A New Algorithm for Skew Detection in Images of Documents

Conference Paper *in* Lecture Notes in Computer Science · September 2004

DOI: 10.1007/978-3-540-30126-4_29 · Source: DBLP

2 authors:

Rafael Dueire Lins
Federal University of Pernambuco

Bruno Avila
Federal University of Pernambuco

A New Algorithm for Skew Detection in Images of Documents

Rafael Dueire Lins and Bruno Tenório Ávila

Universidade Federal de Pernambuco Recife - PE, Brazil
{rdl, bta}@ee.ufpe.br

Abstract. Very frequently the digitalisation process of documents produce images rotated of small angles in relation to the original image axis. The skew introduced makes more difficult the visualisation of images by human users. Besides that, it increases the complexity of any sort of automatic image recognition, degrades the performance of OCR tools, increases the space needed for image storage, etc. Thus, skew correction is an important part of any document processing system being a matter of concern of researchers for almost two decades now. The search for faster and good quality solutions to this problem is still on. This paper presents an efficient algorithm for skew detection and correction of images of documents including non-textual graphical elements, such as pictures and tables. The new algorithm was tested in over 10,000 images yielding satisfactory performance.

Keywords: Document Image Analysis, Skew detection, Rotated Images.

1 Introduction

Organisations are moving at a fast pace from paper to electronic documents. However, large amounts of paper documents inherited from a recent past are still needed. Digitalisation of documents appears as a bridge over the gap of past and present technologies, organizing, indexing, storing, retrieving directly or making accessible through networks, and keeping the contents of documents for future generations. Scanners tend to be of widespread use for the digitalization of documents, today.

One of the important problems in this field is that very often documents are not always correctly placed on the flat-bet scanner either manually by operators or by the automatic feeding device. This very frequent problem yields rotated images. For humans, rotated images are unpleasant for visualisation and introduce extra difficulty in text reading. For machine processing, image skew brings a number of problems that range from needing extra space for storage to making more error prone the recognition and transcription of the image by automatic OCR tools 15. These reasons make skew correction phases a common place in any environment for document processing. This problem has been addressed by researchers for over two decades

A. Campilho, M. Kamel (Eds.): ICIAR 2004, LNCS 3212, pp. 234–240, 2004.

now, but it remains a matter of interest today to find faster and more accurate solutions to it.

The very comprehensive review by Cattoni and his colleagues 6 presents that skew estimation algorithms may be classified into eight different classes, according to the basic approach they adopt, as shown on Table 1.

Table 1. Cattoni et al. 6 classification of skew detection algorithms

Method	Ref.	Input type resolution	Skew range / accuracy	Characteristics of documents
Projection Profile	19	B/w, g.l 160 dpi	±45° / 0.6°	Complex documents
	4	B/w 300 dpi	±45° / 0.6°	Most text
	8	B/w 300 dpi	±45° / 0.6°	Complex documents
	12	B/w 300 dpi	±45° / 0.6°	Complex documents
	3	B/w, JBIG, 300dpi	±3°	Most text
Hough Transform	25	B/w 128 dpi	±90° / 1°	Text only documents
	11	B/w 75 dpi	±15° / 0.5°	Complex documents
	13	B/w 200 dpi	±15° / 0.5°	Complex documents
	16	B/w 300 dpi	±20° / 0.5°	Documents with tables
	18	B/w 160 dpi	±45° / 0.2°	Complex documents
	28	B/w 50-75 dpi	±90° / 0.1°	Complex documents
Nearest Neighbour Clustering	10	B/w 54-63 dpi	±90° / 5°	Simple documents
	17	B/w 300 dpi	±90°	Text only documents
	22	B/w 300 dpi	±15° / 0.05°	One text direction
	18	B/w 160 dpi	±45° / 0.2°	Complex documents
Cross Correlation	2	B/w 200 dpi	±10°	Complex documents
	27	B/w, g.l., colour	±45°	Complex documents
	9	B/w 96-300 dpi	±5° / 0.05°	Complex doc. fast
Gradient Analysis	20	B/w, g.l.	±20° / 1°	Complex documents
	26	G.l.	±90°	Complex documents
Fourier Transform	19	B/w, g.l., 160 dpi	±45°	Dominant text direction
Morphology	7	B/w 300 dpi	±5° / 0.5°	Complex documents
SLIDE	1	B/w, g.l., 144 dpi	±90° / 0.01°	Simple documents

All the referenced methods above work on monochromatic input images and have as underlying assumption that these approaches deal with documents in which text is arranged along parallel straight lines and that text represents most of the document image; performance of most algorithms often decay in the presence of other components like graphics or pictures. Furthermore, the major part of the skew detection schemes assume to deal with documents with a clearly dominant skew angle, and only a few methods can deal with documents containing multiple skews (the very recent paper 24 addresses this problem in the context of correcting the image produced by an uneven feeding device to a scanner, for instance).

This paper presents a new algorithm for skew detection that works with complex documents, with a clear dominant skew direction between –45° to +45°, any image resolution and user defined accuracy up to resolution limitations. The proposed algorithm does not fully meet any of the classifications presented on Table 1. At most, one could say that it brings a vague resemblance to cross correlation algorithms.

2 The New Algorithm

Although the algorithm proposed herein may work in greyscale and colour documents, for the sake of simplicity it is assumed that documents are monochromatic. The basic idea of the new algorithm is to focus the analysis on the leftmost black points of documents. The average of their horizontal coordinate provides the axis taken as basis for rotation. The horizontal distance of the leftmost points to the reference axis (deviation) is used to calculate the rotation angle to be applied. The rotated image undergoes the same process until a rotation error is met (termination condition). In what follows the algorithm is detailed.

Fig. 1. Example of rotated image

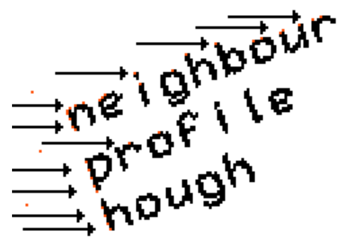


Fig. 2. Image with selected points

The first step of the algorithm scans the image top-down looking for the leftmost points in the image. The points selected are shown in Figure 2.

The distance between the leftmost point and the other selected points is measured. A depth parameter is used to avoid points that do not belong to the leftmost border of the document, but lay on rotated horizontal lines (Figure 3).

The average of the abscissa, called A, of the remaining leftmost points is calculated. Point A determines the axis taken as basis for rotation correction (Figure 4).

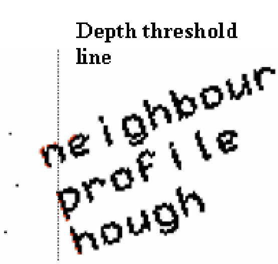


Fig. 3. Points selected by depth parameters

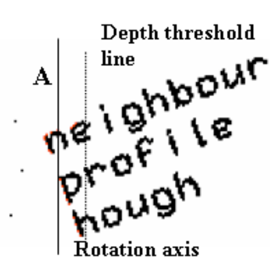


Fig. 4. Calculus of rotation axis A

At this step the ten percent leftmost points are selected and their average, called L_A, is calculated (Figure 5). The distance of L_A to A is taken as the tangent of the rotation angle ▯. The position of L_A in relation to the middle point of A determines if rotation is clockwise (L_A in the superior half of A) or anti-clockwise (L_A in the inferior half of A).

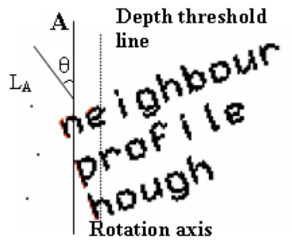


Fig. 5. Calculus of rotation angle θ and direction

The image formed only by the original leftmost points is rotated of angle θ by using the following rotation transformation. Notice that using the image of leftmost points not only saves processing time as fewer points are rotated, but also preserves image quality by avoiding image degradation introduced successive transformation.

$$\begin{aligned} x' &= (x - c_x)\cos\theta - (y - c_y)\sin\theta + c_x \\ y' &= (y - c_y)\cos\theta + (x - c_x)\sin\theta + c_y \end{aligned} \tag{1}$$

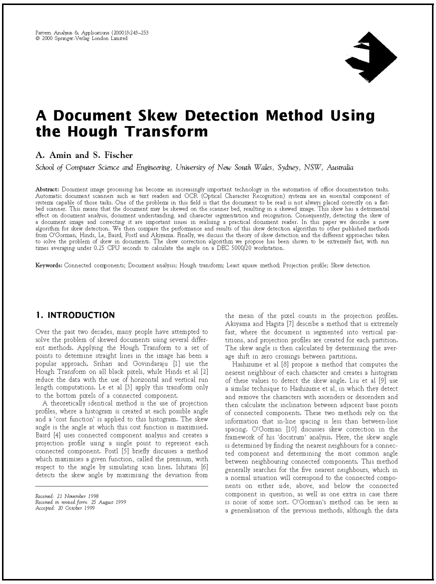
The new image is iterated successively until the new rotation angle calculated is less than the accepted by the application. This parameter depends also on image resolution. Another possibility is to apply the algorithm until it starts to diverge (the trend of the algorithm is to oscillate around a minimal point). The addition of all successive rotation angles yields the total rotation angle to be applied to the original image providing the needed skew correction.

3 Performance

The algorithm presented herein was tested using images such as the ones presented on figures 6, 7, 8, and 9. Altogether, 200 page images were used to benchmark the detection of skew detection quality and space-time performance. These images are different pages extracted from articles on skew detection and correction in the literature, all of them referenced in this paper. The test images were rotated in both directions (clockwise and counter-clockwise) from 0.1º to 0.9º with steps of 0.1º, from 1º to 9º with steps of 1º, and finally from 10º to 45º with steps of 5º. This makes 10,400 test images in the benchmark set used.

The skew detection algorithm was implemented in C and ran on an Intel Pentium IV, 2.4 GHz clock, 512MB RAM and HD IDE. Image files were in TIFF format. No image took longer than 30 ms to have its skew detected. The average processing time for skew detection of the set of benchmark images was 9ms in a left-to right scanning and 8ms in bottom-up scanning. The average number of iterations was 4 with maximum of 14 iterations for both scan orientations. The minimum error obtained by the algorithm was 0.004 degrees.

Some of the tested images with rotation angle greater than 20º, produced unsatisfactory results at first and with only one iteration (the initial one). The threshold depth of the algorithm was widened allowing deeper analysis into the document image. This increases the number of points to handle and makes heavier the computational effort involved in skew calculation. In these cases, increasing the threshold depth from 20 pixels to 100 or 200 pixels yielded acceptable skew recognition.



Pattern Analysis & Applications (2000)3:243–253
© 2000 Springer-Verlag London Limited

A Document Skew Detection Method Using the Hough Transform

A. Amin and S. Fischer

School of Computer Science and Engineering, University of New South Wales, Sydney, NSW, Australia

Abstract: Document image processing has become an increasingly important technology in the automation of office documentation tasks. Automatic document scanners such as text readers and OCR (Optical Character Recognition) systems are an essential component of systems capable of those tasks. One of the problems in this field is that the document to be read is not always placed correctly on a flat-bed scanner. This means that the document may be skewed on the scanner bed, resulting in a skewed image. This skew has a detrimental effect on document analysis, document understanding, and character segmentation and recognition. Consequently, detecting the skew of a document image and correcting it are important issues in realising a practical document reader. In this paper we describe a new algorithm for skew detection. We then compare the performance and results of this skew detection algorithm to other published methods from O'Gorman, Hinds, Le, Baird, Postl and Akiyama. Finally, we discuss the theory of skew detection and the different approaches taken to solve the problem of skew in documents. The skew correction algorithm we propose has been shown to be extremely fast, with run times averaging under 0.25 CPU seconds to calculate the angle on a DEC 5000/20 workstation.

Keywords: Connected components; Document analysis; Hough transform; Least square method; Projection profile; Skew detection

1. INTRODUCTION

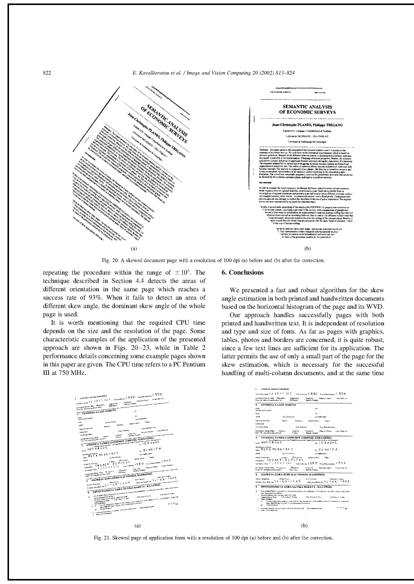
Over the past two decades, many people have attempted to solve the problem of skewed documents using several different methods. Applying the Hough Transform to a set of points to determine straight lines in the image has been a popular approach. Srihari and Govindaraju [1] use the Hough Transform on all black pixels, while Hinds et al [2] reduce the data with the use of horizontal and vertical run length computations. Le et al [3] apply this transform only to the bottom pixels of a connected component.

A theoretically identical method is the use of projection profiles, where a histogram is created at each possible angle and a 'cost function' is applied to this histogram. The skew angle is the angle at which this cost function is maximised. Baird [4] uses connected component analysis and creates a projection profile using a single point to represent each connected component. Postl [5] briefly discusses a method which maximises a given function, called the premium, with respect to the angle by simulating scan lines. Ishitani [6] detects the skew angle by maximising the deviation from the mean of the pixel counts in the projection profiles. Akiyama and Hagita [7] describe a method that is extremely fast, where the document is segmented into vertical partitions, and projection profiles are created for each partition. The skew angle is then calculated by determining the average shift in zero crossings between partitions.

Hashizume et al [8] propose a method that computes the nearest neighbour of each character and creates a histogram of these values to detect the skew angle. Liu et al [9] use a similar technique to Hashizume et al, in which they detect and remove the characters with ascenders or descenders and then calculate the inclination between adjacent base points of connected components. These two methods rely on the information that in-line spacing is less than between-line spacing. O'Gorman [10] discusses skew correction in the framework of his 'docstrum' analysis. Here, the skew angle is determined by finding the nearest neighbours for a connected component and determining the most common angle between neighbouring connected components. This method generally searches for the five nearest neighbours, which in a normal situation will correspond to the connected components on either side, above, and below the connected component in question, as well as one extra in case there is noise of some sort. O'Gorman's method can be seen as a generalisation of the previous methods, although the data

Received: 21 November 1998
Received in revised form: 25 August 1999
Accepted: 20 October 1999

Fig. 6. Image with pictures



822 *E. Kavallieratou et al. / Image and Vision Computing 20 (2002) 813–824*

SEMANTIC ANALYSIS
OF ECONOMIC SURVEYS

Jean-Christophe PLANES, Philippe TRIGANO

(a) (b)

Fig. 20. A skewed document page with a resolution of 100 dpi (a) before and (b) after the correction.

repeating the procedure within the range of $\pm 10°$. The technique described in Section 4.4 detects the areas of different orientation in the same page which reaches a success rate of 93%. When it fails to detect an area of different skew angle, the dominant skew angle of the whole page is used.

It is worth mentioning that the required CPU time depends on the size and the resolution of the page. Some characteristic examples of the application of the presented approach are shown in Figs. 20–23, while in Table 2 performance details concerning some example pages shown in this paper are given. The CPU time refers to a PC Pentium III at 750 MHz.

6. Conclusions

We presented a fast and robust algorithm for the skew angle estimation in both printed and handwritten documents based on the horizontal histogram of the page and its WVD.

Our approach handles successfully pages with both printed and handwritten text. It is independent of resolution and type and size of fonts. As far as pages with graphics, tables, photos and borders are concerned, it is quite robust, since a few text lines are sufficient for its application. The latter permits the use of only a small part of the page for the skew estimation, which is necessary for the successful handling of multi-column documents, and at the same time

(a) (b)

Fig. 21. Skewed page of application form with a resolution of 100 dpi (a) before and (b) after the correction.

Fig. 7. Multi skew documents

4 Conclusions and Lines for Further Work

This paper presents a new algorithm for skew detection that works with complex documents, with a clear dominant skew direction between –45º to +45º, any image resolution and user defined accuracy up to resolution limitations. It is able to handle monochromatic images with non-textual (figures, graphs, tables, etc.) elements. It can be easily adapted to work with greyscale and colour images. The algorithm presented good space-time performance figures as skew calculations are performed in a small subset of points selected from the original image. The choice of a bottom-up scan has proved to be more time efficient and yield more precise results than the left-to-right scan, used to describe the algorithm herein.

The proposed algorithm is being tuned for better accuracy of skew detection. The performance figures obtained so far are extremely optimistic. The choice of the points taken as reference to the skew correction is crucial. We are currently analysing documents to fine-tune the parameters of the algorithm. One of the strategies adopted, is to analyse the mode and variance of the points under consideration. The algorithm presented here only reduces the variance. In the fine-tuning, after variance is minimised, starts a new step where the mode of the chosen points is calculated. Then the chosen points are rotated in steps of 0.1 degrees until reaching the maximum number of points at the mode. The use of component labelling 21 as a way to select the points to fine-tune the algorithm is under consideration. A few anomalous cases found in images rotated of angles greater than 20º were observed in the left-to-right scan of documents. These are being studied also. The results obtained in the fine-tuning of the algorithm presented here, are reported in reference 14.

Acknowledgements. The research reported herein was partly sponsored by CNPq-Conselho Nacional de Desenvolvimento Científico e Tecnológico of the Brazilian Government, to whom the authors express their gratitude.

References

1. H.K.Aghajan, B.H.Khalaj and T.Kailath. Estimation of skew angle in text-image analysis by SLIDE: subspace-based line detection. Machine Vision & Applications, 7:267-276, 1994.
2. T. Akiyama and N. Hagita. Automated entry system for printed documents. Pattern Recognition, 23(11): 1141-1154, 1990.
3. A.Bagdanov and J.Kanai. Projection profile based skew estimation alg. for JBIG comp. Images, 4th Int.Conf.Doc. Analysis and Recog., pp401-405, Aug. 1997.
4. H.S. Baird. The skew angle of printed documents. Proc.Conf. Society of Photographic Scientists and Engineers, vol. 40, pp 21-24, 1987.
5. D.S.Bloomberg, G.E..Kopec and L.Dasani. Measuring document image skew and orientation. Proceedings of IS&T/SPIE EI'95, pp 302-316, February 1995.
6. R.Cattoni, T.Coianiz, S.Messelodi and C.M.Modena, Geometric Layout Analysis Techniques for Document Image Understanding:a Review, ITC_IRST, Jan. 1998.
7. S. Chen and R.M. Haralick. An automatic algorithm for text skew estimation in document images using recursive morphological transforms. Proc.IEEE Int.Conf. on Image Processing, pp 139-143, 1994.
8. G. Ciardiello, G. Scafuro, M.T. Degrandi, M.R. Spada and M.P. Roccotelli. An experimental system for office document handling and text recognition. Proc.9th International Conference on Pattern Recognition, vol. 2, pp 739-743, 1988.
9. B. Gatos, N. Papamarkos and C. Chamzas. Skew detection and text line position determination in digitized documents. Patt. Recognition. 30(9):1505-1519, 1997.
10. A. Hashizume, P.S. Yeh and A. Rosenfeld. A method of detecting the orientation of aligned components. Pattern Recognition Letters, 4:125-132, 1986.
11. S. Hinds, J. Fisher and D. D'Amato. A document skew detection method using run-length encoding and the Hough Transform. Proc.10th. International Conference on Pattern Recognition, pp 464-468, June 1990.
12. Y. Ishitani. Document skew detection based on local region complexity. Proc. of the 2nd Int. Conf. on Document Analysis and Recognition, pp 49-52, Oct. 1993.
13. D.S. Le, G.R. Thoma and H. Wechsler. Automated page orientation and skew angle detection for binary document images. Pattern Recognition, 27(10): 1325-1344, 1994.
14. R.D.Lins and B.T.Ávila. Fast skew detection in images of Documents. In preparation.
15. C.A.B.Mello and R.D.Lins. A Comparative Study on OCR Tools. Vision Interface 99, pp. 700-704, May 1999.
16. Y. Min, S.B. Cho and Y. Lee. A data reduction method for efficient document skew estimation based on Hough Transformation. Proc.13th International Conference on Pattern Recognition, pp 732-736, August 1996.

17. L. O'Gorman. The document spectrum for page layout analysis. IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(11): 1162-1173, 1993.
18. U. Pal and B.B. Chaudhuri. An improved document skew angle estimation technique. Pattern Recognition Letters, 17(8): 899-904, July 1996.
19. W. Postl. Detection of linear oblique structures and skew scan in digitized documents. In Proc. of the 8th Int. Conf.Pattern Recognition, pp. 687-689, 1986.
20. J. Sauvola and M. Pietikäinen. Skew angle detection using texture direction analysis. Proc.9th Scandinavian Conf. Image Analysis, pp 1099-1106, June 1995.
21. L.G.Shapiro and G.C.Stockman, Computer Vision, March 2000. http://www.cse.msu.edu/~stockman/Book/book.html.
22. R. Smith. A simple and efficient skew detection algorithm via text row accumulation. 3rd International Conference on Document Analysis and Recognition. pp 1145-1148, Aug. 1995.
23. A.L. Spitz. Skew determination in CCITT Group 4 compressed document images. Proceedings of Symposium on Document Analysis and Information Retrieval, pp 11-25, 1992.
24. A.L.Spitz. Correcting for variable skew in document images. International Journal on Document Analysis and Recognition, vol 6: 192-200, 2004.
25. S.N. Srihari and V. Govindaraju. Analysis of textual images using the Hough Transform. Machine Vision and Applications, 2(3): 141-153, 1989.
26. C. Sun and D. Si. Skew and slant correction for document images using gradient direction. 4th International Conference on Document Analysis and Recognition, pp 142-146, Aug. 1997.
27. H. Yan. Skew correction of document images using interline cross-correlation. CVGIP: Graphical Models and Image Processing. 55(6): 538-543, Nov. 1993.
28. B. Yu and A.K. Jain. A robust and fast skew detection algorithm for generic documents. Pattern Recognition, 29(10): 1599-1629, 1996.