

## ROTATION OF THE ANISOTROPIC GAUSSIAN KERNEL FOR THE IMPROVEMENT OF THE TEXT SKEW IDENTIFICATION

**Darko Brodić**

*University of Belgrade  
Technical Faculty Bor  
V.J. 12, 19210 Bor, Serbia*

**Zoran Milivojević**

*Technical College Niš  
Aleksandra Medvedeva 20  
18000 Niš, Serbia*

**Dragan R. Milivojević**

*Mining and Metallurgy Institute  
Department of Informatics  
Z. Bulevar bb, 19210 Bor, Serbia*

### Abstract

*In this paper, an approach to algorithm with the implementation of the anisotropic Gaussian kernel is presented. As a result of algorithm, the growing region around text is created. To improve skew identification in text, anisotropic Gaussian kernel is rotated. Furthermore, algorithm is evaluated under different text samples. Results are given and comparative analysis is made. From the obtained results, the optimal value of the parameters defining rotated anisotropic Gaussian kernel dimension is proposed.*

**Keywords:** OCR, Gaussian kernel, region growing algorithm, skew rate.

### INTRODUCTION

Historical documents are built from hand-written text. It is characterized as fully or partially cursive text. Text lines in the handwritten documents are basically curvilinear. Hence, it tends to be differently skewed. Similarly, different text lines are close to each other. Therefore, text lines casually get into each other. Descendants and ascendants from neighbor text lines are occasionally overlapped. Word in text lines aren't formed regularly. Hence, their distance is different. However, inter word spacing is tolerable. But, appearance of different orientation skewed lines and text lines close to each other made the handwritten text to be less readable.

Previous work on text parameter detection can be categorized in few types [1]:

- Histogram analysis,
- Docstrum (k-nearest neighbor clustering),
- Projection profile (Hough transform),
- Fourier transform,
- Cross-correlation,
- Other methods.

In [2] is mentioned previously proposed and accepted technique of reference line extraction

based on identifying valleys of horizontal pixel density histogram. Method failed due to multi-skewed text lines.

The Docstrum method [3] is by product of a larger page layout analysis system, which assumed that only text is being processed. The connected components formed by the nearest neighbors clustering are essentially characters only. The method is suitable for finding skew angle. But, it is limited to Roman languages due to poor text line segmentation.

Another method proposed in [4] deal with "simple" multi-skewed text. It uses as a basis simple type of Hough transform for straight lines. But, it's too specific.

The Fourier transform method is a representation in the Fourier domain of the projection profile method in the pixel domain. The results are mathematically identical, but Fourier transform is only different approach to the same text and document properties that projection profile is based upon [1].

The cross-correlation method calculates both horizontal and vertical projection profiles and then compares the shift inter-line cross-correlation to determine the skew rate. Although

method can handle complex layout structure documents, applied range is limited to  $(-10^\circ, 10^\circ)$  [1].

Algorithm proposed by [5] model text line detection as an image segmentation problem by enhancing text line structure using a Gaussian window and adopting the level set method to evolve text line boundaries. Author specified method as robust, but rotating text by an angle of  $10^\circ$  has an impact on reference line hit rate.

Method of identifying words contour area as a start of detecting baseline point proposed in [6]. The assumptions made on the word elements definition are too specific.

Organization of this paper is as follows. Section II includes proposed algorithm information and description. In Section III text experiment is defined. Further, in Section IV results are examined, compared and discussed. In Section V conclusions are made.

## PROPOSED ALGORITHM

### DOCUMENT TEXT IMAGE

At the beginning of the process, original image is used. Document text image is obtained as product of original image scanning. Document text image is digital text image represented by matrix  $\mathbf{D}$  with  $M$  rows,  $N$  columns, and intensity with  $L$  discrete levels of gray.  $L$  is the integer number from the set  $\{0, \dots, 255\}$ . Currently,  $D(i,j) \in \{0, \dots, 255\}$ , where  $i = 1, \dots, M$  and  $j = 1, \dots, N$ .

After applying intensity segmentation with binarization, intensity function is converted into binary intensity function given by:

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

where  $D_{th}$  is given by Otsu algorithm [7] or equivalent algorithm [8-10]. It represents threshold sensitivity decision value.

Currently, document text image is represented as binary matrix  $\mathbf{B}$  featuring  $M$  rows by  $N$  columns. Consequently, it consists of the only black and white pixels where value 0 represents black pixels and value 1 white pixel.

### ANISOTROPIC GAUSSIAN FILTER

Establishing distinct areas that mutually separate text lines is the primary task of the

optical character recognition (OCR) algorithm. After, definition of the text line representing the baseline of the handwritten text called reference text line is of the major importance.

In this paper, algorithm based on the analogy with Gaussian probability density function (PDF) is established. This function given by [11]:

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

where  $\mathbf{x}$  and  $\boldsymbol{\mu}$  are column vectors and  $\Sigma$  is covariance matrix. For 2D column vector  $\mathbf{x}$  is given as:

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \quad (3)$$

while vector  $\boldsymbol{\mu}$  as:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \quad (4)$$

Furthermore, covariance matrix  $\Sigma$  is given as:

$$\Sigma = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \quad (5)$$

while its determinant  $|\Sigma|$  as:

$$|\Sigma| = \sigma_x^2 \sigma_y^2 \quad (6)$$

In eq. (5) and (6)  $\sigma_x$  and  $\sigma_y$  represent the standard deviation defining curve spread parameter and  $\mu_x$  and  $\mu_y$  represent the mean in  $x$  and  $y$  direction, respectively. However, eq. (2) is starting point for creating anisotropic kernel. Hence, converting Gaussian PDF into point spread function (PSF) creates anisotropic Gaussian kernel. The idea of Gaussian smoothing is to use this 2-D distribution as a PSF. Since the image is stored as a collection of discrete pixels we need to produce a discrete approximation to the Gaussian function  $G(\mathbf{x})$  named  $G(i,j)$  before performing the convolution. However, the Gaussian distribution is non-zero everywhere, which would require an infinitely

large convolution kernel. In practice, it is effectively zero for more than about  $3\sigma_x$  and  $3\sigma_y$  from the mean in  $x$  and  $y$  direction, respectively. These values represent Gaussian threshold sensitivity level  $L_{gtsx}$  and  $L_{gtsy}$ . It truncates  $3\sigma_x$  in  $x$  direction and  $3\sigma_y$  in  $y$  direction of the kernel forming the ellipse. All pixels that belong inside ellipse form the same area with level higher than  $L_{gtsx}$  or  $L_{gtsy}$ . Hence, anisotropic Gaussian kernel  $G(i,j)$  is defined by  $2P+1$  in  $x$  and  $2R+1$  in  $y$  directions.

Converting all these pixels into the same region forms the areas named boundary growing areas. Boundary growing areas form control image with distinct objects that are prerequisite for the text segmentation as well as reference text line extraction of the document image. Matrix  $\mathbf{X}$  is created by convolving the isotropic Gaussian kernel  $\mathbf{G}$  with the image represented by binary matrix  $\mathbf{B}$  as follows [11]:

$$X(i, j) = \sum_{k=-P}^P \sum_{l=-R}^R B(i+k, j+l) G(k, l) , \quad (7)$$

where  $i$  is from  $P$  to  $M-P$  and  $j$  is from  $R$  to  $N-R$ . Further, elements of matrix  $\mathbf{X}$  is obtained as follows: IF  $X(i,j) \neq 0$  THEN  $X(i,j) = 1$ .

#### ROTATED ANISOTROPIC GAUSSIAN FILTER

Rotation of the anisotropic Gaussian kernel forms extended and rotated anisotropic Gaussian kernel  $\mathbf{E}$ . Hence, proposed kernel extension is made by rotating anisotropic Gaussian kernel  $\mathbf{G}$  for the angle  $\phi$ . Due to the nature of rotation, kernel is extended in  $x$  direction and diminished in  $y$  direction. This can be mathematically described by the following relation:

$$\mathbf{E}(\mathbf{x}) = \mathbf{T}\mathbf{G}(\mathbf{x}) , \quad (8)$$

where  $\mathbf{T}$  represents transformation matrix given by:

$$\mathbf{T} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} . \quad (9)$$

Furthermore, new kernel dimensions are given as  $2S+1$  in  $x$  and  $2T+1$  in  $y$  direction, respectively. Difference between two kernels could be illustrated as in Figure 1.

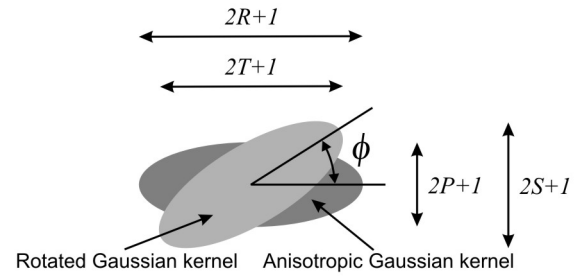


**Fig. 1.** Rotated and anisotropic Gaussian kernel.

Main difference between original algorithm [11] and this approach is in text segmentation domain. Currently, matrix  $\mathbf{Y}$  is defined by convolving the rotated anisotropic Gaussian kernel  $\mathbf{E}$  with the matrix  $\mathbf{B}$  as follows [12]:

$$Y(i, j) = \sum_{k=-S}^S \sum_{l=-T}^T B(i+k, j+l) E(k, l) , \quad (10)$$

where  $i$  is from  $S$  to  $M-S$  and  $j$  is from  $T$  to  $N-T$ . Further, elements of matrix  $\mathbf{Y}$  is obtained as follows: IF  $Y(i,j) \neq 0$  THEN  $Y(i,j) = 1$ .

#### TESTING

Skew rate test experiment is mainly concerned with skew rate identification. It evaluates algorithm performance in skew tracking domain. Although, this experiment is primarily based on printed text, it is good prerequisite for testing handwritten text as well. In our example, it is the only valid test. Practically, due to dynamic skewing of text any other test type will be out of effect.

In this test, sample printed text rotated by the angle  $\alpha$  from  $0^\circ$  to  $90^\circ$  by step of  $5^\circ$  around  $x$ -axis is used [13]. It is presented in Fig. 2. Referent line of the test sample text is represented by:

$$y = ax + b . \quad (11)$$

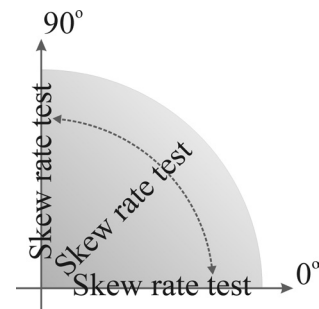


**Fig. 2.** Sample text rotated up to  $90^\circ$  by step of  $5^\circ$

After applying any algorithm to sample text, referent text line estimation implies average position calculation of only black pixels in every column of document text image. It is calculated by [14-15]:

$$x_i = \frac{\sum_{j=1}^U y_j}{U} \quad i=1, \dots, K, \quad (12)$$

where  $x_i$  is the point position of calculated referent text line,  $i$  is the number of column position of calculated referent text,  $y_j$  is the position of black pixel in column  $j$  and  $U$  is the sum of black pixel number in specified column  $j$  of an image.

After calculation, image matrix with only one black pixel per column is obtained. It defines estimated referent text line as well as text line skewness. This referent text line forms partly continuous referent text line. To achieve continuous linear referent text line, least square method is used. Function approximation by first-degree polynomial is given as:

$$y = a'x + b' \quad (13)$$

Furthermore  $m = 1, \dots, V$ , while  $V$  represents the total number of data points. It is used in relation for calculating the slope  $a'$ , and the  $y$ -intercept  $b'$  as follows [16]:

$$a' = \frac{\sum_{m=1}^V y_m \sum_{m=1}^V x_m y_m - V \sum_{m=1}^V x_m y_m}{(\sum_{m=1}^V x_m)^2 - V \sum_{m=1}^V x_m^2} \quad (14)$$

and

$$b' = \frac{\sum_{m=1}^V x_m \sum_{m=1}^V x_m y_m - \sum_{m=1}^V y_m \sum_{m=1}^V x_m^2}{(\sum_{m=1}^V x_m)^2 - V \sum_{m=1}^V x_m^2} \quad (15)$$

For algorithms approximation and evaluation, a quantity called relative error [17] is important. Referent line hit rate i.e.  $RLHR$  incorporates this quantity. It is defined as [13]:

$$RLHR = 1 - \frac{\Delta\beta}{\beta} = 1 - \frac{|\beta_{est} - \beta_{ref}|}{|\beta_{ref}|} \quad (16)$$

where  $\beta_{ref}$  is arctangent from origin (11) i.e.  $a$  and  $\beta_{est}$  is arctangent from estimated (13) i.e.  $a'$ .

Obviously,  $RLHR$  is equal to 1-relative error [16-17]. Now, the root mean square error  $RMSE_{skew}$  is calculated by [13]:

$$RMSE_{skew} = \sqrt{\frac{1}{W} \sum_{w=1}^W (x_{ref}^w - x_{est}^w)^2} \quad (17)$$

where  $w = 1, \dots, W$  is the number of examined text rotating angles up to  $90^\circ$ ,  $x_{ref}^w$  is  $RLHR$  for  $\beta_{est}$  equal to  $\beta_{ref}$ , due to normalization equal to 1, and  $x_{est}^w$  is  $RLHR$ .

## RESULTS AND COMPARATIVE ANALYSIS

In [18], optimized parameter set for the text attributes is proposed. This set is given by two parameters:  $P$  and  $\lambda$ . Furthermore,  $2P+1$  represents  $x$  dimension of the Gaussian kernel. Consequently,  $\lambda$  represents the ratio of the  $y$  and  $x$  dimension of the Gaussian kernel i.e.  $y/x$ . Hence,  $y$  dimension of the Gaussian kernel is  $2K\lambda+1$ . In our case, for the rotated Gaussian kernel,  $P$  and  $R = P\lambda$  are valid. Starting parameter set used for the evaluation and testing are:

- a)  $R = \{10, 15, 20\}$ ;
- b)  $\lambda = \{2, 3, 4\}$ ;
- c)  $\phi = \{0^\circ, 15^\circ, 30^\circ, 45^\circ\}$ ,

where  $\phi$  represents rotation angle. The results are shown in Table I-IV.

TABLE I. REFERENCE LINE HIT RATE FOR  $R=5$  (IN %)\*

| $\phi$              | 0°    | 15°   | 30°   | 45°   | 0°    | 15°   | 30°   | 45°   | 0°    | 15°   | 30°   | 45°   |
|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\lambda$           | 2     |       |       |       | 3     |       |       |       | 4     |       |       |       |
| $\alpha$            | R = 5 |       |       |       |       |       |       |       |       |       |       |       |
| 5°                  | 89.02 | 87.41 | 87.41 | 86.84 | 90.73 | 89.93 | 89.02 | 87.99 | 93.71 | 93.48 | 91.42 | 89.59 |
| 10°                 | 93.25 | 93.36 | 93.31 | 3.02% | 94.55 | 94.61 | 94.27 | 93.70 | 96.09 | 96.31 | 95.46 | 94.61 |
| 15°                 | 95.37 | 95.45 | 95.41 | 95.26 | 96.19 | 96.34 | 96.19 | 95.78 | 96.98 | 97.31 | 96.98 | 96.45 |
| 20°                 | 96.10 | 96.18 | 96.18 | 96.13 | 96.62 | 96.81 | 96.76 | 96.51 | 97.09 | 97.44 | 97.33 | 96.98 |
| 25°                 | 96.70 | 96.78 | 96.78 | 96.80 | 97.00 | 97.23 | 97.23 | 97.13 | 97.25 | 97.60 | 97.64 | 97.47 |
| 30°                 | 96.83 | 96.93 | 96.95 | 7.00% | 97.00 | 97.25 | 97.32 | 97.26 | 97.11 | 97.44 | 97.59 | 97.54 |
| 35°                 | 96.96 | 97.06 | 97.10 | 97.17 | 97.04 | 97.26 | 97.39 | 97.37 | 97.03 | 97.36 | 97.57 | 97.66 |
| 40°                 | 96.87 | 96.96 | 97.02 | 97.10 | 96.85 | 97.09 | 97.26 | 97.28 | 96.78 | 97.09 | 97.37 | 97.53 |
| 45°                 | 96.72 | 96.79 | 96.88 | 96.96 | 96.62 | 96.85 | 97.06 | 97.13 | 96.43 | 96.75 | 97.06 | 97.34 |
| 50°                 | 96.43 | 96.49 | 96.57 | 96.67 | 96.22 | 96.43 | 96.67 | 96.79 | 95.94 | 96.24 | 96.61 | 96.92 |
| 55°                 | 95.98 | 96.05 | 96.14 | 96.27 | 95.69 | 95.90 | 96.15 | 96.32 | 95.29 | 95.58 | 96.02 | 96.41 |
| 60°                 | 95.09 | 95.16 | 95.25 | 95.39 | 94.69 | 94.87 | 95.17 | 95.39 | 94.15 | 94.45 | 94.95 | 95.40 |
| 65°                 | 93.69 | 93.78 | 93.85 | 94.05 | 93.15 | 93.31 | 93.70 | 93.98 | 92.42 | 92.72 | 93.34 | 93.90 |
| 70°                 | 91.43 | 91.54 | 91.65 | 91.89 | 90.72 | 90.90 | 91.32 | 91.74 | 89.74 | 90.07 | 90.85 | 91.56 |
| 75°                 | 93.25 | 93.38 | 93.52 | 93.85 | 92.17 | 92.40 | 92.97 | 93.54 | 90.74 | 91.13 | 92.18 | 93.14 |
| 80°                 | 78.72 | 78.87 | 79.03 | 79.59 | 77.18 | 77.49 | 78.21 | 78.94 | 75.27 | 75.71 | 77.04 | 78.25 |
| RMSE <sub>avg</sub> | 0.284 | 0.281 | 0.278 | 0.273 | 0.297 | 0.291 | 0.281 | 0.274 | 0.319 | 0.309 | 0.290 | 0.276 |

\* columns with  $\phi = 0^\circ$  represents anisotropic Gaussian kernel

TABLE II. REFERENCE LINE HIT RATE FOR R=10 (IN %)\*

| $\phi$              | 0°     | 15°   | 30°   | 45°   | 0°    | 15°   | 30°   | 45°   | 0°    | 15°   | 30°   | 45°   |
|---------------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\lambda$           | 2      |       |       |       | 3     |       |       |       | 4     |       |       |       |
| $\alpha$            | R = 10 |       |       |       |       |       |       |       |       |       |       |       |
| 5°                  | 93.48  | 92.91 | 91.42 | 90.27 | 97.25 | 96.68 | 95.65 | 93.48 | 97.48 | 98.05 | 97.71 | 95.88 |
| 10°                 | 95.75  | 95.97 | 95.35 | 94.84 | 97.33 | 97.62 | 97.33 | 96.48 | 97.56 | 98.13 | 98.13 | 97.45 |
| 15°                 | 96.8%  | 97.0% | 96.7% | 96.49 | 97.65 | 98.06 | 98.02 | 97.57 | 97.65 | 98.28 | 98.43 | 98.17 |
| 20°                 | 96.98  | 97.25 | 97.11 | 96.98 | 97.47 | 97.86 | 97.99 | 97.75 | 97.33 | 98.02 | 98.30 | 98.13 |
| 25°                 | 9.19%  | 97.45 | 97.47 | 97.38 | 97.43 | 97.77 | 98.07 | 98.01 | 97.21 | 97.90 | 98.31 | 98.26 |
| 30°                 | 97.07  | 97.32 | 97.45 | 97.38 | 97.16 | 97.52 | 97.89 | 97.92 | 96.81 | 97.51 | 98.01 | 98.11 |
| 35°                 | 97.02  | 97.26 | 97.43 | 97.46 | 96.93 | 97.36 | 97.74 | 97.99 | 96.53 | 96.67 | 97.71 | 98.21 |
| 40°                 | 96.81  | 97.01 | 97.25 | 97.32 | 96.57 | 96.98 | 97.43 | 97.66 | 96.08 | 96.76 | 97.41 | 97.77 |
| 45°                 | 96.49  | 96.70 | 96.96 | 97.11 | 96.10 | 96.54 | 97.02 | 97.40 | 95.51 | 96.21 | 96.91 | 97.44 |
| 50°                 | 96.00  | 96.20 | 96.52 | 96.68 | 95.44 | 95.89 | 96.43 | 96.88 | 94.69 | 95.42 | 96.22 | 96.86 |
| 55°                 | 95.37  | 95.58 | 95.95 | 96.13 | 94.60 | 95.08 | 95.69 | 96.25 | 93.67 | 94.45 | 95.32 | 96.14 |
| 60°                 | 94.22  | 94.45 | 94.86 | 95.09 | 93.18 | 93.71 | 94.43 | 94.84 | 90.87 | 92.87 | 93.86 | 94.88 |
| 65°                 | 92.51  | 92.76 | 93.25 | 90.00 | 91.14 | 91.69 | 92.53 | 93.41 | 89.64 | 90.59 | 91.71 | 92.97 |
| 70°                 | 89.82  | 90.11 | 90.74 | 91.15 | 88.05 | 88.62 | 89.64 | 90.77 | 86.09 | 87.16 | 88.47 | 90.08 |
| 75°                 | 90.83  | 91.17 | 92.03 | 92.57 | 88.30 | 88.93 | 90.23 | 91.79 | 85.51 | 86.78 | 88.39 | 90.58 |
| 80°                 | 75.38  | 75.73 | 76.80 | 77.46 | 72.09 | 72.73 | 74.20 | 76.11 | 68.65 | 77.79 | 71.67 | 74.25 |
| RMSE <sub>seg</sub> | 0.318  | 0.310 | 0.295 | 0.297 | 0.361 | 0.346 | 0.321 | 0.295 | 0.416 | 0.329 | 0.354 | 0.314 |

\* columns with  $\phi = 0^\circ$  represents anisotropic Gaussian kernel

TABLE III. REFERENCE LINE HIT RATE FOR R=15 (IN %)\*

| $\phi$              | 0°     | 15°   | 30°   | 45°   | 0°    | 15°   | 30°   | 45°   | 0°    | 15°   | 30°   | 45°   |
|---------------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\lambda$           | 2      |       |       |       | 3     |       |       |       | 4     |       |       |       |
| $\alpha$            | R = 15 |       |       |       |       |       |       |       |       |       |       |       |
| 5°                  | 96.22  | 96.57 | 95.65 | 94.28 | 97.37 | 98.17 | 98.05 | 97.14 | 97.14 | 99.08 | 99.89 | 98.86 |
| 10°                 | 97.22  | 97.50 | 97.28 | 96.71 | 97.50 | 98.24 | 98.13 | 97.90 | 97.22 | 98.41 | 99.04 | 98.75 |
| 15°                 | 97.65  | 97.91 | 97.95 | 97.65 | 97.57 | 98.32 | 98.39 | 98.32 | 97.20 | 98.32 | 98.39 | 98.32 |
| 20°                 | 97.44  | 97.75 | 97.88 | 97.72 | 97.28 | 98.02 | 98.24 | 98.19 | 96.81 | 97.91 | 98.54 | 98.65 |
| 25°                 | 97.45  | 97.77 | 98.01 | 97.90 | 97.15 | 97.86 | 98.24 | 98.28 | 96.59 | 97.64 | 98.43 | 98.65 |
| 30°                 | 97.16  | 97.49 | 97.80 | 97.80 | 96.73 | 97.45 | 97.96 | 98.08 | 96.07 | 97.09 | 98.06 | 98.37 |
| 35°                 | 96.97  | 97.32 | 97.66 | 97.76 | 96.40 | 97.14 | 97.70 | 97.99 | 95.64 | 96.67 | 97.71 | 98.21 |
| 40°                 | 96.60  | 96.97 | 97.35 | 97.52 | 95.90 | 96.63 | 97.28 | 97.70 | 94.98 | 96.03 | 97.18 | 97.88 |
| 45°                 | 96.16  | 96.53 | 96.97 | 97.23 | 95.25 | 96.02 | 96.74 | 97.36 | 94.19 | 95.27 | 96.50 | 97.49 |
| 50°                 | 95.49  | 95.90 | 96.37 | 96.69 | 94.36 | 95.16 | 95.97 | 96.70 | 93.09 | 95.16 | 95.97 | 96.70 |
| 55°                 | 94.68  | 95.09 | 95.62 | 96.02 | 93.28 | 94.11 | 95.02 | 95.89 | 91.75 | 92.96 | 94.39 | 95.79 |
| 60°                 | 93.26  | 93.71 | 94.34 | 94.84 | 91.52 | 92.40 | 93.46 | 94.52 | 89.63 | 90.94 | 92.54 | 94.26 |
| 65°                 | 91.22  | 91.70 | 92.43 | 93.05 | 89.01 | 89.94 | 91.16 | 92.47 | 86.66 | 88.05 | 89.86 | 91.91 |
| 70°                 | 88.11  | 88.63 | 89.52 | 90.31 | 85.22 | 86.26 | 87.68 | 90.31 | 82.26 | 83.76 | 85.86 | 88.37 |
| 75°                 | 88.36  | 88.92 | 91.15 | 92.18 | 84.25 | 85.43 | 87.18 | 89.41 | 80.20 | 81.87 | 84.46 | 87.72 |
| 80°                 | 72.12  | 72.63 | 73.97 | 75.27 | 67.04 | 68.17 | 70.04 | 72.71 | 62.29 | 63.82 | 66.55 | 70.20 |
| RMSE <sub>seg</sub> | 0.360  | 0.347 | 0.322 | 0.303 | 0.437 | 0.411 | 0.376 | 0.332 | 0.516 | 0.477 | 0.425 | 0.364 |

\* columns with  $\phi = 0^\circ$  represents anisotropic Gaussian kernel

TABLE IV. REFERENCE LINE HIT RATE FOR R=20 (IN %)\*

| $\phi$              | 0°     | 15°   | 30°   | 45°   | 0°    | 15°   | 30°   | 45°   | 0°    | 15°    | 30°    | 45°    |
|---------------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|
| $\lambda$           | 2      |       |       |       | 3     |       |       |       | 4     |        |        |        |
| $\alpha$            | R = 20 |       |       |       |       |       |       |       |       |        |        |        |
| 5°                  | 97.03  | 97.71 | 97.60 | 96.68 | 97.14 | 98.97 | 99.66 | 99.08 | 96.68 | 100.11 | 102.52 | 102.63 |
| 10°                 | 97.33  | 97.84 | 97.96 | 97.62 | 97.22 | 98.30 | 98.87 | 98.81 | 96.65 | 98.70  | 100.00 | 100.34 |
| 15°                 | 97.61  | 98.02 | 98.28 | 98.10 | 97.24 | 98.17 | 98.81 | 98.88 | 96.60 | 98.28  | 99.40  | 99.85  |
| 20°                 | 97.36  | 97.83 | 98.08 | 97.99 | 96.87 | 97.83 | 98.54 | 98.65 | 96.13 | 97.75  | 98.74  | 99.26  |
| 25°                 | 97.28  | 97.77 | 98.07 | 98.07 | 96.63 | 97.60 | 98.33 | 98.54 | 95.78 | 97.38  | 98.48  | 99.03  |
| 30°                 | 96.90  | 97.32 | 97.45 | 97.38 | 96.12 | 97.52 | 97.89 | 97.92 | 95.13 | 97.51  | 98.01  | 98.11  |
| 35°                 | 96.66  | 97.17 | 97.59 | 97.74 | 95.70 | 96.67 | 97.60 | 98.06 | 94.56 | 96.14  | 97.51  | 98.33  |
| 40°                 | 96.20  | 96.71 | 97.21 | 97.41 | 95.05 | 96.06 | 97.08 | 97.68 | 93.73 | 95.32  | 96.83  | 97.88  |
| 45°                 | 95.62  | 96.16 | 96.74 | 97.02 | 94.27 | 95.29 | 96.42 | 97.21 | 92.70 | 94.32  | 95.93  | 97.31  |
| 50°                 | 94.81  | 95.40 | 96.01 | 96.40 | 93.18 | 94.24 | 95.49 | 96.48 | 91.35 | 92.99  | 94.77  | 96.48  |
| 55°                 | 93.83  | 94.45 | 95.15 | 95.65 | 91.86 | 92.96 | 94.31 | 95.55 | 89.66 | 91.37  | 93.29  | 95.34  |
| 60°                 | 92.20  | 92.87 | 93.65 | 94.30 | 89.75 | 90.95 | 92.47 | 93.97 | 87.13 | 88.92  | 91.07  | 93.48  |
| 65°                 | 89.85  | 90.58 | 91.46 | 92.28 | 86.78 | 88.07 | 89.79 | 91.57 | 83.61 | 85.48  | 87.84  | 90.69  |
| 70°                 | 86.27  | 87.09 | 88.12 | 89.18 | 82.35 | 83.76 | 85.75 | 87.93 | 78.45 | 80.44  | 83.10  | 86.52  |
| 75°                 | 85.64  | 86.64 | 87.91 | 89.33 | 80.25 | 81.83 | 84.29 | 87.15 | 75.12 | 77.32  | 80.43  | 84.79  |
| 80°                 | 68.60  | 69.62 | 70.99 | 72.70 | 62.24 | 63.70 | 66.31 | 69.49 | 56.56 | 58.55  | 61.63  | 66.26  |
| RMSE <sub>seg</sub> | 0.411  | 0.390 | 0.364 | 0.338 | 0.515 | 0.479 | 0.429 | 0.375 | 0.615 | 0.563  | 0.498  | 0.417  |

\* columns with  $\phi = 0^\circ$  represents anisotropic Gaussian kernel

From the all above results it should be noted that skew identification by algorithm incorporating rotated anisotropic Gaussian kernel is more efficient compared to the one without rotation. Consequently, the skew rate identification is efficient up to  $70^\circ$ . Hence, compared to the original algorithm the efficient identification of the angle  $\alpha$  is extended from  $60^\circ$  to  $70^\circ$  (See Table I-IV for reference). This is also evident in Fig. 3 and 4. Fig. 3 shows the incorporation of the rotation angle  $\phi = \{0^\circ, 15^\circ, 30^\circ, 45^\circ\}$  as well as its improvement for the skew rate identification measured by *RLHR*. Furthermore, Fig.4 represents *RMSE* of the skew identification for the  $R = 15$  and  $\lambda = \{2, 3, 4\}$ .

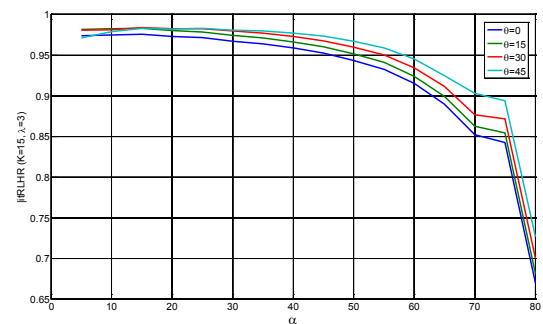


Fig. 3. *RLHR* for the given parameters:  
 $\theta = \{0^\circ, 15^\circ, 30^\circ, 45^\circ\}$ ,  $R = 15$ ,  $\lambda = 3$

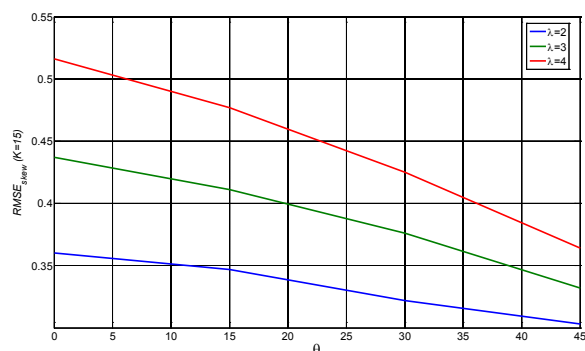


Fig. 4.  $RMSE_{skew}$  for  $R = 15$  and  $\lambda = \{2, 3, 4\}$

## CONCLUSION

In this paper, an approach to Gaussian kernel algorithm for skew identification is presented. The proposed method assumes creation of boundary growing area around text based on Gaussian kernel algorithm extended by rotation. Algorithm quality and robustness is examined by skew rate test [14]. Results are evaluated by  $RMSE$  method. All results are presented and compared with the anisotropic Gaussian kernel method without rotation. The strength of this approach in skew rate identification is mandatory. Its improvement is based on the expansion of the growing areas under specified angle around the text.

## REFERENCE

- [1] Amin A., Wu S.: Robust Skew Detection in mixed Text/Graphics Documents. Proceedings of 8th International Conference on Document Analysis and Recognition, pp. 247–251, Seoul, Korea (2005).
- [2] Said H., Peake G., Tan T.: Baker K. Writer identification from non-uniformly skewed handwriting images. Proceedings of the 9th British Machine Vision Conference, pp. 478–487, Southampton, G. Britain (1998).
- [3] O’Gorman L.: The Document Spectrum for Page Layout Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15 1162–1173 (1993).
- [4] Louloudis G., Gatos B., Pratikakis I., Halatsis C.: Text Line Detection in Handwritten Documents. Pattern Recognition, Vol.41, pp. 3758–3772 (2008).
- [5] Li Y., Zheng Y., Doermann D., Jaeger S.: A New Algorithm for Detecting Text Line in Handwritten Documents. Proceedings of 18th International Conference on Pattern Recognition, Vol.2, pp. 1030–1033. Hong Kong, China (2006).
- [6] Wang J., Mazlor K., Leung H., Hui S.C.: Cursive Word Reference Line Detection. Pattern Recognition, Vol.30, No.3, pp. 503–511 (1997).
- [7] Otsu N.: A Threshold Selection Method from Gray-level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, Vol.9, No.1, pp.62–66 (1979).
- [8] Sauvola L., Pietikainen M.: Adaptive Document Image Binarization, Pattern Recognition, Vol.33, No.2, pp. 225–236 (2000).
- [9] Bukhari S. S., Shafait F., Bruesl T. M.: Adaptive Binarization of Unconstrained Hand-Held Camera-Captured Document Images, Journal of Universal Computer Science, Vol.15, No.18, pp. 3343–3363 (2009).
- [10] Khashman A., Sekeroglu B.: Document Image Binarisation Using a Supervised Neural Network. International Journal of Neural Systems, Vol.18, No.5, pp. 405–418 (2008).
- [11] Jähne B.: Digital Image Processing, Springer-Verlag: Berlin Heidelberg, Germany, (2005).
- [12] Gonzales R.C., Woods R.E.: Digital Signal Processing, 2nd Ed. Prentice-Hall, Upper Saddle River, U.S.A. (2002).
- [13] Brodić D., Milivojević D.R., Milivojević Z.: Basic Test Framework for the Evaluation of Text Line Segmentation and Text Parameter Extraction. Sensors, Vol.10, No.5, pp.5263–5279 (2010).
- [14] Brodić D., Milivojević Z.: An Approach to Modification of Water Flow Algorithm for Segmentation and Text Parameters Extraction. Emerging Trends in Technological Innovation, Camarinha-Matos, L.M., Pereira, P., Ribeiro, L., (Eds.), IFIP AICT, Springer, Boston, U.S.A., Vol.314, pp. 324–331 (2010).
- [15] Basu S., Chaudhuri C., Kundu M., Nasipuri M., Basu D.K.: Text Line Extraction from Multi-Skewed Handwritten Documents. Pattern Recognition, Vol.40, pp. 1825–1839 (2006).
- [16] Bolstad W.M.: Introduction to Bayesian Statistics. John Wiley & Sons, New Jersey, U.S.A. (2005).
- [17] Terell G.R.: Mathematical Statistics: A Unified Introduction. Springer-Verlag, New York, U.S.A. (1999).
- [18] Brodić, D.: Optimization of the Anisotropic Gaussian Kernel for Text Segmentation and Parameter Extraction, Proceedings of Theoretical Computer Science, Calude C.S. & Sassone V., Eds.; IFIP AICT, Springer: Boston, U.S.A., Vol.323, pp. 140–152, (2010).