



**HAL**  
open science

## An effective method for text line segmentation in historical document images

Tien-Nam Nguyen, Jean-Christophe Burie, Thi-Lan Le, Anne-Valérie Schweyer

► **To cite this version:**

Tien-Nam Nguyen, Jean-Christophe Burie, Thi-Lan Le, Anne-Valérie Schweyer. An effective method for text line segmentation in historical document images. 2022 26th International Conference on Pattern Recognition (ICPR), Aug 2022, Montreal, Canada. pp.1593-1599, 10.1109/ICPR56361.2022.9956617 . hal-03922470

**HAL Id: hal-03922470**

**<https://hal.science/hal-03922470v1>**

Submitted on 4 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# An effective method for text line segmentation in historical document images

Tien-Nam Nguyen\*, Jean-Christophe Burie\*, Thi-Lan Le<sup>†</sup>, Anne-Valérie Schweyer<sup>‡</sup>

\* Laboratoire Informatique Image Interaction (L3i)

La Rochelle University, La Rochelle, France

tnguye28@univ-lr.fr, jcburie@univ-lr.fr

<sup>†</sup> School of Electrical and Electronic Engineering (SEEE)

Hanoi University of Science and Technology, Hanoi, Vietnam

<sup>‡</sup>Centre Asie du Sud-Est (CASE), CNRS, Paris, France

anne-valerie.schweyer@cnrs.fr.

**Abstract**—In this paper, we present a text-line segmentation method for historical documents. Historical documents are challenging given their characteristics of highly degradation, writing style variation and diacritics. From these observations, we proposed an effective approach for text line segmentation by analysing the properties of document layouts. We combine the idea of seam carving method with the novel cost functions to accurately split text lines. Experiments were conducted on two challenging datasets of historical documents, namely the DIVA-HisDB dataset and our ChamDoc dataset. Our methods provided good results on the DIVA-HisDB dataset with 99.36% of Line IU and 98.86% of Pixel IU. On the ChamDoc dataset, the proposed method outperformed the two baseline approaches i.e. seam carving-based and A\* path planning by a large margin.

## I. INTRODUCTION

Text line segmentation is a crucial step in the whole pipeline of document image analysis and especially for historical documents. Although, some recent recognition methods can work at paragraph level [1], [2], their applications are still very limited and dependent on the layout of documents. Therefore, text line extraction is still widely used in document analysis pipeline. The main aim of text line segmentation is to split the whole document image line by line to obtain a set of text line images that will be fed to text recognition process. The latest approaches for this problem mainly utilize deep learning based methods along with additional post processing steps to correct the polygon boundary. The main issue of the deep learning based approaches is that they require a burden of training data. Therefore, when working with historical documents where the image dataset is often modest and the annotation process requires the participation of high expertise annotator, deep learning based approaches could not show its superior results. Moreover, in [3], the authors have shown that a good binarization of documents can lead to good segmentation results without requiring the training of complex deep learning models.

Our work mainly focuses on historical documents images, especially Cham inscriptions images. An introduction about Cham language can be found at [4]. These images are captured

from Cham inscriptions that are mainly carved on steles of stone. The text is written from the Cham language system used in Champa (nowadays, Vietnam coastal areas) and some surrounding areas. The challenges of Cham document images come from the particular properties of the writing style such as touching and overlapping characters, multiple size of characters and the huge variety of diacritics and low space between lines. Then skew lines and even partially missing lines (due to damaged inscriptions) made difficult the analysis of these documents. Furthermore, these documents are often highly degraded. The climatic conditions and the running of time have damaged the characters and created bumps. It is therefore difficult to distinguish between texts and noise. Figure 1 shows some examples of Cham inscription images. In this paper, we propose an efficient and simple approach for the text line segmentation. Our approach mainly focus on documents highly degraded with lots of diacritics, ascending and descending characters by analysing the properties of these specific elements. Our approach is based on seam carving method. We propose to use an adapted global cost function in order to split the text lines with a better accuracy. The proposed method takes a binary image as input and considers the aforementioned challenges of historical documents.

The paper is organized as follows: section II gives a brief overview of text line segmentation approaches proposed in the literature. In section III, we detail the proposed method. In section IV, we present the experimental results on two datasets: DIVA-HisDB dataset [5] and our own ChamDoc dataset.

## II. RELATED WORK

Text-line segmentation methods can be mainly divided into three main categories: top-down, bottom-up and deep learning approaches. In the top-down approaches, the pioneer works are based on projection profiles [6]. The projection profile consists of adding the number of pixels projected horizontally or vertically (depending on the studied direction). The peaks of the profile correspond to the text parts (lines of text) while the valleys correspond to the space between the lines. This method works efficiently with documents which have simple structure

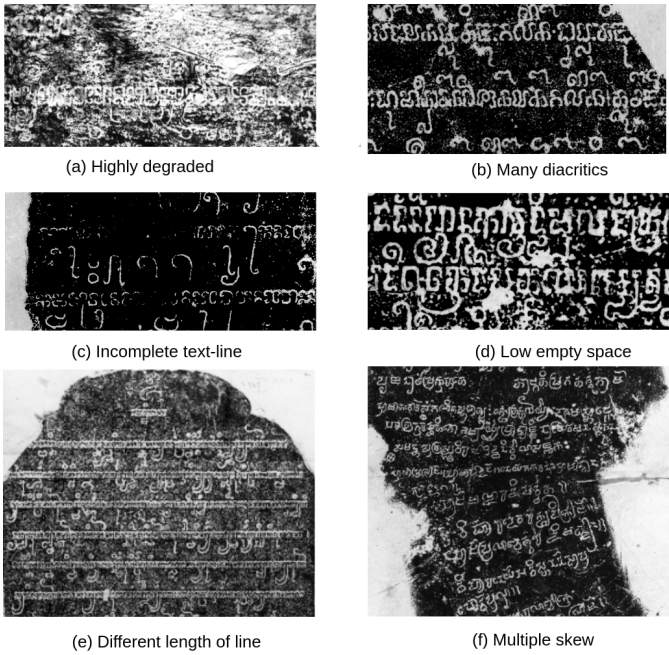


Fig. 1: Samples inscription images of ChamDoc dataset.

such as printed document or document with weakly skewed and curved lines. Different variants of this method have been proposed to deal with documents having skewed and curved lines [7], [8], [9]. Some problems arise with this kind when lines of text have ascending, descending characters or diacritics which may make them appear to belong to the upper / lower line, leading to many local maxima and minima. Another approach belonging to top-down category is seam carving-based [10], [7], [11], [3]. Seam carving method generates the line separation by minimizing the accumulated energy map computed from the image. In [10], the authors proposed a two steps seam carving to extract text lines. First, they choose the seam seeds. Then, in the second step, they extracted the lines by computing the lower and upper boundaries for each seam seed. In [7], the authors used only one step seam carving which split the lines from the constrained regions in the image. In [3] the Labeling, Cutting, Grouping (LCG) method is proposed. The approach is based on a new energy map which is built from background energy and text energy. After that, the seams are casted every 100 pixels. The final separation line is selected from the number of binning of the centroids over these seams. A close approach to the seam carving method is the A\* path planning method [12], it consists in computing the separation lines by minimizing different cost functions at a given list of seed points. In the bottom-up approaches [13], [14], [15], topological features are used such as connected components. The main idea of these approaches is to merge the pixels of foreground part to re-create the text line. Merging pixels can be done through different methods such as image blurring [13], steerable directional filters [14] or isotropic Gaussian filters in multi-scale space [15]. A binarization step is then applied

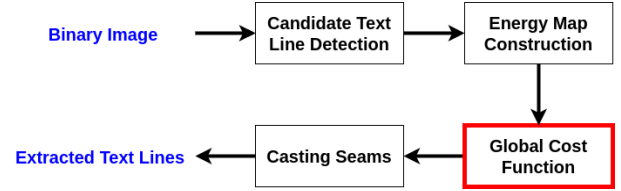


Fig. 2: Overview of the proposed method.

on filtered image to obtain the main components which will be then assigned to the text lines. The problem with this approach is when the line is fragmented, the smearing group can not be merged continuously. In addition, if the space between two consecutive lines is relatively small (ascending and descending characters touching the upper/lower line), this method fails. Recently, deep learning-based methods for text line segmentation have been emerged [16], [17], [18], [19]. In [18] the authors first used a deep convolution network to assign each pixel to the corresponding class (baseline, separator, other), then they calculated super-pixels from the output of the previous step before estimating and assigning them to the text line. In [17], the authors have trained Siamese network to determine whether two patches of image are similar or different, then have employed this network as an embedding network to extract the blobs of image. Energy minimization is used to split these blobs to the appropriate text lines.

### III. PROPOSED METHOD

In this section, we present the details of the proposed method. The proposed method is based on seam carving method which used the energy map proposed in [3]. However, the only energy map used is not enough to guide the seams correctly when working with documents that have more accents or diacritics. Therefore, the idea is to add an appropriate cost function to improve the generation of seams when there are a lot of ascending/descending character between two consecutive lines. The proposed approach is closed to the work of LCG [3] but it is different in two points. First, we proposed an additional cost function which decomposes more accurately the ascenders, descenders and diacritics. Secondly, in our method, seams are generated at every pixel in the constrained regions while in the LCG method, one seam is created at every 100 pixels. The overview of proposed method can be found in Fig. 2.

#### A. Candidate Text Line Detection

The aim of this step is to detect the candidate lines from the input image. Firstly, let  $I \in R^{h \times w}$  denote the input image, and  $\mathcal{L}: \{L_i\}; i = 1 : n$  denote the list of the candidate lines. Each candidate line is defined by  $L_i = (l_i, m_i, u_i)$  where  $l_i, m_i, u_i$  represents respectively the ordinate of lower, middle, upper line of the candidate text line  $i^{th}$ . These positions are computed by the modified version of the projection profile method presented in [7]. The idea of this method is to compute the projection profile on the small columns of the image. However this approach does not work efficiently for Cham

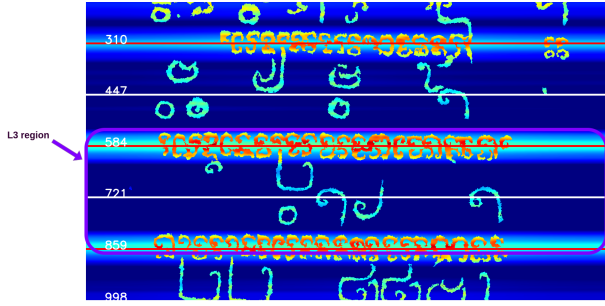


Fig. 3: Heat maps of energy map computed with LCG [3]. The red color (resp. blue) represents high energy (resp. low). Let's note that most of the characters in the main line are in red color. The red lines represent the peaks while the white lines represent the valleys of smoothed histogram from image. One candidate line is represented by two consecutive peaks and one valley. For example, The candidate line  $L_3$  is defined by  $(l_i, m_i, u_i) = (584, 721, 859)$ .

document images due to the presence of a lot of diacritics between the lines. In our work, we applied projection profile on the whole image. First, the Sobel operator is applied on the input image to extract edges. Then the small components which are noisy components still existing after the denoising step, are removed by comparing their areas with the average area of all the components in the image. We hence compute horizontal projection profile histogram  $H_I$  on the edge image.

$$H_I = \sum_{i=0}^h I_i \quad (1)$$

where  $h$  is the height of the image. This histogram is smoothed by cubic spline smoothing filter to reduce the number of local maximum. This smoothing filter ( $f_\sigma$ ) has one parameter  $\sigma$  (smooth parameter) which is defined experimentally.

$$H_{f_I} = f_\sigma(H_I) \quad (2)$$

One advantage of using histogram on the whole image for Cham images is that we can use the property space between lines as a heuristic rule to remove the close candidates. The peaks of the smoothed histogram are considered as the lower and upper position of the candidate line while the valleys are considered as middle position between two consecutive lines.

### B. Energy map construction

To cast the seams, we have to construct an energy map to guide the seams on how to separate the lines. Several methods can be used to compute the energy map including Gray-level Distance Transform (GDT) [20], gradient map (GM) [7], or a combination of background energy and text energy map (LCG) in [3]. Depending on the workflow proposed in these methods, each energy map function has its own advantages. In this paper, we employ the method in [3] as the baseline method and adapt it for Cham documents. Figure 3 visualizes the energy map with LCG method and candidate lines position.

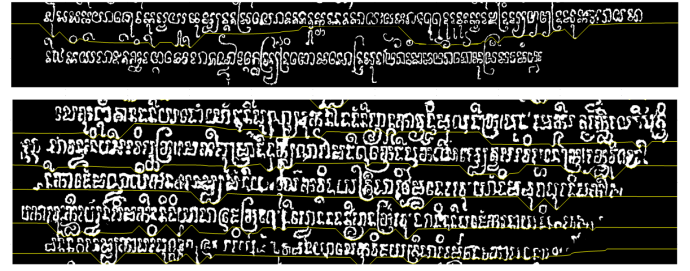


Fig. 4: Separation lines (yellow lines) with the simple seam carving method when using only energy map [3]

### C. Proposed cost function

Why do we need an additional cost function? We run a simple experiment with the seam carving method [3] on our Cham documents. Figure 4 illustrates some results obtained by the method in [3]. We can observe that some accents or diacritics are assigned to the wrong line, the energy map used in these approaches is not enough to guide the seams correctly. We therefore proposed an additional cost function when casting seams to handle this problem. This additional cost function is adapted for the text line segmentation from the cost function found in [12]. Our work is different from the work of [12] in three points. First, we do not need to define the starting point, and its destination point in several terms for the cost function. Second, we propose different cost functions to consider the properties of the text lines. Third, our cost function is combined with the baseline energy map. Let's define  $p$ , a point at position  $(p_y, p_x)$ ,  $V(p)$  the value of pixel  $p$  and  $L_i$  the candidate line. We then defined the three cost functions as follows:

1) *Character cost function*: The character cost function is defined by:

$$E_{C_p} = \begin{cases} 1 & \text{if, } V(p) = 255 \\ 0 & \text{if, } V(p) = 0 \end{cases}$$

The term simply forces the seams to avoid the character pixels when casting in the region.

2) *Middle cost function*: This term forces the seams have to pass closely as much as possible to the middle line and helps seams come back near the middle line after moving away from it.

$$E_{M_p} = \begin{cases} \frac{p_y - m}{l - m} & \text{if } p_y - m < 0 \\ \frac{p_y - m}{u - m} & \text{if } p_y - m > 0 \end{cases}$$

where  $l, m, u$  is the lower, middle, upper ordinate of the candidate line that determines the region where the pixel  $p$  belongs to.  $p_y$  is the height ordinate of the point  $p$ . The value of this cost function reaches 0 when the point  $p$  has the same ordinate of the middle line.

3) *Balance cost function*: If we used only the character cost function, the seams, in some cases, can go up or down very close to character pixels. Moreover, in some cases where the middle line is not accurately determined, using only character

cost function may lead to an incorrect text lines separation. Therefore, we proposed the balance cost function to encourage seams come back at the **middle of gaps** between two lines. This term is computed as follows. First we define  $L(p)$ ,  $U(p)$  respectively the distance from the current point to the nearest white pixel from the lower  $l_i$  and the upper  $u_i$  ordinates of candidate line  $L_i$ . The normalization of  $L(p)$  and  $U(p)$  is computed as:

$$\overline{L(p)} = \frac{L(p)}{L(p) + U(p)}; \overline{U(p)} = \frac{U(p)}{L(p) + U(p)}$$

Then the balance cost function is evaluated by:

$$E_{B_p} = \begin{cases} \frac{1}{\overline{U(p)} * \overline{L(p)}} & \text{if } \frac{\overline{U(p)}}{\overline{L(p)}} < 8 \text{ or } \frac{\overline{L(p)}}{\overline{U(p)}} < 8 \\ 10 & \text{otherwise} \end{cases}$$

This term forces the seams to pass through at the **middle of gaps** of between two white pixels, in order to avoid to jump to the previous or next line. The range of this function is between 4 and 10. The smallest score is obtained when the seams is located at the **middle of gaps** space.

4) *Global cost function*: Finally, the global cost function is a weighted sum of the three terms detailed above:

$$C_p = w_1 * E_{C_p} + w_2 * E_{B_p} + w_3 * E_{M_p} \quad (3)$$

This global cost function allows to balance previous terms to provide better guide for casting seams. An illustration of each term of the global cost function can be found in Fig. 5

#### D. Casting seams

In this step, we generated the separate seams by combining the baseline energy map and the global cost function. Let denote the space from  $l_i$  to  $u_i$  (corresponding to the lower and upper ordinate of the candidate line  $L_i$ ),  $E_{l_i:u_i,w}$  the energy map and  $C_{l_i:u_i,w}$  the global cost function where  $w$  is the width of the image. Seam carving method consists of two steps: a forward step and a backward step. In the forward step, we initialize seams at each vertical pixel from  $l_i$  to  $u_i$ . Each seam move from the current column to next column by computing the scores  $M_{y,x}$  and  $M_{y,x+1}$  as follow :

$$M_{y,x} = E_{y,x} \quad (4)$$

$$M_{y,x+1} = E_{y,x+1} + \alpha * C_{y,x+1} + \min(M_{y-1:y+1,x-1}) \quad (5)$$

with  $y \in [l_i, u_i]$  and  $x \in [1, w]$ . This process is sequentially done from the first column on the left to the last column on the right of the image  $I$ . In the backward step, we construct the separating seam at the position with lowest cumulative score of  $M_{l_i:u_i,w}$  then we go back from right to left by choosing the pixel with the lowest cost on the left column. To create a continuous separation line, only the adjacent pixels are considered (pixels at previous, current and next line).

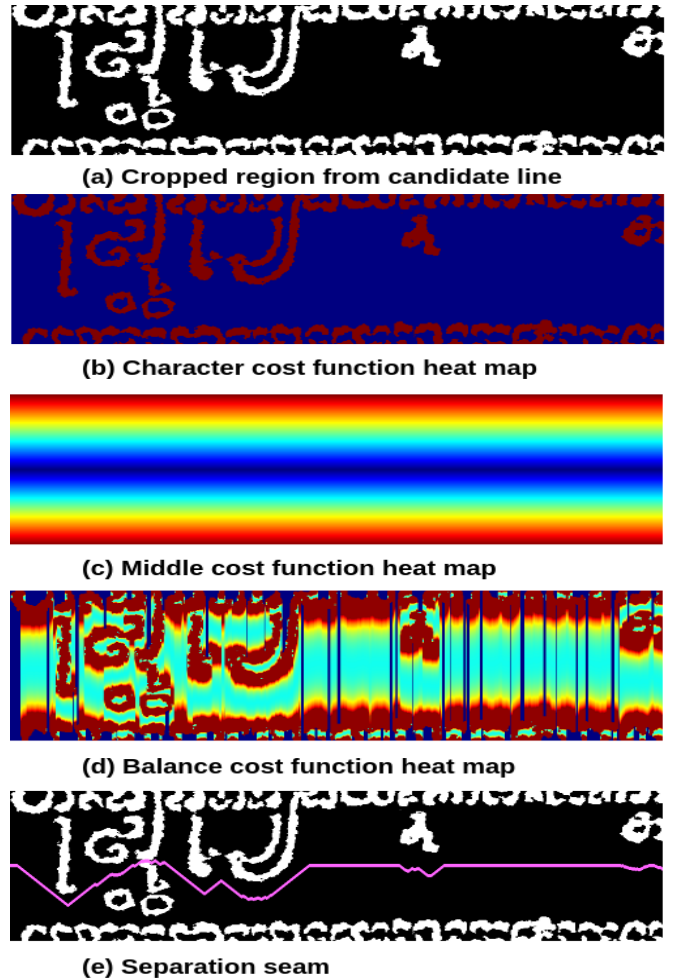


Fig. 5: Heat map of each term of the global cost function. From top to bottom: input image, Character cost, Middle cost and Balance cost functions; Separation seams (in purple). For the character cost function, high values (red) correspond to character pixels while low values (blue) correspond to the background one. In the middle cost function, pixels closed to the middle line are in blue and represent lower values, values increase (cyan, green, yellow and red) for pixels moving away from the middle line. For the last cost function, the middle gaps between two lines has the lowest values (cyan) and the values increases (red) when getting closer to the previous or the next line.

## IV. EXPERIMENTS

In this section, we conducted the experiments on two datasets: the DIVA-HisDB [5] and our own ChamDoc dataset. We use the DIVA-HisDB dataset [5] to compare our approach with other state-of-the-art methods.

### A. Datasets

1) *DIVA-HisBD dataset*: DIVA-HisDB dataset [5] contains 150 pages of three different medieval manuscripts (CB55, CSG863, CSG18) dating from 11th to the 14th century and

digitized with a resolution of 600 dpi camera. The collection consists of complex layout containing main text body, interlinear glosses, additions and corrections. Ground truths are provided at both pixel and text level. The particularity of this dataset includes: interlinear glosses, lettrines decoration, different script sizes and across manuscripts. As some works on this dataset, we used the good binarization provided with the dataset as input of our algorithm. For this experiment, we use the private test set with 30 pages (875 lines) of the ICDAR 2017 competition [5].

2) *ChamDoc dataset*: Our ChamDoc dataset consists of various images which have been obtained from the stamping of the inscriptions. Due to the highly degraded, the original inscription images are cleaned by our previous work [21]. After that, the entire cleaned images are then segmented manually to the line by line image by our expert linguistic at line level. Totally, we have 395 lines from 26 inscriptions images. Dataset can be downloaded at <http://13i-share.univ-lr.fr/2022CHAMDoc/CHAMDoc.html>.

3) *Evaluation metric*: In order to evaluate the results of the experiments and facilitate the comparison with other approaches, we use two sets of metrics. The first one from [22] includes Detection Rate (DR), Recognition Accuracy (RA), F-measure (FM) is used with the ChamDoc dataset. The evaluation metrics are defined as follows:

$$DR = \frac{M}{N_1}; RA = \frac{M}{N_2}; FM = \frac{2 * DR * RA}{DR + RA}$$

where  $N_1$  and  $N_2$  are respectively the number of lines in ground truth and number of segmented lines.  $M$  is the number of one-to-one matches. Matches are obtained when the intersection region between segmented and ground truth line is greater than a threshold  $\lambda$ .  $\lambda$  is defined experimentally and set to 90 for the ChamDoc dataset.

For the second set of metrics used with the Diva-HisBD dataset, we have adopted the metrics Line IU and Pixel IU [5] used in different papers of the literature. Pixel IU is defined by :

$$\text{Pixel IU} = \frac{TP}{TP + FP + FN} \quad (6)$$

$TP$  is the number of pixels that are correctly detected,  $FP$  is the number of extra pixels and  $FN$  is the number of missed pixels. All the metrics are taken into account the only foreground pixels. Line IU is defined by :

$$\text{Line IU} = \frac{CL}{CL + ML + EL} \quad (7)$$

where  $CL$  is the number of lines detected correctly,  $ML$  is the number of missed lines and  $EL$  is the number of extra lines. Lines are considered correct when line precision and line recall are above a given threshold. We used threshold 75 as the same setup of evaluation tools provided in the competition [5].

### B. Parameters determination

The proposed approach needs to define several parameters. For the smoothing filter, the value  $\sigma$  depends on the space between lines in the documents. We set the value to  $1e - 6$

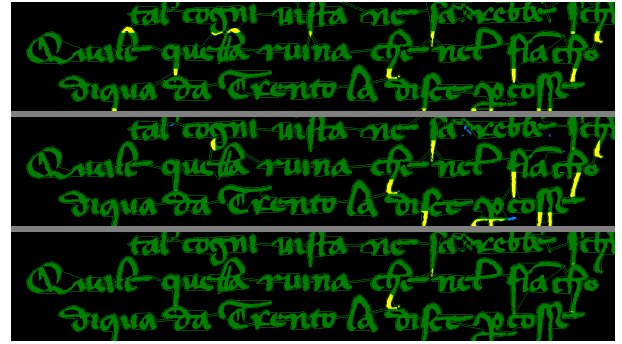


Fig. 6: Qualitative results on DIVA-HisDB. From top to bottom: Seam carving method with only global cost function, LCG [3], the proposed method. The green part correspond to the correct segmentation of foreground pixels while the yellow part corresponds to the wrong segmentation.

on ChamDoc dataset and  $5e - 4$  on the DIVA-HisDB. These values have been defined experimentally. The three weights of the global cost function  $w_1$ ,  $w_2$  and  $w_3$  are set experimentally respectively to 10, 3, 0.02 (character cost function, middle cost function and balance cost function) for the ChamDoc dataset. We found that with the DIVA-HisDB dataset, the weight related to the balance cost function has a minor impact. So we have used the values 10, 1, 0.01.

### C. Results

1) *DIVA-HisDB dataset*: Table I shows the comparison results of our method with different approaches on DIVA-HisDB dataset. We can note that, in comparison with seam carving

TABLE I: Comparison results on the DIVA-HisDB dataset

|                                      | Line IU      | Pixel IU     |
|--------------------------------------|--------------|--------------|
| Wavelength (Seam Carving based) [11] | 97.86        | 97.05        |
| LCG [3]                              | <b>99.42</b> | 96.76        |
| FCN+EM [23]                          | 99.21        | 97.53        |
| UTLS [17]                            | 97.85        | 97.04        |
| Seam carving with only cost function | 99.36        | 96.88        |
| The proposed method                  | <b>99.36</b> | <b>98.86</b> |

based method [11], our approach improved both Line IU and Pixel IU scores. The comparison with deep learning approach [23], [17], shows that our method obtains competitive results for Line IU score and the highest result for the Pixel IU score. Figure 6 illustrates some qualitative results. We observed that the proposed method can avoid most of the ascending and descending parts of the characters.

2) *ChamDoc dataset*: We conducted two experiments in ChamDoc dataset. The first experiment aims to evaluate the role of each component of the proposed method. The obtained

TABLE II: Ablation study on the ChamDoc dataset

|                                      | DR           | RA           | FM           |
|--------------------------------------|--------------|--------------|--------------|
| Only energy map                      | 80.76        | 81.58        | 81.17        |
| Only global cost function            | 90.37        | 91.30        | 90.84        |
| The proposed method w/o Balance cost | 90.63        | 91.56        | 91.09        |
| The proposed method                  | <b>91.14</b> | <b>92.07</b> | <b>91.60</b> |

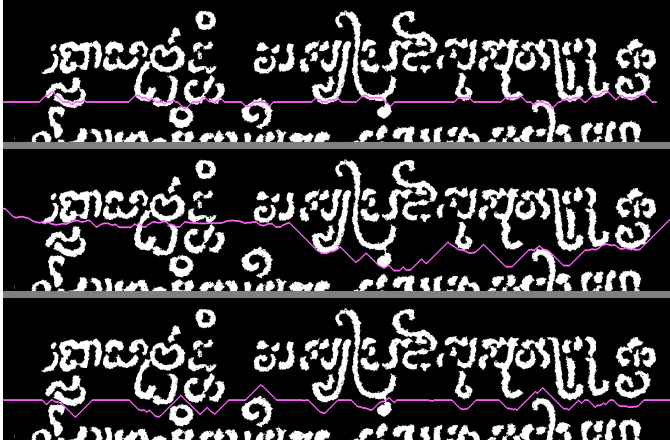


Fig. 7: Qualitative results on ChamDoc dataset. From top to bottom: A\*path planning [12], seam carving based [7], proposed method. Separation seams are in purple

results for ChamDoc dataset are shown in Tab. II.

First, we evaluated results when using only energy map (cf. Section III-B), only global cost function (cf. Section III-C) and the proposed method which combines energy map and global cost function. The proposed method shows a significant improvement in comparison with the experiments using only energy map or only global cost function. We also evaluate the role of balance cost function in the global cost of the proposed method. Experimental results show that thanks to balance cost function, the values of DR, RA and FM are increased from 90.63, 91.56 and 91.09 to 91.14, 92.10, 91.60 respectively. One qualitative example is given in Fig. 8. In the second experiment, we compared the results of the proposed method with three baseline approaches that are seam-carving-based [7] and A\* path planning [12] and LCG [3]. The results are shown in Table III. On ChamDoc dataset, the proposed method obtained 91.14%, 92.10% and 91.60% for DR, RA and FM metrics while those of seam-carving-based method [7] and [3] are 54.68%, 52.17%, 53.40% and 65.31%, 72.47%, 68.70%. In comparison with A\* path planning [12], the score is 84.8%, 85.71%, 85.25% respectively. Some examples results are shown in Fig. 7. We can observe that with the proposed method, separating seams are more flexible and can avoid the characters when going back to the middle line by choosing empty spaces between the two lines.

TABLE III: Comparison with the state-of-the-art methods on ChamDoc dataset

|                        | DR           | RA           | FM           |
|------------------------|--------------|--------------|--------------|
| Seam-carving based [7] | 54.68        | 52.17        | 53.40        |
| LCG [3]                | 65.31        | 72.47        | 68.70        |
| A* path planning [12]  | 84.8         | 85.71        | 85.25        |
| The proposed method    | <b>91.14</b> | <b>92.10</b> | <b>91.60</b> |

#### D. Discussion

The proposed method is based on the combination of three cost functions. In these experiments, the weight assigned to

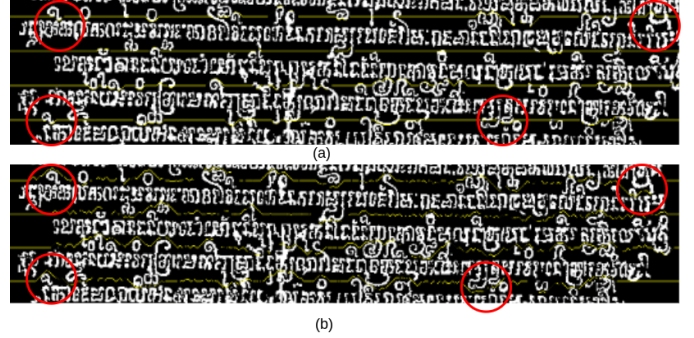


Fig. 8: Text line segmentation results without (a) / with (b) balance cost function. The red circles indicate the improvement between results.

each cost function has to be chosen experimentally according to the characteristics of the used dataset. First of all, the proposed global cost function determines the way to split the lines, so the impact of the structure of a given dataset on the weights has to be studied. The general question is how to appropriate the selection of these different weights for each dataset. You will find below some empirical rules to determine them. In the section IV-C, we can see that Character cost function is relevant when the weight is high (around 10) because the separation line will try to avoid the white pixels as much as possible. This is highest priority. The next priority is the weight of Middle cost function (around 1 to 3) because after avoiding the white pixels the separation line can be far from the middle so it is very difficult to avoid the next ascending or descending character. So if we set a highest weight for this cost function the separation line will go through a lot of white pixels. The last cost function, is designed for documents with diacritics, ascending and descending characters such as Cham documents as shown in Figure 8, it is not suitable with all cases, so the smaller weight to get better results in term of total performance.

## V. CONCLUSION

In this paper, we present a simple and efficient method for text line segmentation of the historical documents. By considering the properties of the documents, we propose an appropriate set of cost functions to improve the accuracy of text line segmentation task. Experiments on our ChamDoc dataset and DIVA-HisDB dataset prove the relevance of our method. We are planning to extend the pre-processing step to tackle some special cases lines that are more skewed or curved. Moreover, our objective is to adapt the approach to obtain a parameter-free model to be able to process different type of documents easily.

## VI. ACKNOWLEDGEMENTS

This work is supported by the French National Research Agency (ANR) in the framework of the ChAMDOC Project, n°ANR-19-CE27-0018-02.

## REFERENCES

- [1] S. S. Singh and S. Karayev, "Full page handwriting recognition via image to sequence extraction," *CoRR*, vol. abs/2103.06450, 2021. [Online]. Available: <https://arxiv.org/abs/2103.06450>
- [2] C. Wigginton, C. Tensmeyer, B. Davis, W. Barrett, B. Price, and S. Cohen, "Start, follow, read: End-to-end full-page handwriting recognition," in *Computer Vision – ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham: Springer International Publishing, 2018, pp. 372–388.
- [3] M. Alberti, L. Vögtlin, V. Pondenkandath, M. Seuret, R. Ingold, and M. Liwicki, "Labeling, cutting, grouping: an efficient text line segmentation method for medieval manuscripts," *CoRR*, vol. abs/1906.11894, 2019.
- [4] M.-T. Nguyen, A.-V. Shweyer, T.-L. Le, T.-H. Tran, and H. Vu, "Preliminary results on ancient cham glyph recognition from cham inscription images," in *2019 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, 2019, pp. 1–6.
- [5] F. Simistira, M. Bouillon, M. Seuret, M. Würsch, M. Alberti, R. Ingold, and M. Liwicki, "Icdar2017 competition on layout analysis for challenging medieval manuscripts," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 01, 2017, pp. 1361–1370.
- [6] R. Manmatha and N. Srimal, "Scale space technique for word segmentation in handwritten documents," in *Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision*, ser. SCALE-SPACE '99. Berlin, Heidelberg: Springer-Verlag, 1999, p. 22–33.
- [7] N. Arvanitopoulos and S. Süsstrunk, "Seam carving for text line extraction on color and grayscale historical manuscripts," in *2014 14th International Conference on Frontiers in Handwriting Recognition*, 2014, pp. 726–731.
- [8] Y. Boulid, A. Souhar, and M. Y. Elkettani, "Arabic handwritten text line extraction using connected component analysis from a multi agent perspective," in *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2015, pp. 80–87.
- [9] D. Valy, M. Verleysen, and K. Sok, "Line segmentation approach for ancient palm leaf manuscripts using competitive learning algorithm," in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 2016, pp. 108–113.
- [10] A. Asi, R. Saabni, and J. El-Sana, "Text line segmentation for gray scale historical document images," in *Proceedings of the 2011 Workshop on Historical Document Imaging and Processing*, ser. HIP '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 120–126. [Online]. Available: <https://doi.org/10.1145/2037342.2037362>
- [11] M. Seuret, D. S. Ben Ezra, and M. Liwicki, "Robust heartbeat-based line segmentation methods for regular texts and paratextual elements," in *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing*, ser. HIP2017. New York, NY, USA: Association for Computing Machinery, 2017, p. 71–76. [Online]. Available: <https://doi.org/10.1145/3151509.3151521>
- [12] O. Surinta, M. Holtkamp, F. Karabaa, J.-P. Van Oosten, L. Schomaker, and M. Wiering, "A path planning for line segmentation of handwritten documents," in *2014 14th International Conference on Frontiers in Handwriting Recognition*, 2014, pp. 175–180.
- [13] A. Nicolaou and B. Gatos, "Handwritten text line segmentation by shredding text into its lines," in *2009 10th International Conference on Document Analysis and Recognition*, 2009, pp. 626–630.
- [14] Z. Shi, S. Setlur, and V. Govindaraju, "A steerable directional local profile technique for extraction of handwritten arabic text lines," in *2009 10th International Conference on Document Analysis and Recognition*, 2009, pp. 176–180.
- [15] R. Cohen, I. Dinstein, J. El-Sana, and K. Kedem, "Using scale-space anisotropic smoothing for text line extraction in historical documents," 10 2014, pp. 349–358.
- [16] B. K. Barakat, A. Drobny, M. Kassis, and J. El-Sana, "Text line segmentation for challenging handwritten document images using fully convolutional network," *CoRR*, vol. abs/2101.08299, 2021. [Online]. Available: <https://arxiv.org/abs/2101.08299>
- [17] B. K. Barakat, A. Drobny, R. Alasam, B. Madi, I. Rabaev, R. Shammes, and J. El-Sana, "Unsupervised text line segmentation," *CoRR*, vol. abs/2003.08632, 2020. [Online]. Available: <https://arxiv.org/abs/2003.08632>
- [18] T. Grüning, G. Leifert, T. Strauß, J. Michael, and R. Labahn, "A two-stage method for text line detection in historical documents," *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 22, no. 3, pp. 285–302, 2019.
- [19] O. Mechi, M. Mehri, R. Ingold, and N. Essoukri Ben Amara, "Text line segmentation in historical document images using an adaptive u-net architecture," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, 2019, pp. 369–374.
- [20] G. Levi and U. Montanari, "A grey-weighted skeleton," *Information and Control*, vol. 17, no. 1, pp. 62–91, 1970. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0019995870800067>
- [21] T.-N. Nguyen, J.-C. Burie, T.-L. Le, and A.-V. Shweyer, "On the use of attention in deep learning based denoising method for ancient cham inscription images," in *Document Analysis and Recognition – ICDAR 2021*, J. Lladós, D. Lopresti, and S. Uchida, Eds. Cham: Springer International Publishing, 2021, pp. 400–415.
- [22] N. Stamatopoulos, B. Gatos, G. Louloudis, U. Pal, and A. Alaei, "Icdar 2013 handwriting segmentation contest," in *2013 12th International Conference on Document Analysis and Recognition*, 2013, pp. 1402–1406.
- [23] B. K. Barakat, A. Drobny, R. Alasam, B. Madi, I. Rabaev, and J. El-Sana, "Text line extraction using fully convolutional network and energy minimization," *CoRR*, vol. abs/2101.07370, 2021. [Online]. Available: <https://arxiv.org/abs/2101.07370>