Transition Thresholds for Binarization of Historical Documents

Marte A. Ramírez-Ortegón* and Raúl Rojas
Institut für Informatik,Freie Universität Berlin, Takustr. 9, 14195 Berlin, Germany
* mars.sasha@gmail.com,rojas@inf.fu-berlin.de

***Abstract*—This paper extends the transition method for binarization based on *transition pixels*, a generalization of edge pixels. This method originally computes transition thresholds using the quantile thresholding algorithm, that has a critical parameter. We achieved an automatic version of the transition method by computing the transition thresholds with the Rosin's algorithm. We experimentally tested four variants of the transition method combining the density and cumulative distribution functions of transition values, with gray-intensity thresholds based on the normal and lognormal density functions. The results of our experiments show that these unsupervised methods yields superior binarization compared with top-ranked algorithms.** [1]

***Index Terms*—binarization; transition; method; documents, historical**

I. Introduction

Binarization classifies each pixel in an image either as *foreground* or *background*. The foreground is a subset of pixels $\mathcal{F}$ containing objects used for further analysis and recognition, while the background $\mathcal{B}$ is the complement of $\mathcal{F}$.

The transition method, proposed by Ramírez-Ortegón et al. [1], is a composite binarization algorithm based on the transition pixel concept, a generalization of edge pixels.

Empirical evidence lead us to propose the *Rosin's threshold* [2] in order to avoid the parameter α that was originally used in the transition method. In this manner, we upgrade the transition method to the category of unsupervised method.

The rest of this paper is organized as follows. Section II introduces the related concepts. The transition method is developed in section III and III-A. In section IV, the experiments results are shown. Conclusions are presented in section V.

II. Preliminary concepts

This paper only considers one input image. An image function F can be defined as the mapping $F : \mathcal{P} \rightarrow \mathcal{A}$ where $\mathcal{P} = \{\boldsymbol{p} \mid \boldsymbol{p} \in \mathbb{N} \times \mathbb{N}\}$ and $\mathcal{A} \subset \mathbb{Z}$. $I : \mathcal{P} \rightarrow \mathcal{G} = \{0, 1, \ldots l\}$ is the gray image function. $B : \mathcal{P} \rightarrow \{0, 1\}$ represents the binarization of I, where one is considered as foreground. Local binarization algorithms compute a threshold surface $T : \mathcal{P} \rightarrow \mathcal{G}$ over the whole image: $B(\boldsymbol{p}) = 1$ if $I(\boldsymbol{p})$ is lower than the threshold $T(\boldsymbol{p})$. The information to compute $T(\boldsymbol{p})$ is gathered from the pixels within a square $\mathcal{P}_r(\boldsymbol{p})$ centered at the pixel $\boldsymbol{p}$ of sides with length $2r + 1$.

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

A pixel $\boldsymbol{p}$ is a *t-transition pixel* if the neighborhood $\mathcal{P}_t(\boldsymbol{p})$ contains foreground and background pixels. The set of these pixels is named $\mathcal{P}_t$. If $t = 1$, then the t-transition pixel is an edge pixel.

Transition pixels can be detected by selecting those pixels that have extreme *transition values*. These values can be computed with *max-min* function [1], [3]

$$V(\boldsymbol{p}) = \max_{\boldsymbol{q} \in \mathcal{P}_t(\boldsymbol{p})} I(\boldsymbol{q}) + \min_{\boldsymbol{q} \in \mathcal{P}_t(\boldsymbol{p})} I(\boldsymbol{q}) - 2I(\boldsymbol{p}). \qquad (1)$$

Pixels in the *positive transition set* $\mathcal{F}_t = \mathcal{P}_t \cap \mathcal{F}$ reach extreme positive transition values. Analogously, pixels in the *negative transition set* $\mathcal{B}_t = \mathcal{P}_t \cap \mathcal{B}$ reach extreme negative transition values, see Fig. 1.

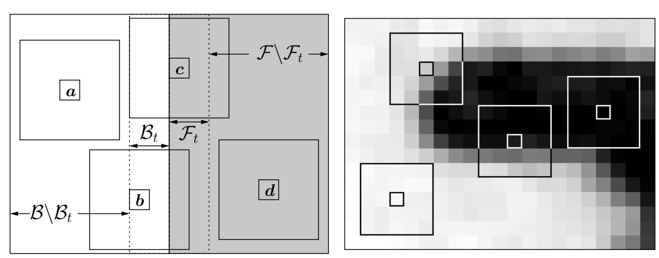

Fig. 1. $\mathcal{N}_t(\boldsymbol{p})$ has four cases considering the pixels contained and its central pixel: $\mathcal{B}\backslash\mathcal{B}_t$, $\mathcal{B}_t$, $\mathcal{F}_t$ and $\mathcal{F}\backslash\mathcal{F}_t$.

We abbreviate $\mathcal{F}_r(\boldsymbol{p}) = \mathcal{F} \cap \mathcal{P}_r(\boldsymbol{p})$, $\mathcal{F}_{t,r}(\boldsymbol{p}) = \mathcal{F}_t \cap \mathcal{P}_r(\boldsymbol{p})$ and $\hat{\mathcal{F}}_{t,r}(\boldsymbol{p}) = \hat{\mathcal{F}}_t \cap \mathcal{P}_r(\boldsymbol{p})$. The sets $\mathcal{P}_{t,r}(\boldsymbol{p})$, $\mathcal{B}_r(\boldsymbol{p})$, $\mathcal{B}_{t,r}(\boldsymbol{p})$, and $\hat{\mathcal{B}}_{t,r}(\boldsymbol{p})$ are defined in a similar way.

III. Overview of the transition method

Since $\mathcal{F}_t$ and $\mathcal{B}_t$ are dual sets, we will only explain models for $\mathcal{F}_t$, leaving out details for $\mathcal{B}_t$.

Consider that Fig. 2 (a) is a neighborhood $\mathcal{P}_r(\boldsymbol{p})$; the left peak of $H_{I,\mathcal{P}_r(\boldsymbol{p})}$ is mainly formed by foreground pixels, while the right peak is formed by background pixels.

20th International Conference on Pattern Recognition, IEEE Computer Society, 2010, 2362 - 2365 August 23-26, 2010. Istambul, Turkey

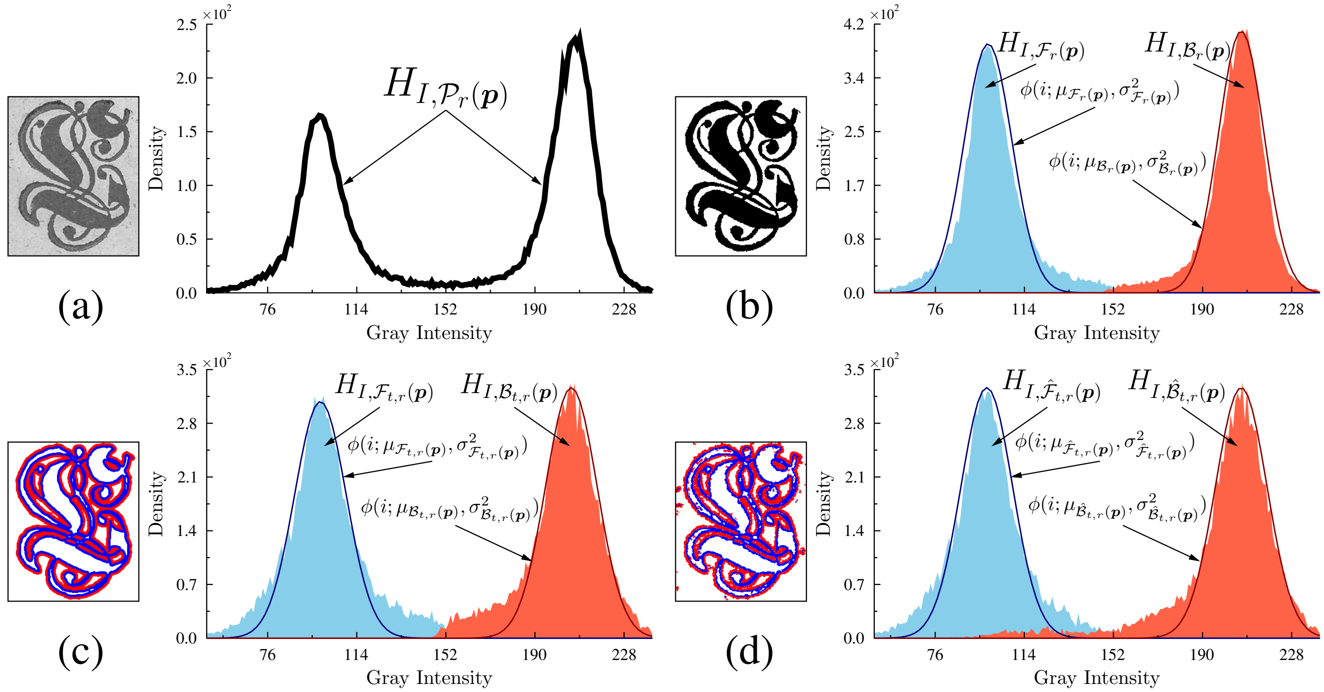

Fig. 2. The gray-intensity distribution of both background and foreground can be approximated by the approximation of transition sets within $\mathcal{P}_r(\boldsymbol{p})$.

If we knew the *class-conditional density* $P(I(\boldsymbol{q}) \mid \boldsymbol{q} \in \mathcal{F}_r(\boldsymbol{p}))$, we could consider the *maximum likelihood estimation* or *Bayesian estimation* approach to solve the binarization problem. Unfortunately, the class-conditional densities are rarely known, nevertheless we can reasonably assume that the gray-intensity density is approximately normally distributed in small neighborhoods $\mathcal{F}_r(\boldsymbol{p})$:

$$H_{I,\mathcal{F}_r(\boldsymbol{p})}(i) \approx c_{\mathcal{F}_r(\boldsymbol{p})}\phi(i;\mu_{\mathcal{F}_r(\boldsymbol{p})},\sigma^2_{\mathcal{F}_r(\boldsymbol{p})}), \tag{2}$$

where $\phi(x;\mu,\sigma)$ is the normal probability density function with mean μ and variance σ^2. $T(\boldsymbol{p})$ is quickly computed if there is an analytic intersection between (2) and the correspondent background function, see Fig. 2 (b). Thus, we can approximate $P(I(\boldsymbol{q}) \mid \boldsymbol{q} \in \mathcal{F}_r(\boldsymbol{p}))$ by drawing a representative sample of $\mathcal{F}_r(\boldsymbol{p})$.

The positive transition set $\mathcal{F}_t$ satisfies

$$P(I(\boldsymbol{q}) \mid \boldsymbol{q} \in \mathcal{F}_r(\boldsymbol{p})) \approx P(I(\boldsymbol{q}) \mid \boldsymbol{q} \in \mathcal{F}_{t,r}(\boldsymbol{p})). \tag{3}$$

Therefore, $\mathcal{F}_{t,r}(\boldsymbol{p})$ is a representative sample of $\mathcal{F}_r(\boldsymbol{p})$, see Fig. 2 (c). Although the transition sets are also unknown, our method provides $\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})$ (Fig. 2(d)) which is an accurate estimate of $\mathcal{F}_r(\boldsymbol{p})$. Then, (3) change to

$$P(I(\boldsymbol{q}) \mid \boldsymbol{q} \in \mathcal{F}_r(\boldsymbol{p})) \approx P(I(\boldsymbol{q}) \mid \boldsymbol{q} \in \hat{\mathcal{F}}_{t,r}(\boldsymbol{p})). \tag{4}$$

We are now able to compute the gray-threshold with usual classification procedures. 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.
2) Calculate the thresholds t_+ and t_-. Take $\hat{\mathcal{F}}_t = \{\boldsymbol{p} \mid V(\boldsymbol{p}) \geq t_+\}$ and $\hat{\mathcal{B}}_t = \{\boldsymbol{p} \mid V(\boldsymbol{p}) \leq -t_-\}$.
3) Restore $\hat{\mathcal{F}}_t$ and $\hat{\mathcal{B}}_t$.
4) Compute the threshold image T and generate the binary image B.
5) Remove noise from B by standard algorithms.

For better understanding of the transition method, the reader is referred to [1].

A. Transition threshold

We found suitable transition thresholds by analyzing functions of transition values: *Empirical scaled density function*

$$u_i = \frac{1}{k}H_{V,\mathcal{P}}(i), \text{ with } k = \max_{i\in[1,l]} H_{V,\mathcal{P}}(i) \tag{5}$$

and *empirical complementary cumulative distribution function*

$$v_i = \frac{1}{t}\sum_{j=i}^{l} H_{V,\mathcal{P}}(j), \text{ with } t = \sum_{j=1}^{l} H_{V,\mathcal{P}}(j). \tag{6}$$

The behavior of (5) and (6) is ideal for Rosin's threshold [2], which proposes a threshold for unimodal histograms.

Rosin's method assumes that one of the two classes produces one dominant peak located at one of the sides of the histogram of gray intensities. The non-dominant class may or may not produce a discernible peak, but needs to be reasonably well separated from the large peak to avoid being swamped by it. A straight line L is drawn from the peak to the lowest non-zero value of the histogram, the threshold point is selected as the histogram index i that maximizes the perpendicular distance between L and the point $(i, H_{F,\mathcal{A}}(i))$ (See Fig. 3).

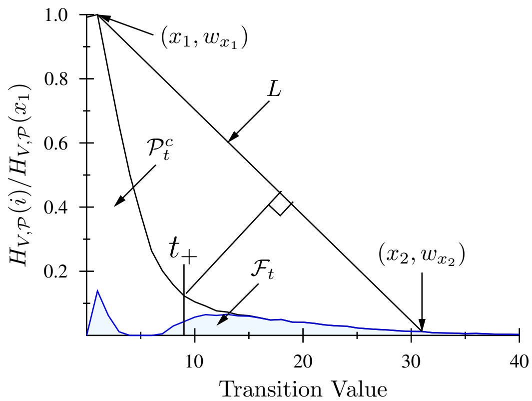

Fig. 3. Rosin's threshold for unimodal histograms.

Let w_i be values computed with either (5), or (6). We compute Rosin' threshold for graph of $w_i's$ which is truncated between x_1 and x_2, where

$$\begin{aligned} x_1 < x_2,\ w_{x_1} > w_i \text{ for } i \neq x_i, \text{ and} \\ \frac{w_{x_2}}{w_{x_1}} \geq \delta > \frac{w_i}{w_{x_1}} \text{ for } i > x_2 \end{aligned} \tag{7}$$

We suggest $\delta = 0.01$). Then, L is defined by the points (x_1, w_{x_1}) and (x_2, w_{x_2}). The distance function and threshold are defined as

$$D(i) = \frac{|(x_2 - x_1)(w_{x_1} - w_i) - (x_1 - i)(w_{x_2} - w_{x_1})|}{\sqrt{(x_2 - x_1)^2 + (w_{x_2} - w_{x_1})^2}} \tag{8}$$

$$t_+ = \underset{i \in [x_1, x_2]}{\arg\max}\, D(i). \tag{9}$$

B. Restoration of transition set

The restoration of the transition set $\hat{\mathcal{P}}_t$ is the process to add and remove pixels from $\hat{\mathcal{P}}_t$ with the aim of increasing the cardinality while reducing the noise.

The *cross* and *diagonal isolate operators* were successfully used on [1] in order to remove false positives of the approximated transition sets. Let us denote $\boldsymbol{p}_{i,j}$ the pixel whose spatial position is the coordinate (i, j). Then given $\boldsymbol{p}_{i,j} \in \hat{\mathcal{F}}_t$, the *cross isolate operator* is defined as

$$\hat{\mathcal{F}}_t \leftarrow \hat{\mathcal{F}}_t \setminus \{\boldsymbol{p}\} \text{ if } \left|\hat{\mathcal{F}}_t \cap \mathcal{N}_\oplus(\boldsymbol{p})\right| = 0 \tag{10}$$

where $\mathcal{N}_\oplus(\boldsymbol{p}_{i,j}) = \{\boldsymbol{p}_{i-1,j}, \boldsymbol{p}_{i+1,j}, \boldsymbol{p}_{i,j-1}, \boldsymbol{p}_{i,j+1}\}$. Instead the cross neighborhood, the *diagonal transition operator* uses the diagonal neighborhood $\mathcal{N}_\otimes(\boldsymbol{p}_{i,j}) = \{\boldsymbol{p}_{i-1,j-1}, \boldsymbol{p}_{i-1,j+1}, \boldsymbol{p}_{i+1,j-1}, \boldsymbol{p}_{i+1,j+1}\}$.

We define the *rectangular transition operator* as:

$$\begin{aligned} \hat{\mathcal{F}}_t \leftarrow \hat{\mathcal{F}}_t \setminus \{\boldsymbol{p}\} \text{ if } \boldsymbol{p} \in \hat{\mathcal{F}}_t \text{ and} \\ \left|\hat{\mathcal{F}}_t \cap \mathcal{N}_{x,y}(\boldsymbol{p})\right| = \left|\hat{\mathcal{F}}_t \cap \mathcal{N}_{x+1,y+1}(\boldsymbol{p})\right|, \end{aligned} \tag{11}$$

where $\mathcal{N}_{y,x}(\boldsymbol{p})$ is defined as the rectangular neighborhood centered at the pixel $\boldsymbol{p}$ of sides with length $2y + 1,\ 2x + 1$.

C. Statistical thresholds

Only pixels in $\hat{\mathcal{P}}_{t,r}(\boldsymbol{p}) = \hat{\mathcal{F}}_{t,r}(\boldsymbol{p}) \cup \hat{\mathcal{B}}_{t,r}(\boldsymbol{p})$ are considered to compute $T(\boldsymbol{p})$. At the same time, outliers are discarded by labeling as background those pixels $\boldsymbol{p}$ that satisfy either $\left|\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})\right| < n_+$ or $\left|\hat{\mathcal{B}}_{t,r}(\boldsymbol{p})\right| < n_-$, where n_+ and n_- depend of r and objects of interest; the higher n_+, the larger the objects that can be removed from the foreground. We suggest $n_+ = n_- = 5$ for detecting small foreground objects.

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

$$\mu_{I,\hat{\mathcal{B}}_{t,r}(\boldsymbol{p})} - \mu_{I,\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})} < c, \tag{12}$$

where c is an integer, which depicts the minimum contrast expected between the foreground and background. We suggest $c = 15$.

Normal threshold

Given $H_{I,\mathcal{F}_{t,r}(\boldsymbol{p})}(i) \approx c_+\phi(i; \mu_+, \sigma_+^2)$ where $\phi(x; \mu, \sigma)$ is the normal probability density function with mean μ and variance σ^2. Then, $H_{I,\mathcal{F}_{t,r}(\boldsymbol{p})}$ is approximated when

$$\begin{aligned} c_+ &= \left|\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})\right| \text{ (complete) or } c_+ = 1 \text{ (simple)} \\ \mu_+ &= \mu_{I,\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})}, \\ \sigma_+^2 &= \max\left(\sigma^2_{I,\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})}, 1\right) \end{aligned} \tag{13}$$

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

$$\begin{aligned} a &= \frac{1}{\sigma_+^2} - \frac{1}{\sigma_-^2}, \quad b = \frac{2\mu_-}{\sigma_-^2} - \frac{2\mu_+}{\sigma_+^2} \\ c &= \frac{\mu_+^2}{\sigma_+^2} - \frac{\mu_-^2}{\sigma_-^2} - 2\ln\left(\frac{\sigma_- \cdot c_-}{\sigma_+ \cdot c_+}\right) \end{aligned} \tag{14}$$

Lognormal threshold

Given $H_{I,\mathcal{F}_{r,t}(\boldsymbol{p})}(i) \approx c_+\lambda\left(i; \mu_+, \sigma_+^2\right)$ where $\lambda\left(\mu, \sigma^2\right)$ denotes the lognormal probability density function.

The intersection of these curves is $\exp(x_0)$, where x_0 is the root of the quadratic equation with coefficients given by (14), but μ_+ and σ_+^2 are estimated based on the estimated mean and variance of the lognormal distribution using the relations:

$$\begin{aligned} \mu_+ &= \ln\left(\mu_{I,\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})}\right) - \frac{1}{2}\sigma_+^2 \text{ and} \\ \sigma_+^2 &= \ln\left(1 + \frac{\sigma^2_{I,\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})}}{\left[\mu_{I,\hat{\mathcal{F}}_{t,r}(\boldsymbol{p})}\right]^2}\right). \end{aligned} \tag{15}$$

IV. Description of experiments

We compare Otsu's, Kittler's and Sauvola's methods(top-ranked algorithms on [4], [5], [6]) with four variants of the transition method: T-DF-L, T-CCD-L, T-DF-N and T-CCD-N.

We implemented both Kittler's and Otsu's methods in their local versions [1] to increase their accuracy. Real applications rarely use more than one parameter's set; this is the main reason we fixed Sauvola's $\alpha = 0.5$ and $\beta = 128$, which are

TABLE I

AC AND PR PAIRWISE COMPARISON. EACH CELL (y-ROW,x-COLUMN) CONTAINS FOUR VALUES n_{yx}, p_{yx}, n'_{yx} AND p'_{yx}, RESPECTIVELY. IN TERMS OF AC, THE NUMBER n_{yx} REPRESENTS THE TIMES THAT THE ALGORITHM y HAS A HIGHER SCORE THAN THE ALGORITHM x, WHILE $p_{yx} = \frac{n_{yx}}{n_{yx}+n_{xy}}$ REPRESENTS THE CONDITIONAL PROBABILITY OF y'S SCORE BEING HIGHER THAN x'S SCORE. SIMILARLY, n'_{yx} AND p'_{yx} ARE DEFINED IN TERMS OF PR MEASURE.

	Kittler	Otsu	Sauvola	T-CCD-L	T-CCD-N	T-DF-L	T-DF-N
Kittler	-	30 (0.39) - 36 (0.44)	49 (0.63) - 54 (0.65)	31 (0.39) - 31 (0.38)	35 (0.44) - 38 (0.46)	29 (0.37) - 26 (0.32)	32 (0.41) - 40 (0.49)
Otsu	46 (0.61)- 45 (0.56)	-	63 (0.84) - 73 (0.90)	25 (0.38) - 29 (0.38)	32 (0.51) - 37 (0.49)	21 (0.32) - 28 (0.35)	31 (0.49) - 37 (0.49)
Sauvola	29 (0.37)- 29 (0.35)	12 (0.16) - 08 (0.10)	-	14 (0.19) - 11 (0.14)	16 (0.21) - 14 (0.17)	10 (0.14) - 09 (0.11)	13 (0.17) - 12 (0.14)
T-CCD-L	49 (0.61)- 51 (0.62)	40 (0.62) - 47 (0.62)	60 (0.81) - 69 (0.86)	-	38 (0.63) - 46 (0.61)	19 (0.40) - 25 (0.43)	38 (0.58) - 51 (0.64)
T-CCD-N	44 (0.56)- 45 (0.54)	31 (0.49) - 39 (0.51)	61 (0.79) - 68 (0.83)	22 (0.37) - 30 (0.39)	-	21 (0.33) - 29 (0.37)	19 (0.44) - 32 (0.50)
T-DF-L	50 (0.63)- 56 (0.68)	45 (0.68) - 51 (0.65)	61 (0.86) - 73 (0.89)	29 (0.60) - 33 (0.57)	42 (0.67) - 49 (0.63)	-	38 (0.66) - 49 (0.70)
T-DF-N	46 (0.59)- 41 (0.51)	32 (0.51) - 38 (0.51)	63 (0.83) - 71 (0.86)	27 (0.42) - 29 (0.36)	24 (0.56) - 32 (0.50)	20 (0.34) - 21 (0.30)	-

the recommended parameters [7]. We set the neighborhood radius to $r = 50$.

Transition methods, denoted by the prefix T, are composite algorithms with the following operations: **A)** Max-min function using $\mathcal{N}_2$. **B)** Rosin's threshold for transition values using as input: (5) denoted by DF and (6) denoted by CCD. **C)** Three *isolation operators*, in order: *Cross*, *diagonal*, and *rectangular isolate transition-operator* ($x = y = 2$). **D)** Gray-intensity thresholds: *Lognormal threshold* denoted by L and *normal threshold* denoted by N. Setting $n_+ = n_- = 25$ and $c = 15$.

The binarized images were post-processed using the following operators, in order: *Cross*, *diagonal* and *rectangular* ($x = y = 2$) isolate operator which are defined as transition operators but over binary images.

Historical documents often are degraded with ink stains and weak ink strokes for mention some kind of degradation. Hence, we tested the binarization algorithms with the historical atlas *"Theatrum orbis terrarum, sive, Atlas novus"* (Blaeu Atlas) [8]. This paper reports the results of 86 text-images, see Fig. 4, extracted from 61 maps.

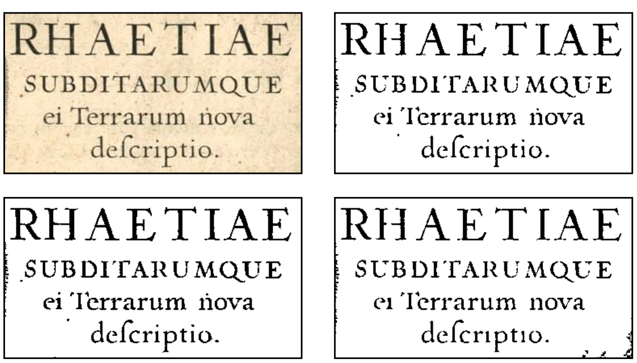

Fig. 4. Sample from the map "Theatrum orbis terrarum, sive, Atlas novus". On the top-right, T-CCD-N; On the bottom-left, T-DF-L; on the bottom-right, Otsu's method.

We measure the segmentation quality by evaluating the performance of an *optical character recognition* (OCR) software. We used *TopOCR* [9] to recognize the text from the binarized images. Our evaluation measures are *accuracy (AC)* and *precision (PR)* computed as

$$AC = \frac{\#T_{\text{match}}}{\#T_{\text{in}}}, \text{ and } PR = \frac{\#T_{\text{match}}}{\#T_{\text{out}}}, \tag{16}$$

where $\#S$ denotes the length of the string S, T_{in} is the original text in the image, T_{out} is the output text from the OCR, and T_{match} is the maximum matching text between T_{in} and T_{out} computed by Needleman-Wuntsh algorithm [10]. AC measure is an important measure for OCR applications, because a high AC value increases the possibility to extract, by further algorithms, relevant information.

TopOCR was tested with four parameter sets. The program tester reports the AC and PR measures from the parameter set that scores higher in terms of AC measure. If there is any draw, PR measure is used.

V. RESULTS AND CONCLUSIONS

We based our observations in the values p_{yx} and p'_{yx} from Table I. We ascertain that algorithm x is better than algorithm y if $n_{yx} \geq 1.33 n_{xy}$.

T-DF-L and T-CCD-L based on lognormal threshold are ranked first and second among the rest of the methods, respectively. T-DF-N and T-CCD-N performed similarly to Otsu's threshold, which is the highest scored among no-transition algorithms, see Fig. 4. Our results also suggest that transition methods based on density function of transition values perform better than those based on complementary cumulative distribution of transition values.

ACKNOWLEDGMENT

This research was partially supported by The National Council on Science and Technology (CONACYT) of Mexico (Grant number:218253).

REFERENCES

[1] M. A. Ramírez-Ortegón, E. Tapia, L. L. Ramírez-Ramírez, R. Rojas, and E. Cuevas, "Transition pixel: A concept for binarization based on edge detection and gray-intensity histograms," *Pattern Recognition*, vol. 43, pp. 1233 – 1243, 2010.

[2] P. L. Rosin, "Unimodal thresholding," *Pattern Recognition*, vol. 34, no. 11, pp. 2083–2096, 2001.

[3] M. Ramirez-Ortegn, E. Tapia, M. Block, and R. Rojas, "Quantile linear algorithm for robust binarization of digitalized letters," in *Ninth International Conference on Document Analysis and Recognition*, vol. 2, September 2007, pp. 1158 –1162.

[4] Ø. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 12, pp. 1191–1201, 1995.

[5] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146–168, January 2004.

[6] P. Stathis, E. Kavallieratou, and N. Papamarkos, "An evaluation technique for binarization algorithms," *Journal of Universal Computer Science*, vol. 14, no. 18, pp. 3011–3030, October 2008.

[7] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern Recognition*, vol. 33, no. 2, pp. 225–236, 2000.
[8] W. Janszoon and J. Blaeu, *Theatrum Orbis Terrarum, Sive, Atlas Novus*. Blaeu Atlas, 1645. [Online]. Available: http://www.library.ucla.edu/yrl/reference/maps/blaeu
[9] T. Soft, *Top OCR*. Top Soft, 2008. [Online]. Available: http://www.topocr.com/
[10] S. B. Needleman and C. D. Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *Journal of Molecular Biology*, vol. 48, no. 3, pp. 443–453, March 1970.