TEXT LINE SEGMENTATION WITH THE ALGORITHM BASED ON THE ORIENTED ANISOTROPIC GAUSSIAN KERNEL

Darko Brodić [*] — Zoran N. Milivojević [**]

The paper presents the algorithm for text line segmentation based on the oriented anisotropic Gaussian kernel. Initially, the document image is split into connected components achieved by bounding boxes. These connected components are cleared from redundant fragments. Furthermore, the binary moments are applied to each of these connected components evaluating local text skewing. According to this information the orientation of the anisotropic Gaussian kernel is set. After the algorithm application the boundary growing areas around connected components are established. These areas are of major importance for the evaluation of text line segmentation. For testing purposes, the algorithm is evaluated under different text samples. Comparative analysis between algorithm with and without orientation based on the anisotropic Gaussian kernel is made. The results show the improvement in the domain of text line segmentation.

K e y w o r d s: moments, connected components, text skew, text line segmentation, anisotropic Gaussian kernel

1 INTRODUCTION

Text line segmentation is one of the most important steps in optical character recognition (OCR). It establishes the text labeling in document text image. These labels separate document text image in foreground and background, *ie* in text line and non text line area. Related works on text line segmentation is categorized as follows [1]:

- projection based methods,
- grouping methods,
- smearing methods,
- methods for processing overlapping and touching components,
- stochastic methods,
- Hough transform methods,
- others.

Many of the proposed methods have been used for text line segmentation. However, smearing methods are distinct due to their robustness, efficiency, and computational inexpensiveness. In these methods, the consecutive black pixels along the horizontal direction are smeared [2]. This way, the enlarged area of black pixels around the text is established. It is called boundary growing area (BGA). Accordingly, the white space area between black pixels is disappeared by filling it with black pixels. BGA represents the labeling regions which enclose the text lines. Determination of these distinct regions that separate text lines represent the primary task of text line segmentation algorithm.

The paper describes the algorithm for text line segmentation based on the analogy with Gaussian probability density function (PDF). This method has been widely used [3–7]. It is based on the pre-assumption that writing text is extending only in the horizontal direction [7].

However, the handwritten text is characterized by variety of skews. In this case, the solution is the incorporation of the local text skew into the algorithm. For the evaluation of the local text skew the moment based method is used. Accordingly, this method is adopted for handwritten text by extracting the connected components with the bounding boxes. Afterwards, the separated connected components are cleared from the redundant fragments. Hence, the handwritten text is split into distinct connected components suitable for processing by the moments. Moment based method identifies the local text skewing of each connected components. These skew angles are used as the inputs into the algorithm based on Gaussian kernel. They establish the kernel orientation by the local text direction. Hence, the algorithm achieves better text line segmentation results.

The paper is organized in the following way. Section 2 explains the procedure of the algorithm. Section 3 describes the process of the experiments. Section 4 contains the results and discusses them. Section 5 makes the conclusion.

2 ALGORITHM (THEORETHICAL PART)

Algorithm based on the oriented anisotropic Gaussian kernel consists of the following steps:

1. Image scanning and binarization,
2. Noise reduction by morphological filter,
3. Connected components extraction by bounding boxes,
4. Connected components clearance from redundant fragments,
5. Skew evaluation by moments,
6. Utilization based on the oriented anisotropic Gaussian kernel.

[*] University of Belgrade, Technical Faculty Bor, V.J. 12, 19210 Bor, Serbia, dbrodic@tf.bor.ac.rs
[**] Technical College Niš, Aleksandra Medvedeva 20, 18000 Niš, Serbia, zoran.milivojevic@vtsnis.edu.rs
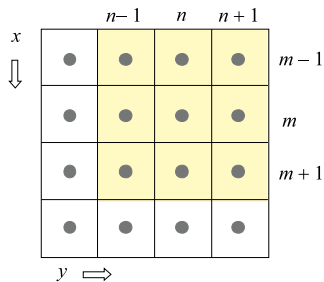
Fig. 1. Basic structuring element **S**

2.1 Image scanning and binarization

Document text image, which is a product of image scanning, represents a digital image given by matrix **D**. It consists of M rows, N columns, and L discrete intensity levels of gray, where $L \in \{0, \ldots, 255\}$. After binarization process, the intensity function is converted into binary function

$$B(I, j) = \begin{cases} 1, & D(i,j) \geq D_{th}(i,j) \\ 0, & D(i,j) < D_{th}(i,j), \end{cases} \quad (1)$$

where D_{th} is given by any local binarization method as threshold sensitivity value [8, 9]. Currently, document image is represented as binary matrix **B** featuring M rows and N columns. Hence, it consists of only black and white pixels, where 0 represents black, and 1 represents white pixel.

2.2 Noise reduction by morphological filter

Noise is by product of scanning as well as the binarization process. However, the document text image can contain noise as well. Separation of data from noise represents the splitting of information from misinformation, respectively. Hence, the noise reduction is an important stage in document image processing. It implies the process of the region extraction which represents the noise. In this step all black pixels of the displaced regions have to be removed. Typically, it is accomplished by the various combinations of morphological operations called morphological filter. In our case, the morphological filter CO, which represents morphological closing upon opening, is used [10]. It is defined as

$$\mathbf{B} = (\mathbf{B} \circ \mathbf{S}) \bullet \mathbf{S}, \quad (2)$$

where **S** is the basic structuring element [10]. It is illustrated in Fig. 1.

After the application of the filter, the noise is almost eliminated. The whole process is shown in Fig. 2.

2.3 Connected components extraction by bounding boxes

Upon the noise reduction step, the connected components extraction is accomplished. It is made by the extraction of bounding boxes over connected components. The bounding box represents a rectangular area whose edges are parallel to the coordinate axes x and y. Hence, each pixel $B(i, j)$ belongs to the bounding box, if it fulfills the following inequalities [11]

$$B(I, j) \big| (x_{\min} \leq i \leq x_{\max}) \wedge (y_{\min} \leq j \leq y_{\max}), \quad (3)$$

where $x_{\min}$, $y_{\min}$, $x_{\max}$, and $y_{\max}$ represent the endpoints of the bounding box. The illustration of the bounding box of connected components from the initial text is shown in Fig. 3.

Hence, each connected components representing the separate text object has been extracted by the bounding boxes. It is assigned as CC_k, where $k = 1, \ldots, m$, and m is the total number of detected objects.

2.4 Connected components clearance from redundant fragments

During the bounding boxes extraction of connected components, some redundant fragments are included as well. It is shown in Fig. 4.

The connected components are cleared from the redundant fragments by extracting only the biggest connected components in each bounding box.

2.5 Skew evaluation by moments

Moment defines the measure of the pixels distribution in the image. Its role is to identify global information of the image concerning its contour. Moments of digital image $f(i, j)$ are defined as [12]

$$M_{pq} = \sum_i \sum_j i^p j^q f(i, j), \quad (4)$$

where p and $q = 0, 1, 2, 3, \ldots, n$, and n represents the order of the moment. From (4) central moments μ_{pq} for digital image can be calculated as [13]

$$\mu_{pq} = \sum_i \sum_j (i - \overline{x})^p (j - \overline{y})^q f(i, j). \quad (5)$$

Accordingly, for the binary image $B(i, j)$ the following is valid

$$m_{pq} = \sum_i \sum_j i^p j^q, \quad (6)$$

and

$$\mu_{pq} = \sum_i \sum_j (i - \overline{x})^p (j - \overline{y})^q. \quad (7)$$

In (5) and (7), $\overline{x}$ and $\overline{y}$ represent the coordinates of the object gravity center. They are defines as follows [12]

$$\overline{x} = \frac{m_{10}}{m_{00}}, \quad \overline{y} = \frac{m_{01}}{m_{00}}. \quad (8)$$
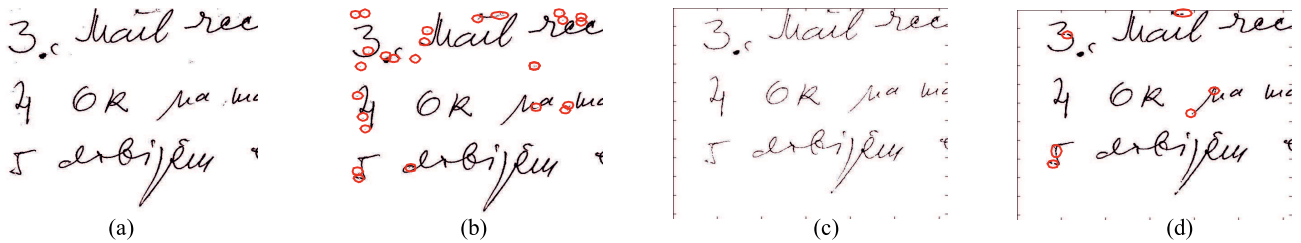
Fig. 2. Noise reduction process: (a) – initial text with noise, (b) – initial text with marked noise region (red circles), (c) – text after morphological opening, and (d) – text after morphological closing upon opening (see circles)
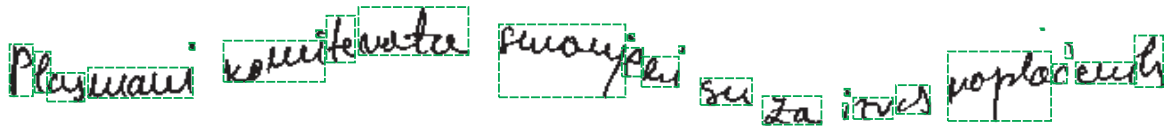

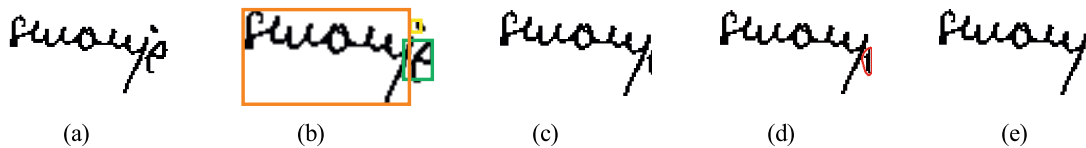**Fig. 3.** Bounding box extraction over connected components (CC)

Fig. 4. Connected component (CC) clearance: (a) – initial text, (b) – overlapping bounding boxes, (c) – CC with redundant fragment, (d) – marked redundant fragment, and (e) – Cleared CC
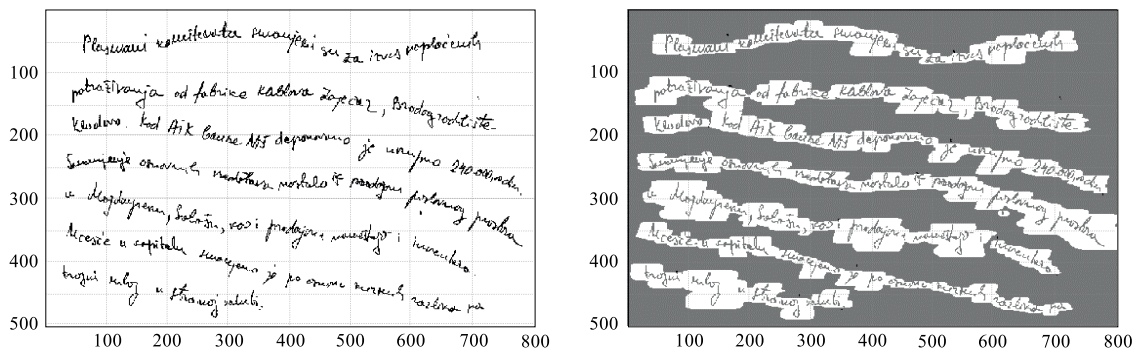

Fig. 5. The application of the algorithm based on the oriented anisotropic Gaussian kernel: (a) – initial text, and (b) – text after the application of the algorithm (BGA are around the text)

Some of the important image features can be obtained from the moments. Among them, the object orientation is of our interest. The object orientation angle θ represents the angle between the object and the horizontal axis. It is defined as follows [12]

$$\theta = \frac{1}{2}\arctan\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}, \qquad (9)$$

where μ_{11} is the first order moment, while μ_{20} and μ_{02} are the second order moments [12].

According to the moments, the local text skew angle θ_k for each of connected components CC_k is calculated. Hence, each connected component CC_k is characterized by the correspondent skew angle θ_k.

2.6. Utilization based on the oriented anisotropic Gaussian kernel

Creating distinct regions that mutually separate text lines is the primary role of the text line segmentation algorithm. Algorithm based on the oriented Gaussian kernel is established on the analogy with Gaussian probability density function (PDF). Gaussian PDF is [13]

$$G(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k|\mathbf{\Sigma}|}}\exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right], \qquad (10)$$

where $\mathbf{x}$ and $\boldsymbol{\mu}$ are column vectors, $\mathbf{\Sigma}$ is covariance matrix, and $|\mathbf{\Sigma}|$ is determinant of the covariance matrix.

This algorithm smears black pixels along the horizontal direction according to the probability values of the Gaussian PDF. Afterwards, the binarization process is applied. Hence, all "probability values" that are inside the region bounded by $\pm 3\sigma_x$ and $\pm 3\sigma_y$ from the mean (in x and y direction) are converted into 1. Accordingly, it represents the binarized anisotropic Gaussian kernel $G(i, j)$, which is equal $2P + 1$ in x and $2R + 1$ in y direction (in pixels). Currently, all these pixels are converted into the same region that is called boundary-growing region (BGA). They represent control image which incorporates connected components. These region are the basis for the text line segmentation. Hence, the obtained matrix **X** is established by convolving the binary matrix **B** with the binarized anisotropic Gaussian kernel **G** [10]

$$X(i, j) = \sum_{k=-P}^{P} \sum_{k=-R}^{R} B(i + k, j + l)G(k, l), \quad (11)$$

where $i = P, \ldots, M - P$ and $j = R, \ldots, N - R$. Elements of the matrix **X** are obtained as follows

```
IF X(i,j) ≠ 0
      THEN X(i,j) = 1
      ELSE X(i,j) = 0
ENDIF.
```

Furthermore, the oriented anisotropic Gaussian kernel is assigned as **E**. Hence, the proposed oriented kernel is made by rotating anisotropic Gaussian kernel **G** for the angle ϕ as

$$\mathbf{E}(\mathbf{x}) = \mathbf{T}G(\mathbf{x}), \quad (12)$$

where **T** is transformation matrix given as [14]

$$\mathbf{T} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}. \quad (13)$$

The angle ϕ is identified by by the angle θ_k for each connected component CC_k. Accordingly, the local angle θ_k represents the input data for the orientation of the each anisotropic Gaussian kernel. The application of the proposed algorithm, which is based on the oriented anisotropic Gaussian kernel on document image is shown in Fig. 5.

3 EXPERIMENTS (EXPERIMENTAL PART)

All experiments are related to the evaluation process of the algorithms efficiency in text line segmentation domain. This way, testing of the algorithm represents the process of its application to the proposed text samples. These text samples called initial ones have theirs connected components configuration. The result of the algorithm application to the initial text samples is new connected components configuration. In an ideal situation the new connected components configuration corresponds to the correct number of the segmented text lines. For the correct evaluation of the algorithms, the following elements are of importance [15]:

- Initial connected components – CC_{init},
- Detected connected components – CC_{det},
- Referent connected components – CC_{ref}.

In the initial text, the number of connected components is given as CC_{init}. After the algorithms application, the number of connected components is changed. It is represented by CC_{det}. Furthermore, assume the referent connected components CC_{ref} as the ideal or correct real number of text lines in the initial sample text.

Afterwards, the comparison between referent and detected number of connected components per each line of text would be accomplished.. If the number of the detected connected components in each text line is equal to one, then CC_{det} will be equal to CC_{ref}. This fulfills the criteria of the correctly segmented text line. Hence, it is assigned as CC_{corrdet}. All others present the text line segmentation errors marked as CC_{errdet}. However, the characterization of the errors is as follows: 1. Over-segmentation text lines CC_{ovdet}, 2. Under-segmentation text lines CC_{undet}, and 3. Mixed-segmentation text lines CC_{mixdet}.

Over-segmentation text lines are defined as text lines which are wrongly divided by the algorithm in two or more connected components [16]. On contrary, under-segmentation text lines incorporate the mutually joined lines. In this case, the sequence of n consecutive text lines is considered as a unique line [16]. It is counted as 1 correctly segmented text line and $n - 1$ under-segmented text lines [16]. Mixed-segmentation text lines are linked with incorrectly assigned words between two adjacent text lines. In this case both text lines are assigned as mixed one.

From these elements the common evaluation measures are *SLHR, OSLHR, USLHR, MSLHR* and $RMSE_{\text{seg}}$ [15] are defined. *SLHR* represents the segmented line hit rate [17] *ie* precision [18] of the correctly segmented text lines. It is defined as [15]

$$SLHR(\%) = \big(1 - |RE_{\text{corrdet}}|\big)*100\% =$$
$$\Big(1 - \Big|\frac{O_{\text{ref}} - O_{\text{corrdet}}}{O_{\text{ref}}}\Big|\Big) * 100\%. \quad (14)$$

OSHLR represents the over-segmented line hit rate [17] *ie* the precision [17] of the over-segmented text lines. It is defined as [15]

$$OSLHR(\%) = \big(1 - |RE_{\text{ovdet}}|\big)*100\% =$$
$$\Big(1 - \Big|\frac{O_{\text{ref}} - O_{\text{ovdet}}}{O_{\text{ref}}}\Big|\Big) * 100\%. \quad (15)$$

USHLR represents the under-segmented line hit rate [17] *ie* the precision [17] of the under-segmented text lines. It is defined as [15]

$$USLHR(\%) = \big(1 - |RE_{\text{undet}}|\big)*100\% =$$
$$\Big(1 - \Big|\frac{O_{\text{ref}} - O_{\text{undet}}}{O_{\text{ref}}}\Big|\Big) * 100\%. \quad (16)$$

Table 1. Text line segmentation results for the parameters pair $(P, \lambda) = \{8, 10\}$

$(P, \lambda) = \{8, 10\}$			
	$\theta = 0°$	$\theta = \theta_M/2$	$\theta = \theta_M$
SLHR	63.64 %	63.64 %	66.67 %
OSLHR	18.18 %	18.18 %	12.12 %
USLHR	18.18 %	18.18 %	21.21 %
MSLHR	0.00 %	0.00 %	0.00 %
$RMSE_{\text{seg}}$	0.674	0.674	0.651

Table 2. Text line segmentation results for the parameters pair $(P, \lambda) = \{10, 10\}$

$(P, \lambda) = \{10, 10\}$			
	$\theta = 0°$	$\theta = \theta_M/2$	$\theta = \theta_M$
SLHR	66.67 %	69.70 %	63.64 %
OSLHR	12.12 %	6.06 %	6.06 %
USLHR	21.21 %	24.24 %	30.30 %
MSLHR	0.00 %	0.00 %	0.00 %
$RMSE_{\text{seg}}$	0.627	0.577	0.696
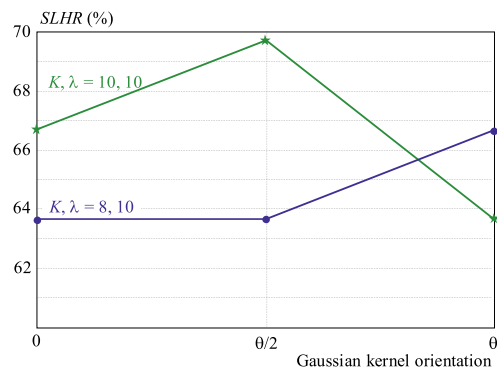
Fig. 6. *SLHR vs* Gaussian kernel orientation (parameter θ)
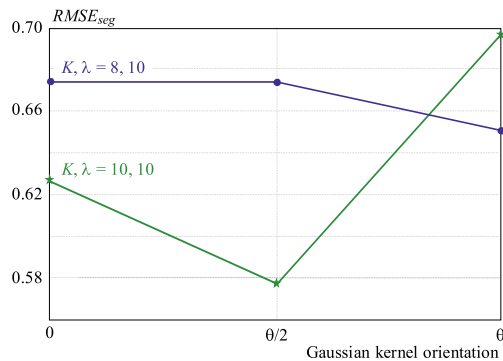
Fig. 7. $RMSE_{\text{seg}}$ *vs* Gaussian kernel orientation (parameter θ)

MLHR represents the mixed line hit rate [17] *ie* precision [17] of the mixed text lines. It is defined as [15]

$$MLHR(\%) = \left(1 - |RE_{\text{mixdet}}|\right) * 100\% =$$
$$\left(1 - \left|\frac{O_{\text{ref}} - O_{\text{mixdet}}}{O_{\text{ref}}}\right|\right) * 100\% . \quad (17)$$

At the end, the number of detected and referent connected components per each text line is compared. The number of referent connected components per line is equal to 1. The variance evaluation is given by $RMSE_{\text{seg}}$ [15]

$$RMSE_{\text{seg}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (CC_{i,\text{ref}} - CC_{i,\text{det}})} , \quad (18)$$

where N is the total number of lines in the reference sample text, $CC_{i,\text{ref}}$ is the number of referent connected components in the text line i (equal to one per each line), and $CC_{i,\text{det}}$ is the number of detected connected components in the text line i.

The values of each introduced measures, individually or together, characterize the quality and effects of performed segmentation process by the testing algorithm.

4 RESULTS AND COMPARATIVE ANALYSIS (DISCUSSION)

The main purpose of the testing process is the algorithm's efficiency evaluation as well as its parameters optimization. For the proposed algorithm based on the oriented anisotropic Gaussian kernel, the parameter values that define kernel size, *ie*, P and $\lambda = P/R$ (sometimes K/L) and the kernel's orientation angle θ are of main interest. Handwritten text samples from our database [15] include letters with height around 50 pixels. According to that, P should be chosen from 10% to 20% of the letter height [19]. The excerpt of the best results for different text samples according to eqs. (14)-(18) are given in Tab. 1 and Tab. 2.

From Tab. 1, the best segmentation results of $SLHR = 66.67\%$ is obtained for parameters set $(P, \lambda, \theta) = \{8, 10, \theta_M\}$. This is obtained for the full moment based evaluated orientation. It is improvement of 3.07 % over the algorithm based non-oriented Gaussian kernel. From $\theta = 0°$ to $\theta = \theta_M$, the value of $RMSE_{\text{seg}}$ is diminished, which confirms the improvement.

From Tab. 2, the best segmentation results of $SLHR = 69.70\%$ is obtained for parameters set $(P, \lambda, \theta) = \{8, 10, \theta_M/2\}$. This is obtained for the half of the moment based evaluated orientation. It is improvement of 3.23 % over the algorithm based non-oriented Gaussian kernel. For $\theta = \theta_M/2$, the value of $RMSE_{\text{seg}} = 0.577$, which is the smallest one and confirms the best testing result.

Figure 6 shows *SLHR vs* Gaussian kernel orientation defined by parameter θ $(K = P)$.

Figure 7 shows $RMSE_{\text{seg}}$ *vs* Gaussian kernel orientation defined by parameter θ $(K = P)$.

Optimal parameter set $\{P, \lambda' \theta\}$ represents the values $\{10, 10, \theta_M/2\}$. The obtained testing results for the measures *SLHR* and $RMSE_{\text{seg}}$ confirm it. Although this result is unusual, it is not completely unexpected. Due to descender and ascender influence in connected components skewing, the full moment based angles are overwhelming. In some circumstances this leads to incorrect skew evaluation. Hence, the half of that value is much

more suitable which leads to better text line segmentation results.

Comparing the obtained results with those which incorporates no orientation, the improvement margin of 3.23 % seems to be modest. However, the direction of research in order to improve the algorithm based on the anisotropic Gaussian kernel is the correct one. Still, it needs some additional fine adjustment. Further adaptation should be made by eliminating the ascender and descender elements in connected components before the application of the moment based skew evaluation method.

5 CONCLUSION

This paper introduces the extension of the classical algorithm based on the anisotropic Gaussian kernel. It is made by the incorporation of moment based method for the evaluation of the connected components skewing. This step represents the preprocessing job for the text line segmentation. It is input as an algorithm parameter θ, which orients the Gaussian kernel. Testing process was performed on the handwritten text database. The algorithm evaluation is based on the errors type classification. The obtained results show slight improvement in the domain of text line segmentation and confirm our starting assumptions. Hence, further adaptation should be investigated as well.

References

[1] LIKFORMAN SULEM, L.—ZAHOUR, A.—TACONET, B.: Text Line Segmentation of Historical Documents: a Survey, International Journal on Document Analysis and Recognition **9** No. 2 (2007), 123-138.

[2] RAZAK, Z.—ZULKIFLEE, K. *et al*: Off-Line Handwriting Text-Line Segmentation: a Review, International Journal of Computer Science and Network Security **8** No. 7 (2008), 12–20.

[3] LI, Y.—ZHENG, Y.—DOERMANN, D.—JAEGER, S.: A New Algorithm for Detecting Text Line in Handwritten Documents, In: 10th International Workshop on Frontiers in Handwriting Recognition 2006 IWFHR 2006, La Baule, France, pp. 35–40.

[4] LI, Y.—ZHENG, Y.—DOERMANN, D.—JAEGER, S.: Script-Independent Text Line Segmentation in Freestyle Handwritten Documents, IEEE Transactions on Pattern Analysis and Machine Intelligence **30** No. 8 (2008), 1313–1329.

[5] BUKHARI, S. S.—SHAFAIT, F.—BREUEL, T. M.: Script-Independent Handwritten Text-lines Segmentation using Active Contours, In: 10th International Conference on Document Analysis and Recognition 2009. ICDAR 2009, Barcelona, Spain, pp. 446–450.

[6] LI, Y.—ZHENG, Y.—DOERMANN, D.—JAEGER, S.: Script-Independent Text Line Segmentation in Freestyle Handwritten Documents, Technical report TR: LAMP-136/CS-4836/ UMIACS-2006-51/CFAR-1017, University of Maryland, College Park, 2006.

[7] BRODIĆ, D.—MILIVOJEVIĆ, Z.: Optimization of the Gaussian Kernel Extended by Binary Morphology for Text Line Segmentation, Radioengineering **19** No. 4 (2010), 718–724.

[8] SAUVOLA, L.—PIETIKAINEN, M.: Adaptive Document Image Binarization, Pattern Recognition **33** No. 2 (2000), 225–236.

[9] KHASHMAN, A.—SEKEROGLU, B.: Document Image Binarisation Using a Supervised Neural Network, The International Journal of Neural Systems **18** No. 5 (2008), 405–418.

[10] GONZALEZ, R. C.—WOODS, R. E.: Digital Image Processing, 2nd ed., Prentice-Hall, USA, 2002.

[11] PREPARATA, F. P.—SHAMOS, M. I.: Computational Geometry: An Introduction, Springer-Verlag, Germany, 1985.

[12] RAMTEKE, R. J.—IMRAN, K. P.—MEHROTRA, S. C.: Skew Angle Estimation of Urdu Document Images: A Moments Based Approach, International Journal of Machine Learning and Computing **1** No. 1 (2011), 7–12.

[13] JÄHNE, B.: Digital Image Processing, 6th ed., Springer-Verlag, Germany, 2005.

[14] TERELL, G. R.: Mathematical Statistics: A Unified Introduction, Springer-Verlag, USA, 1999.

[15] BRODIĆ, D.—MILIVOJEVIĆ, D. R.—MILIVOJEVIĆ, Z.: An Approach to a Comprehensive Test Framework for Analysis and Evaluation of Text Line Segmentation Algorithms, Sensors **11** No. 9 (2011), 8782–8812.

[16] SANCHEZ, A.—SUAREZ, P. D.—MELLO, C. A. B.—OLIVEIRA, A. L. I.—ALVES, V. M. O.: Text Line Segmentation in Images of Handwritten Historical Documents, In: First International Workshops on Image Processing. Theory, Tools and Applications 2008, IPTA 2008, Sousse, Tunisia, pp. 1–6.

[17] BRODIĆ, D.—MILIVOJEVIĆ, D. R.—MILIVOJEVIĆ, Z.: Basic Test Framework for the Evaluation of Text Line Segmentation and Text Parameter Extraction, Sensors **10** No. 5 (2010), 5263–5279.

[18] POPOV, V. S.: Principles of Symmetry and Relative Errors of Instrumentation and Transducers, Automation and Remote Control **62** No. 5 (2001), 842–846.

[19] BRODIĆ, D.: The Evaluation of the Initial Skew Rate for Printed Text, J. Electrical Engineering **62** No. 3 (2011), 142–148.

Darko Brodić (Prof, PhD) was born in Konjic, BiH, Yugoslavia, in 1963. He received his BEE and MEE from the Faculty of Electrical Engineering, University of Sarajevo in 1987 and 1990, respectively, as well as PhD from the Faculty of Electrical Engineering, University of Banja Luka in 2011. Currently, he is with the University of Belgrade, Technical Faculty in Bor. His research interests include document image processing. He is the author and coauthor of over 40 journal (over 15 in journals indexed by Thomson JCR) and conference papers as well as few books and the textbooks.

Zoran N. Milivojević (Prof, PhD) was born in Svrljig, Serbia, in 1959. He received the BE, ME, and PhD degrees in electrical engineering from the Faculty of Electronic Engineering, University of Niš, Niš, in 1984, 1994 and 2002, respectively. He is a Professor at Technical Faculty Bor University of Belgrade and Technical College in Niš. His primary research interests are digital signal processing and speech processing: algorithms and applications. He is the author and coauthor of over 130 journal and conference papers.