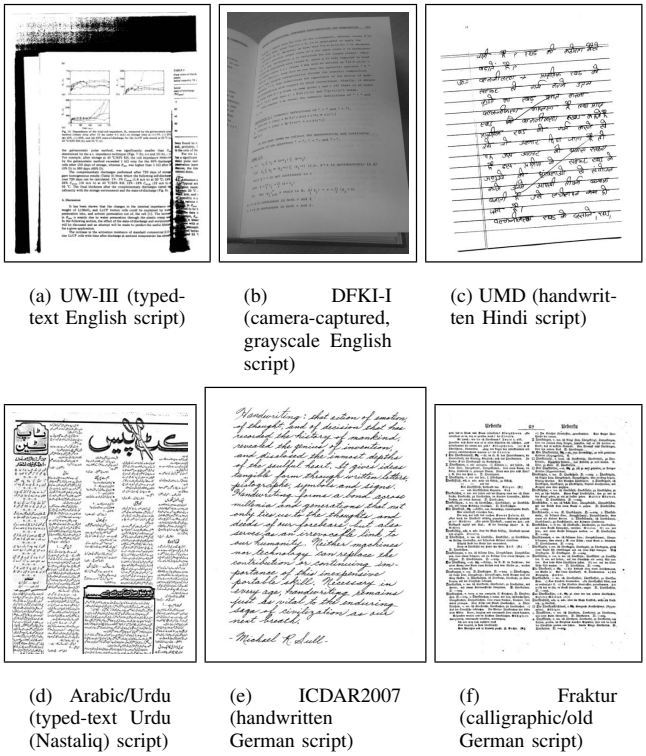
2013 12th International Conference on Document Analysis and Recognition

Towards Generic Text-Line Extraction

Syed Saqib Bukhari*, Faisal Shafait†, and Thomas M. Breuel*
*Technical University of Kaiserslautern, Germany
†The University of Western Australia, Perth, Australia
bukhari@informatik.uni-kl.de, faisal.shafait@uwa.edu.au, tmb@informatik.uni-kl.de

***Abstract*—Text-line extraction is the backbone of document image analysis. Since decades, a large number of text-line finding methods have been proposed, where these methods rely on certain assumptions about a target class of documents with respect to writing styles, digitization methods, intensity values, and scripts. There is no generic text-line finding method that can be robustly applied to a large variety of simple and complex document images. We introduced the ridge-based text-line finding method, and published its initial results for curled text-line detection on camera-captured document images. In this paper, we demonstrates our ridge-based method as a generic text-line finding approach that can be robustly applied on a diverse collection of simple and complex document images. The comprehensive performance evaluation of the ridge-based method and its comparison with several state-of-the-art methods is presented in the paper. For this purpose, diverse categories of publicly available and standard datasets have been selected: UWIII (scanned, printed English script), DFKI-I (camera-captured, printed English script), UMD (handwritten Chinese, Hindi, and Korean scripts), ICDAR2007 handwritten segmentation contest (handwritten English, French, German and Greek scripts), Arabic/Urdu (scanned, printed script), and Fraktur (scanned, calligraphic German script). Experiments on these datasets show that the ridge-based method achieves better text-line extraction results as those of the best performing, domain-specific text-line finding methods. Firstly, these results show that the ridge-based method is a generic text-line extraction method. Secondly, these results are also helpful for the community to assess the advantages of this method.**

***Keywords*—Generic Layout Analysis; Generci Text Line Extraction Method; Ridge-based Text-Line Extraction Method; Collection of Diverse Documents; Performance Evaluation and Benchmarking**



(a) UW-III (typed-text English script) (b) DFKI-I (camera-captured, grayscale English script) (c) UMD (handwritten Hindi script)

(d) Arabic/Urdu (typed-text Urdu (Nastaliq) script) (e) ICDAR2007 (handwritten German script) (f) Fraktur (calligraphic/old German script)

Fig. 1. A collection of diverse document images

I. Introduction

Text-line is the most dominant geometrical layout structure in the context of diverse collection of document images [1]. A sample collection of diverse document images is shown in Figure 1. Text-line extraction is a challenging task, and its difficultly is based on underlying categories of document images, which are composed of following features: digitization methods (scanner or camera-imaging), intensity values (binary, grayscale, or color), scripts (Latin, Chinese, Arabic, etc.), and writing styles (typed-text or handwritten). The inherent difficulties in text-line extraction process with respect to each of these features are briefly described as follows. *Digitization methods:* documents are usually digitized by scanner or camera-imaging. Scanned document images consist of straight text-line (Figure 1(a)), whereas camera-captured document images usually consist of high degree of curled text-lines (Figure 1(b)) due to geometric and perspective distortions. Curled text-lines are more difficult to extract than straight text-lines. *Intensity values:* document images are usually present in binary, grayscale, or color forms, among them the most common ones are binary or grayscales. Extraction of text-line directly from grayscale (camera-captured or historical) document images is a difficult task as compared to binary images mainly because of non-uniform illumination and degradations (camera-captured documents, Figure 1(b))). *Scripts:* English script is one of the simplest scripts with respect to document image processing. However, some of the scripts, like Arabic (Figure 1(d)) and Urdu (Figure 1(d)), are composed of diacritics, lots of dots and interline overlapping, which are difficult to handle for text-line finding [2]. *Writing style:* handwritten/historical text-lines finding can be considered as the most difficult task, as compared to typed-text straight, skewed or curled text-lines extraction, because of irregular layout, lack of a well-defined baseline, variability in skew angle between different text-lines and along a single text-line, interline overlap and touching, smudges, smears, faded print, and bleed-through. Some of these challenging problems are shown in Figure 1(c) and 1(e).

Over the last four decades, several text-line extraction

1520-5363/13 $26.00 © 2013 IEEE
DOI 10.1109/ICDAR.2013.153

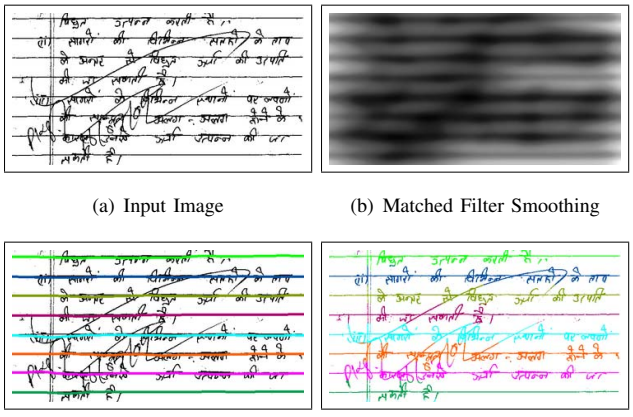
methods have been proposed in the literature. Comprehensive overview of the state-of-the-art page segmentation methods, which are also widely used for text-line extraction, has been provided in [3], [4]. Most of these algorithms rely on certain assumptions about the structure of target class of document images for which they are designed, and fail on other categories of document images where the underlying assumptions are not satisfied. For example these methods do not perform well on complex document images because of their specific challenging problems such as camera-captured warped document images [5], complex scripts document images [6], and handwritten/historical document images [1], [7]. Therefore, a large number of text-line finding algorithms have been proposed in the literature for solving the specific challenging problems in these types of complex document images [8], [5]. However, there is no universal or generic text-line finding method that can be robustly applied to a diverse collection of simple and complex document images [1], [9].

In this paper, we demonstrate that our ridge-based text-line extraction method is a generic text-line finding algorithm and it can be robustly applied on a large variety of simple as well as complex document images that are composed of different intensity values (binary or grayscale), different scripts (Latin, Arabic, Chinese, Hindi, etc.) and different text-line structures (typed-text straight, skewed and curled text-lines, and free-style handwritten text-lines). The ridge-based text-line extraction method was initially designed for curled text-line extraction [10], [11]. It is composed of two standard image processing techniques (filter bank smoothing followed by ridge detection), and it does not necessarily require any preprocessing or post-processing step, although zone segmentation preprocessing step can further improved its results especially in case of multicolumn documents. Here, we present the performance evaluation and benchmarking of the ridge-based text-line finding method on a collection of diverse document images, and its comparison with several domain-specific state-of-the-art text-line extraction methods.

The rest of the paper is organized as follows. The ridge-based text-line extraction method is briefly described in Section II for the completeness of this paper. The diverse collection of standard datasets, that have been selected for performance evaluation and comparison, is presented in Section III. Performance evaluation results of ridge-based text-line finding method and its comparison with a large number of state-of-the-art text-line detection methods are shown in Section IV. Section V presents our conclusions.

II. The Generic Text-Line Extraction Method

We introduced the ridge-based text-line finding method in [10], [11], which is a combination of two standard image processing techniques: *matched filtering* and *ridge detection*. In this method, a document image is first processed by by matched filter bank smoothing for enhancing text-line structure. After smoothing, the regions of text-lines are extracted on the smoothed image using ridge detection method. Finally, the detected ridges are used for text-lines labeling. The processing flow of ridge-based text-line extraction method is shown in Figure 2. We introduced two different approaches for matched filtering: i) *anisotropic Gaussian filter bank smoothing* [10], ii) *line averaging filter bank smoothing* [11]. Both of these



(a) Input Image (b) Matched Filter Smoothing

(c) Ridge Detection (d) Text-Line Labeling

Fig. 2. Processing flow of the ridge-based text-line extraction method [10], [11].

approaches, which are briefly described below, contain well-defined free/tunable parameters, where as the ridge detection method does not contain any free/tunable parameter.

In case of *anisotropic Gaussian filter bank smoothing* [10], a set of Gaussian filters is first generated with different combinations of σ_x, σ_y and θ from their predefined ranges, where σ_x is x-axis standard deviation, σ_y is y-axis standard deviation, and θ is orientation. Then the set of filters is applied to each pixel of an input image, and a maximum filter response at each pixel is selected for the smoothed image. For binary document images, the ranges for σ_x and σ_y are defined with respect to relative values of average width (w) and average height (h) of connected components in a document image, such that $\sigma_x := r_w \times w \rightarrow (r_w + 2) \times w$; $\sigma_y := r_h \times h \rightarrow (r_h + 2) \times h$. For grayscale document images, these ranges are: $\sigma_x := r_w \rightarrow r_w \times 2$ and $\sigma_y := r_h \rightarrow r_h \times 2$. For both binary and grayscale images, a general-purpose range of θ could be set from -45° to 45°. In this way, there are two main free parameters for anisotropic Gaussian filter bank smoothing: r_w and r_h.

However, anisotropic Gaussian filter bank takes a large number of computational operations for a large number of filters. In order to overcome this problem, we introduced a novel concept of *line averaging filter bank smoothing* [11] for enhancing text-line structure that requires fewer computational operations for a large number filters as compared to anisotropic Gaussian filter bank smoothing. In case of line filter bank smoothing, 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, with varying lengths (L) and orientations (θ), is applied and the maximum filter response at each pixel is selected for the smoothed image. For binary document images, the value of (σ) is defined relatively with respect to the average height (h) of connected components in a document image, i.e $\sigma := r_h \times h$. Similarly, the range for L is also defined relatively with respect to average width (w), i.e $L := r_w \times w \rightarrow (r_w + 2) \times w$. For grayscale document images, $\sigma := r_h$ and $L := r_w$. For both binary and grayscale images, the range for slope (θ) is defined similarly like orientation for anisotropic Gaussian filter bank smoothing (i.e -45° to 45°).

In this way, line filter bank smoothing also contains two main free/tunable parameters: r_w and r_h.

III. Datasets : Collection of Diverse Documents

A large number of publicly available and standard datasets have been selected for performance evaluation and comparison. These datasets have been selected on the basis of following criteria: i) collectively all document images in these dataset can represent a collection of diverse document images with respect to writing styles (typed-text or handwritten), scripts (English, Arabic, Chinese, Hindi, etc.), digitization methods (scanned or camera-captured), and intensity values (binary or grayscale), as shown in Figure 1, and ii) these dataset would have been used by the researchers for the performance evaluation of state-of-the-art text-line finding methods. Empirically, we have found a common dataset-independent/default values of free parameters of Gaussian filter bank smoothing, i.e $r_w = 3$ and $r_h = 0.4$ for anisotropic Gaussian smoothing; and and $r_w = 5$ and $r_h = 0.3$ for line filter bank smoothing. These values can be used equally for any type of binary document images. However, We have also investigated dataset-specific/optimized values of these parameters for each of the selected dataset.

The main characteristics of the datasets are described as follows. **UW-III** [12]: it contains typed-text, English script, scanned binary document images. A subset of 100 document images is selected for evaluation. A sample document image is shown in Figure 1(a). The optimized values of free parameters for this dataset are: $r_w = 2$ and $r_h = 0.5$ for anisotropic Gaussian smoothing, and $r_w = 2$ and $r_h = 0.25$ for line filter bank smoothing. **DFKI-I** [13]: it contains typed-text, English script, camera-captured binary and grayscale document images. It consists of 102 (grayscale/binarized) images of pages from several technical books captured by an off-the-shelf hand-held digital camera in a normal office environment. A sample document image is shown in Figure 1(b). The optimized values of free parameters for this dataset are: $r_w = 3$ and $r_h = 0.5$ for anisotropic Gaussian smoothing, and $r_w = 2$ and $r_h = 0.3$ for line filter bank smoothing. **UMD** [7]: it contains handwritten Chinese, Hindi, Korean, Japanese, Persian, Arabic, Cyrillic, Greek, Hebrew, and Thai scripts scanned binary document images. The publicly available UMD dataset consists of around 300 documents of Chinese, Hindi, and Korean scripts. A sample document image is shown in Figure 1(c). The optimized values of free parameters for this dataset are: $r_w = 2$ and $r_h = 0.5$ for anisotropic Gaussian smoothing, and $r_w = 5$ and $r_h = 0.5$ for line filter bank smoothing. **ICDAR2007** handwritten segmentation contest [8]: it contains handwritten English, French, German and Greek scripts scanned document images. It had been used in ICDAR 2007 handwritten segmentation contest and it consists of 80 document images. The optimized values of free parameters for this dataset are: $r_w = 5$ and $r_h = 0.4$ for anisotropic Gaussian smoothing, and $r_w = 6$ and $r_h = 0.25$ for line filter bank smoothing. **Arabic/Urdu** [14]: it contains typed text Arabic and Urdu scripts scanned document images. It consists of 25 Arabic and 20 Urdu document images. A sample document image is shown in Figure 1(d). The optimized values of free parameters for this dataset are: $r_w = 4$ and $r_h = 0.4$ for anisotropic Gaussian smoothing, and $r_w = 2$ and $r_h = 0.3$ for line filter bank smoothing. **Fraktur** [14]: it contains calligraphic German script scanned document images. It consists of 22 document images. The optimized values of free parameters for this dataset are: $r_w = 2$ and $r_h = 0.3$ for anisotropic Gaussian smoothing, and $r_w = 6$ and $r_h = 0.25$ for line filter bank smoothing. For achieving better performance on the complex Arabic and Fraktur documents datasets, whitespace cover [15] based column segmentation is used as a postprocessing step. The ground-truth of all these datasets are presented in color coded pixel form, as described in [4].

IV. Performance Evaluation and Comparison

The performance evaluation metrics for text-line detection accuracy are defined in [4], where a text-line is said to be correctly detected if it does not fall into any of the following categories of errors: over-segmentation, under-segmentation, missed text-lines, and false-alarms. Let, N_g : ground-truth text-lines; N_s : segmented text-lines; N_{o2o} : one-2-one correctly detected text-lines. The one-to-one text-line detection accuracy is represented by $P_{o2o}\% = N_{o2o}/N_g$.

As mentioned earlier, there is no general purpose text-line finding method in the literature. There are some state-of-the-art methods that come to wide spread use, like, smearing [18], constrained text-line extraction (RAST) [15], x-y cut [17], but these method can not be applied on complex documents like camera-captured documents or free-style handwritten documents. Researchers have proposed different solutions for solving different domain specific problems of complex documents. Therefore, for comparison, we have selected some widely used text-line finding methods for those datasets where they can be applied, and some domain-specific state-of-the-art methods for complex document datasets. The state-of-the-art methods that are used for comparison for each datasets are listed as follows: i) smearing [18] and constrained text-line extraction (RAST) [15] for UW-III dataset; ii) nearest-neighbor [19], rule-based [20] and Couple-Snakelets [21] methods for DFKI-I dataset; iii) adapted levelset [7] method for UMD dataset; iv) 5 participants methods [8] for ICDAR2007 handwritten segmentation contest dataset. v) x-y cut [17] and RAST [15] for Arabic/Urdu dataset; vi) RAST [15] for Fraktur dataset.

We have discussed about the dataset-specific as well as dataset-independent free/tunable parameters of the ridge-based method in Section II and III. For all of the datasets, the text-line detection accuracy of the ridge-based method using both types of matched filtering techniques for dataset-specific parameter values are shown in Table I. For comparison, the performance evaluation results of state-of-the-art methods are also shown in Table I. From Table I, the aggregate result of the best performing state-of-the-art methods is 73.51%. For our ridge-based method, the aggregate results for dataset-specific parameter values are 85.37% and 87.62% for anisotropic Gaussian smoothing and line filter bank smoothing, respectively. For the dataset-independent/default values of free parameters, the aggregate results are 84.89% and 86.10% for anisotropic Gaussian smoothing and line filter bank smoothing, respectively. All these aggregate results are also are shown in Figure 3. From these results, firstly, it is important to note that the text-line detection accuracy for both versions of the ridge-base method as well as for both dataset-specific and dataset-independent values of parameters are nearly same. Secondly, the text-line detection accuracy of the ridge-based method is much better

TABLE I. TEXT-LINE EXTRACTION ACCURACY OF THE RIDGE-BASED METHOD (FOR BOTH ANISOTROPIC GAUSSIAN FILTER BANK SMOOTHING AND LINE FILTER BANK SMOOTHING WITH DATASET-SPECIFIC PARAMETER VALUES) ON A LARGE NUMBER OF STANDARD DATASETS THAT BELONG TO A DIVERSE COLLECTION OF DOCUMENTS AND ITS COMPARISON WITH A VARIETY OF DOMAIN-SPECIFIC STATE-OF-THE-ART METHODS BY USING PERFORMANCE EVALUATION METRICS THAT ARE DEFINED IN [4]. PERFORMANCE EVALUATION METRICS: N_g: GROUND-TRUTH COMPONENTS; N_s: SEGMENTED COMPONENTS; N_{o2o}: ONE-TO-ONE MATCHED COMPONENTS; TEXT-LINE EXTRACTION ACCURACY: $P_{o2o}\% = N_{o2o}/N_g$

Dataset	**Method**	Performance Evaluation Metrics		
		N_s	N_{o2o}	$P_{o2o}\%$
UMD[16]	Adapted Levelset[16]	6242	4595	52.85%
(Docs: 300)	**Ridge-based (Aniso/Line)**	**7981**	**6063**	**69.74%**
(N_g: 8694)		**8408**	**6461**	**74.32%**
ICDAR2007 [8]	ILSP-LWSeg [8]	1773	1713	96.73%
(Docs: 80)	**Ridge-based (Aniso/Line)**	**1767**	**1719**	**97.06%**
(N_g: 1771)		**1807**	**1731**	**97.74%**
Arabic/Urdu[14]	X-Y cut [17]	3836	2611	72.63%
	RAST [2]	3564	3058	85.06%
(Docs: 45)	**Ridge-based (Aniso/Line)**	**3648**	**3373**	**93.83%**
(N_g: 3595)		**3782**	**3377**	**93.94%**
Fraktur [14]	RAST [15]	2827	1545	90.38%
(Docs: 22)	**Ridge-based (Aniso/Line)**	**2857**	**2760**	**98.01%**
(N_g: 2816)		**2858**	**2761**	**98.05%**
UW-III [12]	Smearing [18]	3281	2952	77.77%
	RAST [15]	3812	3618	95.31%
(Docs: 100)	**Ridge-based (Ansio/Line)**	**3725**	**3566**	**93.94%**
(N_g: 3796)		**3879**	**3609**	**95.07%**
DFKI-I [13]	Nearest-Neighbor [19]	3293	2753	89.07%
	Rule-Based [20]	2924	2816	91.10%
(Docs: 102)	Couple-Snakelets [21]	3106	2940	95.12%
(N_g: 3091)	**Ridge-based (Aniso/Line)**	**3032**	**2805**	**90.75%**
		3296	**2882**	**93.24%**
Aggregate	*best of state-of-the-art methods*	*21324*	*17469*	*73.51%*
(Docs: 650)	***Ridge-based (Aniso/Line)***	***23010***	***20286***	***85.37%***
(N_g: 23763)		***24030***	***20821***	***87.62%***

than the aggregate accuracy of the best performing domain-specific methods.

The ridge-based text-line finding method can also be equally used for grayscale document images. We have evaluated it on grayscale camera-captured document images in DFKI-I dataset, but we could not find any state-of-the-art method for grayscale camera-captured document images for comparison. For grayscale images of DFKI-I dataset, we have empirically selected absolute values of free parameters for both anisotropic Gaussian filter bank smoothing ($r_w = 40$ and $r_h = 8$) and line filter bank smoothing ($r_w = 50$ and $r_h = 10$), and their corresponding text-line detection accuracies are computed as 91.17% and 92.75%, respectively.

V. CONCLUSION

A large number of text-line finding algorithms have been introduced in the literature. Most of them are designed for document images that hold certain assumptions about writing styles, scripts, digitization methods, intensity values and text-line structures, and fails when these assumptions are not satisfied. We introduced the ridge-based text-line detection method in [10], [11] that was initially tested for curled text-line extraction from typed-text camera-captured document images. In this paper, we demonstrated that the ridge-based text-line extraction method is a generic text-line finding method. It can be robustly applied to simple as well as complex document images containing different challenging problems such as skewed and/or curled text-lines, touching and/or overlapping text-lines, free style handwritten text-lines, irregular layout, noise and distortions. For an extensive performance evaluation, we have compared the ridge-based method with several domain-specific state-of-the-art methods for a large number of publicly available datasets that belong to both simple and complex types of document images. The performance evaluation and comparison results are shown in Table I and Figure 3. The text-line detection accuracy of ridge-based method is significantly better than the aggregate results of the best performing state-of-the-art methods.

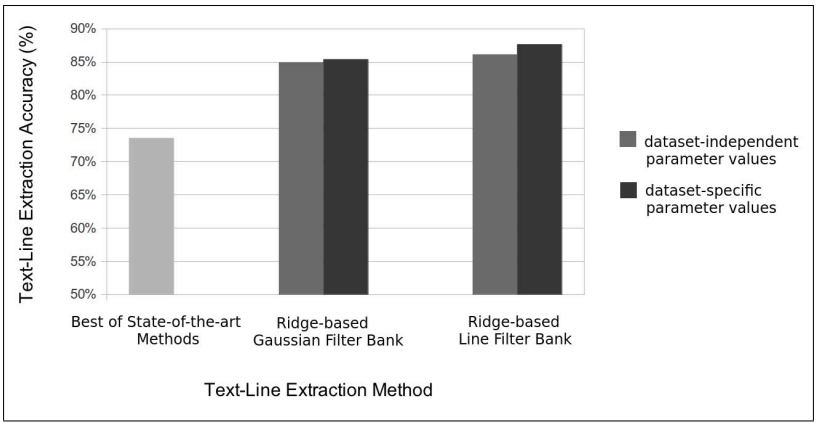


Fig. 3. The aggregate results of the best performing domain-specific state-of-the-art methods and both versions of the ridge-based method (for both dataset-independent and dataset-specific values of parameters).

REFERENCES

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

[2] F. Shafait, A. ul Hasan, D. Keysers, and T. M. Breuel, "Layout analysis of Urdu document images," in *IEEE, INMIC '06*, Islamabad, Pakistan, 2006, pp. 293–298.

[3] R. Cattoni, T. Coianiz, S. Messelodi, and C. M. Modena, "Layout analysis techniques for document image understanding: a review," in *available from http://citeseer.nj.nec.com/, IRST, Trento, Italy, Tech. Rep. 9703-09*, 1998.

[4] F. Shafait, D. Keysers, and T. M. Breuel, "Performance evaluation and benchmarking of six page segmentation algorithms," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 941–954, 2008.

[5] S. S. Bukhari, F. Shafait, and T. M. Breuel, "Performance evaluation of curled textlines segmentation algorithms on CBDAR 2007 dewarping contest dataset," in *Proceedings 17th Int. Conf. on Image Processing*, China, 2010.

[6] K. S. S. Kumar, S. Kumar, and C. V. Jawahar, "On segmentation of documents in complex scripts," in *Proceedings of the 9th Int. Conf. on Document Anal. and Recog.*, Washington, DC, USA, 2007, pp. 1243–1247.

[7] Y. Li, Y. Zheng, D. Doermann, and S. Jaeger, "Script-independent text line segmentation in freestyle handwritten documents," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 8, pp. 1313–1329, 2008.

[8] B. Gatos, A. Antonacopoulos, and N. Stamatopoulos, "ICDAR2007 handwriting segmentation contest," in *Proceedings of the 9th Int. Conf. on Document Anal. and Recog.*, Curitiba, Brazil, 2007, pp. 1284–1288.

[9] D. J. Kennard and W. A. Barrett, "Separating lines of text in free-form handwritten historical documents," in *2nd Int. Conf. on Document Image Analysis for Libraries.*, Los Alamitos, CA, USA, april 2006, pp. 12 – 23.

[10] 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*, ser. Lecture Notes in Computer Science, Muenster, Germany, 2009, vol. 5702, pp. 173–180.

[11] ——, "Text-line extraction using a convolution of isotropic gaussian filter with a set of line filters," in *Proceedings 11th International Conference on Document Analysis and Recognition*, Beijing, China, 2011, pp. 579–583.

[12] I. Guyon, R. M. Haralick, J. J. Hull, and I. T. Phillips, "Data sets for OCR and document image understanding research," in *Handbook of character recognition and document image analysis*. World Scientific, Singapore, 1997, pp. 779–799.

[13] F. Shafait and T. M. Breuel, "Document image dewarping contest," in *Proceedings 2nd Int. Workshop on Camera-Based Document Anal. and Recog.*, Curitiba, Brazil, Sep 2007.

[14] Arabic, Urdu, and Fraktur datasets online. [Online]. Available: https://sites.google.com/a/iupr.com/bukhari/

[15] T. M. Breuel, "Two geometric algorithms for layout analysis," in *Proceedings of the 5th Int. Workshop on Document Anal. Systems*. London, UK: Springer-Verlag, 2002, pp. 188–199.

[16] 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*, vol. 30, no. 8, pp. 1313 –1329, aug. 2008.

[17] G. Nagy, S. Seth, and M. Viswanathan, "A prototype document image analysis system for technical journals," *Computer*, vol. 25, no. 7, pp. 10–22, 1992.

[18] K. Y. Wong, R. G. Casey, and F. M. Wahl, "Document analysis system," *IBM Journal of Research and Development*, vol. 26, no. 6, pp. 647–656, 1982.

[19] B. Gatos, I. Pratikakis, and K. Ntirogiannis, "Segmentation based recovery of arbitrarily warped document images," in *Proceedings 9th International Conference on Document Analysis and Recognition*, Curitiba, Brazi, 2007, pp. 989–993.

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

[21] S. S. Bukhari, F. Shafait, and T. M. Breuel, "Coupled snakelets for curled text-line segmentation from warped document images," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 16, no. 1, pp. 33–53, 2013.