ITC 3/46
Journal of Information Technology and Control
Vol. 46 / No. 3 / 2017
pp. 308-318
DOI 10.5755/j01.itc.46.3.15797

Text Skew Detection Using Combined Entropy Algorithm

Received 2016/07/26 | Accepted after revision 2017/07/20

http://dx.doi.org/10.5755/j01.itc.46.3.15797

Text Skew Detection Using Combined Entropy Algorithm

Darko Brodić

University of Belgrade, Technical Faculty in Bor, Vojske Jugoslavije 12, 19210 Bor, Serbia
e-mail: dbrodic@tfbor.bg.ac.rs

Zoran N. Milivojević

College of Applied Technical Sciences, Aleksandra Medvedeva 20, 18000 Niš, Serbia
e-mail: zoran.milivojevic@vts.edu.rs

Corresponding author: dbrodic@tfbor.bg.ac.rs

This paper proposes the algorithm for text skew estimation based on the combined entropy calculation. The method consists of three steps. In the first step, it calculates the horizontal and vertical projection profiles. In the second step, the horizontal and vertical entropy for the rough angles is calculated. In the last step, the horizontal and vertical entropy calculation for the smooth angles is performed. The calculated entropy creates the two cost functions: horizontal and vertical. The position where each of cost functions has an absolute minimum represents horizontal and vertical estimated text skew angle. In the last step, it estimates the text skew angle as a mean of horizontal and vertical estimated text skew angles. The functionality, correctness and robustness of the proposed algorithm is investigated by the experiment, which is based on DISEC'13 document database. The obtained results are promising. They confirm that the method correctly estimates text skew angle.

KEYWORDS: algorithm, document image processing, entropy, optical character recognition, text skew.

1. Introduction

Document digitalization is typically carried out by a scanner. As a consequence of this process, the text skew is introduced [30]. Accordingly, the text skew is an unavoidable in digitized document. Humans are very sensitive to any presence of text skew. They can register the text skew of at least 0.1° [13].

Optical character recognition (OCR) is a computer system that can recognize characters from the image that contains text. Any existence of the residual skew can lead to improper character recognition by OCR. Hence, the presence of the text skew is undesirable in any stage of OCR system. If the text skew in document image is not corrected before further processing, then the accuracy of the OCR will be significantly reduced.

Hence, the text skew detection and correction is the prerequisite to any recognition step in the OCR system. As a consequence, the text skew detection represents a mandatory stage in the OCR system [2].

Many different methods have been proposed for the text skew detection in document images [4]. They are usually classified into the following groups: (i) methods based on projection profiles, (ii) clustering of nearest neighbor's methods, (iii) cross-correlation methods, (iv) methods based on the Hough transform, (v) Fourier transforms methods, (vi) moment based methods, (vii) straight-line fitting methods, and (viii) other methods. However, some researchers classify text skew estimation algorithms into the four categories [15]: (i) projection profiles, (ii) Hough transforms, (iii) nearest neighbor clustering, and (iv) interline cross-correlation. Essentially, the above categorization can be modified in the following way: (i) Transformation methods, (ii) Feature extraction methods, (iii) Clustering methods, and (iv) Hybrid methods.

The algorithms based on the mathematical transformations, which extract text features into the space different from image one, are considered as transformation methods. Essentially, they convert image space into some other space. After finding a solution in that space, the information is translated into image space in the form of the text skew estimation. They include methods like Hough transform [2], [31], [33], Fourier transform [26], Radon transform [1], [16], log-polar transform [6].

Feature extraction methods perform some geometrical and morphological transformation as well as image processing in the image space in order to extract the feature(s), which are used for the estimation of the text skew. They include methods like morphology [22], [35], moments [5], cross-correlation [7], linear regression [32], straight-line fitting [8], [20], connected-components [11], [29], projection profiles [17], [23], [27], [28].

Clustering methods perform the extraction of particular type of data characterized by some feature(s) from the database. Typically, these methods can be used for any kind of data given in a big database. According to their methodology and given criteria, the extraction of certain data is performed. These types of methods need to be adjusted for solving the text skew estimation problem. They include methods like K-nearest neighbor [12], [14], [19], fuzzy c-regression models [18].

Hybrid methods combine different aforementioned techniques using additional preprocessing steps. Preprocessing steps are used to improve the efficiency of the certain algorithm by solving some of its disadvantages, which are present in its basic form. In this way, they represent complex algorithms or systems that are computer time intensive, but efficient in the estimation of the text skew [9], [10], [21], [34].

From the aforementioned methods, we choose to explore the extension of the projection profiles method, which is classified as a feature extraction method. A method based on the projection profiles is widely used to detect the text skew angle. It is an easy method which sums up the pixels in the rotated document image creating a histogram at each possible angle. Then a cost function is applied to these histograms. The optimization of the cost function detects the text skew angle. The extension of this method is given in [3], by introducing the entropy concept. Furthermore, the combination of the vertical and horizontal projection profiles method is introduced in [25].

The aim of this paper is to describe an algorithm for the text skew estimation based on the entropy. Accordingly, we suggest the optimization of previously proposed entropy method [3], which leads to speeding up the algorithm by the factor up to 4 times. Furthermore, the method is extended by calculating entropy from vertical projection profiles, too. As a result, the two different methods estimate the text skew angle: (i) horizontal based entropy and (ii) vertical based entropy. Finally, their combination in the form of the mean is calculated. It gives the best estimation result. This approach is called a combined entropy algorithm. It is similar to the one given in [25], which used both horizontal and vertical projection profiles to calculate the text skew of the image. However, our approach uses an entropy calculation instead of projection profile ones, which represents the extension to the method given in [3], [25]. The proposed combined entropy as well as each of the entropy algorithms is tested using a document database DISEC'13 during the experiment [24]. The results of the experiment are compared to the well-known Chou's algorithm [8], Mascaro's algorithm [20], and validated according to the widely accepted methodology [24].

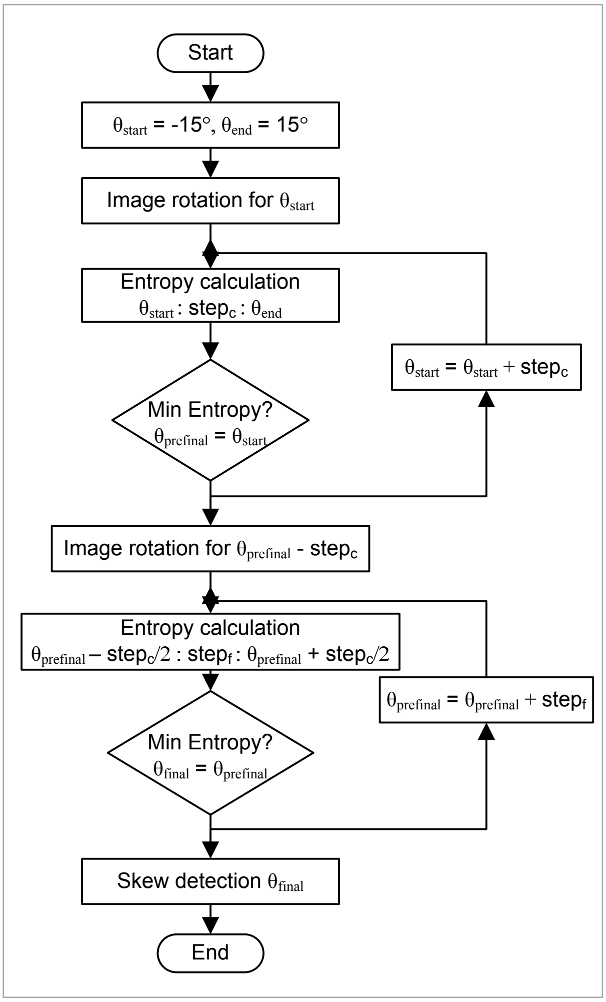
The paper is organized as follows. Section 2 describes the proposed method. Sections 3 defines the experiment. Section 4 defines the evaluation measures that quantify the tested algorithm. Section 5 shows the experimental results and discusses them. Section 6 draws conclusions.

2. Entropy Based Method

The proposed algorithm detects the text skew angle in a document image by calculating entropy of the projection profiles. The preprocessing stage, which includes cleaning of documents, extracting blocks and denoising images [3] is out of the scope of our approach. Figure 1 shows the flow diagram of the proposed algorithm in a general manner.

Figure 1

The flow diagram of the entropy based algorithm

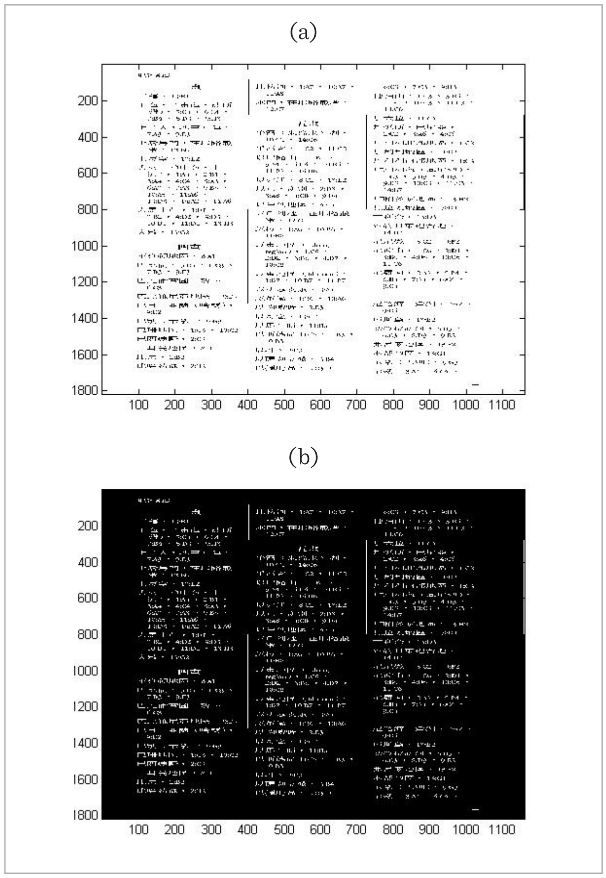


The proposed method is applied to the binary image $B(i,j)$ with M rows and N columns.

Figure 2 shows the initial binary image **B** ((a) skewed for 0.11°, (b) complementary image).

Figure 2

The Initial binary image: (a) skewed for the angle of 0.11°, (b) complementary image



The horizontal projection profile of the binary image **B** is calculated as:

$$HPP(i) = \sum_{j=1}^{N} B(i,j) \tag{1}$$

where $i = 1, ..., M$ represents the row and $j = 1, ..., N$ represents the column of the image **B**. Each element of the horizontal projection profiles $HPP(i)$ represents the sum of the pixel values in the corresponding row. Similarly, the vertical projection profile of the binary image **B** is calculated as:

$$VPP(j) = \sum_{i=1}^{M} B(i, j) \quad (2)$$

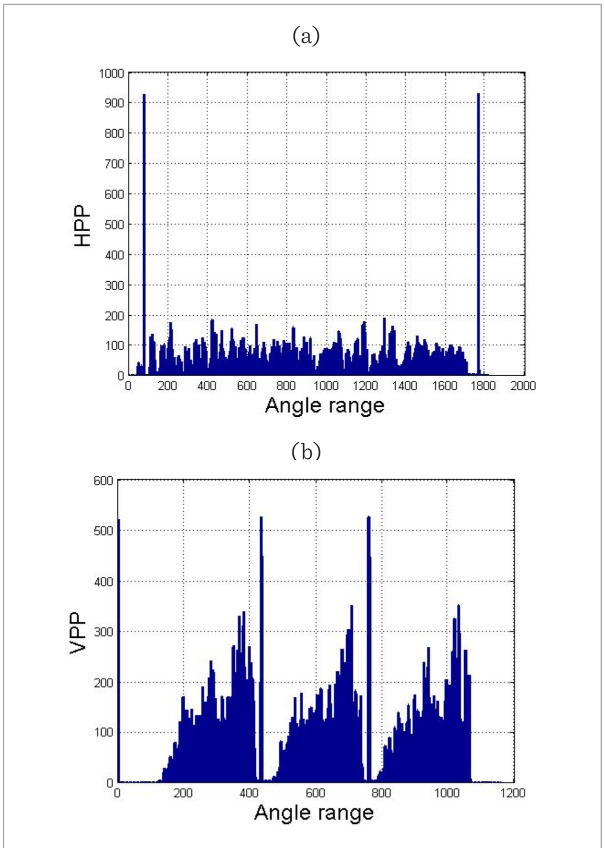
where $i = 1, ..., M$ represents the row and $j = 1, ..., N$ of the image **B**. Currently, each element of the vertical projection profiles $VPP(j)$ represents the sum of the pixel values in the corresponding column.

Figure 3 (a) and (b) shows a corresponding horizontal and vertical projection profiles (*HPP* and *VPP*) for document image given in Figure 2 (b).

Currently, the number of pixels per row (horizontal projection profiles) is under observation. It is clear that the highest value will be received when the text is unskewed. Hence, the procedure of rotating skewed image in the range of angles with certain angle step will establish the *HPP*.

Figure 3

Projection profiles for skewed image (0.1°): (a), Horizontal Projection Profile (HPP) and (b) Vertical Projection Profile (VPP)



Using horizontal projection profiles as a basis, we can introduce the concept of horizontal entropy E_h. The entropy is determined as a statistical measure of image randomness. The horizontal entropy, i.e. the entropy per row is calculated as [3]:

$$E_h(i) = \sum_{i=1}^{M} -HPP(i) \cdot \log\left(HPP(i)\right) \quad (3)$$

If the text-line has more randomness in its texture, then E_h (i) will receive higher values. In the original algorithm [3], the full range of $E_h(i)$ is used for calculation of the cost function E_h.

Similarly, vertical entropy E_v can be defined. The vertical entropy, i.e. the entropy per column is calculated as:

$$E_v(j) = \sum_{j=1}^{N} -VPP(j) \cdot \log\left(VPP(j)\right) \quad (4)$$

Hence, the full range of $E_v(j)$ is used for calculation of the cost function E_v.

Typically, in order to detect the text skew with the accuracy of 0.1° in the text skew range from -15° to +15°, it is necessary to calculate all entropy values for the text images rotated by the step of 0.1°, i.e. 300 times. To reduce the search space for the text skew angle detection, a simple optimization technique is proposed, which is less computer time intensive than the one given in [25]. In our case the number of calculations will be only 40 times.

Our optimized approach consists of the two-stage calculation of the entropy function E_h *or* E_v. In the first stage, the image is rotated from θ_{start} to θ_{end}, where θ_{start} is equal to -15° and θ_{end} is equal to +15°. The step of the rotation is $step_c$, which is equal to 1°. As a result, the entropy cost function E_h *or* E_v is obtained. The lowest value of entropy means that it can be considered as the starting (prefinal) estimated text skew angle, i.e. $\theta_{h,prefinal}$ or $\theta_{v,refinal}$. The obtained accuracy is given by the step of rotation equal to 1°, which is too coarse. If we use this step, the accuracy of the detected text skew will be at the level of 1°. Figure 4 illustrates the horizontal entropy cost function E_h for document image given in Figure 2 (b) with the accuracy of 1°.

Figure 4

The horizontal entropy cost function E_h for the skewed image (0.11°) established by rotating the image using angles from -15° to +15° with the step 1° (minimum determines $\theta_{h,prefinal}$)

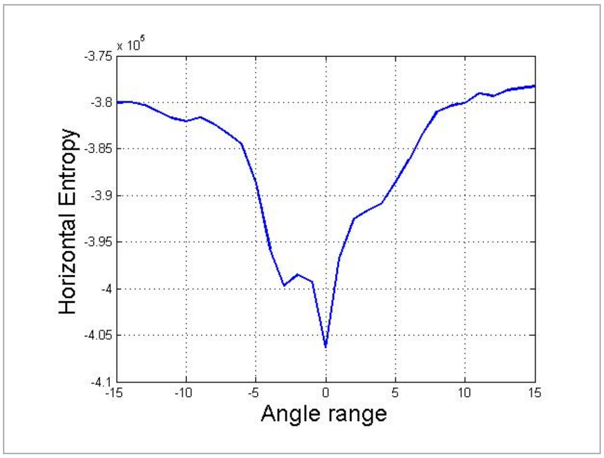
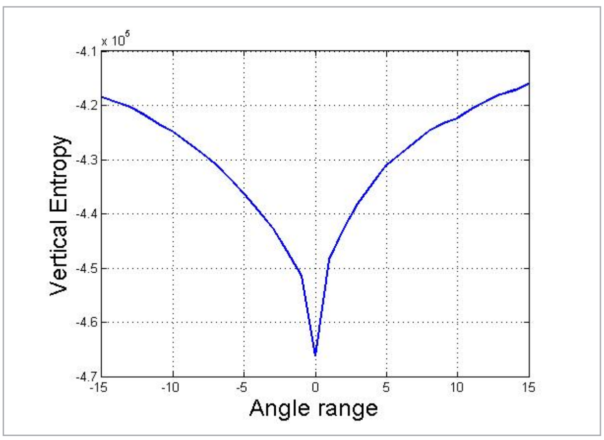


Figure 5 illustrates the vertical entropy cost function E_v for document image given in Figure 2 (b) with the accuracy of 1°.

To correctly detect the text skew angle (with the accuracy of 0.1°), we have to explore the neighborhood of the cost function minimum E_h and E_v, respectively. Hence, in the second stage the subrange around the prefinal text skew angle $\theta_{h,prefinal}$ and $\theta_{v,prefinal}$ given by E_h and E_v, which represent the current minimum of the cost function is examined.

Figure 5

The vertical entropy cost function E_v for the skewed image (0.11°) established by rotating the image using angles from -15° to +15° with the step 1° (minimum determines $\theta_{v,prefinal}$)



The final text skew should be in its nearby of 1° around the cost function minimum. Hence, we divide the step $step_c$ of 1° equally to the left and right side around the minimum. Accordingly, the image is rotated from $\theta_{prefinal} - step_c/2$ to $\theta_{prefinal} + step_c/2$ with the finer step $step_f$. The step $step_f$ is equal to 0.1°, in order to achieve the accuracy of 0.1°. The newly obtained result for the minimal entropy E_h *or* E_v represents the final estimated text skew, i.e. $\theta_{h,final}$ or $\theta_{v,final}$, respectively, with the accuracy of 0.1°.

Figure 6 illustrates the horizontal entropy cost function E_h for document image given in Figure 2 (b) with the accuracy of 0.1°.

Figure 6

The horizontal entropy cost function E_h for the skewed image (0.11°) established by rotating the image using angles around the minimum cost function from $-step_c/2$ to $+step_c/2$ with the step of 0.1° (minimum determines $\theta_{h,final}$)

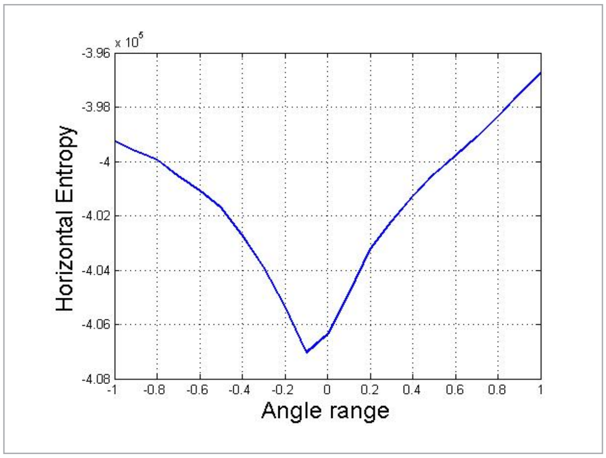


Figure 7 illustrates the vertical entropy cost function E_v for document image given in Figure 2 (b) with the accuracy of 0.1°.

The final detected minimum of the horizontal and vertical entropy cost function E_h and E_v represents the estimated text skew $\theta_{h,final}$ and $\theta_{v,final}$ obtained by the algorithm with the accuracy of 0.1°.

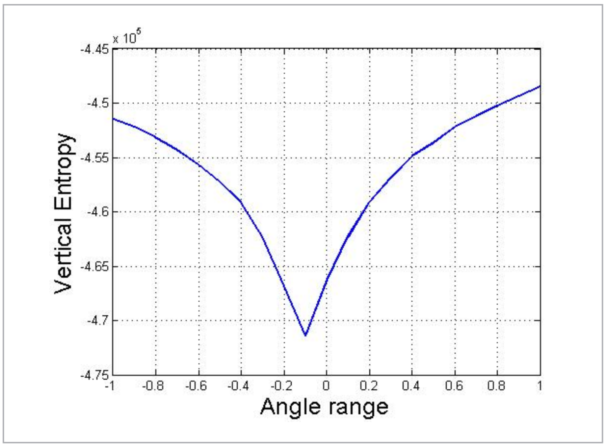
At the end, we additionally calculate so-called combined test skew value $\theta_{c,final}$, which represent the mean of: (i) estimated text skew $\theta_{h,final}$ obtained by the horizontal entropy cost function E_h, and (ii) estimated text skew $\theta_{v,final}$ obtained by the vertical entropy cost function E_v with the accuracy of 0.1. Hence, its value is given as:

$$\theta_{c,\,final} = \frac{\theta_{h,\,final} + \theta_{v,\,final}}{2} \tag{5}$$

In the current case, it is also equal to 0.1°, because both $\theta_{h,final}$ and $\theta_{v,final}$ are equal to 0.1° with an error of 0.1°. However, the mean of these values achieves better accuracy of detected text skew, because it is less sensitive to noise and easy to follow multi-column disposition of the document [25].

Figure 7

The vertical entropy cost function E_v for the skewed image (0.11°) established by rotating the image using angles around the minimum cost function from $-step_c/2$ to $+step_c/2$ with the step of 0.1° (minimum determines $\theta_{v,final}$)



3. Experiment

The experiment consists of two parts. In the first part, a few text documents are rotated from -15° to +15° by the step of 0.1°. This experiment is used only to prove that our optimized approach is better than the original entropy based algorithm [3].

In the second part, the Document Image Skew Estimation Contest (DISEC'13) is used [24]. It is recognized as a standard tool for unbiased testing of the algorithms for the text skew detection. DISEC'13 consists of 20 different images that are ten times differently skewed, which creates the total of 200 images. Furthermore, all these 200 images are used for exploring the quality of the text skew algorithm in predicting the document inclination. It includes plain text documents, newspaper samples, dictionary pages, book excerpts, documents with text, images and/or drawings. The documents are written in different languages, like Chinese, English and Greek or their multilingual combination. Documents from the database are randomly rotated at different angles, ranging from -15° to +15°. The result of the experiment shows the deviation of the estimated text skew obtained by the tested algorithm and correct text skew value from DISEC'13 database (truth-table).

Figure 8 illustrates a few document images from DISEC'13 database.

Figure 8

Document image samples from DISEC'13 database

4. Evaluation Measures

For each document d which is taken from the DISEC'13 database, the distance $Dist(d)$ is calculated. It is determined as a difference between the text skew, which is obtained from testing algorithm and the ground-truth. To quantify a performance of the results obtained from testing algorithm, the following three criteria are proposed [24]: (i) the Average Error Deviation (*AED*), (ii) the average error deviation of the Top 80 % of the results (*TOP80*), and (iii) the percentage of Correct Estimations (*CE*). *AED* is calculated as:

$$AED = \frac{\sum_{d=1}^{X} Dist(d)}{X} \quad (6)$$

where d = 1, ..., X is the number of documents taken from the DISEC'13 database, and $Dist(d)$ is the distance between the estimated text skew and truth-table. In our case, we calculate three versions of the estimated text skew, i.e. $\theta_{h, final}$, $\theta_{v, final}$ and $\theta_{c, final}$. At the end, the comparison will be made and the best version of the entropy based algorithm will be proposed. To calculate *TOP80* measures, the distance $Dist(d)$ should be sorted in ascending order. Then, the first 80% of the documents with the smallest $Dist(d)$ was chosen to represent $sDist(d)$ in order to calculate *TOP80* as:

$$TOP80 = \frac{\sum_{d=1}^{Y} sDist(d)}{Y} \quad (7)$$

where $d = 1, ..., Y$ is the number of the first 80% of DISEC'13 documents, and $sDist(d)$ is the list of the 80% documents with the smallest distance between the estimated text skew and truth-table. *TOP80* characterizes 80% of the best text skew estimations produced by tested algorithm.

The last proposed measure represents *CE*. It is calculated as:

$$CE = \frac{\sum_{d=1}^{X} K(d)}{X} \quad (8)$$

where $K(d) = \begin{cases} 1 & \text{if} \quad Dist(d) \leq 0.1° \\ 0 & \text{otherwise} \end{cases}$

This measure is given in percent. It is worth noting that 0.1° may be visible to a human observer [13]. Hence, it is chosen as the threshold element in eq. (8).

5. Results and Discussion

The results from the first part of the experiment are used to compare the original entropy based algorithm [3] with our optimized approach. Accordingly, the optimized horizontal entropy algorithm is compared to the original one, because the original algorithm is using only horizontal entropy approach. These results are given in the min, max and average error manner. If we draw a parallel to the three aforementioned measures, then min, max and average are calculated in accordance to $Dist(d)$. The obtained results are given in Table 1.

Table 1

The results of the original and optimized entropy based algorithms

Error	Original Entropy	Horizontal Entropy (Optimized)
min	0.15	0.10
max	0.30	0.25
average	0.192	0.166

It is clear that the original algorithm is outperformed by our simplest horizontal entropy based algorithm in terms of error correctness. The second part of the experiment is much more important for the real validation of the proposed versions of the entropy based algorithm. According to the measures *AED*, *TOP80* and *CE*, all three versions of the entropy based algorithm for text skew detection are evaluated. The results are given in Table 2.

Table 2

The results of different entropy based algorithms

	Horizontal Entropy	Vertical Entropy	Combined Entropy
AED	0.118	0.079	0.098
TOP80	0.050	0.051	0.060
CE	72.00	70.00	76.00

It should be noted that these results are obtained without any preprocessing steps before applying different versions of the entropy based algorithm, i.e. using the plain entropy calculation. Furthermore, their comparison is made with the well-known algorithms developed by Chou et al. [8] and Mascaro et. al. [20]. The results obtained by these algorithms are shown in Table 3.

We should use into account that Mascaro et al.'s algorithm is the optimized version of the original Chou et al.'s algorithm.

Table 3

The results obtained from Chou et al.'s and Mascaro et al.'s algorithms

	Chou et al.'s algorithm	Mascaro et al.'s algorithm
AED	1.198	0.556
TOP80	0.103	0.060
CE	45.50	57.00

Table 4

The results of the algorithm based on Hough space derivatives

	vertical estimation	horizontal estimation	equal weights	tuned on experimental set
AED	1.088	0.165	0.103	0.115
TOP80	0.115	0.055	0.048	0.049
CE	53.10	68.71	73.61	73.74

The interpretation of the results linked to the *AED* measure shows that Mascaro et al.'s algorithm compared to Chou et al.'s algorithm makes twice lower accumulated errors in estimating text skew. However, it is 5-6 times higher than any of our versions of the entropy based algorithm.

Furthermore, the *CE* measure shows the following: (i) Chou et al.'s algorithm estimated only 45.50% of all images with the error below or equal to 0.1°, (ii) Mascaro et al.'s algorithm estimated 57.00% of the all images with the error below or equal to 0.1°. It is an obvious step forward in terms of optimization and correctness of the algorithm. Still, the comparison with our simplest version of the entropy based algorithm shows that our algorithm has clear advantages. It can estimate the text skew with an error below or equal to 0.1° in 72.00% of cases. Even better, the combined entropy algorithm obtained the results of 76.00%.

At the end, *TOP80* measure shows that Chou et al.'s algorithm has many big misses in the text skew estimation. On the contrary, Mascaro et al.'s algorithm is much better in estimating text skew on the best 80.00% results. It can be said that it is almost on a par with different versions of the entropy based algorithm.

From all above, it is clear that the proposed entropy based versions of the algorithm have the real potential for text skew detection. Hence, it can be noted that any of our versions of the algorithm outperforms by a large margin Chou et al. and Mascaro et. al. algorithms. Obviously, the entropy based algorithm has a clear advantage over both of these algorithms.

Furthermore, the comparison is made with the newly developed accurate text skew estimation algorithm based on Hough space derivatives [31], which includes many steps and complex calculations. Various derivatives of the same algorithm with different values of parameters are shown in Table 4.

If we compare our combined entropy algorithm with the best version of the algorithm based on Hough space derivatives (which is tuned on experimental set), then our algorithm has a better *AED* of 0.098 compared to 0.115 (margin of approximately 17%), and better *CE* of 76.00% compared to 73.74% (margin of approximately 3%), while *TOP80* is slightly worse, i.e. 0.060 compared to 0.049 (margin of approximately 12%).

From above results, it can be noted that our algorithm has an advantage in correctly estimating text skew angles in the whole database of document. Hence, the accumulated text skew error is smaller (*AED* measure). In addition, it makes smaller errors in predicting document text skew by the margin of 0.1° (*CE* measure). Still, the given algorithm has a better result if we take into account 80% of database documents with the smallest errors. Hence, it can be noted that in other 20% of database documents predicting errors is high.

At the end, the comparison is made with the high performance complex algorithms, which includes many stages with complex calculations. These algorithms have won the best three places in the ICDAR contest. Their results are shown in Table 5.

If we compare these results with different versions of the entropy based algorithm, it is clear that the pro-

Table 5

The results of best algorithms on ICDAR contest

	LRDE-EPITAa (1) [10]	Ajou-SNU (2)	LRDE-EPITAb (3)
AED	0.072	0.085	0.097
TOP80	0.046	0.051	0.053
CE	77.48	71.23	68.32

posed algorithm is on a par with these algorithms. In fact, it will be on the second place in this contest. However, it is worth noting that in some cases the entropy based algorithm (2-3 document images out of all) creates a rather high error in text skew estimation. Obviously, some kind of preprocessing would be helpful to reduce such errors. If it it going to be solved in further research, the proposed versions of the entropy based algorithm will outperform the best of the given complex algorithms.

6. Conclusion

The paper proposed the optimized entropy based algorithm for the text skew estimation as well as its extension by introducing the horizontal, vertical and combined entropy versions of the algorithm. The basic entropy method was optimized by incorporating two stages with so-called rough and fine step rotation of the analyzed image. In this way, the proposed algorithm was speeding up introducing a better accuracy. Furthermore, the extended version of the entropy algorithm based on the horizontal and vertical projection profiles is developed. As a final step, a combined algorithm is created using horizontal and vertical entropy algorithm. All versions of the proposed algorithm were tested using DISEC'13 database, which contains a variety of document images written in different scripts, including different elements like images, comics and so on. Then, the proposed algorithm was compared to the well-known text skew estimation methods proposed by Chou et al. and Mascaro et al. Any version of our method shows clear advantages over these algorithms. At the end, the versions of the entropy based algorithm are compared to a newly developed complex algorithm as well as the complex algorithms used in the ICDAR contest. The obtained results of the experiment show that the proposed algorithm is on a par with these algorithms.

Future research will be toward improving the algorithm by adding some preprocessing techniques in order to further reduce the text skew estimation error.

Acknowledgments

This work was supported in part by the Grant of the Ministry of Education, Science and Technological Development of the Republic Serbia, as a part of the project TR33037.

References

1. Aithal, P. K., Karnataka, M., Rajesh, G., Siddalingaswamy, P. C., Acharya, D. U. A Novel Skew Estimation Approach Using Radon Transform. International Conference on Hybrid Intelligent Systems (HIS), Mellaca, 2011, 1-4. https://doi.org/10.1109/HIS.2011.6122070
2. Amin, A., Fischer, S. A Document Skew Detection Method Using the Hough Transform. Pattern Analysis and Applications, 2000, 3(3), 243-253. https://doi.org/10.1007/s100440070009
3. Arvind, K. R., Kumar, J., Ramakrishnan, A. G. Entropy Based Skew Correction of Document Images. Pattern Recognition and Machine Intelligence (PReMI 2007), LNCS 4815, Kolkata, 2007, 495-502. https://doi.org/10.1007/978-3-540-77046-6_61
4. Brodić, D., Milivojević, D. R. An Algorithm for the Estimation of the Initial Text Skew. Information Technology and Control, 2012, 41(3), 211-219. https://doi.org/10.5755/j01.itc.41.3.1249
5. Brodić, D., Milivojević, Z. N. Estimation of the Handwritten Text Skew Based on Binary Moments. Radioengineering, 2012, 21(1), 162-169.
6. Brodić, D., Milivojević, Z. N. Log-Polar Transformation as a Tool for Text Skew Estimation. Elektronika Ir Elektrotechnika, 2013, 19(2), 61-64. https://doi.org/10.5755/j01.eee.19.2.3471
7. Chaudhuri, A., Chaudhuri, S. Robust Detection of Skew in Document Images. IEEE Transactions on Image Processing, 1997, 6(2), 344-349. https://doi.org/10.1109/83.551708
8. Chou, C. H., Chu, S. Y., Chang, F. Estimation of Skew Angles for Scanned Documents Based on Piecewise Covering by Parallelograms. Pattern Recognition, 2007, 40(2), 443-455. https://doi.org/10.1016/j.patcog.2005.10.030
9. Dhanya, D., Ramakrishnan, A. G., Pati, P. B. Script Identification in Printed Bilingual Documents. Sad-

hana, 2002, 27(1), 73-82. https://doi.org/10.1007/BF02703313

10. Fabrizio, J. A Precise Skew Estimation Algorithm for Document Images Using KNN Clustering and Fourier Transform. International Conference on Image Processing (ICIP), Paris, 2014, 2585-2588. https://doi.org/10.1109/ICIP.2014.7025523
11. O'Gorman, L., Kasturi, R. Executive Briefing: Document Image Analysis. IEEE Computer Society Press, 1997, 1-127.
12. Hashizume, A., Yeh, P. S., Rosenfeld, A. A Method of Detecting the Orientation of Aligned Components. Pattern Recognition Letters, 1986, 4(2), 125-132. https://doi.org/10.1016/0167-8655(86)90034-6
13. Hull, J. J. Document Analysis Systems II. S. L. Taylor (Eds.), World Scientific, 1998, 40-64. https://doi.org/10.1142/9789812797704_0003
14. Jiang, X., Bunke, H., Widmer-Kljajo, D. Skew Detection of Document Images by Focused Nearest-Neighbor Clustering. 5th International Conference on Document Analysis and Recognition, Bangalore, 1999, 629-632.
15. Jung, K., Kim, K., Jain, A. K. Text Information Extraction in Images and Video: A Survey. Pattern Recognition, 2004, 37(5), 977-997. https://doi.org/10.1016/j.patcog.2003.10.012
16. Kapoor, R., Bagai, D., Kamal, T. S. A New Algorithm for Skew Detection and Correction. Pattern Recognition Letters, 2004, 25(11), 1215-1229. https://doi.org/10.1016/j.patrec.2004.03.020
17. Li, S., Shen, Q., Sun, J. Skew Detection Using Wavelet Decomposition and Projection Profile Analysis. Pattern Recognition Letters, 2007, 28(5), 555-562. https://doi.org/10.1016/j.patrec.2006.10.002
18. Lin, C.-T., Fan, K.-W., Yeh, C.-M., Pu, H.-C., Wu, F.-Y. High-Accuracy Skew Estimation of Document Images. International Journal of Fuzzy Systems, 2006, 8(3), 119-126.
19. Lu, Y., Tan, C. L. Improved Nearest Neighbor Based Approach to Accurate Document Skew Estimation. 7th International Conference on Document Analysis and Recognition (ICDAR'03), Edinburgh, 2003, 503-507.
20. Mascaro, A. A., Cavalcanti, G. D. C., Mello, C. A. B. Fast and Robust Skew Estimation of Scanned Documents Through Background Area Information. Pattern Recognition Letters, 2010, 31(11), 1403-1411. https://doi.org/10.1016/j.patrec.2010.03.016
21. Meng, G., Pan, C., Zheng, N., Sun, C. Skew Estimation of Document Images Using Bagging. IEEE Transactions on Image Processing, 2010, 19(7), 1837-1846. https://doi.org/10.1109/TIP.2010.2045677
22. Najman, L. Using Mathematical Morphology for Document Skew Estimation. SPIE Document Recognition and Retrievals XI, SPIE Proceedings, 2004, 5296, 182-191.
23. Papandreou, A., Gatos, B. A Novel Skew Detection Technique Based on Vertical Projections. 11th International Conference on Document Analysis and Recognition (ICDAR'11), Beijing, 2011, 384-388. https://doi.org/10.1109/ICDAR.2011.85
24. Papandreou, A., Gatos, B., Louloudis, G., Stamatopouos, N. ICDAR2013 Document Image Skew Estimation Contest (DISEC'13). 12th International Conference on Document Analysis and Recognition, (ICDAR'13), Washington, 2013, 1444-1448.
25. Papandreou, A., Gatos, B., Perantonis, S. J., Gerardis, I. Efficient Skew Detection of Printed Document Images Based on Novel Combination of Enhanced Profiles. International Journal on Document Analysis and Recognition (IJDAR), 2014, 17(4), 433-454. https://doi.org/10.1007/s10032-014-0228-5
26. Peake, G. S., Tan, T. N. A General Algorithm for Document Skew Angle Estimation. 4th International Conference in Image Processing (ICIP), Santa Barbara, 1997, 2, 230-233. https://doi.org/10.1109/ICIP.1997.638728
27. Postl, W. Detection of Linear Oblique Structures and Skew Scan in Digitized Documents. International Conference on Pattern Recognition (ICPR), Paris, 1986, 687-689.
28. Sadri, J., Cheriet, M. A New Approach for Skew Correction of Documents Based on Particle Swarm Optimization. 10th International Conference on Document Analysis and Recognition (ICDAR'09), Barcelona, 2009, 1066-1070. https://doi.org/10.1109/ICDAR.2009.268
29. Sarfraz, M., Mahmoud, S. A., Rasheed, Z. On Skew Estimation and Correction of Text. International Conference on Computer Graphics, Imaging and Visualisation (CGIV'07), Bangkok, 2007, 308-313. https://doi.org/10.1109/CGIV.2007.63
30. Sarfraz, M., Rasheed, Z. Skew Estimation and Correction of Text Using Bounding Box. 5th International Conference on Computer Graphics, Imaging and Visualization (CGIV'08), Penang, 2008, 259-264. https://doi.org/10.1109/CGIV.2008.10
31. Singh, C., Bhatia, N., Kaur, A. Hough Transform Based Fast Skew Detection and Accurate Skew Correction Methods. Pattern Recognition, 2008, 41(12), 3528-3546. https://doi.org/10.1016/j.patcog.2008.06.002

32. Shivakumara, P., Hematha Kumar, G., Guru, D. S., Nagabhushan, P. A Novel Technique for Estimation of Skew in Binary Text Document Images based on Linear Regression Analysis. Sadhana, 2005, 30(1), 69-85. https://doi.org/10.1007/BF02710080

33. Srihari, S. N., Govindaraju, V. Analysis of Textual Images Using the Hough Transform. Machine Vision and Applications, 1989, 2(3), 141-153. https://doi.org/10.1007/BF01212455

34. Stahlberg, F., Vogel, S. Document Skew Detection Based on Hough Space Derivatives. 13th International Conference on Document Analysis and Recognition, (ICDAR'15), Tunis, 2015, 366-370.

35. Sun, C., Si, D. Slant Correction for Document Images Using Gradient Direction. 4th International Conference on Pattern Recognition (ICPR), Santa Barbara, 1997, 1, 142-146.

Summary / Santrauka

This paper proposes the algorithm for text skew estimation based on the combined entropy calculation. The method consists of three steps. In the first step, it calculates the horizontal and vertical projection profiles. In the second step, the horizontal and vertical entropy for the rough angles is calculated. In the last step, the horizontal and vertical entropy calculation for the smooth angles is performed. The calculated entropy creates the two cost functions: horizontal and vertical. The position where each of cost functions has an absolute minimum represents horizontal and vertical estimated text skew angle. In the last step, it estimates the text skew angle as a mean of horizontal and vertical estimated text skew angles. The functionality, correctness and robustness of the proposed algorithm is investigated by the experiment, which is based on DISEC'13 document database. The obtained results are promising. They confirm that the method correctly estimates text skew angle.

Straipsnyje pristatomas kombinuotos entropijos skaičiavimu pagrįstas teksto iškreiptumo lygį įvertinantis algoritmas. Metodas susideda iš trijų žingsnelių. Pirmajame žingsnyje skaičiuojami horizontalieji ir vertikalieji projekcijų profiliai. Antrajame žingsnyje skaičiuojama horizontalioji ir vertikalioji apytikslių kampų entropija. Paskutiniajame žingsnyje atliekami horizontaliosios ir vertikaliosios lygių kampų entropijos skaičiavimai. Apskaičiuota entropija sukuria dvi sąnaudų funkcijas: horizontaliąją ir vertikaliąją. Pozicija, kurioje kiekviena sąnaudų funkcija turi absoliutų minimumą, atspindi horizontalaus ir vertikalaus vertinamo teksto iškreiptumo kampą. Paskutiniajame žingsnelyje teksto iškreiptumo kampas įvertinamas kaip horizontalių ir vertikalių vertinamų teksto iškreiptumo kampų vidurkis. Siūlomo algoritmo funkcionalumas, teisingumas ir stiprumas patikrinamas atliekant eksperimentą, pagrįstą DISEC'13 dokumentų duomenų baze. Gauti rezultatai yra daug žadantys. Jie patvirtina, kad autorių siūlomas metodas teisingai įvertina teksto iškreiptumo kampą