

Extraction of Virtual Baselines from Distorted Document Images Using Curvilinear Projection

Gaofeng MENG^{*†}, Zuming HUANG[†], Yonghong SONG[‡],
Shiming XIANG[†] and Chunhong PAN[†]

[†] National Laboratory of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences, Beijing P.R.China, 100190

Email: {gfmeng, smxiang, chpan}@nlpr.ia.ac.cn, huangzuming2014@ia.ac.cn

[‡] Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University
Xi'an, Shaanxi Province, P.R.China 710049

Email: songyh@mail.xjtu.edu.cn

Abstract

The baselines of a document page are a set of virtual horizontal and parallel lines, to which the printed contents of document, e.g., text lines, tables or inserted photos, are aligned. Accurate baseline extraction is of great importance in the geometric correction of curved document images. In this paper, we propose an efficient method for accurate extraction of these virtual visual cues from a curved document image. Our method comes from two basic observations that the baselines of documents do not intersect with each other and that within a narrow strip, the baselines can be well approximated by linear segments. Based upon these observations, we propose a curvilinear projection based method and model the estimation of curved baselines as a constrained sequential optimization problem. A dynamic programming algorithm is then developed to efficiently solve the problem. The proposed method can extract the complete baselines through each pixel of document images in a high accuracy. It is also scripts insensitive and highly robust to image noises, non-textual objects, image resolutions and image quality degradation like blurring and non-uniform illumination. Extensive experiments on a number of captured document images demonstrate the effectiveness of the proposed method.

1. Introduction

The baselines of a document page consist of a set of virtual horizontal and parallel lines, to which the printed contents of document, e.g., text lines, tables or inserted photos, are aligned. Many methods for document image anal-

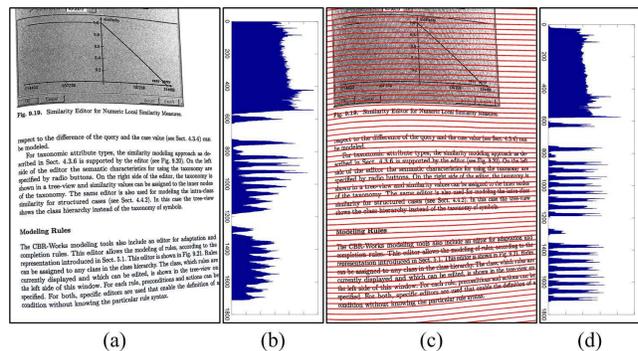


Figure 1. An example of virtual baseline extraction from a curved document image by our method. (a) the curved document image, (b) the projection profiles along a fixed direction, (c) the estimated virtual baselines, (d) the curvilinear projection profiles along the estimated baselines.

ysis, for example, the recursive XY-cut method [18], often assume that the image contents are well aligned to a group of horizontal and parallel straight baselines. However, this is often not the case when a document image gets distorted. Estimating the baselines and realigning the image contents with them is thus an indispensable preprocessing step in many tasks of document image processing and analysis, such as page layout segmentation, optical character recognition (OCR) and so on.

For a flat document image, its baselines are usually straight lines. The estimation of these baselines has been extensively studied in the context of document image deskewing [17] and image perspective rectification [14, 22]. The problem, however, gets worse when a document image is distorted due to nonlinear page curl. This commonly happens when one captures the images of an opened thick and

bound book by a hand-held camera.

The extraction of baselines is of great importance to many methods for the rectification of geometric distortion in camera-captured document images [4, 28, 7, 8, 12, 13, 15, 16, 24]. To yield a desirable result, these methods require the curved baselines to be estimated reliably in a high accuracy. However, this is often a quite challenging task, due to image distortion, various types of non-textual objects and the image quality degradation introduced during the imaging process such as image blurring, low resolutions and non-uniform illumination.

Typically, the baselines of a document image can be estimated by fitting the horizontal text lines. To this end, horizontal text lines in document images have to be firstly extracted. According to how the text lines are obtained, these methods can be roughly classified into four major categories, i.e., the tracing based methods [3, 4, 24, 25], the clustering based methods [9, 10, 22, 27], the segmentation based methods [11, 19, 23] and the projection based methods [20, 26, 2, 21].

Early methods for text lines extraction apply a tracing strategy to the connected components (CCs) of a binarized image [4, 16, 24]. These methods firstly pick up a connected component from the image as a seed, then perform seed growing by linking the seed to its nearest neighbors. Tracing on CCs is generally scripts sensitive and very unstable. For characters with multiple components, *e.g.*, Chinese characters, tracing on them often fails to yield the correct text lines.

Later improvements directly implement tracing on the gray-scale images. Tian and Narasimhan [25] observe that patches extracted from a set of points along a text line are similar to each other. They thereby propose an interesting line tracing method based on self-similarity measure between image patches. A very similar idea is also used in Liang *et al.*'s work [12], where texture flow is introduced for tracing curved text lines. However, the tracing based methods are vulnerable to page layouts, changes of font sizes and non-textual image contents, leading to an inaccurate tracing on curved text lines.

In most cases, it is beneficial to view text lines extraction as a clustering problem of CCs. From this perspective, Yin and Liu [27] propose a bottom-up method for text lines segmentation in unconstrained handwritten Chinese documents. The method first designs a distance metric between CCs by supervised learning. Based upon it, CCs of the document image are grouped into a tree structure, from which text lines can be extracted by dynamically cutting the edges. Koo and Cho [9, 10] formulate text lines extraction as an energy minimization problem on the states of CCs. A cost function that encodes the interactions between text lines and the curvilinearity of each text line is proposed. Their method is robust to the interference between text lines, spa-

tially varying skew and irregular inter-character distance. In comparison to the tracing based approaches, the methods using CCs clustering are generally more robust to the curling of text lines. However, these methods often suffer from heuristic merging rules, artificial parameters and topological changes of CCs [10].

The segmentation based methods [19, 23, 11] treat text lines extraction as an image segmentation problem. Inspired by the great success of seam carving in image resizing [1], Raid *et al.*[23] propose to use seam carving to automatically segment text lines from binary or gray-scale images of handwritten documents. However, their method often produces seams that cut through words and line components. Nikolaos and Sabine [19] later improve the method by incorporating some constraints into the optimization procedure to yield more robust separating seams. The segmentation based methods are more general-purpose and require less knowledge about document layouts and scripts. However, these methods share the similar limitations that most segmentation methods may have. They are sensitive to image noises, changes of image resolutions and overlapping of adjacent text lines.

The projection based methods [20, 6, 2, 21] have been extensively studied to segment text lines in handwritten documents. These methods firstly compute the projection profiles over the entire image [6] or vertical strips [20, 2, 21], and then find the peaks and valleys of the projection to locate the text lines. Despite the effectiveness of these methods in segmenting handwritten text lines, they generally cannot be directly used to extract baselines from a distorted document image, due to the complicated page layouts and existing of large areas of non-textual objects in images.

In this paper, we propose a novel curvilinear projection based method for accurate extraction of curved baselines from a distorted document image. Our method comes from two basic observations that the baselines of a curved document image do not intersect with each other and that within a narrow vertical strip, the baselines can be well approximated by linear segments. Based upon these observations, we model baseline extraction as a constrained sequential optimization problem on the projection map of vertical strips. A dynamic programming algorithm is also developed to efficiently solve the problem.

Our method is a segment-free method and can directly extract the curved baselines from the images. It has a much lower computation complexity and can be implemented very efficiently. Moreover, our method can exploit more general types of visual cues for curved baselines estimation, including curved text lines, horizontal lines in tables, boundaries of pages and inserted photos. It shows great robustness to various scripts, different page layouts, non-textual objects, changes of image resolutions and image quality degradation like image blurring and non-uniform il-

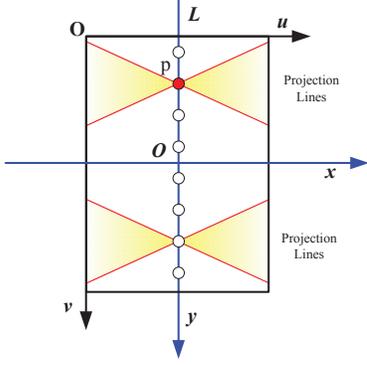


Figure 2. The basic idea for optimal strip projections. We can determine the optimal projection lines in sequence at each point on the central line of the strip. In the figure, all the possible projection lines passing through p are illustrated in the shading area between two red lines. For the convenience of computation, two coordinate systems are also defined on the vertical strip.

lumination. Figure 1 illustrates an example of baseline extraction from a curved document image by our method.

2. Approach

2.1. Basic Assumptions

To begin with, we have to make two basic assumptions on the documents. First, we assume that the document page is not being significantly bent so that a small segment of its baseline can be well approximated by a straight line. Second, we require that the printed contents of a document page are aligned in horizontal lines. This means that some complicated hybrid layouts, which allow the mixture of horizontal and vertical alignments of text lines, are not allowed.

2.2. Curvilinear Projection for Baseline Extraction

To extract the curved baselines, we first compute the edges of an input document image, and then divide the edge map into a sequence of overlapping vertical strips. According to the above assumptions, the baselines of a vertical strip can be approximated by a group of nonintersecting straight line segments, given that the width of the strip is not too large. Generally, these straight lines are not parallel to each other due to the curling of documents and the perspective effects of camera. The problem is then to find a group of optimal projection lines that do not intersect within the boundaries of a vertical strip.

Our basic idea to address this problem is to determine these projection lines in sequence for each point on the central line of the strip, as illustrated in Figure 2. An objective function on these projection lines is defined to determine the optimal projection. The resulted problem turns out to be a typical sequence optimization problem. More details will be described in the following sections.

2.2.1 Projection map from Radon transform

For a point p on the central line L of a strip, consider all the possible projection lines passing through this point. To decide which line is finally chosen from these candidate lines, we have to compute the projection profiles of the strip along any possible lines. This requires us to compute the projection map, from which all possible projection profiles of the strip can be directly obtained. This projection map can be easily obtained from the Radon transform of the strip.

Denote the width and height of an image strip S by w and H , respectively. Further define a projection coordinate system on S , with its origin locating at the center of the strip and its x and y axes parallel to the corresponding axes in the image coordinate system, as illustrated in Figure 2. The Radon transform of S in a given angle θ is defined by the line integral along each of the following projection lines:

$$x \cos \theta - y \sin \theta = \rho, \quad (1)$$

where (x, y) is defined in the projection coordinate system of S , θ is the angle from the normal vector of the line to the x axis in clockwise, and ρ is the distance of the line from the origin.

Radon transform will yield a 2D array associated with θ and ρ . We denote this array by $\mathfrak{R}(\rho, \theta)$ for simplicity. Every pair of (ρ, θ) defines a projection line and the corresponding $\mathfrak{R}(\rho, \theta)$ gives the projection value. Now, rewriting the equations of the projection lines of Eq.(1) in the image coordinate system, we have

$$\left(u - \frac{w}{2}\right) \cos \theta - \left(v - \frac{H}{2}\right) \sin \theta = \rho, \quad (2)$$

where (u, v) is defined in the image coordinate system of S . Using this equation, we can quickly obtain all the candidate projection lines passing through a point $p(\frac{w}{2}, k)$ on the central line of S , i.e.,

$$\left(\frac{H}{2} - k\right) \sin \theta = \rho, \quad (k = 1, \dots, H, \quad \theta \in [\alpha, \beta]) \quad (3)$$

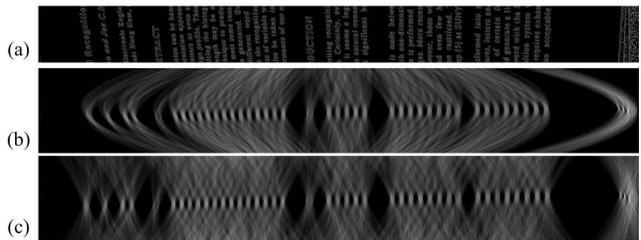


Figure 3. Computing the projection map from Radon transform. (a) the edge map of a strip. For the convenience of illustration, we rotate the strip from the vertical direction to the horizontal direction. (b) the Radon transform, (c) the computed projection map, which is a coordinate transformation of Radon transform from the ρ - θ coordinates system to the k - θ coordinates system.

where k is the row index of p in the strip, α and β are the lower and upper bounds of the possible angles of projection lines, respectively.

According to Eq.(3), the Radon transform $\mathfrak{R}(\rho, \theta)$ can be easily converted from the (ρ, θ) coordinates system to the (k, θ) coordinates system. This transformed array, denoted by $\mathbf{R}(k, \theta)$, is called the *projection map*. One great benefit of using the projection map $\mathbf{R}(k, \theta)$ is that it can explicitly give the projection value of the strip along any line passing through a point on L . Figure 3 illustrates the obtained projection map from Radon transform.

2.2.2 Optimized strip projection

We can estimate the optimal projection lines across the central line L of a strip by solving the following constrained optimization problem, i.e.,

$$\max_{\theta_1, \dots, \theta_H} \sum_{k=1}^H \mathbf{R}^p(k, \theta_k) + \lambda \phi(\theta_1, \dots, \theta_H), \quad (4)$$

given the constraints that two adjacent projection lines defined by θ_k and θ_{k+1} do not intersect between the boundaries of the strip. The explicit expression of these constraints will be given in next section. In Eq.(4), p is a prescribed exponent (typically $p \geq 3$) and $\phi(\theta_1, \dots, \theta_H)$ is a smoothness measure that penalizes sharp changes in the angle sequence, and λ is a weight for balancing the two terms.

Geometrically, the optimization of Eq.(4) suggests to find an optimal path on the projection map $\mathbf{R}(k, \theta)$, ($k = 1, \dots, H$, $\theta \in [\alpha, \beta]$), which passes through \mathbf{R} from its left side to the right side under the non-intersection constraints. This problem can be efficiently solved by a dynamic programming method.

There are many ways to define the smoothness measure. Here we give an example of the first-order smoothness terms, which is defined on two adjacent angles, i.e.,

$$\phi(\theta_1, \dots, \theta_H) = \sum_{k=2}^H \exp\left(-\frac{(\theta_k - \theta_{k-1})^2}{2\sigma^2}\right), \quad (5)$$

where σ is used to control the sensitivity of the smoothness term to the angle differences. Similarly, high-order smoothness term that involves more adjacent angles can also be defined. However, it has to be pointed out that although high-order terms may have better performance in comparison with the low-order terms, optimization on them will result in an explosion in computation and storage.

2.2.3 Optimization

Discretizing θ uniformly between $[\alpha, \beta]$ by a fixed angle resolution, yields an angle sequence:

$$\alpha = \theta^1 < \theta^2 < \dots < \theta^{m-1} < \theta^m = \beta. \quad (6)$$

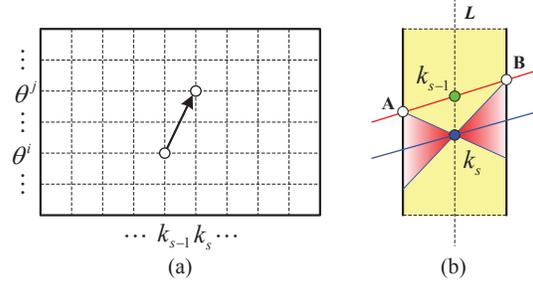


Figure 4. Construction of the weighted directed graph G . (a) only two vertices having adjacent k -coordinates can be connected by a directed edge, (b) deriving the non-intersection condition of two adjacent projection lines.

Similarly, taking samples of k between $[1, H]$ by a fixed interval Δ_k gives

$$1 = k_1 < k_2 < \dots < k_{n-1} < k_n = H. \quad (7)$$

We then construct a weighted directed graph G with totally $m \times n$ vertices on the discretized k - θ plane, as illustrated in Figure 4(a). Each vertex locates at the grid point (k_s, θ^j) ($1 \leq s \leq n, 1 \leq j \leq m$). In the graph G , only points that have adjacent k -coordinates can be connected by a directed edge. Two points (k_{s-1}, θ^i) and (k_s, θ^j) are connected if and only if their corresponding projection lines satisfy the non-intersection condition.

For shortness, we rewrite θ^i of vertex (k_{s-1}, θ^i) as $\theta_{k_{s-1}}^i$ and further use it to denote the vertex. We now derive the explicit form of the non-intersection condition on θ_{k_s} given $\theta_{k_{s-1}}^i$. Suppose the projection line defined by $\theta_{k_{s-1}}^i$ intersects with the two vertical boundaries of a strip at \mathbf{A} and \mathbf{B} respectively, as illustrated in Figure 4(b). The v -coordinates of \mathbf{A} and \mathbf{B} in the image coordinate system of strip can be computed respectively as:

$$v_A = k_{s-1} - \frac{w}{2} \cot \theta_{k_{s-1}}^i, \quad (8)$$

$$v_B = k_{s-1} + \frac{w}{2} \cot \theta_{k_{s-1}}^i, \quad (9)$$

where w is the width of the strip.

As illustrated in Figure 4(b), to avoid the intersection of the projection line with its previous one in the strip, the line must fall in the red angular area determined by \mathbf{A} , \mathbf{B} and point $(\frac{w}{2}, k_s)$. Thus, to satisfy the non-intersecting condition, θ_{k_s} has to be limited within:

$$\left[\cos^{-1}\left(\frac{k_s - v_A}{\sqrt{\frac{w^2}{4} + (k_s - v_A)^2}}\right), \cos^{-1}\left(\frac{v_B - k_s}{\sqrt{\frac{w^2}{4} + (v_B - k_s)^2}}\right) \right]. \quad (10)$$

In the construction of G , every directed edge is also assigned a weight. For an edge connecting two vertices $\theta_{k_{s-1}}^i$

and $\theta_{k_s}^j$, the weight on this edge, denoted by w_s^{ij} , is given as:

$$w_s^{ij} = \sum_{t=1}^{\Delta_k} \mathbf{R}^p(k_{s-1} + t, \theta_{k_{s-1}}^i + th) + \lambda \Delta_k \exp\left(-\frac{h^2}{2\sigma^2}\right), \quad (11)$$

where p is a predefined exponent in Eq.(4), Δ_k is the fixed interval in Eq.(7) for sampling k and

$$h = \frac{\theta_{k_s}^j - \theta_{k_{s-1}}^i}{\Delta_k}. \quad (12)$$

After G is constructed, the optimization of Eq.(4) turns out to be finding an optimal path that passes through G and maximizes the total sum of the weights of edges on the path. By adding a virtual starting node and ending node to the left and right side of G respectively, the problem turns to be a classic longest path problem, which can be efficiently solved by the Dijkstra's algorithm [5].

Figure 5 illustrated an example of the solved optimal path on \mathbf{R} and the estimated baselines of the strip. In the example, θ is uniformly discretized between 45° and 135° by an angle step 0.5° , and totally 30 samples of k are used to solve the path. The comparisons of strip projections along a fixed direction and the estimated baselines are also illustrated in the figure. From the results, we can see that the projections get mixed on the left side along the fixed direction. In comparison, the projections along the estimated baselines are much better in separability.

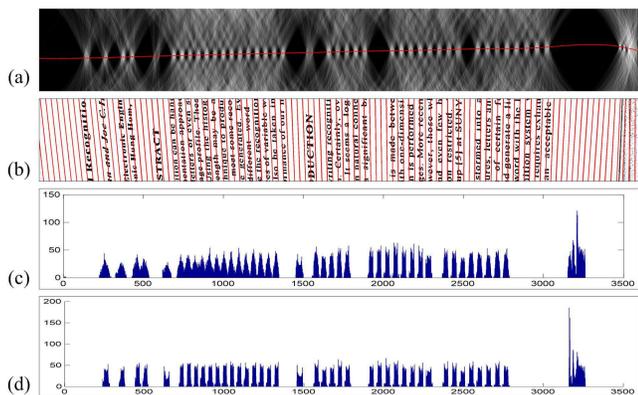


Figure 5. Results of the estimated baselines on the strip in Figure 3(a). (a) the solved optimal path (red curve) on \mathbf{R} , (b) the estimated baselines on the strip, (c) the projection profiles of strip along a fixed direction, (d) the projection profiles of strip along the estimated baselines.

2.3. Baselines Connection Between Strips

After the baselines on each strip are estimated, we further need to link them up to obtain the entire baselines on the whole document page. The connection of baselines between neighboring strips is much straightforward. In this

process, a basic step is to determine the ending point of a baseline giving its starting point. This can be quickly computed by using the equations of baselines in Eq.(3). Recall that two neighboring strips overlap each other. We thereby simply set the starting and ending points of all baselines of a strip at the center lines of the overlapping areas. Finally, every extracted baseline is further smoothed by a cubic spline.

3. Experimental Results

To test the performance of the proposed method, we carried out a series of experiments on a great number of curved document images. These testing images include samples of scanned book pages and pages captured from books, magazines and paper sheets by a hand-held camera. In the experiments, we empirically set the exponent $p = 3$ and $\lambda = 1$ in Eq.(4) and $\sigma = 1$ for the smoothness term in Eq.(5). In constructing the directed weighted graph, we took totally 30 samples of k and discretized θ uniformly in $[45^\circ, 135^\circ]$ by a fixed angle step 1° . Canny operator is used to compute the edge map of each input image. A 3×3 closing operation followed by a morphological removing and dilation operation is implemented on the edge map to remove some dotted noises. The edge map is then divided into 20-30 strips of same size according to image width, with every two neighboring strips sharing 50% overlapping.

Figure 6 illustrates some example results of baseline extraction by our method on curved binary document images from the DFKI dataset¹. This publicly available dataset is specially designed for the evaluation of various methods on geometric distortion rectification of curved document images. From the results, we can see that the proposed method can extract the curved baselines in a high accuracy. The estimated baselines are also robust to marginal noises and non-textual objects in the images, such as inserted formulas and photos. This is mainly because the proposed method can well exploit the available visual cues, such as text lines, horizontal lines in tables and page boundaries, to infer the baselines in the blank areas and non-textual regions.

We also give more results of our method on gray-scale document images in Figure 7. These images include pages captured by a hand-held camera or scanned by a flat-bed scanner from books, magazines and paper sheets. Note that some of the images contain large areas of non-textual objects. As can be seen from the results, our method works well on the gray-scale images. It can correctly extract the curved baselines in the blank and non-textual object regions. In the figure, the left-most three images come from English documents and the right-most four images come from Chinese documents. The method can extract the curved baselines for the both cases in high accuracy. This indicates that our method is not sensitive to different scripts.

¹The dataset can be downloaded from <http://www.csse.uwa.edu.au/~shafait/downloads.html>

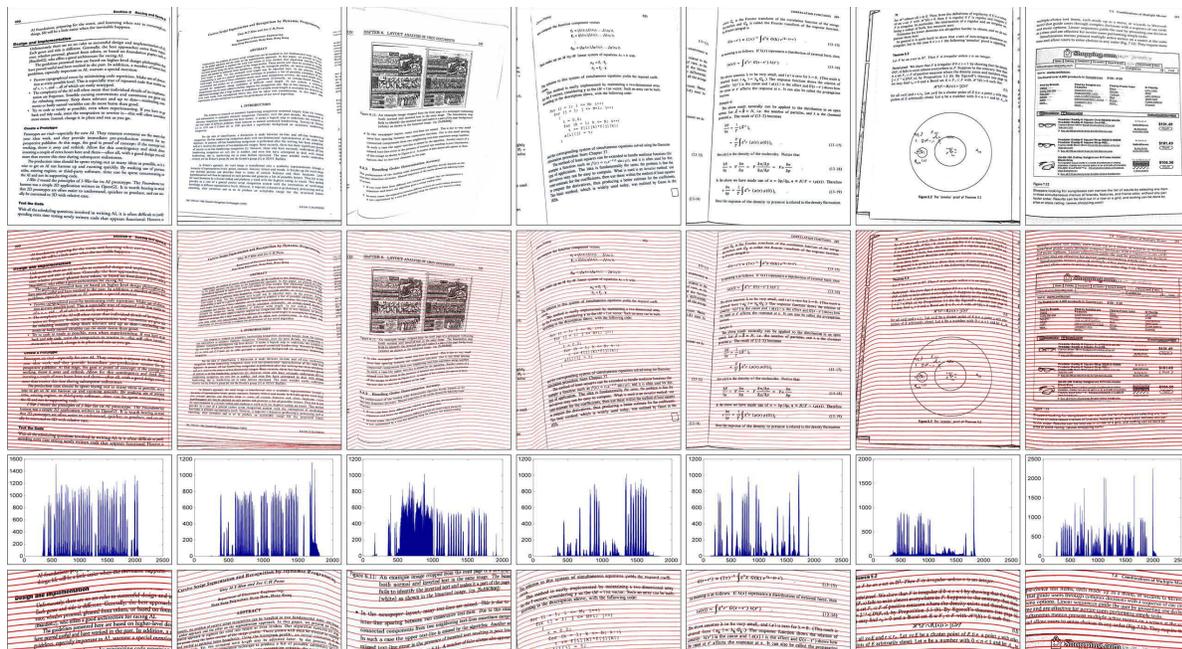


Figure 6. Example results of baseline extraction from the curved document images in DFKI dataset. From top to bottom: the curved document images, the extracted virtual baselines, the projections of image along the estimated baselines, the close-up image patches.

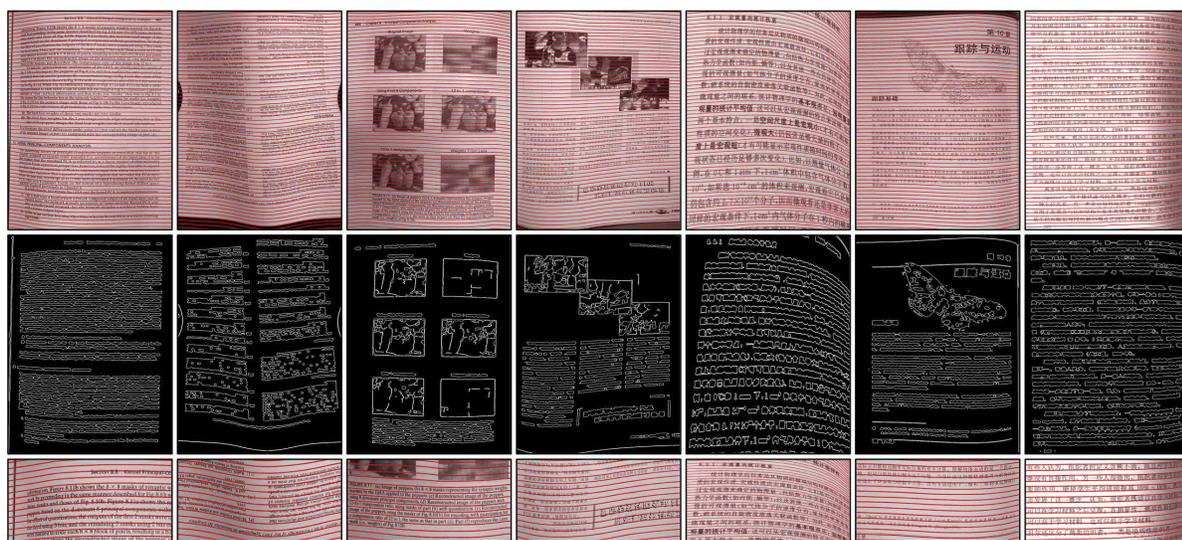


Figure 7. More results of baseline extraction of our method on some gray-scale images. From top to bottom are: the gray-scale images with the extracted baselines overlapped, the edge maps and the close-up image patches, respectively. The left-most three images come from English documents and the right-most four images come from Chinese documents. The last image is a scanned book page.

Out-of-focus blurring commonly happens to a camera-captured document image. This is because that the camera is hard to focus accurately on the document page when it gets distorted. To test the robustness of our method to such kind of blurring, we make a sequence of blurring images by filtering a clear document image with different Gaussian kernels. Figure 8 shows the results of baseline extraction on these images. As can be seen from the results, our

method yields comparably stable results for the images with different levels of blurring. This demonstrates that the proposed method is quite robust to image blurring. Since image down-sampling will result in a similar blurring, this also means that the proposed method is not sensitive to changes of image resolutions. The robustness to image blurring will benefit the process of image noise removal, since we can choose a larger kernel for image smoothing without signif-

icantly losing the accuracy in baseline extraction. The robustness to image blurring also means that the method is insensitive to the touched neighboring characters and text lines. These cases are very common to a curved document image captured by a hand-held camera.

Figure 9 shows the comparisons of our method with Koo and Cho's method [9]. Koo and Cho's method is a CCs grouping based method. Their method can extract the text lines from a curved document image and estimate the corresponding baselines by smooth curve fitting. In comparison with their method, our method can exploit the available visual cues in the image to estimate the complete baselines through each pixel, not merely limited to the baselines of horizontal text lines regions. This feature may greatly benefit many subsequent procedures of document images processing, such as geometric distortion correction and page layout analysis. In these cases, a complete baseline is much preferred [16, 24].

In Figure 10, we show the comparisons of our method with Tian and Narasimhan's method [25] and Nikolaos and Sabine's method [19]. Tian and Narasimhan's method is a line tracing based approach. This method relies on a self-similarity measure between image patches to trace the points on a text line. However, this similarity measure is very sensitive to document layouts, changes of image resolutions and non-textual objects, which often leads to the failure of tracing process, resulting in incomplete or even erroneous extraction of baselines. Although an elaborated refinement algorithm that further considers the parallelism between text lines is implemented to refine the estimations, the final results are not yet desirable in some cases. Nikolaos and Sabine's method is a segmentation based approach. They use the seam carving method to segment the text lines. This method can correctly divide the horizontal text lines when the image is not seriously distorted. However, if a document image is of poor qualities, for example, with low image resolutions and large geometric distortion (these factors are very common to a camera-captured document image), the performance of the method degrades rapidly. Moreover, large errors may also be introduced in the process of text lines fitting, even when the text lines are correctly segmented. In comparison, our method performs quite satisfactorily due to its great robustness to low image resolutions and image blurring. Our method is a segment-free approach and can directly extract the baselines without segmenting the curved horizontal text lines.

Our method takes about 6-8 seconds to process an image of 6M pixels. All the experiments are implemented on a PC with a 2.6GHz Intel(R) Core(TM) CPU and 4GB RAM. The implementation codes are written in Matlab without specialized code optimization. The running speed may be further increased by using an image down-sampling technique, since reduction of image resolutions will not greatly

decrease the accuracy of baseline extraction.

4. Discussions and Conclusion

4.1. Limitations

Our method relies on the local linearity assumption of the baselines. That is, the curved baselines can be well approximated by linear segments within a narrow image strip. However, if this assumption does not hold, the method may fail. This commonly happens to the captured document images with severe geometric distortions. Due to this reason, the method cannot be applied to the document images with non-smooth distortion, for example, folding distortion, in which the baselines have many abrupt turnings and cannot be well approximated by linear segments without any prior knowledge of the turning points.

Our method also shares the common limitations of most projection based methods. To estimate the baselines on an image strip, it requires the available visual cues within the strip to be dominant. As a result, the projection may fail if an image strip consists of very sparse visual cues or large areas of non-textual objects that have too little available information to infer the correct projection direction. This case generally occurs to page margins, where insufficient cues are available for the correct estimation of baselines.

4.2. Conclusion

We have proposed a curvilinear projection based method in this paper for virtual baseline extraction from a curved image of printed documents. Our method is motivated from two basic observations that the baselines of a curved document image do not intersect with each other and that in a narrow image strip, the baselines can be well approximated by straight line segments. Based on these observations, a constrained optimal curvilinear projection is proposed to estimate the baselines.

In comparison with the existing approaches, the proposed method has a much low computation complexity and can be implemented very efficiently. It is a segment-free method and can extract the complete baselines in a high accuracy for every pixel in the image, not merely limited to the baselines of textual regions. Moreover, our method is scripts insensitive and can well exploit more general types of visual cues in the image for baselines estimation, such as horizontal text lines, lines in tables and boundaries of inserted photos and pages. It shows great robustness to image noises, non-textual objects, changes of image resolutions and image quality degradation like image blurring and non-uniform illumination.

Acknowledgments

The authors would like to thank the anonymous reviewers and area chairs for their valuable remarks and sugges-

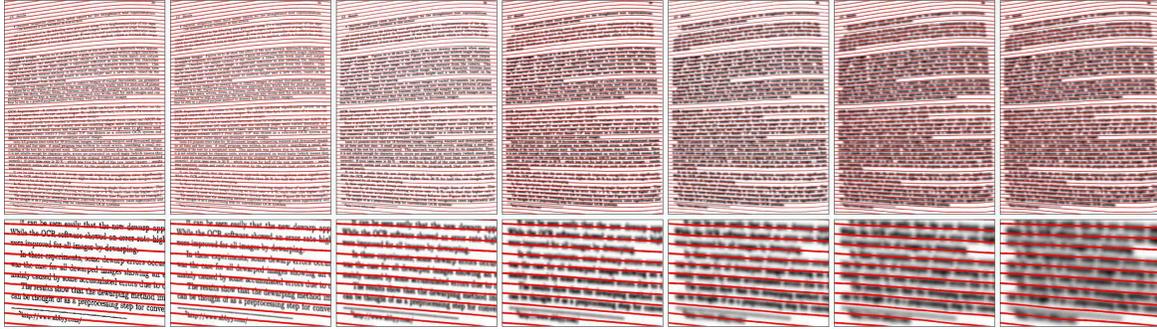


Figure 8. Results of baseline extraction on a sequence of blurring document images. These images are made by filtering a clear image of size 1611×2428 using a 75×75 Gaussian kernels with standard deviations σ ranging from 0 to 12 by a fixed step 2. (top) the extracted baselines overlapping on the blurring images, (bottom) the close-up patches for details illustration.

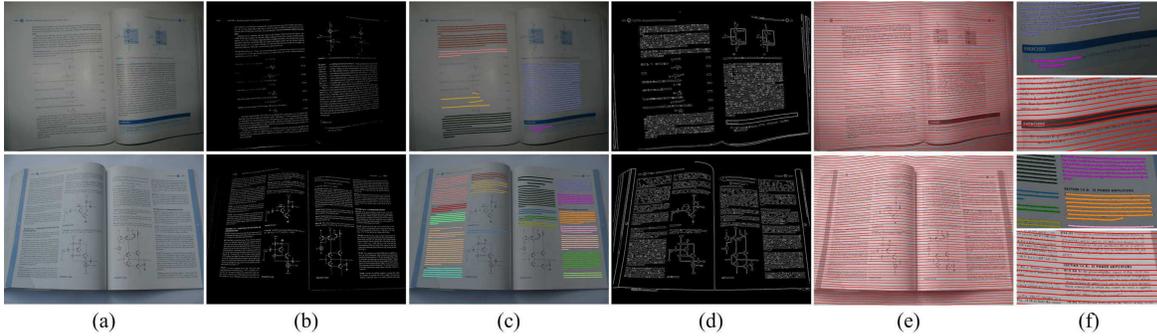


Figure 9. Comparisons of our method with Koo and Cho's method [9]. (a) the curved document images, (b) edge maps used by Koo and Cho's method, (c) text-lines extraction and fitting by Koo and Cho's method, (d) edge maps used by our method, (e) baseline extraction by our method (totally 25 strips used), (f) the close-up image patches of results by Koo and Cho's method (top) and our method (bottom), respectively.

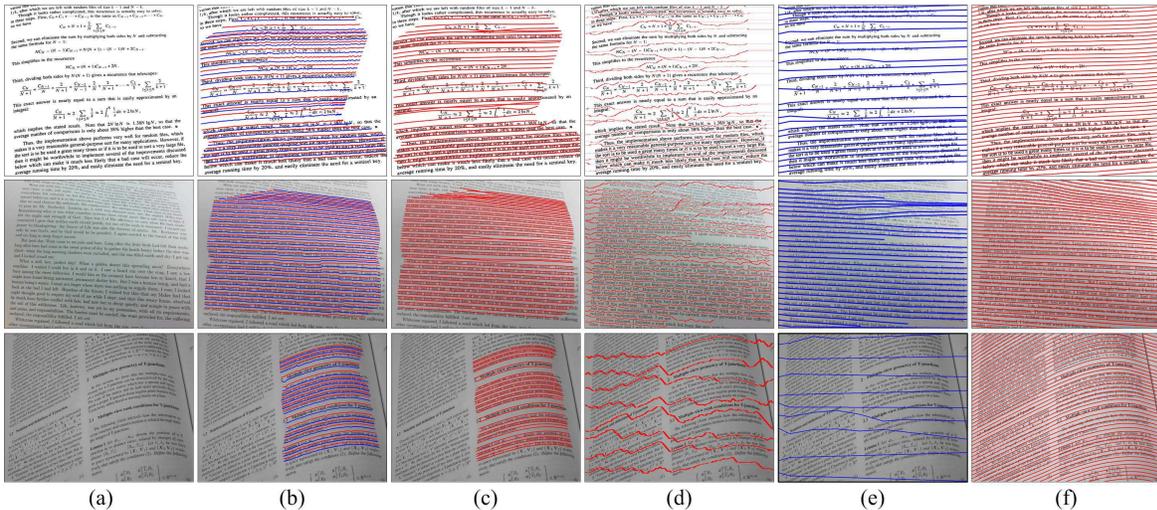


Figure 10. Comparisons of our method with Tian and Narasimhan's method [25] and Nikolaos and Sabine's method [19]. (a) the curved images, (b) the coarse estimation of baselines by Tian and Narasimhan's method, (c) the refinement results of Tian and Narasimhan's method, (d) the separating seams produced by Nikolaos and Sabine's method, (e) the estimated baselines by Nikolaos and Sabine's method, (f) the extracted baselines by our method.

tions on improving this paper. This work was supported in part by the projects of the National Natural Science Founda-

tion of China (Grant No. 61370039, 61272331, 91338202).

References

- [1] S. Avidan and A. Shamir. Seam carving for content-aware image resizing. *ACM Transactions on graphics (TOG)*, 26(3):10, 2007.
- [2] I. Bar-Yosef, N. Hagbi, K. Kedem, and I. Dinstein. Line segmentation for degraded handwritten historical documents. In *Proceedings of the 10th International Conference on Document Analysis and Recognition*, pages 1161–1165, July 2009.
- [3] S. Bukhari, F. Shafait, and T. Breuel. Text-line extraction using a convolution of isotropic gaussian filter with a set of line filters. In *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR)*, pages 579–583, Sept 2011.
- [4] H. Cao, X. Ding, and C. Liu. A cylindrical surface model to rectify the bound document image. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 228–233, 2003.
- [5] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959.
- [6] R. dos Santos, G. Clemente, T. I. Ren, and G. Cavalcanti. Text line segmentation based on morphology and histogram projection. In *Proceedings of the 10th International Conference on Document Analysis and Recognition*, pages 651–655, July 2009.
- [7] H. I. Koo. Segmentation and rectification of pictures in the camera-captured images of printed documents. *IEEE Transactions on Multimedia*, 15(3):647–660, 2013.
- [8] H. I. Koo and N. I. Cho. Rectification of figures and photos in document images using bounding box interface. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3121–3128, June 2010.
- [9] H. I. Koo and N. I. Cho. State estimation in a document image and its application in text block identification and text line extraction. In *Proceedings of European Conference on Computer Vision (ECCV)*, volume 6312, pages 421–434, 2010.
- [10] H. I. Koo and N. I. Cho. Text-line extraction in handwritten chinese documents based on an energy minimization framework. *IEEE Transactions on Image Processing*, 21(3):1169–1175, March 2012.
- [11] Y. Li, Y. Zheng, D. Doermann, and S. Jaeger. Script-independent text line segmentation in freestyle handwritten documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(8):1313–1329, 2008.
- [12] J. Liang, D. DeMenthon, and D. Doermann. Flattening curved documents in images. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 338–345 vol. 2, 2005.
- [13] J. Liang, D. DeMenthon, and D. Doermann. Geometric rectification of camera-captured document images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4):591–605, 2008.
- [14] S. Lu, B. M. Chen, and C. Ko. Perspective rectification of document images using fuzzy set and morphological operations. *Image and Vision Computing*, 23(5):541 – 553, 2005.
- [15] S. Lu, B. M. Chen, and C. C. Ko. A partition approach for the restoration of camera images of planar and curled document. *Image and Vision Computing*, 24:837–848, 2006.
- [16] G. Meng, C. Pan, S. Xiang, J. Duan, and N. Zheng. Metric rectification of curved document images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 34(4):707–722, 2012.
- [17] G. Nagy. Twenty years of document image analysis in pami. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):38–62, 2000.
- [18] G. Nagy and S. Seth. Hierarchical representation of optically scanned documents. In *Proceeding of International Conference on Pattern Recognition (ICPR)*, volume 1, pages 347–349, 1984.
- [19] A. Nikolaos and S. Sabine. Seam carving for text line extraction on color and grayscale historical manuscripts. In *Proceedings of the 14th International Conference on Frontiers in Handwriting Recognition*, pages 726–731, Crete, Greece, 2014.
- [20] U. Pal and S. Datta. Segmentation of bangla unconstrained handwritten text. In *Proceedings of the 7th International Conference on Document Analysis and Recognition*, pages 1128–1132, Aug 2003.
- [21] V. Papavassiliou, T. Stafylakis, V. Katsourous, and G. Carayannis. Handwritten document image segmentation into text lines and words. *Pattern Recognition*, 43(1):369 – 377, 2010.
- [22] M. Pilu. Extraction of illusory linear clues in perspectively skewed documents. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 363–363, 2001.
- [23] S. Raid, A. Abedelkadir, and E.-S. Jihad. Text line extraction for historical document images. *Pattern Recognition Letters*, 35(0):23 – 33, 2014.
- [24] N. Stamatopoulos, B. Gatos, I. Pratikakis, and S. J. Perantonis. Goal-oriented rectification of camera-based document images. *IEEE Transactions on Image Processing*, 20(4):910–920, 2011.
- [25] Y. Tian and S. Narasimhan. Rectification and 3d reconstruction of curved document images. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 377–384, June 2011.
- [26] A. Ulges, C. H. Lampert, and T. Breuel. Document image de-warping using robust estimation of curled text lines. In *Proceedings of the 8th International Conference on Document Analysis and Recognition*, pages 1001–1005, Aug 2005.
- [27] F. Yin and C. L. Liu. Handwritten chinese text line segmentation by clustering with distance metric learning. *Pattern Recognition*, 42(12):3146 – 3157, 2009.
- [28] Z. Zhang and C. L. Tan. Correcting document image warping based on regression of curved text lines. In *Proceedings of the 7th International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 589–593, 2003.