Text-Line Extraction using a Convolution of Isotropic Gaussian Filter with a Set of Line Filters

Syed Saqib Bukhari[1], Faisal Shafait[2], and Thomas M. Breuel[1]
[1]Technical University of Kaiserslautern, Germany
[2]German Research Center for Artificial Intelligence (DFKI), Kaiserslautern, Germany
bukhari@informatik.uni-kl.de, faisal.shafait@dfki.de, tmb@informatik.uni-kl.de

Abstract—**Text-line extraction is a key task in document analysis. Methods based on anisotropic Gaussian filtering and ridge detection have shown good results. This paper describes performance improvements to these technique based on the use of a convolution of isotropic Gaussian filter with line filters. These new filter banks are motivated by a matched filter approach to text-lines and, in addition, require fewer operations to compute. We evaluate the performance of the new filter bank in combination with ridge detection on the public DFKI-I (CBDAR 2007 dewarping contest) dataset, which contains camera-captured document images and demonstrate improvements in performance to previous state-of-the-art techniques.**

Keywords-**Text-Line Extraction, Camera-Captured Documents, Filter Bank**

I. INTRODUCTION

TEXT-LINE EXTRACTION is one of the important layout analysis steps in document image understanding systems. It is a challenging task, and its difficulty is based on writing styles, scripts, digitization methods, and intensity values. In the literature, a large number of text-line finding methods are proposed [1], [2], [3]. Most of these methods are designed with certain assumptions and fail where these assumptions are not satisfied. Kumar et al. [4] demonstrated that successful page segmentation algorithms for Latin scripts [1] do not give good results for complex Indic scripts. Bukhari et al. [3] showed that text-line extraction methods for scanned document images cannot be directly used for camera-captured (warped/curled) document images. Li et al. [5] also described that the well-know layout analysis algorithms for machine-printed document images do not perform well for freestyle handwritten text-line detection. For a literature survey of text-line detection, please refer to [1] for typed-text scanned documents, [3] for typed-text camera-captured documents, and [2] for handwritten documents.

We developed a text-line finding algorithm [6] applying two well-know image processing techniques: anisotropic Gaussian filter bank smoothing [7], [8] and ridge detection[9], [10]. It was initially tested for curled text-lines detection on warped camera-captured document images [6], and then for handwritten text-lines [11]. Our ridge based

text-line finding method can be considered as a general-purpose text-line extraction technique that can be robustly used for a large variety of document image classes. However, it takes a large number of computational operations because of a big filter bank for document image smoothing.

In this paper, we introduce a novel technique for oriented filter bank smoothing that can be used with our ridge based text-line extraction method and also for other image processing tasks. Instead of using a bank of multi-oriented, multi-scale anisotropic Gaussian filters for smoothing text-lines, we can achieve nearly similar smoothing using a single isotropic Gaussian filter followed by a bank of multi-oriented, multi-scale line averaging filters. Our new line filter bank smoothing technique is motivated by the concept of steerable filters [12], [13]. Besides the novelty of this new filter bank smoothing technique, it requires fewer computational operations for a large number of orientations as compared to anisotropic Gaussian filter bank smoothing technique.

The rest of this paper is organized as follows. Ridge based text-line finding method [6] is briefly described in Section II for the completeness of this paper. The basic concept of oriented anisotropic Gaussian filter bank smoothing is explained in Section III. Our new filter bank smoothing technique is described in Section IV. Performance evaluation results are discussed in Section V. Section VI presents our conclusions.

II. RIDGE BASED TEXT-LINE EXTRACTION

We have developed a ridge based text-line finding method [6], in which a document image is first smoothed by anisotropic Gaussian filter bank, and then text-lines are extracted over the smoothed image using ridge detection technique. The main reason of using multi-oriented, multi-scale anisotropic Gaussian filter bank approach for smoothing document images is that, a document image may contain a diversity of font sizes, text-line orientations, and spaces between characters, words and text-lines. Therefore, a single isotropic or anisotropic Gaussian filter is not sufficient for enhancing/smoothing the structure of text-lines. For a sample document image, the smoothing results of isotropic,

(a) Image (b) Isotropic Gaussian Smoothing (c) Anisotropic Gaussian Smoothing [5]

(d) Oriented Gaussian Filter Bank Smoothing [6] (e) Ridge Detection (f) Labeled Text-Lines
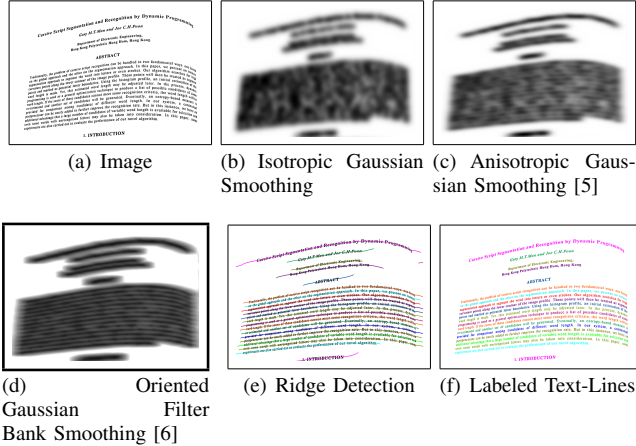
Figure 1. Different possible ways of smoothing a sample camera-captured warped document image and their effects. b) Isotropic Gaussian filter smoothing (Equation 1) lost the structure of text-lines, firstly because text-lines are usually horizontal in nature and secondly they are not straight. c) Anisotropic Gaussian filter smoothing, where $\sigma_x > \sigma_y$ and $\theta = 0°$ in Equation 2, also lost the details of text-lines structure mainly because they are not straight. d) Gaussian filter bank smoothing enhanced text-lines structure well without mixing them with their neighboring text-lines. e) Detected ridges from the smoothed image (d) are mapped over the input image. f) Text-lines are labeled using detected ridges.

anisotropic [5] and filter bank [6] Gaussian smoothing are shown in Figures 1(b), 1(c), and 1(d), respectively. In the smoothed document image of Figure 1(d), the text-lines look like ridges. Therefore, the regions of text-lines are extracted from the smoothed image applying the ridge detection method [9]. The detected ridges are shown in Figures 1(e), and the corresponding labeled text-lines image is shown in Figure 1(f).

III. BASIC CONCEPT OF FILTER BANK SMOOTHING

Filter bank is the concept of applying a set of filters, instead of one, for a given data processing task. It is being used in computer vision for various image processing tasks, like image smoothing for enhancing multi-oriented and/or multi-scale structures [14], [8], [6]. A common way of using a set of Gaussian filters for image smoothing [14], [8], [6] is described here.

The standard formulas of isotropic Gaussian and oriented anisotropic Gaussian filters are shown in Equation 1 and 2, respectively.

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} exp\{-\frac{1}{2}\frac{(x^2 + y^2)}{\sigma^2}\} \quad (1)$$

$$\begin{aligned} G(x, y; \sigma_x, \sigma_y, \theta) &= \frac{1}{2\pi\sigma_x\sigma_y} exp\{-\frac{1}{2}(\frac{(x\cos\theta + y\sin\theta)^2}{\sigma_x{}^2} \\ &+ \frac{(-x\sin\theta + y\cos\theta)^2}{\sigma_y{}^2})\} \end{aligned} \quad (2)$$

(a) Block diagram of oriented anisotropic Gaussian filter bank smoothing

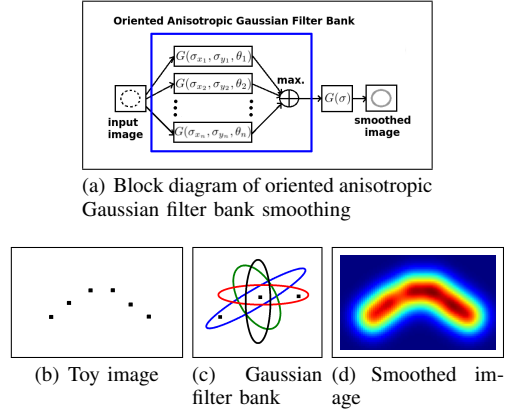(b) Toy image (c) Gaussian filter bank (d) Smoothed image

Figure 2. a) Processing flow of image smoothing using oriented anisotropic Gaussian filter bank approach [6]. b) Toy image for illustrating the basic principle of filter bank smoothing, c) illustration of multi-oriented, multi-scale Gaussian filter bank on a pixel, d) result of multi-oriented, multi-scale Gaussian filter bank smoothing.

Isotropic Gaussian filter contains a single parameter–standard deviation (σ). Oriented anisotropic Gaussian filter contains three parameters: x-axis standard deviation (σ_x), y-axis standard deviation (σ_y) and angle of orientation (θ). For filter bank smoothing, a set of Gaussian filters is first generated for different combinations of σ_x, σ_y and θ. The set of filters is then applied to each pixel of an input image, and a maximum filter response at each pixel is selected for the smoothed image. For blending the fine discontinuities in the smoothed image, the smoothed image can be finally processed by an isotropic Gaussian filter with a small value of standard deviation. The block diagram of Gaussian filter bank smoothing is shown in Figure 2(a). A simple toy image to illustrate the concept of Gaussian filter bank smoothing is also shown in the bottom row of Figure 2.

IV. A NEW FILTER BANK SMOOTHING TECHNIQUE

In this paper, we would like to represent an oriented anisotropic Gaussian filter (Equation 2) by a linear combination of basic filters for implementation ease and execution time speedup. We introduce that an oriented anisotropic Gaussian filter can be approximated by a linear combination of an isotropic Gaussian filter (Equation 1) and a simple *line averaging filter*. Our line averaging filter is defined as a function of length (L pixels) and orientation (θ degrees) as slope of a line, and it works as follows. For a pixel (x, y) in an image, the center point of a line filter is first placed over the pixel. By doing that, the line filter coincides with that pixel and its neighboring pixels. Now, the line averaging filter simply returns an average intensity value corresponding to all of these pixels. The line averaging filter is represented as $A_{vg}(x, y; L, \theta)$.

For approximating the smoothing affect of an oriented anisotropic Gaussian filter ($G(x, y; \sigma_x, \sigma_y, \theta)$) on an image,
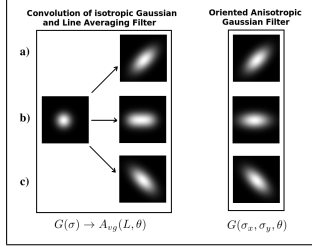
Figure 3. Sample results for the approximation of an oriented anisotropic Gaussian filter ($G(\sigma_x, \sigma_y, \theta)$) using the presented linear combination of an isotropic Gaussian filter ($G(\sigma)$) and a line averaging filter ($A(L, \theta)$). a) $\theta = 45°$, b) $\theta = 0°$, c) $\theta = -45°$.

(a) Block diagram of isotropic Gaussian filter with line averaging filter bank

(b) Isotropic Gaussian smoothing   (c) Line filter bank   (d) Smoothed image
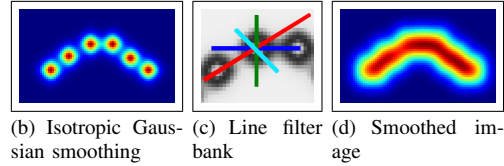
Figure 4. a) Processing flow of image smoothing using our new line filter bank smoothing technique. b) The toy image (Figure 2(b)) after Isotropic Gaussian filter smoothing. c) illustration of multi-oriented, multi-scale line averaging filters on a point, d) result of our line filter bank smoothing technique.

the image is first smoothed by an isotropic Gaussian filter ($G(x, y; \sigma)$), and then the smoothed image is processed by a line averaging filter ($A_{vg}(x, y; L, \theta)$). Some of approximation examples are shown in Figure 3.

The concept of our filter bank smoothing using the presented combination of isotropic Gaussian filter and line averaging filter is described as follows. An input image is first smoothed by an isotropic Gaussian filter with a predefined value of standard deviation (σ). Then, a set of line averaging filters is generated with varying lengths (L) and slopes (θ). The set of line averaging filters is applied to the smoothed image, and for each pixel a maximum filter response is selected for the resulting smoothed image. The resulting image is further smoothed by an isotropic Gaussian filter with a small value of σ. The block diagram and a simple illustration of our new filter bank smoothing technique are shown in Figure 4. In contrast to anisotropic Gaussian filter bank smoothing technique, it is interesting to note that our new filter bank smoothing method requires a single isotropic Gaussian filter followed by a set of line averaging filters.

A sample camera-captured document and it corresponding smoothed text-lines results for anisotropic Gaussian filter bank smoothing and our new line filter bank smoothing are shown in Figure 5. The corresponding detected ridges on the smoothed images are also shown in this figure. It is clearly visible that our smoothing technique enhances the text-lines structure and helps in finding the regions of text-lines through ridge detection approach similar to anisotropic Gaussian smoothing. It is, however, important to note that the outputs of both smoothing methods may not be exactly the same, even for similar parameter settings.

To compare the computational complexities of the filter bank approaches (oriented anisotropic Gaussian filters and newly presented oriented line averaging filters), we consider their basic implementations without any optimization. Computational complexity of a filter bank, with F number of oriented anisotropic Gaussian filters of $W \times W$ window size, for an $N \times N$ image is equal to $O(F \times W^2 \times N^2)$. Similarly,
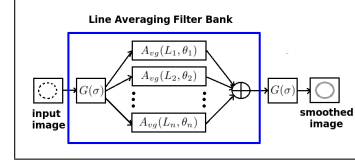
computational complexity of a isotropic Gaussian filter of $W \times W$ window size with a set of F number of oriented line averaging filters of L pixels, for an $N \times N$ image is equal to $O(W^2 \times N^2 + F \times L \times N^2)$. This means, comparatively a large number of computational operations are required for oriented anisotropic Gaussian filter bank than our new line averaging filter bank.

V. PERFORMANCE EVALUATION

In this paper, we evaluate the performance of ridge based text-line finding method for the presented filter bank smoothing technique on DFKI-I (CBDAR 2007 dewarping contest) dataset [15], and compared its performance with other state-of-the-art methods. DFKI-I dataset contains 102 grayscale and binarized document images of pages from several technical books captured by an off-the-shelf hand-held digital camera in a normal office environment. Document images in this dataset consist of warped text-lines. Here, performance evaluation of a text-line finding method is based on vectorial performance evaluation metrics presented in Shafait et al. [1].

Our new line filter bank smoothing technique mainly contains three free parameters; σ for an isotropic Gaussian smoothing, and a set of L and θ for line averaging filters. The value of σ and the range of values of L can be selected automatically for a binary document image using connected components statistics, such as $\sigma = \sigma_r \times h_{cc}$ and $L = L_{rs} \times w_{cc} \rightarrow L_{re} \times w_{cc}$ (where σ_r and L_{rs} and L_{re} are the relative values with respect to median height (h_{cc}) and median width (w_{cc}), respectively). The range of values of θ can be selected empirically. If target documents do not contain a large variety of font sizes like DFKI-I dataset, single value can be selected for $L = Lr \times w_{cc}$ instead of a range of values. Similarly, if target documents do not contain
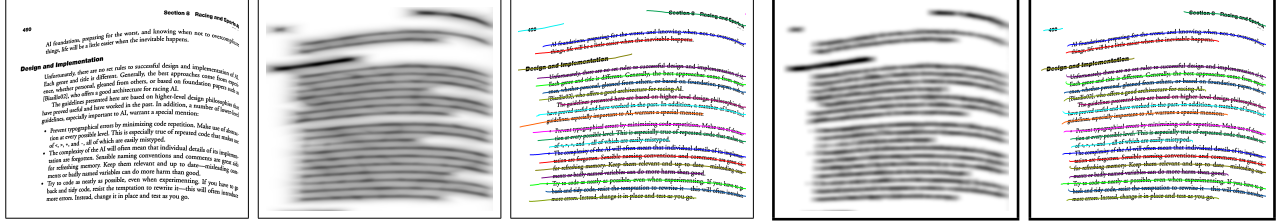
Figure 5. A sample camera-captured document image and it corresponding text-line detection results for our previous anisotropic Gaussian filter bank and new line filter bank smoothing methods. Column 1: sample image, Column 2: result of anisotropic Gaussian filter bank smoothing. Column 3: result of ridge detection on the smoothed image in column 2. Column 4: result of line filter bank smoothing. Column 5: result of ridge detection on the smoothed images in column 4.
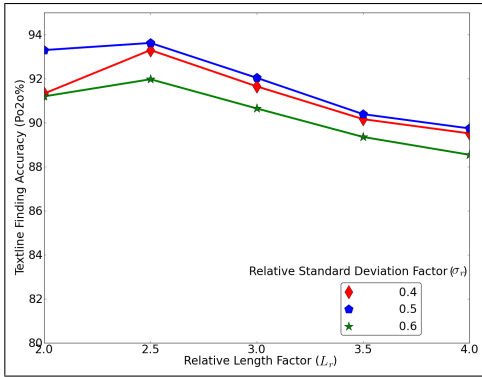
The ridge based text-line extraction method with anisotropic Gaussian filter bank smoothing (without any computational optimization) takes around 17 min. per image on the DFKI-I dataset, where the size of an image is approximately 8 Mega pixels. On the other hand, the ridge based method with our new line filter bank smoothing technique (also without any computational optimization) takes around 3.24 min. per image. The machine, that was used for the execution time analysis, had following specifications: 2.53 GHz processor and 40 GB RAM with Linux (Ubuntu) operating system. This means, our new filter bank smoothing technique is competitively faster than oriented anisotropic Gaussian filter bank smoothing. The execution time of oriented anisotropic Gaussian smoothing can be reduced by using fast implementations of anisotropic Gaussian filter [7], [8], and the execution time of line averaging filter bank smoothing can also be reduced by using integral images [18].

Figure 6. Text-line detection accuracy (P_{o2o}) of the presented method for different values of its free/tunable parameters (σ_r and L_r).

different direction of curls, single value can also be selected for θ. For the grayscale document images, values of the free parameters can be selected empirically.

For DFKI-I binary document dataset, we have tested our presented method for different values of L_r and σ_r with a empirically selected range of values for $\theta = \pm 5°$ as shown in Figure 6. The optimized values of σ_r and L_r for the dataset are 0.5 and 2.5, respectively. Text-line detection accuracy of the new method in comparision with our previous method [6], rule-based [16], and nearest neighbor based [17] methods are shown in Table I. For grayscale documents of DFKI-I dataset, text-line detection accuracy of new and previous methods with optimized values of their free parameters are also shown in Table I. The vectorial performance evaluation metrics, which are shown in Table I, help in analyzing the performance of these text-lines detection methods in detail. Our new method has achieved the best one-to-one text-line detection accuracy for both binary (93.79%) and grayscale (92.75%) document images.

We have also conducted a controlled experiment for analyzing the execution times of anisotropic Gaussian filter bank smoothing and line filter bank smoothing methods. In this experiment, we have used the same number of filters and similar range of values of parameters for both of them.

VI. CONCLUSION

In this paper, we have introduced a novel oriented line filter bank smoothing technique for our ridge based text-line finding method [6], in contrast to oriented anisotropic Gaussian filter bank smoothing. Our new line filter bank smoothing is a linear combination of an isotropic Gaussian filter and a set of simple line averaging filters. We evaluated our line filter bank smoothing technique with ridge based text-line finding method for extracting text-lines on DFKI-I (CBDAR 2007 dewarping contest) dataset [15]. As shown in Table I, we have achieved better one-to-one text-line finding accuracy for ridge based method with the presented line filter bank smoothing as compared to state-of-the-art methods. Our new filter bank smoothing technique also requires comparatively fewer computational operations than anisotropic Gaussian filter bank smoothing. The presented line filter bank smoothing technique can also be used in various other image smoothing and segmentation related computer vision applications.

Table I
PERFORMANCE EVALUATION RESULTS OF THE PRESENTED METHOD IN COMPARISON WITH OTHER STATE-OF-THE-ART METHODS ON DFKI-I (CBDAR 07 DEWARPING CONTEST) DATASET [15] BY USING VECTORIAL PERFORMANCE EVALUATION METRICS [1]: N_g:GROUND-TRUTH COMPONENTS; N_s:SEGMENTED COMPONENTS; N_{o2o}:ONE-TO-ONE MATCHED COMPONENTS; N_{falarm}: FALSE ALARMS; N_{useg}: UNDERSEGMENTATIONS; N_{oseg}: OVERSEGMENTATIONS; N_{ucomp}: UNDERSEGMENTED COMPONENTS; $P_{ucomp} = N_{ucomp}/N_g$; N_{ocomp}:OVERSEGMENTED COMPONENTS; $P_{ocomp} = N_{ocomp}/N_g$; N_{mcomp}: MISSED COMPONENTS; $P_{mcomp} = N_{mcomp}/N_g$; $P_{o2o} = N_{o2o}/N_g$: ONE-TO-ONE SEGMENTATION ACCURACY;

Algorithm	Performance Evaluation Metrics									
	N_g	N_s	N_{o2o}	N_{falarm}	N_{useg}	N_{oseg}	$P_{ucomp}\%$	$P_{ocomp}\%$	$P_{mcomp}\%$	$P_{o2o}\%$
Nearest-Neighbor modified [17] (binary)	3091	3293	2753	4264	103	241	3.20%	7.02%	0.03%	89.07%
Rule-Based [16] (binary)	3091	2924	2816	785	57	682	1.81%	21.71%	4.43%	91.10%
Ridge with Anisotropic Gaussian Filter Bank [6] (binary)	3091	3041	2820	2041	118	90	3.46%	2.46%	0.71%	91.23%
Ridge with Anisotropic Gaussian Filter Bank [6] (gray)	3091	3032	2818	1690	113	74	3.46%	2.23%	0.65%	91.17%
Ridge with Line Filter Bank (binary)	**3091**	**3204**	**2894**	**2062**	**59**	**194**	**1.49%**	**3.66%**	**0.71%**	**93.63%**
Ridge with Line Filter Bank (gray)	**3091**	**3133**	**2867**	**1826**	**70**	**136**	**1.81%**	**3.98%**	**0.78%**	**92.75%**

REFERENCES

[1] F. Shafait, D. Keysers, and T. M. Breuel, "Performance evaluation and benchmarking of six page segmentation algorithms," *IEEE TPAMI*, vol. 30, no. 6, pp. 941–954, 2008.

[2] L. Likforman-Sulem, A. Zahour, and B. Taconet, "Text line segmentation of historical documents: a survey," *Int. Journal on Document Analysis and Recognition*, vol. 9, pp. 123–138, 2007.

[3] S. S. Bukhari, F. Shafait, and T. M. Breuel, "Performance evaluation of curled textlines segmentation algorithms on CBDAR 2007 dewarping contest dataset," in *Proc. Int. Conf. on Image Processing*, Hong Kong, China, 2010.

[4] K. S. S. Kumar, S. Kumar, and C. V. Jawahar, "On segmentation of documents in complex scripts," in *Proc. Int. Conf. on Document Analysis and Recognition*, Washington, DC, USA, 2007, pp. 1243–1247.

[5] Y. Li, Y. Zheng, D. Doermann, and S. Jaeger, "Script-independent text line segmentation in freestyle handwritten documents," *IEEE TPAMI*, vol. 30, no. 8, pp. 1313–1329, 2008.

[6] S. S. Bukhari, F. Shafait, and T. M. Breuel, "Ridges based curled textline region detection from grayscale camera-captured document images," in *Int. Conf. on Computer Analysis of Images and Patterns*, Muenster, Germany, 2009, vol. 5702, pp. 173–180.

[7] J.-M. Geusebroek, A. W. M. Smeulders, and J. V. D. Weijer, "Fast anisotropic gauss filtering," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 938–943, 2003.

[8] C. H. Lampert and O. Wirjadi, "An optimal nonorthogonal separation of the anisotropic gaussian convolution filter," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3501–3513, 2006.

[9] M. D. Riley, "Time-frequency representation for speech signals," *PhD Thesis, MIT*, 1987.

[10] D. Eberly, R. Gardner, B. Morse, S. Pizer, and C. Scharlach, "Ridges for image analysis," *J. Math. Imaging Vis.*, vol. 4, no. 4, pp. 353–373, 1994.

[11] S. S. Bukhari, F. Shafait, and T. M. Breuel, "Script-independent handwritten textlines segmentation using active contours," in *Proc. Int. Conf. on Document Analysis and Recognition*, Barcelona, Spain, 2009, pp. 446 – 450.

[12] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE TPAMI*, vol. 13, no. 9, pp. 891 –906, Sep. 1991.

[13] M. Benjelil, S. Kanoun, R. Mullot, and A. Alimi, "Complex documents images segmentation based on steerable pyramid features," *Int. Journal on Document Analysis and Recognition*, vol. 13, pp. 209–228, 2010.

[14] R. Manmatha and J. L. Rothfeder, "A scale space approach for automatically segmenting words from historical handwritten documents," *IEEE TPAMI*, vol. 27, no. 8, pp. 1212–1225, 2005.

[15] F. Shafait and T. M. Breuel, "Document image dewarping contest," in *Proc. 2nd Int .Workshop on Camera-Based Document Analysis and Recognition*, Curitiba, Brazil, Sep 2007.

[16] D. M. Oliveira, R. D. Lins, G. Torreo, J. Fan, and M. Thielo, "A new method for text-line segmentation for warped document," in *Proc. of Int. Conf. on Image Analysis and Recognition*, Povoa de Varzim, Portugal, 2010, pp. 398–408.

[17] B. Gatos, I. Pratikakis, and K. Ntirogiannis, "Segmentation based recovery of arbitrarily warped document images," in *Proc. Int. Conf. on Document Analysis and Recognition*, Curitiba, Brazi, 2007, pp. 989–993.

[18] F. Shafait, D. Keysers, and T. M. Breuel, "Efficient implementation of local adaptive thresholding techniques using integral images," in *Proc. SPIE Document Recognition and Retrieval XV*, San Jose, CA, USA, Jan. 2008, pp. 101–106.