Hough and Fourier Transforms in the Task of Text Lines Detection

Conference Paper · May 2013

CITATIONS
0

READS
986

4 authors, including:

Alexey Potapov
ITMO University
81 PUBLICATIONS **150** CITATIONS

SEE PROFILE

Hough and Fourier Transforms in the Task of Text Lines Detection

I.P. Gurov, A.S. Potapov, O.V. Scherbakov and I.N. Zhdanov
Department of Computer Photonics and Videomatics
Saint Petersburg National Research University of Information Technologies, Mechanics and Optics
Saint Petersburg, Russia
Email: avenger15@yandex.ru

***Abstract* — Text line segmentation is one of important tasks in automatic text analysis. In this paper, we propose two different approaches to estimate a text block orientation. Presented approaches are applicable under several assumptions based on the text lines structure.**

Keywords—text line detection; Fourier transform; Hough transform; rapid skew estimation

I. INTRODUCTION

Skew correction is one of important steps in the layout control of machine printed text. Although several methods of text line segmentation already exist, not many of them are used in practice [1-9]. Thereby skew correction based on the text line segmentation presents a challenging task.

Most of the existing methods involve complex algorithms of character recognition [9], which are not necessary in the case of express skew analysis. Such express analysis would be more time-efficient in printing. It should be noted that different approaches to text line segmentation use different assumptions about text structure such as text style format, capability of character parameterization, etc. and do not work very well, when these assumptions are not satisfied.

In this paper, we present two algorithms of detecting text lines in images: finding lines through the centroids of character contours of text line and Fourier spectrum analysis of initial text.

II. ALGORITHMS OVERVIEW

This paper is concerned with research of several approaches to detecting text lines in images (text lines are further called “strings” to avoid confusion with the lines as graphical objects). One of the approaches is to reveal centroids of chained components, which correspond to characters and then to apply the Hough transform for finding lines, along which the strings are placed.

The second approach is to search for local maxima in Fourier spectrum of initial text.

A. Hought Transform for Text Lines Detection

The developed method of detecting strings by revealing centroids of letters consists of the following steps:

- Smoothing and detecting edges of letters using the Canny edge detector.
- Distinguishing connected components.
- Calculation of connected components’ centroids.
- Applying the Hough transform to the obtained mass centers.
- Determining θ- and *r*-parameters of lines, along which strings are placed at the image. Here the parameter θ is the angle of the vector from the origin to the closest point of the line; *r* is the distance between the line and the origin.

  This method is based on the assumption that there is insignificant deviation of centroids from the lines which give direction to string, as well as on the assumption that there is small average distance between symbols and different words of the recognized text.

B. Fourier Spectrum Analysis for Text Lines Detection

The method of search for local maxima in spectrum of a text image in order to detect strings involves the following stages:

- Calculating the Fourier transform of initial image.
- Suppressing a high-frequency component of the obtained Fourier spectrum.
- Searching for local maxima of the residual spectrum.
- Restoring a periodic component of the text by local maximum using the inverse Fourier transform.
- Detecting areas containing strings using the Otsu automated threshold binarization.
- Determining θ- and *r*-parameters of text lines, either using the Hough transform or by distinguishing connected components and computing the second-order row statistical moment and the second-order column statistical moment parameters.

The main feature of this method is that it assumes the existence of constant frequency (periodicity) in the arrangement of strings, which enables the search for local maximum in the Fourier spectrum. This approach can be compared with the text recognition by a defocused optical system (e.g. by a person with myopia), which distinguishes strings quite well, though being incapable to determine each symbol with sufficient reliability.

III. Algorithm Application for Typescript

Let us consider in detail peculiarities of application of the methods listed above and obtained results.

A. Hough Transform and Character Centroid

As almost no letters exist in majority of alphabets with mass centre located in the interline interval, one point can be assigned to each symbol and then the obtained points can be connected to form a string using the existing detection methods such as the Hough transform.

The centroid of chained components can be computed using the following formulas:

$$x = 1/A \sum_{(x,y)\in R} x, \quad (1)$$

$$y = 1/A \sum_{(x,y)\in R} y, \quad (2)$$

where x and y are appropriate coordinates in the set R, A is a number of elements in the set R. The centroid presents average location of pixels in the set R.

After obtaining the centroids it is necessary to find the lines that connect all centroids of one string to each other. The Hough transform is very appropriate for solving this task.

The normal parameterization for this transform can be described by the equation

$$r = x\cos(\theta) + y\sin(\theta), \quad (3)$$

where θ is the angle of the vector from the origin to this closest point and r shows the distance between the line and the origin.

Thus, strings are found as a set of maxima within the Hough space. Figs. 1–4 show an image example, the centroids found in the image, grouped centers, and resulting detected text lines.

It can be seen that centroids have to be grouped before application of the Hough transform.

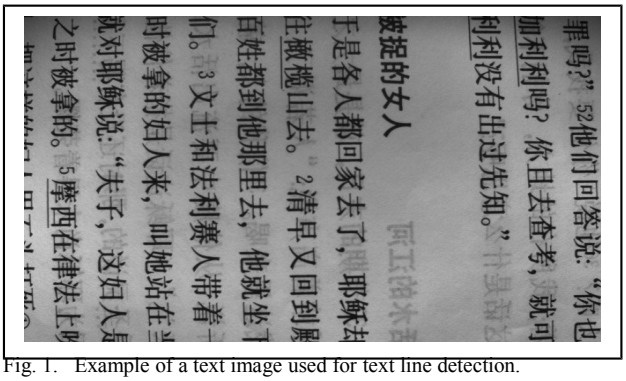

Fig. 1. Example of a text image used for text line detection.

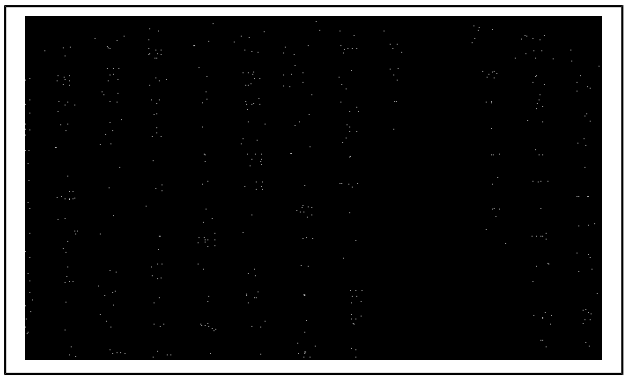

Fig. 2. Centroids found in the image.

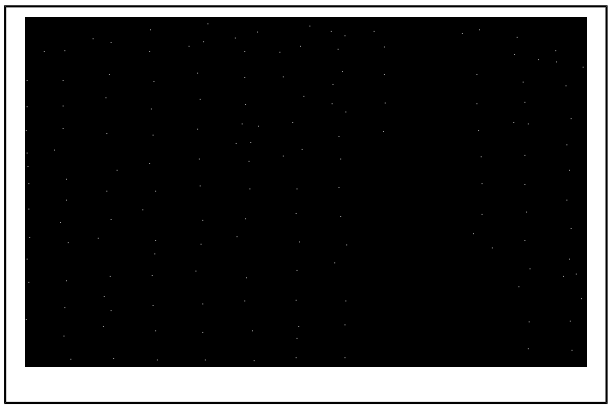

Fig. 3. Image with grouped centers.

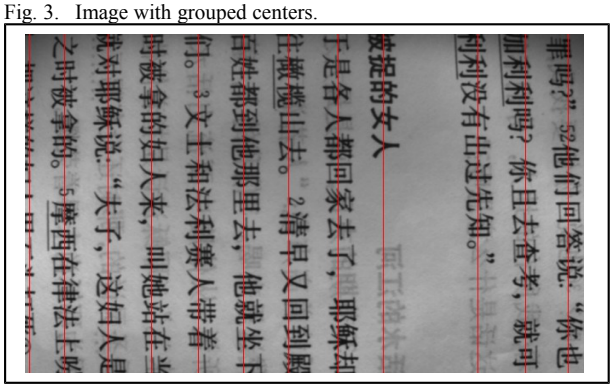

Fig. 4. Detected text lines.

B. Text Lines Periodicity and Fourier Transform

It seems possible to mark out some periodicity associated with line-spacing metric for many text blocks. The Fourier transform decomposes an image into a set of spatial frequencies, one of which represents spatial periodicity. Thus, corresponding area of spatial spectrum indicates strings within a source image. Mathematically it is represented as follows:

$$G(u,v) = F\{g_1(x,y) - g_2(x,y)\}(u,v)\cdot F\{I(x,y)\}(u,v)\cdot G_0(u,v), \quad (4)$$

$$g_1(x,y) = 1/(2\pi\sigma_1^2)\cdot\exp(-(x^2+y^2)/(2\sigma_1^2)), \quad (5)$$

$$g_2(x,y,K) = 1/(2\pi K^2\sigma_1^2)\cdot\exp(-(x^2+y^2)/(2K^2\sigma_1^2)), \quad (6)$$

$$G_0(u,v) = 1/(2\pi\sigma_2^2)\cdot\exp(-(u^2+v^2)/(2\sigma_2^2)), \quad (7)$$

where $G(u,v)$ is the function that represents resulting spectral area, $F\{\}(u,v)$ is the Fourier transform operator, $g_1(x,y)$, $g_2(x,y,K)$ are the Gaussian functions with K as a scale coefficient, $G_0(u,v)$ is the Gaussian function representation in Fourier domain.

The obtained results are presented in Figs. 5–10.

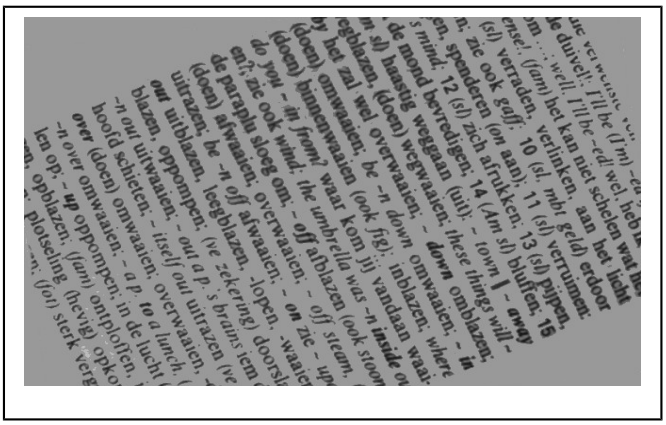

Fig. 5. Example of a text part with skew.

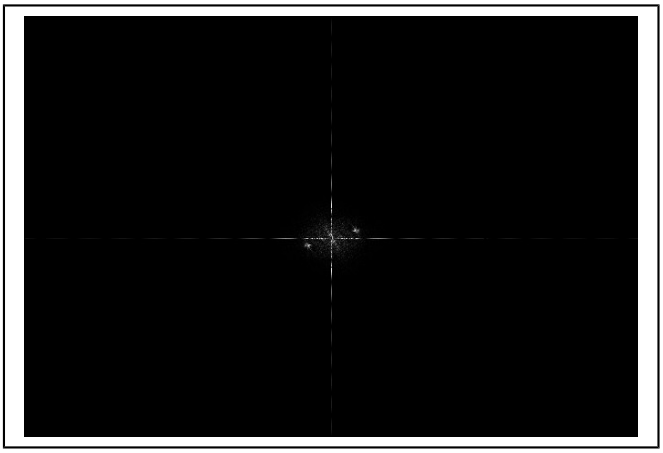

Fig. 6. Spatial Fourier spectrum of the source text image.

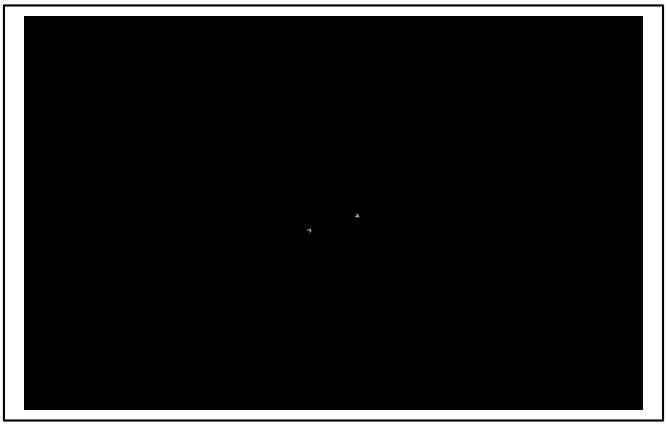

Fig. 7. Detected local maxima in the spatial frequence domain.

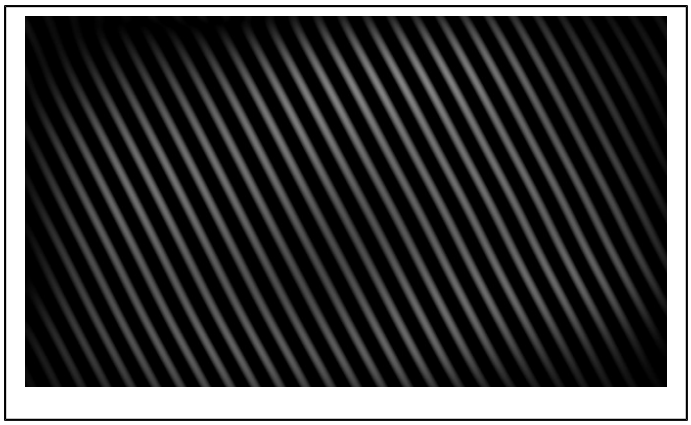

Fig. 8. Result of the inverse Fourier transformation of the spectral maxima in Fig. 7.

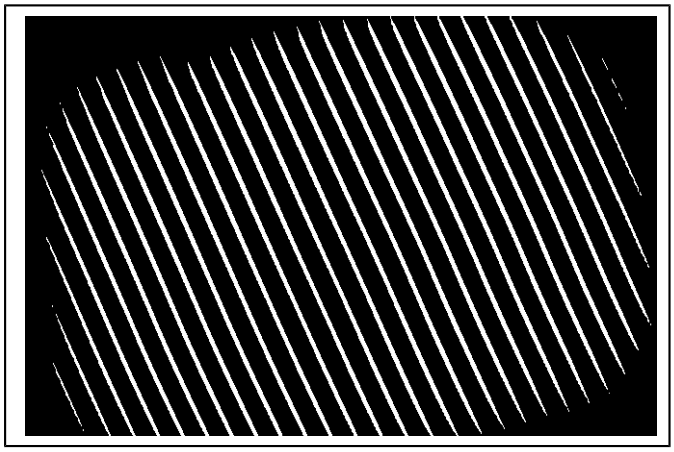

Fig. 9. Binarization result.

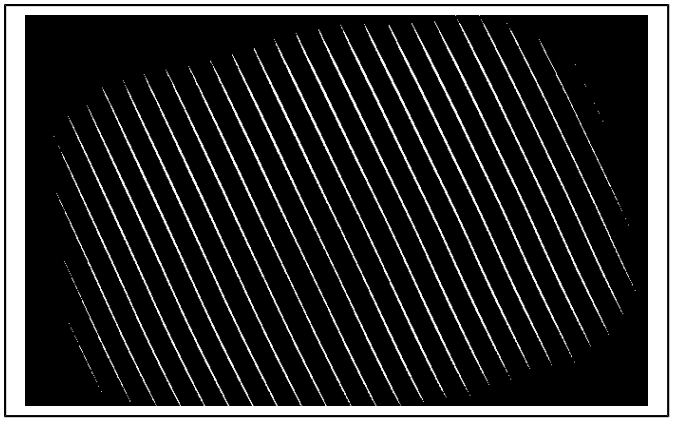

Fig. 10. Detected text lines.

The skew parameter can be described by the second-order row, second-order column and second-order mixed statistical moments.

IV. Algorithm Application for Penscript

It is interesting to consider possibility of skew detection of pen script using the proposed algorithm based on the Fourier transform. Test samples used consist of script and pen script text as illustrated in Fig. 11 and Fig. 12. Fig. 11 shows

approximately horizontal text lines position, while one can see essentially skewed text samples in Fig. 12.

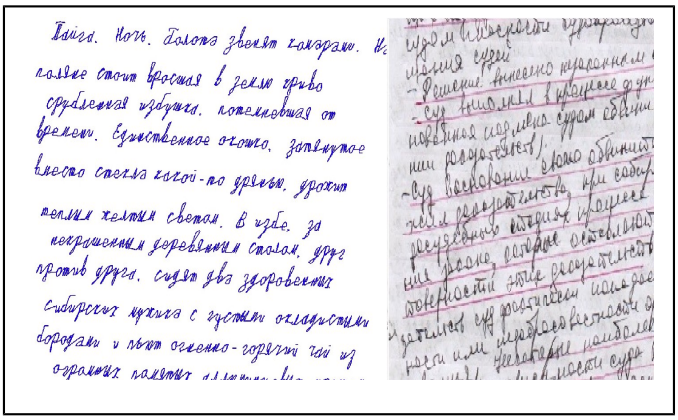

Fig. 11. Samples of non-skewed pen script images.

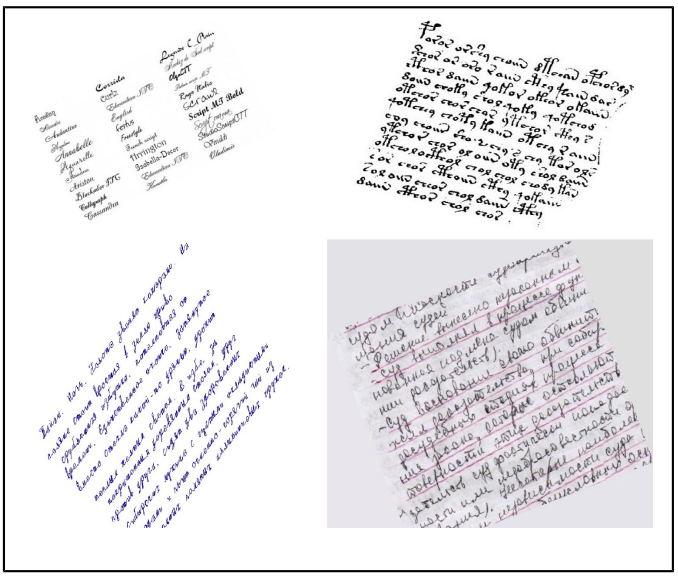

Fig. 12. Samples of skewed script and pen script images (sample 1 upper-left, sample 2 – upper-right, sample 3 bottom-left, sample 4 bottom-right)

Since the samples 1 and 2 are script samples, it is not necessary to show detection results of their non-skewed variations. In spite of the property of pen script, which consists of irregular text lines, as well as irregular words and characters positions, Fourier transform provides useful information about text orientation contained in the phase spectrum in spatial frequency domain.

Most appropriate results can be obtained by varying the two parameters: smoothing level (preprocessing stage) and size of a spectrum local maximum area. The following results (see Figs. 13–18) are obtained using various values of these two parameters.

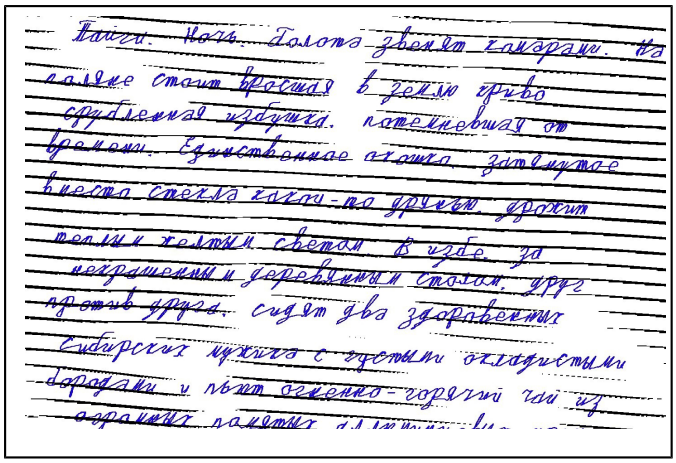

Fig. 13. The result of line detection. Sample 3. Approximately horizontal text lines.

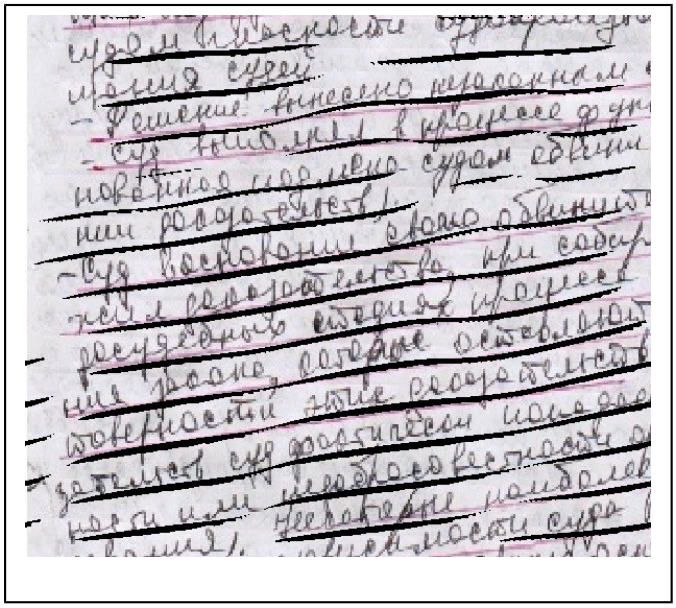

Fig. 14. The result of line detection. Sample 4. Approximately horizontal text lines.

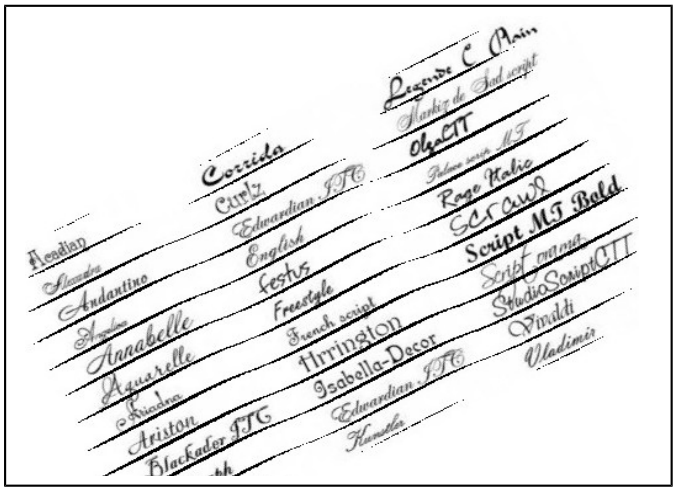

Fig. 15. The result of line detection. Sample 1. Skewed text lines.

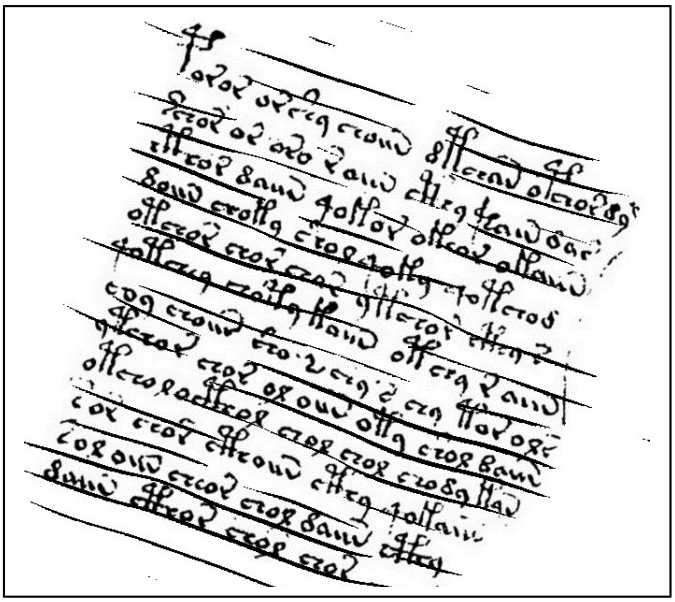

Fig. 16. The result of line detection. Sample 2. Skewed text lines.

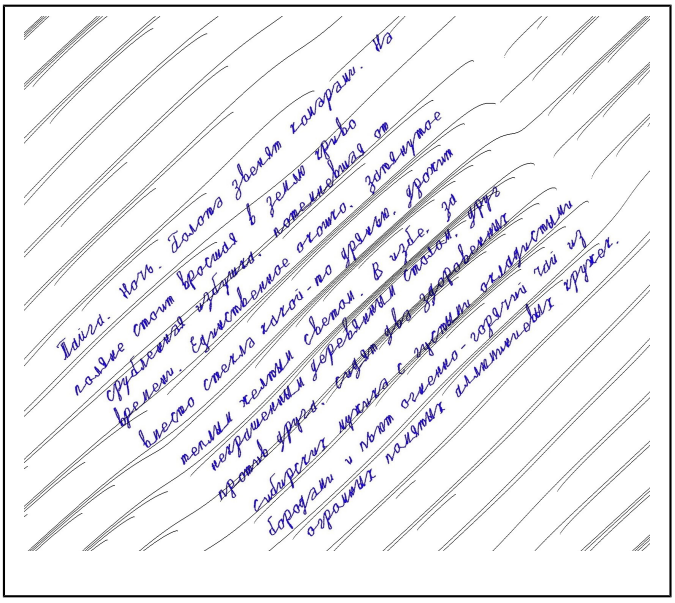

Fig. 17. The result of line detection. Sample 3. Skewed text lines.

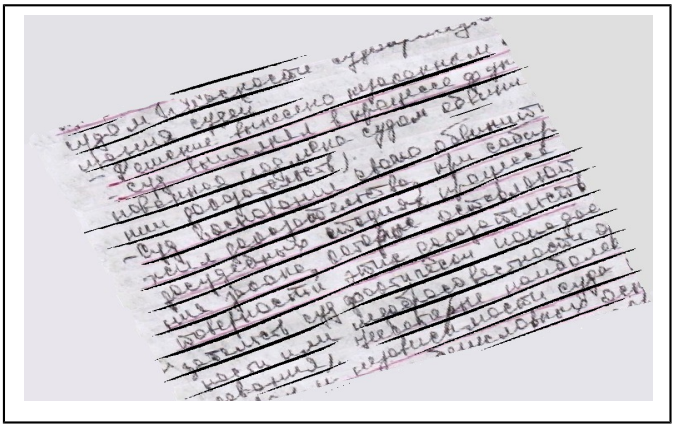

Fig. 18. The result of line detection. Sample 4. Skewed text lines.

It is seen in Figs. 13-18 that the suggested method provides successful evaluation of pen script orientation similar to the results obtained when evaluating typescript.

V. CONCLUSION

Two different approaches to detecting text lines in a text image were considered and investigated based on finding lines through the centroids of character contours of text line using Hough transform and Fourier spectrum analysis of periodicity in text.

Using initial assumptions about centroids of characters and existence of periodicity in line-space metric, estimation of skew parameter for text block was provided.

Fig. 3 shows that the lines detected by the Hough transform are not persistent to large deviations of centroids of the characters. It is not persistent as well to the multiple noise factors such as stamps, stains, paper texture, etc.

The approach based on Fourier spectrum analysis is more resistant with respect to influence of noise factors, but it is more sensitive to large deviations (much larger than line-space metric) of gap between words.

The results for penscript line detection has also been obtained by applying the method based on Fourier transform. It has been shown that it is possible to detect skew of penscript, but by variation of two parameters.

Although the methods which involve character recognition are more stable, the suggested approaches enable fast estimation of the text block slope guided by general view of text lines.

ACKNOWLEDGMENT

This work was supported by the Russian Federation President's grant Council (MD-1072.2013.9) and the Ministry of Education and Science of the Russian Federation.

REFERENCES

[1] V. Papavassiliou, T. Stafylakis, V. Katsouros, and G. Carayannis "Handwritten document image segmentation into text lines and words," Pattern Recognition, 2010, vol. 43, pp. 369-377.

[2] U.-V. Marti and H. Bunke "Text Line Segmentation and Word Recognition in a System for General Writer Independent Handwriting Recognition," In Proc. 6th Int. Conference on Document Analysis and Recognition, 2001, pp. 159–163.

[3] M. Liwicki, E. Indermuhle, and H. Bunke, "On-Line Handwritten Text Line Detection Using Dynamic Programming," In Proc. Ninth International Conference on Document Analysis and Recognition, 2007, vol. 1, pp. 447-451.

[4] J. S. Cardoso, "Stable Text Line Detection," In Proc. Workshop on Applications of Computer Vision (WACV), 2009, pp. 1-5.

[5] S. S. Bukhari, F. Shafait, and T. M. Breuel "Text-Line Extraction using a Convolution of Isotropic Gaussian Filter with a Set of Line Filters," In Proc. International Conference on Document Analysis and Recognition, 2011, pp. 579-583.

[6] G. Louloudis, B. Gatos, I. Pratikakis, and K. Halatsis, "A Block-Based Hough Transform Mapping for Text Line Detection in Handwritten Documents." International Workshop on Frontiers in Handwriting Recognition (IWFHR'06), Oct. 23-26, 2006, pp. 515-520

[7] G. Gluhchev and F. Thomas “Automatic Slope Evaluation of Handwritten Text,” In Proc. International Conference on Computer Systems and Technologies−CompSysTech, 2005, pp. III.A.10-1–III.A.10-4.

[8] S. F. Rashid, F. Shafaity, and T. M. Breuel “Scanning Neural Network for Text Line Recognition,” In Proc. 10th IAPR International Workshop on Document Analysis Systems (DAS), 2012, pp. 105-109.

[9] S. Srinivasan, L. Zhao, L. Sun, Z. Fang, P. Li, T. Wang, R. Iyer, R. Illikkal, D. Liu “Performance Characterization and Acceleration of Optical Character Recognition on Handheld Platforms,”IEEE International Symposium on Workload Characterization (IISWC), 2010, p. 1-10.