A New Scheme for Text Line and Character Segmentation from Gray Scale Images of Palm Leaf Manuscript

Made Windu Antara Kesiman, Jean-Christophe Burie, Jean-Marc Ogier

To cite this version:

Made Windu Antara Kesiman, Jean-Christophe Burie, Jean-Marc Ogier. A New Scheme for Text Line and Character Segmentation from Gray Scale Images of Palm Leaf Manuscript. 15th International Conference on Frontiers in Handwriting Recognition 2016, At Shenzhen, China, Oct 2016, Shenzhen, China. pp.325-330, 10.1109/ICFHR.2016.63 . hal-01389850

HAL Id: hal-01389850

https://hal.science/hal-01389850

Submitted on 30 Oct 2016

A New Scheme for Text Line and Character Segmentation from Gray Scale Images of Palm Leaf Manuscript

Made Windu Antara Kesiman, Jean-Christophe Burie, Jean-Marc Ogier
Laboratoire Informatique Image Interaction (L3i)
University of La Rochelle, Avenue Michel Crépeau 17042, La Rochelle Cedex 1, France
{made_windu_antara.kesiman, jcburie, jean-marc.ogier}@univ-lr.fr

Abstract— **Most of text line and character segmentation methods for handwritten document image basically still depend on the binary image of the document. Unfortunately, for palm leaf manuscript images, the binarization process is a real challenge. We proposed a new binarization free scheme for text line and character segmentation for palm leaf manuscript images. Our scheme consists of 4 sub-tasks: brushing character area of gray level images with minimum filtering, average block projection profile of gray level images, selection of candidate area for segmentation path, and construction of nonlinear segmentation path. For evaluation, we compared our method with the shredding method which is applied in three different schemes of experiment. The experimental results showed that the proposed method performed optimal on the palm leaf manuscript images which contain discolored parts, with low intensity variations or poor contrast, random noises, and fading.**

Keywords-text line segmentation; character segmentation; Balinese script; palm leaf manuscript; image

I. INTRODUCTION

Segmentation of document image into text lines, words, and characters is oftenly performed prior to recognition step of an OCR system [1–7]. In general, text recognition method vary in the need of segmentation process. Therefore, each text recognition method can be categorized as segmentation based or segmentation free method. The segmentation based text recognition method need prior segmentation process of the document image into text line segments, word segments, or character segments. The performance of OCR system is greatly influenced by the result of the segmentation process.

Many methods of text line and character segmentation for handwritten document image were already proposed [1,8–13]. Some works deal directly with the text line and character segmentation and recognition [2–4]. But most of those methods basically still depend on the binary image of the document. Some other methods used the combined information from both binary and gray scale image [2,3]. In this case, a good initial binarization process is required. Unfortunately, for some type of historical document image, for example the palm leaf manuscript images from Southeast Asia, the binarization process to separate the ancient text from the background is a real challenge [14–16]. A review of evaluation of optimal binarization technique for character segmentation in historical manuscripts was presented in [14]. In our previous work [15], we already experimented and compared several alternative well-known binarization algorithms on the palm leaf manuscript images. We showed that those binarization methods do not give a good binary image for palm leaf manuscript images. All methods extract unrecognizable characters on palm leaf manuscripts with noise. In addition of the varying space between letters, and varying space between lines, the binary image of the manuscript contains the merges, fractures and other deformations of the character shapes. Consequently, the text line and character segmentation method which based on the binary image will not provide a good result for this kind of document image. Therefore, a new scheme for text line and character segmentation for palm leaf manuscript images is required.

Some methods for text line or character segmentation from gray scale image were already proposed [4,7,17,18]. A survey of text line segmentation methods for historical documents was given in [19]. In this paper, we propose a new scheme for text line and character segmentation from gray scale images of palm leaf manuscript. This scheme is based on the work of [4], [2] and [3]. It consists of 4 sub-tasks: brushing character area of gray level images with minimum filtering, average block projection profile of gray level images, selection of candidate area for segmentation path, and construction of nonlinear segmentation path. But instead of using the projection profile described in [4], our scheme proposed to use the average block projection profile. We also defined the rule to select candidate areas for segmentation path, instead of using the topographic features described in [4] or using the local maxima of the outer counter of the binary image described in [2] and [3]. Our scheme is a binarization free scheme for text line and character segmentation.

This paper is organized as follow: Section II gives a brief description about Balinese palm leaf manuscripts and the challenges for text line and character segmentation. Section III presents the detail description about the proposed scheme. The result and evaluation of the proposed scheme is presented in Section IV. Conclusions with some prospects for the future works are given in Section V.

II. PALM LEAF MANUSCRIPTS

A. *The Balinese Palm Leaf Manuscripts*

Many literary texts of the Balinese were written on dried and treated palm leaves, called *Lontar*. The Balinese palm leaf manuscripts were written in Balinese script in Balinese language. Writing in Balinese script, there is no space between words in a text line. Some characters are written on upper baseline or under the baseline of text line (Fig. 1). *Lontars* were written in the ancient literary texts composed in the old Javanese language of Kawi and Sanskrit. Balinese script is considered to be one of the complex scripts from Southeast Asia. The alphabet and numeral of Balinese script is composed of ±100 character classes including consonants, vowels, diacritics, and some other special compound characters. The number of compound characters can not be counted precisely because it depends on the writing style of the writers. Two basic characters are sometimes written in their compound character.

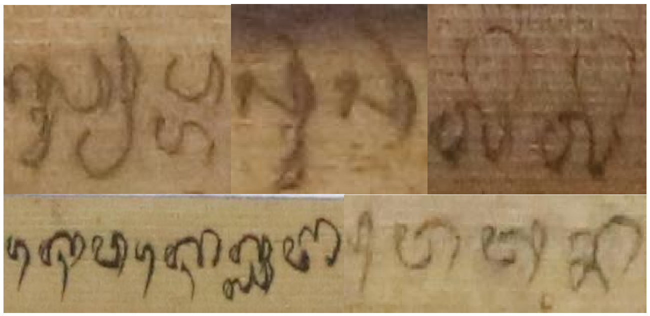


Figure 1. Balinese script on palm leaf manuscripts

There is only a limited access to the content of the manuscripts, because of the linguistic difficulties and the fragility of the documents. The majority of Balinese has never read any lontar because of language obstacles as well as tradition which perceived them as a sacrilege. Many discovered lontars are now part of collections of museums and private families. They are in a state of disrepair due to age and due to inadequate storage conditions. Therefore, the digitization and indexing projects for palm leaf manuscripts were proposed. They work not only to digitize the palm leaf manuscripts, but also to develop an automatic analysis, transcription and indexing system for the manuscripts, to make palm leaf manuscripts more accessible, readable and understandable to a wider audience and to scholars and students all over the world.

B. *The Challenges for Text Line and Character Segmentation*

Due to its specific characteristics, palm leaf manuscripts are providing new challenges in document analysis. Usually, palm leaf manuscripts are of poor quality since the documents have degraded over time due to storage conditions (Fig. 2). Natural materials from palm leaves certainly cannot fight against time, and therefore the processes of digitizing and indexing *Lontars* are very important. The palm leaf manuscripts contain discolored parts and artefacts due to aging and low intensity variations or poor contrast, random noises, and fading [15]. Several deformations in the character shapes are visible due to the merges and fractures of the characters, varying space between letters, and varying space between lines. The overlaps, and interconnection of the neighboring characters further complicate the work of the OCR systems [3]. The [illegible] so segmentation, for example skewed and fluctuating text lines, and irregularity in geometrical properties of the line, such as line width, height, and distance in between lines [1]. These characteristics provide a suitable challenge for text line and character segmentation.

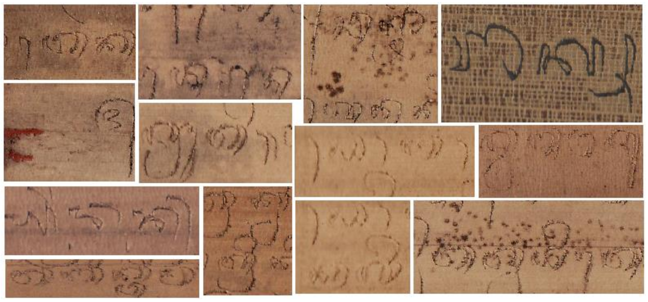
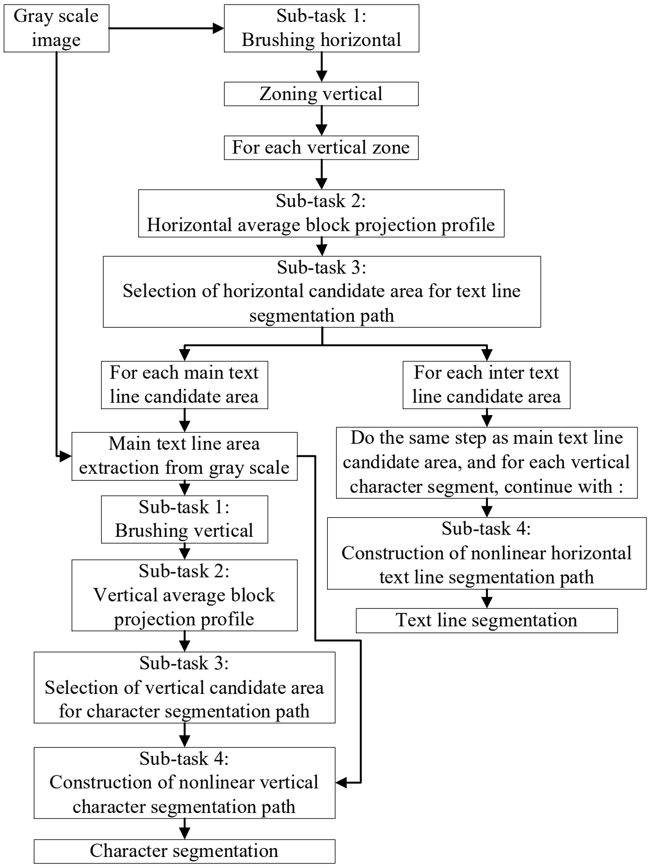


Figure 2. The degradations on palm leaf manuscripts [15]



Gray scale image
Sub-task 1: Brushing horizontal
Zoning vertical
For each vertical zone
Sub-task 2: Horizontal average block projection profile
Sub-task 3: Selection of horizontal candidate area for text line segmentation path
For each main text line candidate area
For each inter text line candidate area
Main text line area extraction from gray scale
Do the same step as main text line candidate area, and for each vertical character segment, continue with :
Sub-task 1: Brushing vertical
Sub-task 4: Construction of nonlinear horizontal text line segmentation path
Sub-task 2: Vertical average block projection profile
Text line segmentation
Sub-task 3: Selection of vertical candidate area for character segmentation path
Sub-task 4: Construction of nonlinear vertical character segmentation path
Character segmentation

Figure 3. Diagram of the proposed scheme

III. The Proposed Scheme

Our proposed scheme (Fig. 3) consists of 4 sub-tasks: brushing character area of gray level images with minimum filtering, average block projection profile of gray level images, selection of candidate area for segmentation path, and construction of nonlinear segmentation path. To deal with some slight skew angles on the text line, first, we divided the manuscript into vertical zones. The approximated zone width was calculated from the average width of 11,496 word segments (patch image of group of characters) which are obtained from our previous work on word annotation ground truthing process and fixed to a value of 400 pixels.

A. Brushing Character Area of Gray Level Images

Most of the Balinese character forms consist of some holes and some semi-spaced areas between their two or more main curve strokes. These semi-space areas are normally found between two vertical strokes within a character and they have almost the same width with the real space between two characters. To segment the character in a text line, it is not easy to differentiate these semi-spaced intra-character areas with the real spaced inter-character areas (Fig. 4).

Even though that, under assumption we have a good binarized image for the characters. The method which try to detect the local minima and maxima on the characters, will fail to differentiate the candidate area for segmentation between the two local minima or maxima of intra-character and inter-character areas.

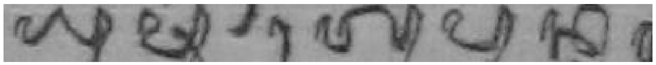


Figure 4. Similar semi-spaced intra-character areas and real spaced inter-character areas

To overcome this problem, we applied a minimum filtering to cover the holes and the intra-character areas by brushing the areas. We used the minimum filtering with a horizontal or vertical filter. This brushing technique with minimum filtering is the application of a morphological operator dilation with a horizontal/vertical structuring element. It replaces the pixel value with the minimum pixel value within the vertically/horizontally defined neighbor pixels (defined as brush width). Minimum filter is used because in our manuscripts the character pixels are always darker than the background pixels. The gray values in our gray scale images are defined from 0 (for black pixels) to 255 (for white pixels).

Brushing for gray level images is defined as follows: Let *g(r,c)* be the intensity of a pixel *(r,c)* in a gray scale image, and *w* be the brush width.

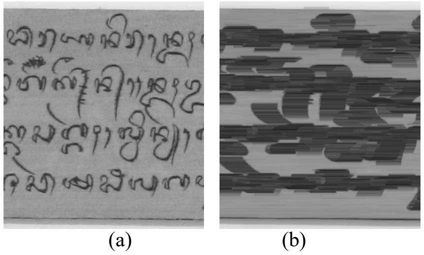
Brushing vertical:

$$g_{brush\_v}(r,c) = \min\{g(i,c)\}, r - \frac{w}{2} \le i \le r + \frac{w}{2}$$

Brushing horizontal:

$$g_{brush\_h}(r,c) = \min\{g(r,j)\}, c - \frac{w}{2} \le j \le c + \frac{w}{2}$$

The spaces between text lines are also difficult to determine because many characters in Balinese script have ascender or descender parts, written in the area between text lines. To detect the main text line areas in the manuscript, first, we applied the brushing horizontal. The brush width is defined as the approximated character width in the manuscript, so it can cover horizontally the holes, the semi-spaced intra-character, and also the real-spaced inter-character areas, to connect the main text line (Fig. 5). The approximated character width and height was calculated from the average width and height of 4,975 character segments which are obtained from our previous work on character annotation ground truthing process and fixed respectively to 50 and 35 pixels.



(a) (b)

Figure 5. a) Sample of original image, b) Brushing horizontal of a)

To segment the characters, first, we applied the brushing vertical (Fig. 6). The brush width is simply defined as the height of the text line segment (or the approximated character height) because this step is normally done in the main text line areas (see Section IIIC).

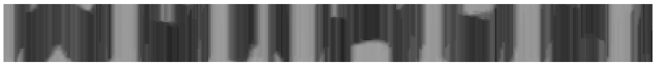


Figure 6. Brushing vertical of image in Fig. 4

B. Average Block Projection Profile of Gray Level Images

A definition of projection profile for gray level images was described in [7,17]. But inspired by the work of [4], the projection profile for gray level images in our work is defined as follows: Let *g(r,c)* be the intensity of a pixel *(r,c)* in a gray-scale image. Then *g(r,c)* is in the following range : $0 \le g(r,c) \le L-1$, where *L* is the level of intensity. Let $H_r(g)$ and $H_c(g)$ be the histograms of row *r* and column *c* with intensity of *g*, respectively. The vertical projection profile in column *c*, *P(c)* can be defined as follows:

$$P(c) = \sum\nolimits_{g=0}^{L-1} H_c(g)c(g)$$

where $c(g) = \frac{g}{L}$ is a ratio contributing to the projection with intensity of *g*, $0 \le c(g) \le 1$. In similar way, the horizontal projection profile in row *r*, *P(r)* can be defined by

$$P(r) = \sum\nolimits_{g=0}^{L-1} H_r(g)c(g)$$

To detect a zone area of the characters or the text lines, we calculated the average of projection profile in some equal sized overlapped blocks. The average block projection profile is defined as follows: Let *bw* be the block width, *AP(c)* be the vertical average block projection profile in column *c*, and *AP(r)* be the horizontal average block projection profile in row *r*.

$$AP(c) = average\{P(j)\}, c \leq j \leq c + bw - 1$$

$$AP(r) = average\{P(i)\}, r \leq i \leq r + bw - 1$$

The average block projection profile is finally normalized into range 0..1.

$$AP_{norm}(c) = \frac{AP(c) - \min}{\max - \min}$$

where $\min = \min\{AP(j)\}, 1 \leq j \leq nb\_col$, and $\max = \max\{AP(j)\}, 1 \leq j \leq nb\_col$.

$$AP_{norm}(r) = \frac{AP(r) - \min}{\max - \min}$$

where $\min = \min\{AP(i)\}, 1 \leq i \leq nb\_row$, and $\max = \max\{AP(i)\}, 1 \leq i \leq nb\_row$.

C. Selection of Candidate Area for Segmentation Path

To detect the candidate area for segmentation path, the average block projection profile is used. The horizontal average block projection profile is used to find the horizontal candidate area to segment the main text line areas and inter text line areas. The horizontal average block projection profile *$AP_{norm}(r)$* with block width *bw* is firstly sorted for all row *r*, $1 \leq r \leq nb\_row$, from the minimum to the maximum value. In this case, we consider the minimum value because in our palm leaf manuscript images, the character strokes are represented in darker pixels, so the main text line area which contains more character pixels has a smaller value of horizontal average block projection profile. Let *r1,r2,r3...rnb_textline* be the sorted row position of *$AP_{norm}(r)$*. The *nb_textline* candidate areas for main text line are then constructed with *bw* rows on the manuscript, which started from row *r^{th}*. To detect the main text line areas, *bw* is defined as the approximated character height in the manuscript. The inter text line areas are simply extracted from the areas between the main text line areas (Fig. 7).

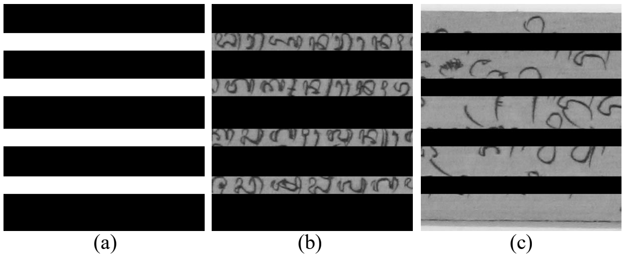
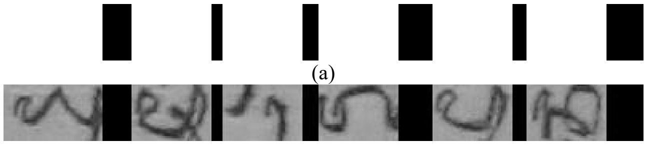

(a) (b) (c)

Figure 7. a) Candidate areas (white zones) for main text line areas of Fig. 5, b) Blending of a) & Fig. 5a, c) Blending with candidate areas for inter text line areas of Fig. 5a

By using the same definition, the vertical average block projection profile is used to find the vertical candidate area to segment the characters in text line. The vertical average block projection profile *$AP_{norm}(c)$* with block width *bw* is firstly sorted for all column c, $1 \leq c \leq nb\_col$. Here, *bw* is defined as the approximated character width in the manuscript (Fig. 8).


(a)

(b)

Figure 8. a) Candidate areas (white zones) for characters of Fig. 4 and 6, b) Blending of a) & Fig. 4

D. Construction of Nonlinear Segmentation Path

The objective is to find a nonlinear segmentation path of one pixel wide in each candidate area of segmentation. Finding a nonlinear segmentation path can be defined as a problem of finding the shortest path which (in our case) maximizes the accumulated intensity in the candidate area of segmentation. We implemented the multistage graph search algorithm to find the nonlinear segmentation path. The multistage graph search algorithm was described and used in [2–4].

Let c_{area} be a candidate area of segmentation. To find a horizontal minimum nonlinear segmentation path from left to right, we consider each pixel in c_{area} as a graph node (Fig. 9). A node of pixel *$c_{area}(r,c)$* has (at most) three vertices directed to the right, which connect to the next three neighbor pixels: node of pixel *$c_{area}(r\text{-}1,c+1)$, $c_{area}(r,c+1)$, and $c_{area}(r+1,c+1)$.*

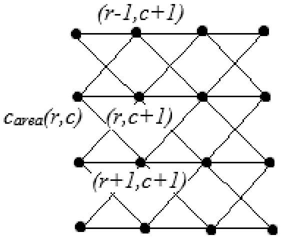
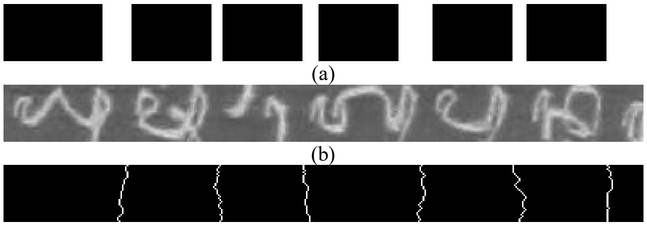


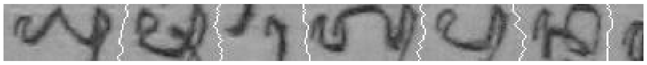
Figure 9. Multistage graph representation

From each pixel in first column (left most pixel), we generate a nonlinear segmentation path to the right most pixel, by consecutively choosing the pixel with minimum intensity value between the three neighbor nodes of pixel. The minimum nonlinear segmentation path in c_{area} is finally defined as the nonlinear segmentation path which is started from a certain pixel, with a minimum accumulated intensity value of pixels on their path.


(a)

(b)

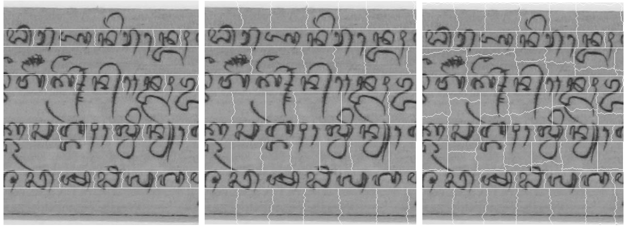
(c)



(d)

Figure 10. a) Candidate areas for character segmentation (white zones) of Fig. 4, b) Complement image of Fig. 4, c) Nonlinear character segmentation path of b) based on a), e) Blending of Fig. 4 & c)

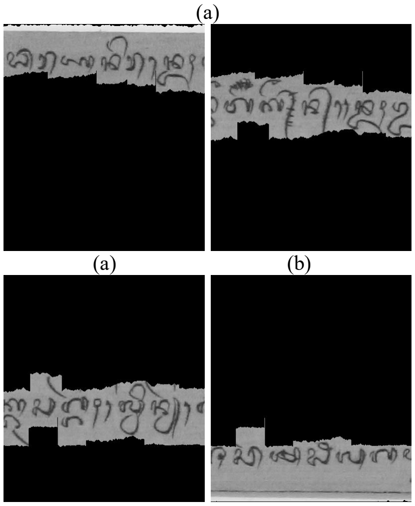
To segment the characters, we can apply the vertical nonlinear segmentation path on each segment of text line, in similar way. The vertical nonlinear segmentation path should be calculated on the complement of original image to find the minimum path between the characters (Fig. 10 & 11).



(a) (b) (c)

Figure 11. a) Nonlinear character segmentation path of Fig. 7b, b) Nonlinear character segmentation path of Fig. 7c, c) Horizontal nonlinear segmentation path in the inter text line areas of b)

(a)

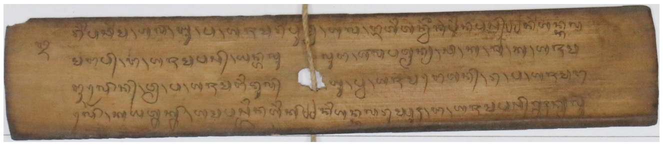


(a) (b)

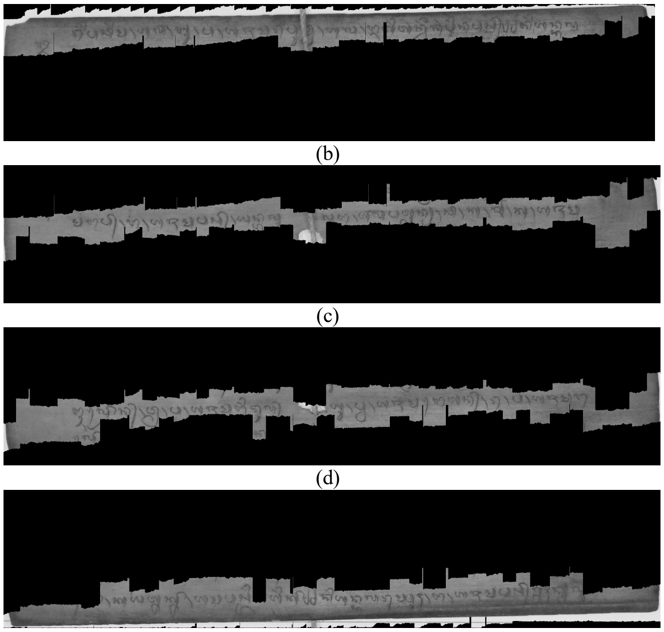
(c) (d)

Figure 12. a) Extracted text line 1, b) Extracted text line 2, c) Extracted text line 3, d) Extracted text line 4.

We finally apply the horizontal nonlinear segmentation path on each vertical segment in the inter text line areas. This path will segment each inter text line areas into two separated areas which belong to two different text lines (Fig. 12).



(a)



(b)

(c)

(d)

(e)

Figure 13. a) Original image, b) Extracted text line 1, c) Extracted text line 2, d) Extracted text line 3, e) Extracted text line 4

IV. Results and Evaluations

We evaluated our proposed method on a set of 35 palm leaf manuscript images which contains 140 text lines in total. Those images were randomly collected and selected from corpus of palm leaf manuscript images from different locations (regions) in Bali, Indonesia. The text line ground truth images were manually segmented from the binary ground truth images of the manuscripts. The binary ground truth images were obtained from our previous ground truthing process [16].

We also compared our method with the shredding method proposed in [1]. This method works on binary image. We performed three different experiments to evaluate this method. First, we tested this method directly on our ground truth binary image (Experiment 1). In this case, this method was evaluated with the best binary image of our manuscript. To see the real challenge for our manuscripts, secondly, we tested this method with the binary image which were obtained from Otsu's binarization method (Experiment 2). And thirdly, we tested a small proposed modification for this method by replacing the bluring process from binary image with the brushing gray scale image used in our proposed method (Experiment 3).

We used the evaluation criteria and tool provided by ICDAR2013 Handwritting Segmentation Contest [20]. First, the one-to-one (*o2o*) match score is computed for a region pair based on the evaluator's acceptance threshold. With the *o2o* score, the three metrics are calculated: detection rate (*DR*), recognition accuracy (*RA*), and performance metric (*FM*). The result is presented on Table I.

We can see that the method of [1] depends greatly on the quality of the binary image. Eventhough this method was tested on the best binary image from our ground truth binary image, it only achieved 80.13% for *FM* score, and it decreased significantly when it was applied to the real binarization method results (Exp. 2). Our small proposed modification gave a promising result from gray scale images (Exp. 3). It means that our brushing images already provide a preliminary information about the main text line positions. Our proposed method achieved a good *FM* score by using gray scale processing scheme. It is performed optimally on the palm leaf manuscript images which contain discolored parts, with low intensity variations or poor contrast, random noises, and fading (Fig. 13).

TABLE I. RESULT OF PERFORMANCE EVALUATION

Exp.	M	o2o	DR (%)	RA (%)	FM (%)
Exp. 1	167	123	87.85	73.65	80.13
Exp. 2	1418	14	10	0.98	1.79
Exp. 3	199	70	50	35.17	41.29
Proposed method	140	110	78.57	78.57	78.57

V. CONCLUSIONS AND FUTURE WORKS

We presented a new binarization free scheme for text line and character segmentation for palm leaf manuscript images. The experimental results showed that the proposed method performed optimal on the palm leaf manuscript images which contain discolored parts, with low intensity variations or poor contrast, random noises, and fading. For the future works, we will integrate this method with a character or text recognition module so it can perform mutually as the verification and correction module for the segmentation and also for the final text recognition task.

ACKNOWLEDGMENT

The authors would like to thank Museum Gedong Kertya, Museum Bali, and all families in Bali, Indonesia, for providing us the samples of palm leaf manuscripts, and the students from the Department of Informatics Education and the Department of Balinese Literature, Ganesha University of Education for helping us in ground truthing process for this research project. This work is also supported by the DIKTI BPPLN Indonesian Scholarship Program and the STIC Asia Program implemented by the French Ministry of Foreign Affairs and International Development (MAEDI).

REFERENCES

[1] A. Nicolaou, B. Gatos, Handwritten Text Line Segmentation by Shredding Text into its Lines, in: IEEE, 2009: pp. 626–630. doi:10.1109/ICDAR.2009.243.

[2] N. Arica, F.T. Yarman-Vural, A new scheme for off-line handwritten connected digit recognition, in: IEEE Comput. Soc, 1998: pp. 1127–1129. doi:10.1109/ICPR.1998.711893.

[3] N. Arica, F.T. Yarman-Vural, Optical character recognition for cursive handwriting, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 801–813. doi:10.1109/TPAMI.2002.1008386.

[4] Seong-Whan Lee, Dong-June Lee, Hee-Seon Park, A new methodology for gray-scale character segmentation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 18 (1996) 1045–1050. doi:10.1109/34.541415.

[5] R.J. Ramteke, Invariant Moments Based Feature Extraction for Handwritten Devanagari Vowels Recognition, Int. J. Comput. Appl. 1 (2010) 1–5. doi:10.5120/392-585.

[6] M. Blumenstein, B. Verma, H. Basli, A novel feature extraction technique for the recognition of segmented handwritten characters, in: IEEE Comput. Soc, 2003: pp. 137–141. doi:10.1109/ICDAR.2003.1227647.

[7] Zhixin Shi, S. Setlur, V. Govindaraju, Text extraction from gray scale historical document images using adaptive local connectivity map, in: IEEE, 2005: p. 794–798 Vol. 2. doi:10.1109/ICDAR.2005.229.

[8] G. Louloudis, B. Gatos, I. Pratikakis, C. Halatsis, Text line and word segmentation of handwritten documents, Pattern Recognit. 42 (2009) 3169–3183. doi:10.1016/j.patcog.2008.12.016.

[9] R. Chamchong, C.C. Fung, Character segmentation from ancient palm leaf manuscripts in Thailand, in: ACM Press, 2011: p. 140. doi:10.1145/2037342.2037366.

[10] R.P. dos Santos, G.S. Clemente, T.I. Ren, G.D.C. Cavalcanti, Text Line Segmentation Based on Morphology and Histogram Projection, in: IEEE, 2009: pp. 651–655. doi:10.1109/ICDAR.2009.183.

[11] X. Zhang, C.L. Tan, Text Line Segmentation for Handwritten Documents Using Constrained Seam Carving, in: IEEE, 2014: pp. 98–103. doi:10.1109/ICFHR.2014.24.

[12] F. Yin, C.-L. Liu, Handwritten Chinese text line segmentation by clustering with distance metric learning, Pattern Recognit. 42 (2009) 3146–3157. doi:10.1016/j.patcog.2008.12.013.

[13] J. Kumar, W. Abd-Almageed, L. Kang, D. Doermann, Handwritten Arabic text line segmentation using affinity propagation, in: ACM Press, 2010: pp. 135–142. doi:10.1145/1815330.1815348.

[14] Chun Che Fung, R. Chamchong, A Review of Evaluation of Optimal Binarization Technique for Character Segmentation in Historical Manuscripts, in: IEEE, 2010: pp. 236–240. doi:10.1109/WKDD.2010.110.

[15] M.W.A. Kesiman, S. Prum, J.-C. Burie, J.-M. Ogier, An Initial Study On The Construction Of Ground Truth Binarized Images Of Ancient Palm Leaf Manuscripts, in: 13th Int. Conf. Doc. Anal. Recognit. ICDAR, Nancy, France, 2015.

[16] M.W.A. Kesiman, S. Prum, I.M.G. Sunarya, J.-C. Burie, J.-M. Ogier, An Analysis of Ground Truth Binarized Image Variability of Palm Leaf Manuscripts, in: 5th Int. Conf. Image Process. Theory Tools Appl. IPTA 2015, Orleans, France, 2015: pp. 229–233.

[17] I. Bar-Yosef, N. Hagbi, K. Kedem, I. Dinstein, Line Segmentation for Degraded Handwritten Historical Documents, in: IEEE, 2009: pp. 1161–1165. doi:10.1109/ICDAR.2009.191.

[18] A. Garz, A. Fischer, R. Sablatnig, H. Bunke, Binarization-Free Text Line Segmentation for Historical Documents Based on Interest Point Clustering, in: IEEE, 2012: pp. 95–99. doi:10.1109/DAS.2012.23.

[19] L. Likforman-Sulem, A. Zahour, B. Taconet, Text line segmentation of historical documents: a survey, Int. J. Doc. Anal. Recognit. IJDAR. 9 (2007) 123–138. doi:10.1007/s10032-006-0023-z.

[20] N. Stamatopoulos, B. Gatos, G. Louloudis, U. Pal, A. Alaei, ICDAR 2013 Handwriting Segmentation Contest, in: IEEE, 2013: pp. 1402–1406. doi:10.1109/ICDAR.2013.283.