponding transition function.

We describe the *quantile method for transition values*, which selects pixels with extreme transition values to approximate the transition set. We prove that the *positive transition set*, that is, the intersection of foreground and transition set, is approximated by the set of pixels with high positive transition values, while the *negative transition set*, the intersection of background and transition set, is approximated by the set of pixels with high negative transition values. Hence, the positive transition set is a sample of the foreground and the negative transition set is a sample of the background. The transition set approximation is refined with morphological operators which remove isolated pixels and add pixels surrounded by pixels within the transition set approximation. Our last contribution is to use the lognormal and normal distributions, among other methods, to model the gray-intensity histograms of the transition sets.

Also, we derive expressions for gray-intensity mean and variance based on *integral images* [20] in order to quickly compute statistical thresholds.

Even though the transition method has the potential to deal with uneven illumination, this paper will focus only on images without sudden illumination changes in small neighborhoods.

The rest of this paper is organized as follows. Section 2 introduces the main concepts and introduces the concept of transition. The transition method is developed in Section 3. Section 4 deals with the algorithm's complexity. In Section 5 experimental results are shown. Conclusions are presented in Section 6.

2. Preliminary concepts

2.1. Notation

We denote images with capital italics, such as I and B . An ordered pair (i, j) represents the pixel located in the i -th row and j -th column. $I(i, j) \in [0, l]$ is an integer that represents the gray intensity of (i, j) . $B(i, j) \in \{0, 1\}$ represents the binarized image, where one is considered as foreground. Local binarization algorithms compute a threshold surface T over the whole image: $B(i, j) = 1$ if $I(i, j)$ is lower than the threshold $T(i, j)$. The information to compute $T(i, j)$ is gathered from the pixels within a square $\mathcal{N}_r(i, j)$ centered at the pixel (i, j) of sides with length $2r + 1$. For simplicity, we refer to (i, j) just as p and it will be used when none of the indexes can be confused.

$P(p \in \mathcal{A})$ denotes the probability that a pixel p belongs to \mathcal{A} . The mean and variance of any variable v are denoted as μ_v and σ_v^2 , respectively.

2.2. Image model

The ideal situation for any binarization algorithm occurs when the gray-intensity difference between foreground and background pixels is large and the gray-intensity difference of a pair of pixels within the same class is small (Fig. 2). In small neighborhoods we expect the foreground to correspond to pixels with maximum gray intensity, and the background to correspond to pixels with the minimal gray intensity. Under these assumptions, *Bersen's method* [21] localize the regions in the image with large gray-intensity differences by computing a threshold that lays between their maximum and minimum gray intensities.

We summarize the above ideas with the following concepts:

- **Foreground tendency:** Locally, the pixel's probability of being foreground increases when its gray-intensity value gets closer to zero. Conversely, the pixel's probability of being background tends to increase when its gray-intensity value gets closer to l .
- **Smoothness:** The gray-intensity difference between two pixels from the same set is close to zero in small neighborhoods \mathcal{N}_s . Mathematically, I is a smooth image if

$$\max_{q \in B \cap \mathcal{N}_s(p)} I(q) - \min_{q \in B \cap \mathcal{N}_s(p)} I(q) < d_s$$

and

$$\max_{q \in F \cap \mathcal{N}_s(p)} I(q) - \min_{q \in F \cap \mathcal{N}_s(p)} I(q) < d_s,$$

with probability close to 1 where d_s is a small value with respect to l .

- **Local contrast:** Locally, the difference between a pair of pixels from different set is greater than the difference between a pair

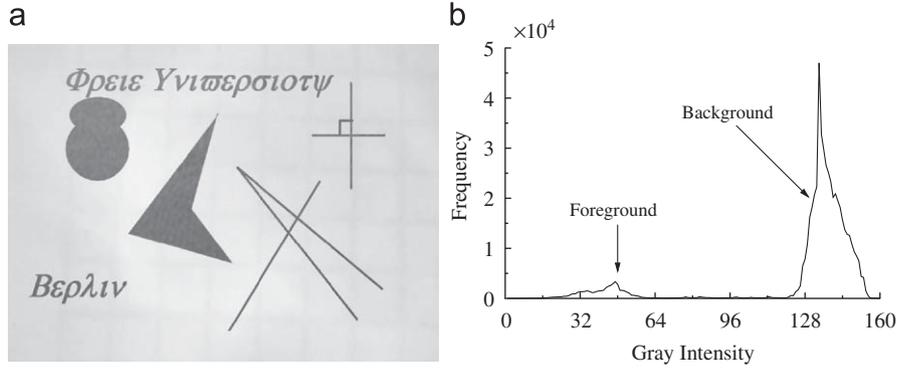


Fig. 2. The image (a) is an ideal image whose gray-intensity histogram is shown in (b).

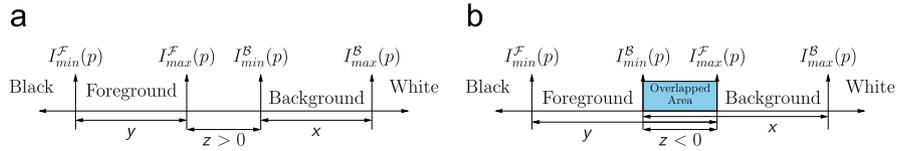


Fig. 3. (a) Representation of the random variables x , y and z into $\mathcal{N}_s(p)$. (b) In a neighborhood with low contrast, \mathcal{F} and \mathcal{B} have a gray-intensity range of overlap between their gray intensities.

of pixels from the same set. Mathematically

$$\min_{q \in \mathcal{B} \cap \mathcal{N}_s(p)} I(q) - \max_{q \in \mathcal{F} \cap \mathcal{N}_s(p)} I(q) > d_c,$$

with probability close to 1, where d_c is a large positive number with respect to d_s .

It should be emphasized that, compared to $\mathcal{N}_t(p)$, $\mathcal{N}_s(p)$ is a small neighborhood. Typically, $s < 5$. We simplify our notation

$$I_{max}^B(p) = \max_{q \in \mathcal{B} \cap \mathcal{N}_s(p)} I(q) \text{ and } I_{min}^B(p) = \min_{q \in \mathcal{B} \cap \mathcal{N}_s(p)} I(q).$$

Similarly $I_{max}^F(p)$ and $I_{min}^F(p)$ are defined.

The previous concepts can be expressed statistically. Let x_s be a random variable with mean μ_{x_s} and variance $\sigma_{x_s}^2$ representing the gray-intensity difference between any pair of pixels $p, q \in \mathcal{B} \cap \mathcal{N}_s(p)$. From now on, we will omit the subindex s . A low mean μ_x , combined with a small variance σ_x^2 , represents a smooth *background surface*. Likewise, a smooth *foreground surface* is obtained when y has a low mean μ_y and a small variance σ_y^2 where y is defined analogously to x but considering foreground pixels, see Fig. 3(a). Finally, the random variable z represents the difference between the minimum background and maximum foreground gray intensities in same neighborhood $\mathcal{N}_s(p)$. We will refer to x , y and z as the random variable of *background differences*, *foreground differences* and *contrast differences*, respectively.

If z is negative, the foreground and background histograms overlap. Therefore, a misclassification may occur in any thresholding method (Fig. 3(b)). For example, the image in Fig. 4(a) is split in dark and bright areas. The foreground and background histograms are bimodal curves. The first mode of the foreground (background) histogram is formed by the foreground (background) in the light area while the second is formed by pixels in the dark area. Misclassifications stem from an overlapping of the second foreground peak with the background modes. Thus, any thresholding method misclassifies either the pixels in the first background mode or the pixels in the second foreground peak Fig. 4(b).

2.3. Transition set

Pixel p is a *t-transition pixel* if its neighborhood $\mathcal{N}_t(p)$ contains foreground and background pixels. The set of those pixels is named \mathcal{T}_t and this group extends along the whole foreground contour. If $t = 1$, then the t -transition pixel is an edge pixel.

In particular, a neighborhood that contains a dense subset of \mathcal{T}_t also contains a significant subset of the *foreground contour*. Furthermore, the statistical distribution of $\mathcal{F} \cap \mathcal{T}_t$ approximates the distributions of \mathcal{F} since it is a large foreground sample. Analogously, the distribution of $\mathcal{B} \cap \mathcal{T}_t$ approximates the distributions of \mathcal{B} .¹

Four neighborhood types help to characterize the transition pixels. Fig. 5 shows neighborhoods of type 1 ($\mathcal{NT1}$) which have only background pixels. Neighborhoods of type 2 ($\mathcal{NT2}$) have their central pixels in the background and has foreground pixels. Analogously, $\mathcal{NT3}$ and $\mathcal{NT4}$ correspond to $\mathcal{NT2}$ and $\mathcal{NT1}$.

The most outstanding feature of transition pixels is easy to appreciate in a binary image. The difference between the central pixel of neighborhoods $\mathcal{NT1}$ ($\mathcal{NT4}$) and any pixel contained within is zero, because all pixels have the same value (either all zero or all one). However, the neighborhoods $\mathcal{NT2}$ and $\mathcal{NT3}$ contain, besides pixels with differences of 0, pixels with differences of -1 and 1 , respectively. The value -1 is reached in $\mathcal{NT2}$ when the central pixel is compared with a foreground pixel. Likewise, 1 is reached in $\mathcal{NT3}$ when the central pixel is compared numerically with a background pixel. Extending the same argument to non-binary but ideal images, $\mathcal{NT1}$ and $\mathcal{NT4}$ have differences close to zero, unlike $\mathcal{NT2}$ and $\mathcal{NT3}$ whose differences are large in absolute magnitude. Table 1 is constructed taking d_s and d_c from the ideal image definitions and the fact that, the pixel of maximal gray intensity in $\mathcal{NT2}$ is a background pixel while the pixel of minimal gray intensity in $\mathcal{NT3}$ is a foreground pixel.

A *transition function* F is a discriminant function taking extreme values only when a transition pixel is evaluated: Positive for foreground pixels and negative for background pixels. Moreover,

¹ The analysis of the sampling bias is beyond the scope of this paper.

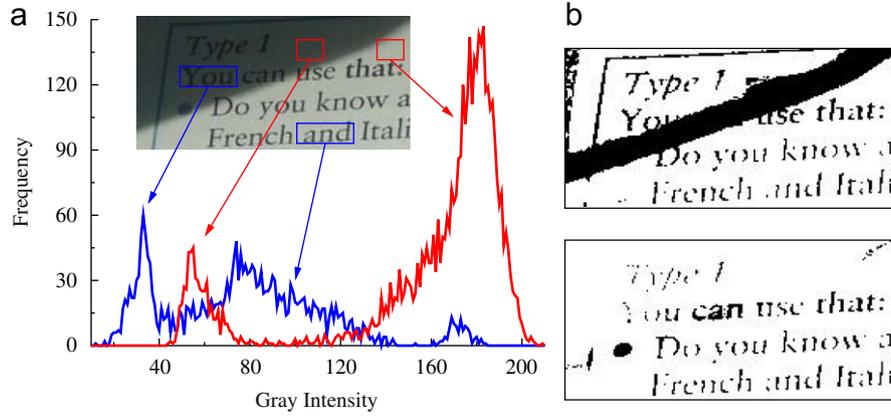


Fig. 4. (a) The histograms of foreground and background pixels are separately drawn. (b) Otsu's method (above) and Sauvola's method with parameters $\alpha = 0.25$ and $\beta = 128$ (below).

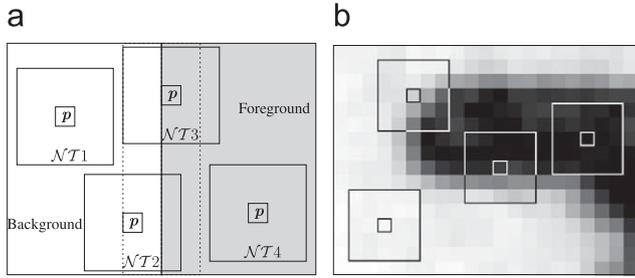


Fig. 5. $\mathcal{N}_t(p)$ has four cases considering the pixels contained and its central pixel.

Table 1
Differences in an ideal image.

Difference	$\mathcal{NT1}$	$\mathcal{NT2}$	$\mathcal{NT3}$	$\mathcal{NT4}$
$\max_{q \in \mathcal{N}_s(p)} I(q) - I(p)$	$< d_s$	$< d_s$	$> d_c$	$< d_s$
$I(p) - \min_{q \in \mathcal{N}_s(p)} I(q)$	$< d_s$	$> d_c$	$< d_s$	$< d_s$
$\max_{q \in \mathcal{N}_s(p)} I(q) + \min_{q \in \mathcal{N}_s(p)} I(q) - 2I(p)$	$> -d_s$ and $< d_s$	$< -d_c + d_s$	$> d_c - d_s$	$> -d_s$ and $< d_s$

pixels in \mathcal{T}_t^c take values close to zero. In terms of conditional probabilities:

$$P(p \in \mathcal{F} \cap \mathcal{T}_t | F(p) \geq t^+) > 1 - \varepsilon_+, \quad (1)$$

$$P(p \in \mathcal{B} \cap \mathcal{T}_t | F(p) \leq -t^-) > 1 - \varepsilon_-, \quad (2)$$

$$P(p \in (\mathcal{T}_t)^c | -t^- < F(p) < t^+) \approx 1 - \varepsilon, \quad (3)$$

where $\varepsilon_+ < 0.5, \varepsilon_- < 0.5$, and $\varepsilon < 0.5$, but the closer they are to zero, the better. Eqs. (1)–(2) mean p is pre-classified as foreground when $F(p)$ is greater than t^+ . In contrast, p is pre-classified as background when $F(p)$ is lower than $-t^-$. Note that there is no information to pre-classify p if $-t^- < F(p) < t^+$.

The previous equations suggest some functions we can use to measure a transition value:

$$\text{Max-min} : V(p) = \max_{q \in \mathcal{N}_s(p)} I(q) + \min_{q \in \mathcal{N}_s(p)} I(q) - 2I(p), \quad (4)$$

$$\text{Discrete Laplace} : L(i, j) = \frac{1}{4}(I(i-1, j) + I(i+1, j) + I(i, j-1) + I(i, j+1)) - I(i, j) \quad \text{and} \quad (5)$$

$$\text{Linear kernel} : G(p) = \sum_{q \in \mathcal{N}_t} \{w(q) \cdot I(q)\} - I(p), \quad \text{with} \quad \sum_{q \in \mathcal{N}_t} w(q) = 1. \quad (6)$$

Note that $V : \{\mathcal{N}_s\} \rightarrow [-1, 1]$. Figs. 6(b)–(d) are computed from Fig. 6(a) applying the transition function. Pixels with negative transition value are shown in red, a pixel with a $-x$ transition value is associated with a x -red intensity. The pixels with positive transition value are shown in blue.

2.4. Max-min function

Proposition 1. Given an image I , suppose that their random variables of background differences x , foreground differences y and contrast differences z have probability densities approximately Gaussian in \mathcal{N}_s , such that $\mu_z > 15\sigma$ where $\sigma = \max\{\sigma_x, \sigma_y, \sigma_z\}$. Then, the max-min function is a transition function in \mathcal{N}_t when $t \leq s$.

Proof. To prove the theorem is sufficient to find t^- and t^+ such that

- $P(V(p) < -t^-) \approx 1$ if $p \in \mathcal{NT2}$,
 $P(V(p) > -t^-) \approx 1$ if $p \in (\mathcal{NT2})^c$,
- $P(V(p) > t^+) \approx 1$ if $p \in \mathcal{NT3}$ and
 $P(V(p) < t^+) \approx 1$ if $p \in (\mathcal{NT3})^c$,

where $(\mathcal{NTi})^c$ represents pixels in all type of neighborhoods, except neighborhood of type \mathcal{NTi} .

We know that practically all the observations drawn from x are within $(\mu_x - 3\sigma_x, \mu_x + 3\sigma_x)$. Explicitly:

$$\begin{aligned} P(-3\sigma_x < x < 3\sigma_x) &= 1 - \varepsilon_x, \\ P(-3\sigma_y < y < 3\sigma_y) &= 1 - \varepsilon_y, \quad \text{and} \\ P(-3\sigma_z < z < 3\sigma_z) &= 1 - \varepsilon_z, \end{aligned} \quad (7)$$

where $\varepsilon_x, \varepsilon_y$, and ε_z are close to zero. Then, in the cases described by Fig. 5 we have:

Neighborhood type 1. All the pixels within $\mathcal{N}_t(p)$ are background. We can rewrite (4) as

$$V(p) = \underbrace{I_{\max}^B(p) - I(p)}_{a_1} - \underbrace{(I(p) - I_{\min}^B(p))}_{a_2}. \quad (8)$$

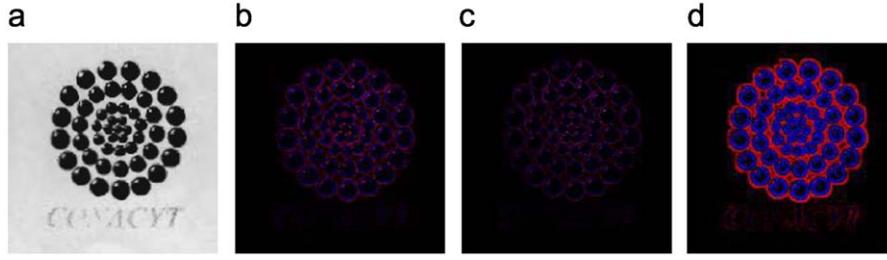


Fig. 6. (a) Original image. (b) $G(p)$ with Gaussian weights ($\sigma^2 = 1$ in $\mathcal{N}_2(p)$). (c) Laplace operator. (d) Max–min in $\mathcal{N}_2(p)$.

Observe that

$$P(a_1 \leq 6\sigma_x) \geq P(|I_{\max}^B(p)| < 3\sigma_x, |I(p)| < 3\sigma_x) = (1 - \varepsilon_x)^2 > 1 - 2\varepsilon_x$$

and $P(a_2 \leq 6\sigma_x) > 1 - 2\varepsilon_x$. Then

$$P(-6\sigma_x \leq V(p) \leq 6\sigma_x) \geq 1 - 4\varepsilon_x. \quad (9)$$

Neighborhood type 2. Both background and foreground pixels within $\mathcal{N}_t(p)$ and p is background. Regardless of outliers, we can assume that the pixel of maximal gray intensity is background and the pixel of minimal gray intensity is foreground. Splitting (4) as:

$$V(p) = \underbrace{I_{\max}^B(p) - I(p)}_{a_1} - \underbrace{(I(p) - I_{\min}^B(p))}_{a_2} - \underbrace{(I_{\min}^B(p) - I_{\max}^F(p))}_{a_3} - \underbrace{(I_{\max}^F(p) - I_{\min}^F(p))}_{a_4}. \quad (10)$$

Notice that $P(a_1 \leq 6\sigma_x) > 1 - 4\varepsilon_x$, $P(a_2 \geq 0) = 1$, $P(a_3 \geq \mu_z - 3\sigma_z) > 1 - \varepsilon_z$, and $P(a_4 \geq 0) = 1$. Then

$$P(V(p) \leq -\mu_z + 6\sigma_x + 3\sigma_z) > 1 - 4\varepsilon_x - \varepsilon_z. \quad (11)$$

Replacing $\mu_z \geq 15\sigma$ in (11)

$$P(V(p) \leq -6\sigma) > 1 - 4\varepsilon_x - \varepsilon_z. \quad (12)$$

From (9) and (12) we conclude that there exists $t^- \leq 6\sigma$ that satisfies (2). The proof of cases three and four are analogous to the proof of case one and two. \square

3. Description of the transition method

The gray-intensity histogram's distribution of \mathcal{B} and \mathcal{F} are approximated from the gray-intensity histogram's distribution of $\mathcal{B} \cap \mathcal{T}$ and $\mathcal{F} \cap \mathcal{T}$. The problem lies in defining a criterion to identify the sets $\mathcal{B} \cap \mathcal{T}$ and $\mathcal{F} \cap \mathcal{T}$. Since $\mathcal{B} \cap \mathcal{T}$ and $\mathcal{F} \cap \mathcal{T}$ are dual sets, we will explain only the criterion for $\mathcal{F} \cap \mathcal{T}$ leaving out details for $\mathcal{B} \cap \mathcal{T}$.

We simplify the notation by omitting r as subindex of \mathcal{T} , writing $\mathcal{T}(p) = \mathcal{N}_r(p) \cap \{\mathcal{T}^+ \cup \mathcal{T}^-\}$, $\mathcal{T}^+(p) = \mathcal{N}_r(p) \cap \mathcal{T}^+$ and $\mathcal{T}^-(p) = \mathcal{N}_r(p) \cap \mathcal{T}^-$.

Our method computes the transition values for all pixels. Once it has managed to find a proper threshold t^+ , it denotes by \mathcal{T}^+ the set of pixels for which transition values are above t^+ . Then, \mathcal{T}^+ is refined by morphological operators in $I(p)$ and $V(p)$ which are applied to gray-intensity images and the image of transition values. Finally, the refined $\mathcal{T}^+(p)$ is used to approximate $\mathcal{F} \cap \mathcal{N}_r(p)$ if its set cardinality is equal or greater than n^+ .

The complete method consists of the following steps:

1. Compute the transition values for each pixel with a transition function. We suggest the max–min function using neighborhoods of radius 2 (Fig. 7(b)).
2. Calculate the thresholds t^+ and t^- (Section 3.1). Take $\mathcal{T}^+ = \{p \mid V(p) \geq t^+\}$ and $\mathcal{T}^- = \{p \mid V(p) \leq t^-\}$ (Figs. 7(c) and (d)).

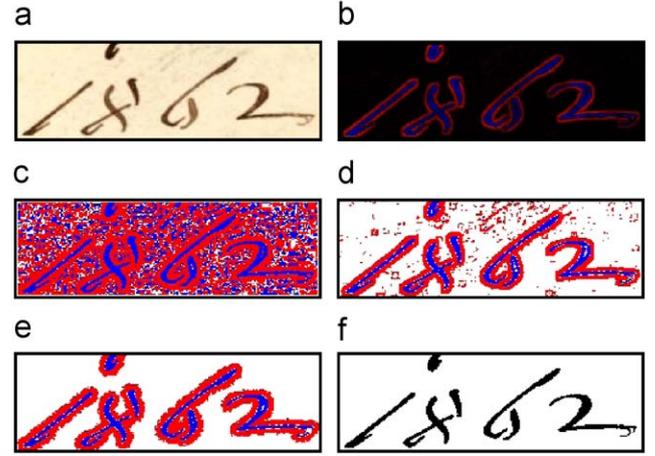


Fig. 7. (a) Original image. (b) Transition image. (c) In red the pixels p that satisfy $V(p) \leq -1$ and in blue $V(p) \geq 1$. (d) The transition image after filtering by $t^+ = 22$ and $t^- = 10$. (e) Transition image restored by morphological operators. (f) Binary image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Refine \mathcal{T}^+ and \mathcal{T}^- (Section 3.2), see Fig. 7(e).
4. Label p using information from \mathcal{T}^+ and \mathcal{T}^- (Section 3.3).
5. Compute $\mathcal{T}(p)$ (Section 3.3.2).
6. Remove noise by standard algorithms.

3.1. Transition threshold

The significance of a high transition value depends on the statistical distributions of background, foreground and contrast differences. Unfortunately, those distributions are only known when a prior information is available. Even if we know the distribution's model, their parameters are hardly fixed.

Several kind of histograms $H_{F,S}$ will be used to keep track of the pixels in relation to their transition value or gray intensity. We have to specify the function F that is used on the pixel set \mathcal{S} . Then, $H_{F,S}(x)$ is the cardinality of $\{p \in \mathcal{S} \mid F(p) = x\}$.

To compute a positive transition threshold, only the positive part of $H_{V,B \cup F}$ is considered. Since the frequency of positive transition values into $\mathcal{B} \cap \mathcal{T}$ is not significant (most of them have high negative values), $H_{V,B \cup F}$ is made out of pixels within $(\mathcal{B} \cup \mathcal{F}) \cap \mathcal{T}$ and $\mathcal{F} \cap \mathcal{T}$ set (see Fig. 8). The histogram $H_{V,B \cup F}$ in Fig. 8 shows the frequency of positive transition values in the whole image, observe that $(\mathcal{B} \cup \mathcal{F}) \cap \mathcal{T}$ constitutes most of the first histogram values while $\mathcal{F} \cap \mathcal{T}$ forms most of the last histogram values.²

² An explanation of the small gray peak in the left of $H_{V,B \cup F}$ is beyond the scope of this paper.

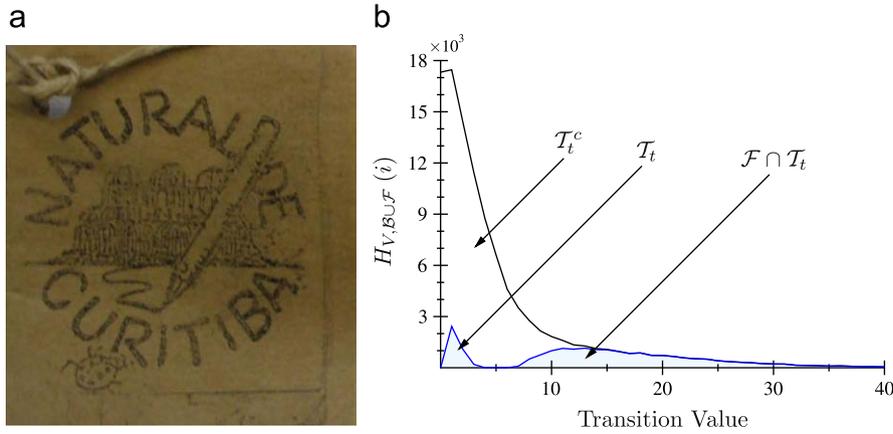


Fig. 8. (a) Original image. (b) Histogram $H_{V,B \cup F}$ of positive transition values.

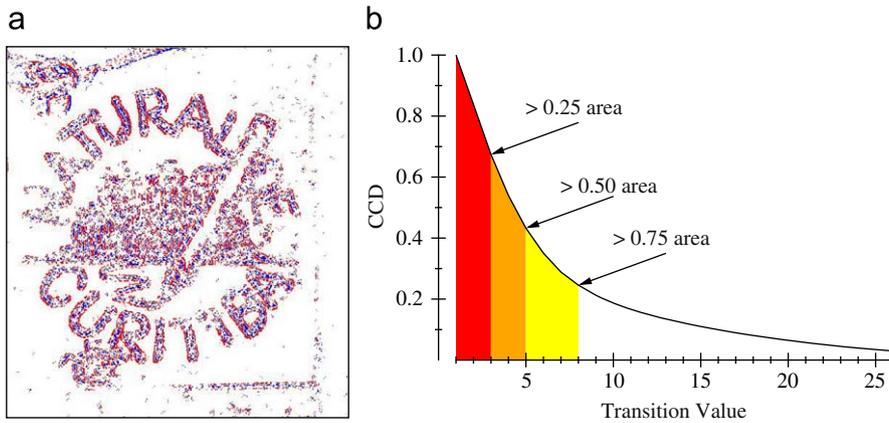


Fig. 9. Using $\alpha^+ = 0.75$ and $\alpha^- = 0.75$ Fig. 8(a) is filtered and displayed as image (a). (b) Complementary cumulative distribution (CCD) function of the positive transition histogram.

3.1.1. Quantile model for transition threshold

We showed in the previous section that values from $\mathcal{F} \cap \mathcal{T}$ are shifted toward the right of the transition-values histogram, see Fig. 8(b). Hence, discarding the pixels with the lowest α^+ percent of transition values approximates $\mathcal{F} \cap \mathcal{T}$. It implies that an $1-\alpha^+$ percent of the highest transition values remain in \mathcal{T}^+ (Fig. 9(c)). This thresholding is called *Quantile threshold*.

Given a value α^+ and $H_{V,B \cup F}$, we choose t^+ as the minimum value that satisfies

$$\frac{1}{k} \sum_{i=1}^{t^+} H_{V,B \cup F}(i) \geq \alpha^+, \quad \text{with } k = \sum_{i=1}^l H_{V,B \cup F}(i). \quad (13)$$

3.2. Restoration of transition set

The restoration of transition set is the process to add and remove pixels from \mathcal{T} with the aim of increasing the cardinality while reducing the noise. In the following sections, the morphological operators use \mathcal{N}_k neighborhood, *cross neighborhood* (left, right, top, and bottom neighbors), *diagonal neighborhood* (left-top, right-top, left-bottom, and right-bottom neighbors), two positive integer a and b , and two positive real values c and d .

Isolation operator:

Given $p \in \mathcal{T}^+$, set $\mathcal{T}^+ \leftarrow \mathcal{T}^+ \setminus \{p\}$ if $|\mathcal{T}^+(p)| < a$.

Given $p \in \mathcal{T}^-$, set $\mathcal{T}^- \leftarrow \mathcal{T}^- \setminus \{p\}$ if $|\mathcal{T}^-(p)| < b$.

Expansion operator:

Given $p \notin \mathcal{T}^+$ and $p \notin \mathcal{T}^-$ set:

- (1) $\mathcal{T}^+ \leftarrow \mathcal{T}^+ \cup \{p\}$ if $|\mathcal{T}^+(p)| \geq a$ and $|\mathcal{T}^-(p)| < b$.
- (2) $\mathcal{T}^- \leftarrow \mathcal{T}^- \cup \{p\}$ if $|\mathcal{T}^+(p)| < a$ and $|\mathcal{T}^-(p)| \geq b$.

Relative operator:

Given $p \notin \mathcal{T}^+$ and $p \notin \mathcal{T}^-$ set:

- (1) $\mathcal{T}^+ \leftarrow \mathcal{T}^+ \cup \{p\}$ if $V(p)/\mu_{1,\mathcal{N}_r(p)} \geq \mu_{V,\mathcal{T}^+}/\mu_{1,B \cup F}$ and $V(p) \geq c$.
- (2) $\mathcal{T}^- \leftarrow \mathcal{T}^- \cup \{p\}$ if $V(p)/\mu_{1,\mathcal{N}_r(p)} \leq \mu_{V,\mathcal{T}^-}/\mu_{1,B \cup F}$ and $V(p) \leq -d$.

Fig. 10(c) was computed with seven transition operators: expansion ($k=2, a=3, b=3$), isolation (cross neighborhood), expansion ($k=2, a=3, b=13$), isolation (diagonal neighborhood), isolation (cross neighborhood), expansion ($k=1, a=5, b=5$), and expansion ($k=2, a=13, b=2$). All the isolate operators set $a=b=1$.

3.3. Positive and negative sets

Only the pixels in $\mathcal{T}^+(p)$ and $\mathcal{T}^-(p)$ are considered to compute $\mathcal{T}(p)$. It is expected that $\mathcal{T}(p) = \emptyset$ if $\mathcal{N}(p)$ is either $\mathcal{NT}1$ or $\mathcal{NT}4$ while $\mathcal{T}(p)$ may have a large cardinality if p is a transition pixel. In

the same way, most of the outliers and small stains are discarded by labeling as background those pixels p that satisfies either $|\mathcal{T}^+(p)| < n^+$ or $|\mathcal{T}^-(p)| < n^-$. The integer values n^+ and n^- depend on the size of r and objects of interest. In general, as high n^+ as large objects from the foreground will be removed. For instance, consider a horizontal line with height 1 as the foreground. The line's extremes are evaluated if $n^+ \leq r+1$. Otherwise, the line's extremes are labeled as background without even compute the $T(p)$. We suggest $n^+ = n^- = 5$, if we have small foreground objects.

A second criterion to discard outliers uses the difference between the gray-intensity means of the transition set. The pixel p is labeled as background if

$$\mu_{I, \mathcal{T}^-(p)} - \mu_{I, \mathcal{T}^+(p)} < c, \quad \text{with } \mu_{I, \mathcal{T}^+(p)} = \frac{1}{|\mathcal{T}^+(p)|} \sum_{q \in \mathcal{T}^+(p)} I(q). \quad (14)$$

The positive value c depicts the minimum contrast expected between the background and foreground. We suggest $c = 15$.

3.3.1. Classic thresholding with transition set

Algorithms like Otsu's [11] and Kittler's [10] increase their accuracy as much as the gray-intensity histogram approximates a bimodal curve. Those algorithms compute $T(p)$ using $H_{I, \mathcal{N}_r(p)}$, we propose using $H_{I, \mathcal{T}(p)}$ instead. Moreover, keeping a track of $H_{I, \mathcal{T}^+(p)}$ and $H_{I, \mathcal{T}^-(p)}$ we are able to determine classification error functions (see Fig. 11). The constant

$$k^+ = \sum_{i=0}^l H_{I, \mathcal{T}^+(p)}(i) \quad \text{and} \quad k^- = \sum_{i=0}^l H_{I, \mathcal{T}^-(p)}(i) \quad (15)$$

are used in both of the following thresholds.

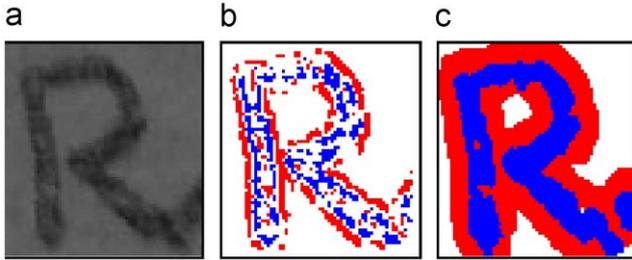


Fig. 10. (a) Original image. (b) Filtered image using $t^+ = 14$ and $t^- = 16$. (c) Restored image by isolation and expansion transition-operators. Blue pixels depict pixels with positive transition value above t^+ . In red, those pixels with high negative transition value below t^- . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- **Minimum-error-rate threshold:**

Given a threshold t , its error function is defined as

$$Error(p, t) = \frac{1}{k^-} \sum_{i=0}^t H_{I, \mathcal{T}^-(p)}(i) + \frac{1}{k^+} \sum_{i=t+1}^l H_{I, \mathcal{T}^+(p)}(i) \quad (16)$$

and its optimal threshold is

$$T(p) = \underset{t \in [1, l-1]}{\operatorname{argmin}} \{Error(p, t)\}. \quad (17)$$

- **Maximum-rate threshold:**

Given a threshold t , its rate classification is defined as

$$Accu(p, t) = \frac{1}{k^+} \sum_{i=0}^t H_{I, \mathcal{T}^+(p)}(i) + \frac{1}{k^-} \sum_{i=t+1}^l H_{I, \mathcal{T}^-(p)}(i) \quad (18)$$

and its optimal threshold is

$$T(p) = \underset{t \in [1, l-1]}{\operatorname{argmax}} \{Accu(p, t)\}. \quad (19)$$

3.3.2. Statistical thresholds

The probability density function which models the gray-intensity histogram of the positive transition pixels in $\mathcal{T}^+(p)$, must approximate the gray-intensity histogram of the foreground pixels in $\mathcal{N}_r(p)$ (see Fig. 11(b)):

$$\frac{H_{I, \mathcal{T}^+(p)}(i)}{|\mathcal{T}^+(p)|} \approx \frac{H_{I, \mathcal{F} \cap \mathcal{N}_r(p)}(i)}{|\mathcal{F} \cap \mathcal{N}_r(p)|}. \quad (20)$$

- **Normal threshold:**

Given $D_{I, \mathcal{T}^+(p)}(i) \approx c_+ \phi(i; \mu_+, \sigma_+^2)$ where $\phi(i; \mu_+, \sigma_+^2)$ denotes the normal probability density function. Then, $D_{I, \mathcal{T}^+(p)}$ approximates $H_{I, \mathcal{T}^+(p)}$ when

$$c_+ = |\mathcal{T}^+(p)|,$$

$$\mu_+ = \mu_{I, \mathcal{T}^+(p)},$$

$$\sigma_+^2 = \max(\sigma_{I, \mathcal{T}^+(p)}^2, 1), \quad (21)$$

where

$$\sigma_{I, \mathcal{T}^+(p)}^2 = \frac{1}{|\mathcal{T}^+(p)| - 1} \sum_{q \in \mathcal{T}^+(p)} (I(q) - \mu_{I, \mathcal{T}^+(p)})^2. \quad (22)$$

The intersection of those curves is the root $\mu_+ < x_0 < \mu_-$ of a quadratic equation with coefficients a , b and c given by

$$a = \frac{1}{\sigma_+^2} - \frac{1}{\sigma_-^2},$$

$$b = \frac{2\mu_-}{\sigma_-^2} - \frac{2\mu_+}{\sigma_+^2},$$

$$c = \frac{\mu_+^2}{\sigma_+^2} - \frac{\mu_-^2}{\sigma_-^2} - 2 \ln \left(\frac{\sigma_- c_-}{\sigma_+ c_+} \right). \quad (23)$$

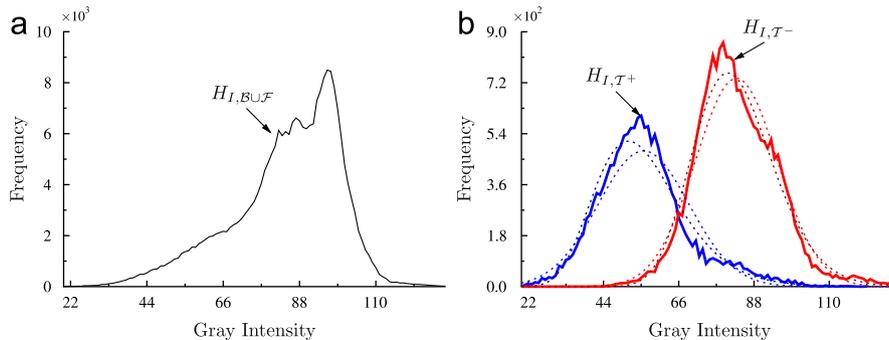


Fig. 11. (a) $H_{I, B \cup F}$ of Fig. 9(a). (b) Histograms of the image filtered by $t^+ = 11$ and $t^- = 12$. $H_{I, \mathcal{T}^+ \cup \mathcal{T}^-}$ is split in H_{I, \mathcal{T}^+} (blue curve) and H_{I, \mathcal{T}^-} (red curve). Each histogram is modeled by normal and lognormal functions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

• **Lognormal threshold:**

Given $D_{I,T^+(p)}(i) \approx c_+ \lambda(i; \mu_+, \sigma_+^2)$ where $\lambda(i; \mu, \sigma^2)$ denotes the lognormal probability density function:

$$\lambda(x; \mu, \sigma) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right).$$

The intersection is computed using the root x_0 of the quadratic equations with coefficients given by (23). This corresponds to $\exp(x_0)$.

In practice, the parameters μ_+ and σ_+^2 are estimated based on the estimated mean and variance of the lognormal distribution using the relations:

$$\mu_+ = \ln(\mu_{I,T^+(p)}) - \frac{1}{2}\sigma_+^2 \quad \text{and} \quad \sigma_+^2 = \ln\left(1 + \frac{\sigma_{I,T^+(p)}^2}{[\mu_{I,T^+(p)}]^2}\right). \quad (24)$$

• **Linear mean-variance threshold:**

$$T(p) = \frac{\mu_{I,T^+(p)} + \alpha\sigma_{I,T^+(p)} + \mu_{I,T^-(p)} - \beta\sigma_{I,T^-(p)}}{2}, \quad (25)$$

where α and β are two tuning parameters.

• **Autolinear threshold:**

$$T(p) = \mu_{I,T^+(p)} + \frac{\sigma_+}{\sigma_+ + \sigma_-} (\mu_{I,T^-(p)} - \mu_{I,T^+(p)}), \quad (26)$$

where σ_+ and σ_- are computed as (21).

4. Complexity

Since the transition method requires a standard storage, we will only show the running-time complexity of the transition method. Hence, we estimate the running time of the algorithms in terms of the number of pixels of the image that is input. If $F(A)$ denotes the total amount of computational labor that the algorithm does in order to compute A . Then, using the conventional Big-O notation:

$$F(V), F(G) \in O(|B \cup \mathcal{F}| \cdot |\mathcal{N}_s|),$$

where G and V are defined by (6) and (4). In the practice, V has a higher running-time than G . It is a consequence of the conditional operations that are involved in the calculation of V .

The quantile threshold has a computational order of $O(|B \cup \mathcal{F}|)$ because it needs read I once to compute $H_{V,\mathcal{F} \cup B}$. All morphological our transition-operators and statistical gray thresholds, have complexity $O(|B \cup \mathcal{F}|)$. An outline proof is shown in Section 4.2.

If a classification rate thresholding is used, the complexity raises to $O(l \cdot |B \cup \mathcal{F}|)$ or $O(r \cdot |B \cup \mathcal{F}|)$ (Section 4.2). In conclusion,

Transition method complexity =

$$\begin{cases} O(l \cdot |B \cup \mathcal{F}|) & \text{if } l-2 > 4r+2, |\mathcal{N}_s|, \\ O(r \cdot |B \cup \mathcal{F}|) & \text{if } 4r+2 > l-2, |\mathcal{N}_s|, \\ O(|B \cup \mathcal{F}| \cdot |\mathcal{N}_s|) & \text{otherwise.} \end{cases}$$

4.1. Integral matrix for a pixel sample

We extend the *integral image* [20] to compute any statistical moment efficiently. To compute the mean and variance of any neighborhood given $S \subset B \cup \mathcal{F}$. Define the image S

$$S(i, j) = \begin{cases} 1 & \text{if } (i, j) \in S, \\ 0 & \text{otherwise.} \end{cases}$$

The integral image of S is

$$\tilde{S}(i, j) = S(i, j) + \tilde{S}(i, j-1) + \tilde{S}(i-1, j) - \tilde{S}(i-1, j-1) \text{ (Fig. 12).}$$

Then, the cardinality of $\mathcal{N}_r(i, j)$ is computed as

$$|\mathcal{N}_r(i, j)| = \tilde{S}(i+r, j+r) - \tilde{S}(i-r-1, j) - \tilde{S}(i, j-r-1) + \tilde{S}(i-r-1, j-r-1).$$

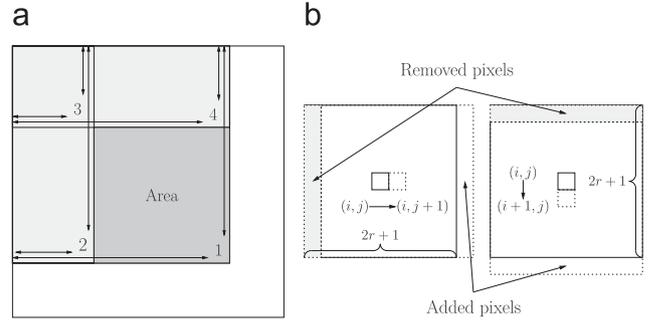


Fig. 12. (a) Using the integral image the area is computed as $1+3-2-4$. (b) $H_{I,T^+(i+1,j)}$ is computed adding and removing pixels from $H_{I,T^+(i,j)}$.

Similarly, considering only the pixels in S , \tilde{I} is the integral image of gray intensities of I given by

$$\tilde{I}(i, j) = I(i, j) \cdot S(i, j) + \tilde{I}(i, j-1) + \tilde{I}(i-1, j) - \tilde{I}(i-1, j-1).$$

The gray-intensity mean $\mu(i, j)$ of $\mathcal{N}_r(i, j)$ is given by

$$\mu(i, j) = \frac{\tilde{I}(i+r, j+r) - \tilde{I}(i-r-1, j) - \tilde{I}(i, j-r-1) + \tilde{I}(i-r-1, j-r-1)}{|\mathcal{N}_r(i, j)|}.$$

The second integral matrix

$$\tilde{I}^2(i, j) = [I(i, j)]^2 \cdot S(i, j) + \tilde{I}^2(i, j-1) + \tilde{I}^2(i-1, j) - \tilde{I}^2(i-1, j-1)$$

is needed to compute the gray-intensity variance of $\mathcal{N}_r(i, j)$. The computational formula for the variance follows in a straightforward manner from the linearity of expected values

$$\sigma^2(X) = E(X^2) - [E(X)]^2,$$

where E is the expected value. Then

$$\sigma^2(p) = \frac{1}{|\mathcal{N}_r(i, j)|} (\tilde{I}^2(i+r, j+r) - \tilde{I}^2(i-r-1, j) - \tilde{I}^2(i, j-r-1) + \tilde{I}^2(i-r-1, j-r-1)) - [\mu(i, j)]^2. \quad (27)$$

In fact, we can extend the integral image's definition as

$$\tilde{I}^n(i, j) = [I(i, j)]^n \cdot S(i, j) + \tilde{I}^n(i, j-1) + \tilde{I}^n(i-1, j) - \tilde{I}^n(i-1, j-1)$$

and compute any statistical moment.

4.2. Gray-threshold complexity

Statistical thresholds use the mean and variance of local gray-intensity histogram. In order to achieve those values, six integral matrices are computed over the binary images of T^+ and T^- (three for each one). Then, all of them have a complexity $O(|B \cup \mathcal{F}|)$.

We can compute $H_{I,\mathcal{N}_r(i,j+1)}$ from $H_{I,\mathcal{N}_r(i,j)}$ with effort $4r+2$. This is: $2r+1$ additions to add the pixels next to the right side of $\mathcal{N}_r(i, j)$ plus $2r+1$ subtractions to remove the first column of $\mathcal{N}_r(i, j)$, see Fig. 12(b). Thus, keep a track of the histograms $H_{I,T^+(p)}$ has a cost of $O(r|B \cup \mathcal{F}|)$. However, for each pixel, the minimum-error rate and maximum rate thresholds search the optimal threshold between $l-2$ values having both a complexity $O(l \cdot |B \cup \mathcal{F}|)$.

5. Experimental results

We compare Otsu's, Sauvola's and Kavallieratou's methods with three variants of the transition method: quantile autolinear, quantile lognormal and quantile normal. It must be remarked we implemented Otsu's in the local version to increase the accuracy,

despite this implementation raised dramatically the running-time. We implemented all the algorithms with integral matrix method to compute local values but Otsu's method, which uses histogram tracking as in Fig. 12(b).

Table 2 presents the parameter's values used in our tests. We post-process all binarized images removing from the foreground small stains (connected components containing four or less pixels) before computing any comparison measure. Only the highest measure score is reported for each pair image-algorithm.

Quantile autolinear, quantile lognormal and quantile normal are composite algorithms with the following operations:

- Max–min function using \mathcal{N}_2 because it obtains the largest and sharply defined transition set among the transition functions described.
- Quantile threshold.
- Two isolation operators ($a = b = 1$). The former using cross neighborhood, the later using diagonal neighborhood.
- Autolinear or lognormal or normal thresholding. Setting $n_+ = n_- = 5$ and $c = 15$.

5.1. Description of the experiments

Historical documents present several kind of image degradations, such as ink stains, burned areas, weak ink strokes and wide variations in background. That is why, we tested the binarization algorithms with the historical atlas “*Theatrum orbis terrarum, sive, Atlas novus*” (Blaeu Atlas) [22]. Those maps contain no-text areas, headers, region's labels, map's comments and digits. The non-text areas are measurable by segmentation criteria, while the rest may be recognized by an optical character recognition (OCR) system. This paper reports three experiments: the first measures the segmentation quality in no-text areas, the second measures the recognition rate of an OCR system, the third experiment compares the algorithm's running-time.

The first and third experiments use 61 maps that were scanned from the Blaeu Atlas at 150dpi resolution. The second uses 83 text-images extracted from the 61 maps. The text-images are

Table 2

Parameter's range.

Algorithm	From/to	Increment
Kavallieratou's	0/9	1
Q. Autolinear	$\alpha : 0.1/0.975$	0.025
Q. Lognormal	$\alpha : 0.1/0.975$	0.025
Q. Normal	$\alpha : 0.1/0.975$	0.025
Sauvola's	$\alpha : 0.025/0.6, \beta : 128$	$\alpha : 0.025$

composed mainly by map headers, map comments and region labels lacking of city labels or handwriting text, see Fig. 13.

Since there is not an available standard benchmark for binarization, this article measures the segmentation accuracy with a variant of the *region non-uniformity* (NU) measure [9]. We named our measure *uniform variance* (UV) which is defined as

$$UV = \sum_{p \in \mathcal{B} \cup \mathcal{F}} \frac{[\sigma_{\mathcal{B}}(p) \cdot |\mathcal{B} \cap \mathcal{N}_r(p)| + \sigma_{\mathcal{F}}(p) \cdot |\mathcal{F} \cap \mathcal{N}_r(p)|]}{|\mathcal{N}_r(p)|},$$

where $\sigma_{\mathcal{B}}(p)$ is the gray-intensity standard deviation of pixels within $\mathcal{B} \cap \mathcal{N}_r(p)$ or zero if the intersection is empty. Similarly $\sigma_{\mathcal{F}}(p)$ is defined. Unlike NU, UV measure is locally well-defined and penalize the gray-intensity variance in both foreground and background. Notice the UV measure is zero in binary images if the foreground match with the set of black pixels. If any black pixel is classified as background, the UV measure increases its value. It is easy to see UV measure behaves similarly in binary images as in gray images. Nevertheless, the UV measure cannot be compared between two different images, because there is not uniform compensation between them. For that reason, only tables of pairwise comparison are reported.

Besides of the UV measure, we used TopOCR [23] to recognize the text from the text-images. Our evaluation's measures are *accuracy* (AC) and *precision* (PR) [24] computed as

$$AC = \frac{\#(\text{characters of } T_{\text{match}})}{\#(\text{characters of } T_{\text{in}})} \quad \text{and} \quad PR = \frac{\#(\text{characters of } T_{\text{match}})}{\#(\text{characters of } T_{\text{out}})},$$

where T_{in} is the original text in the image, T_{out} denotes the recognized text from the OCR, and T_{match} denotes the maximum matching text between T_{in} and T_{out} . T_{match} is computed using Needleman–Wuntsh algorithm [25]. AC measure is an important measure for OCR's system, because a high AC's value increases the possibility to extract, by further algorithms, relevant information.

We implemented the algorithms in C++ and ran our tests on a computer with 3.2GHz Pentium IV Dual core processor and 2GB in RAM. Table 4 presents the 95% confidence intervals for algorithm's running-time expressing the interval limits on millisecond/megapixel.

5.2. Pairwise tables

Quantile lognormal performed the best in the first experiment (Table 3) having a reasonable running-time lower than 3s/megapixel (Table 4). Otsu's and quantile normal performed in 2nd and 3rd place. However, Otsu's running-time (over 30 min) is a disadvantage when the system's running-time is limited like in reader systems.

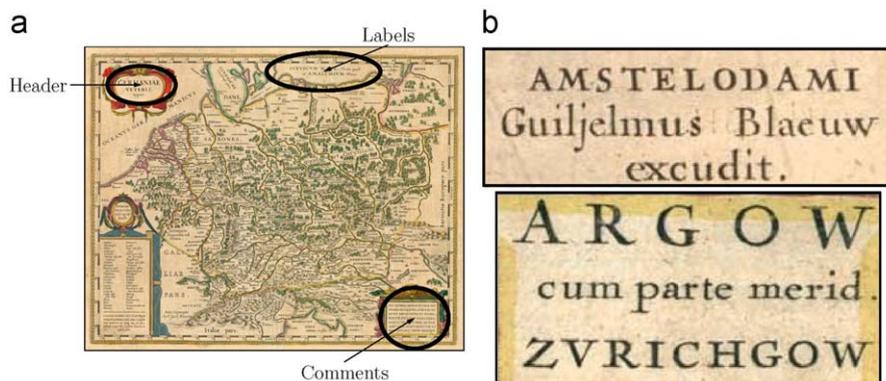


Fig. 13. Example of header, labels a comment in a historical map.

Table 3
Pairwise comparison for 61 maps.

	Kavallieratou	Otsu	Q. Autolinear	Q. Lognormal	Q. Normal	Sauvola
Kavallieratou	–	8	14	2	4	0
Otsu	53	–	45	22	36	35
Q. Autolinear	47	16	–	10	1	7
Q. Lognormal	59	39	51	–	42	40
Q. Normal	57	25	60	19	–	34
Sauvola	61	26	54	21	27	–

Table 4
95% confidence intervals for binarization-running time.

	Raw	Normalized
Kavallieratou	(1718,1727)	(4.0,4.0)
Otsu	(265 757,266 908)	(630.9,630.9)
Q. Autolinear	(1802,1813)	(4.2,4.2)
Q. Lognormal	(2568,2580)	(6.1,6.1)
Q. Normal	(2039,2051)	(4.8,4.8)
Sauvola	(421,423)	(1.0,1.0)

The intervals are normalized respect Sauvola's running-time (millisecond/megapixel) which is the fastest.

Table 5
UV's pairwise comparison for 83 text-images.

	Kavallieratou	Otsu	Q. Autolinear	Q. Lognormal	Q. Normal	Sauvola
Kavallieratou	–	12	1	1	1	0
Otsu	71	–	6	3	3	0
Q. Autolinear	82	77	–	2	19	3
Q. Lognormal	82	80	81	–	56	8
Q. Normal	82	80	64	27	–	4
Sauvola	83	83	80	75	79	–

Table 6
Pairwise comparison for 83 text-images.

	Kavallieratou	Otsu	Q. Autolinear	Q. Lognormal	Q. Normal	Sauvola
Kavallieratou	–/–	55/ 36	8/7	10/3	4/6	10/5
Otsu	18/44	–/–	1/4	2/6	1/3	1/1
Q. Autolinear	58/74	70/ 74	–/–	25/23	13/22	33/26
Q. Lognormal	56/76	71/ 71	29/41	–/–	17/39	31/31
Q. Normal	61/74	71/ 74	27/28	25/25	–/–	33/21
Sauvola	54/75	69/ 76	21/41	22/39	16/39	–/–

In each cell, the left number is the AC measure, while PR measure is given by the right value. The sum of complementary cells are not always the total of evaluated images because some algorithms have even scores in several images.

Unexpectedly, Sauvola's has the highest UV's score for text-images (Table 5) but it scored, at the same time, poorly in the AC measure. Quantile methods have performed better than other algorithms in AC measure. Regarding PR measure, quantile methods have performed only below Sauvola's method (Table 6).

6. Conclusions

We proposed a new flexible framework capable of solving the binarization problem more efficiently than classic binary threshold methods. This methodology is based on the new concepts of transition pixel, transition set and transition values. Transition values are computed by transition functions that can be specialized in any specific image. Particularly using \mathcal{N}_2 neighborhoods, max–min function (4) associates high values to transition pixels in comparison with no-transition pixels. So that, the quantile method is able to compute an accurate threshold, which selects a representative samples of both background and foreground sets. We presented a wide survey of gray-level thresholding which were adapted to use the transition set approximation. Even more, the transition method accepts new models easily for each method's step.

Our test used a large database of historical maps containing text and pictorial document images under different levels of degradations. Quantile lognormal performed the best in both quality tests. On the other hand, quantile normal had a similar performance compared with standard binarization algorithms. The transition method has satisfactory running time processing 1 megapixel in 3 s or faster.

Acknowledgments

We thank the anonymous referees for helpful and relevant comments, and Edgar Duñez for useful discussions. This research was partially supported by The National Council on Science and Technology (CONACYT) of Mexico (Grant number: 218253).

References

- [1] H. Kohmura, T. Wakahara, Determining optimal filters for binarization of degraded characters in color using genetic algorithms, in: ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition, IEEE Computer Society, Washington, DC, USA, 2006, pp. 661–664.
- [2] J. He, Q.D.M. Do, A.C. Downton, J.H. Kim, A comparison of binarization methods for historical archive documents, in: ICDAR '05: Proceedings of the Eighth International Conference on Document Analysis and Recognition, IEEE Computer Society, Washington, DC, USA, 2005, pp. 538–542.
- [3] J.S. Noh, K.H. Rhee, Palmprint identification algorithm using Hu invariant moments and Otsu binarization, in: ICIS '05: Proceedings of the Fourth Annual ACIS International Conference on Computer and Information Science, IEEE Computer Society, Washington, DC, USA, 2005, pp. 94–99.
- [4] C. Mello, A. Sanchez, A. Oliveira, A. Lopes, An efficient gray-level thresholding algorithm for historic document images, Journal of Cultural Heritage 9 (2) (2008) 109–116.
- [5] R. Milewski, V. Govindaraju, Binarization and cleanup of handwritten text from carbon copy medical form images, Pattern Recognition 41 (4) (2008) 1308–1315.
- [6] P.K. Sahoo, S. Soltani, A.K. Wong, Y.C. Chen, A survey of thresholding techniques, Computer Vision, Graphics and Image Processing 41 (2) (1988) 233–260.
- [7] J. Sauvola, M. Pietikinen, Adaptive document image binarization, Pattern Recognition 33 (2) (2000) 225–236.
- [8] Ø.D. Trier, A.K. Jain, Goal-directed evaluation of binarization methods, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (12) (1995) 1191–1201.
- [9] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging 13 (1) (2004) 146–168.
- [10] J. Kittler, J. Illingworth, Minimum error thresholding, Pattern Recognition 19 (1) (1986) 41–47.
- [11] N. Otsu, A threshold selection method from grey-level histograms, IEEE Transaction on Systems, Man, and Cybernetics 9 (1) (1979) 62–66.
- [12] W. Niblack, An Introduction to Digital Image Processing, Prentice Hall, Birkeroed, Denmark, Denmark, 1985.
- [13] Y. Li, C.Y. Suen, M. Cheriet, A threshold selection method based on multiscale and graylevel co-occurrence matrix analysis, in: ICDAR '05: Proceedings of the Eighth International Conference on Document Analysis and Recognition, IEEE Computer Society, Washington, DC, USA, 2005, pp. 575–579.

- [14] Q. Chen, Q.-s. Sun, P. Ann Heng, D.-s. Xia, A double-threshold image binarization method based on edge detector, *Pattern Recognition* 41 (4) (2008) 1254–1267.
- [15] J. Canny, A computational approach to edge detection, *Pattern analysis and machine intelligence* 8 (1) (1986) 679–698.
- [16] E. Kavallieratou, A binarization algorithm specialized on document images and photos, in: *ICDAR '05: Proceedings of the Eighth International Conference on Document Analysis and Recognition*, IEEE Computer Society, Washington, DC, USA, 2005, pp. 463–467.
- [17] 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.
- [18] J.M. White, G.D. Rohrer, Image thresholding for optical character recognition and other applications requiring character image extraction, *IBM Journal of Research and Development* 27 (4) (1983) 400–411.
- [19] S.J. Lu, C.L. Tan, Binarization of badly illuminated document images through shading estimation and compensation, in: *ICDAR '07: Proceedings of the Ninth International Conference on Document Analysis and Recognition*, IEEE Computer Society, Washington, DC, USA, 2007, pp. 312–316.
- [20] P. Viola, M. Jones, Robust real-time object detection, *International Journal of Computer Vision* (2001) 1–25.
- [21] J. Bernsen, Dynamic thresholding of grey-level images, in: *Proceedings of the Eighth International Conference on Pattern Recognition (ICPR)*, Paris, France, 1986, pp. 1251–1255.
- [22] W. Janszoon, J. Blaeu, *Theatrum orbis terrarum, sive, Atlas novus*, Blaeu Atlas, 1645, URL: <<http://www.library.ucla.edu/yrl/reference/maps/blaeu>>.
- [23] T. Soft, Top OCR, Top Soft, 2008, URL <<http://www.topocr.com/>>.
- [24] M. Junker, A. Dengel, R. Hoch, On the evaluation of document analysis components by recall, precision, and accuracy, in: *Proceedings of the Fifth International Conference on Document Analysis and Recognition (ICDAR '99)*, IEEE Computer Society, Los Alamitos, CA, USA, 1999, p. 713.
- [25] S.B. Needleman, C.D. Wunsch, A general method applicable to the search for similarities in the amino acid sequence of two proteins, *Journal of Molecular Biology* 48 (3) (1970) 443–453.

About the Author—MARTE ALEJANDRO RAMÍREZ ORTEGÓN is a Ph.D. student at the Department of Mathematics and Computer Science of the Freie Universität Berlin. In 2002, he received a BCS degree (Computer Science) from the University of Guanajuato (UG)/the Centre for Mathematical Research (CIMAT), Guanajuato, Mexico (1998–2002). His current research fields include pattern recognition, edge detection, binarization, multithresholding and text recognition.

About the Author—ERNESTO TAPIA is a Researcher of Artificial Intelligence group at the Department of Mathematics and Computer Science of the Freie Universität Berlin. He received a Ph.D. degree from the Freie Universität Berlin for his work on handwriting recognition. Dr. Tapia holds an M.Sc. degree in Applied Mathematics from the Centro de Investigación en Matemáticas (CIMAT), Mexico, and a B.Sc. degree from the Universidad Veracruzana (UV), Mexico. His current research fields include pattern recognition, image processing neural networks, robotics and e-learning.

About the Author—LILIA LETICIA RAMÍREZ RAMÍREZ is a Postdoctoral Fellow at the University of Waterloo. She received the Ph.D. degree in Statistics at the University of Waterloo, Canada, in 2008. Lilia holds an M.S. degree 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 stochastic ‘epidemic’ models, random networks and inference for complex systems.

About the Author—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 (Free University of Berlin), Berlin, Germany, in 2006. Currently, he is a full professor in the Computer Science Department at the University of Guadalajara. His research interests include computer vision, learning systems, and soft computing.

About the Author—RAUL ROJAS received the B.S. and M.S. degrees in Mathematics from the National Polytechnic Institute (IPN), Mexico City, Mexico. He obtained the Ph.D. degree at the Freie Universität Berlin (Free University of Berlin), Berlin, Germany. Currently, 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.