

Document Skew Detection and Correction Algorithm using Wavelet and Radon Transforms

Hilal M. Yousif Al-Bayatti*

Abdul Monem S. Rahma**

F.A. Al-adhadh***

*** Amman Arab University - For Graduate Studies - Amman, Jordan**

**** Computer Science Dept.-University of Technology - Baghdad, Iraq**

*****Computer department -Mustansriah University-Baghdad, Iraq**

Received:27/10/2008 Accepted:22/3/2009

Abstract:As part of research into document image processing, an algorithm has been developed to detect and correct the degree of skew in a scanned image. The principal components of the algorithm are the wavelet and radon transforms. The wavelet transform is used as a form of down-sampling process to reduce the amount of data to be processed. The performance of the algorithm has been evaluated using a sample size of several images taken from different types of document. Evaluation shows that a skew angle may be detected and corrected.

Keywords: Document Skew , Detection , Correction Algorithm ,Wavelet and Radon Transforms

Introduction

The conversion of paper-based document electronic image format is important system for automated document delivery, document presentation and other applications document conversion includes scanning, displaying processing, text recognition, and creating image and text databases. The previous type of applications requires the detection and correction of skew of the document image before further processing. Separation of lines; words and characters requires the text to be free of any skew.

Skew estimation is a process which aims to detect the deviation of the document orientation angle from the horizontal or vertical direction. First document reading system assumed that documents were printed with a single direction of the text and that the acquisition process did not introduce relevant skew.

The advent of flatbed scanner and the need to process large amount of documents at high rates made the above assumption unreliable and the introduction of the skew estimation phase became mandatory.

Most of the skew estimation techniques can be divided into the following main classes according to the basic approach they adopt: analysis of projection [1,2,3,4,5], Hough transform [6, 7, 8, 9, 10, 11, 12,], use of the radon transform with ad hoc down sampling

[13] and on the use of features of the script of document [14].

Despite many efforts, which have been spent on the development of skew estimation algorithms are proposed in the literatures. This is mainly due to the need for:

- (a) Accurate and computationally efficient algorithms.
- (b) Methods that do not make strong assumptions about the class of document they can deal with.

Computational efficiency may be pursued by reducing the number of points to be processed, by under sampling the image, or by selecting representative points. A common solution, which is applied to binary input images, is to use the connected components as the bases for further computation: an approach which is convenient only if connected components are needed in subsequent step but has the disadvantage of one has to run the connected component labeling module again after deskewing the document.

Another approach is to consider only sub-regions of the image under the assumption that only a single skew is present. A general assumption which, at different extents, underlies skew estimation techniques is that text represents the most relevant part of the document image; performance often decay in the presence of other components like graphics or pictures .The main advantage of the

projection profile methods is that the skew angle accuracy can be controlled by changing the angle step size and the accumulator resolution. Obviously, a finer accuracy requires a longer computation time. A priori information about the maximum expected skew angle could be straightforwardly exploited by the method.

Important features to be considered for a comparison of skew algorithms are: assumption on the document image (input type, resolution, maximum amount of skew, layout structure), accuracy of the estimated angle and computational time.

Unfortunately, a common database of skewed documents is not adopted, making it difficult to do a rigorous comparison of features such as accuracy and computational time. Furthermore, some papers don't provide a detailed description of the experimental results.

In the current research the radon transform has been borrowed from the field of tomography and used to detect the skew angle. It is a novel and efficient way in dealing with the skew of text lines. None of the reviewed papers have used this method to detect the text skew. The wavelet transform is used to reduce the amount of points to be processed. It is based on more rigorous methods for under sampling the image instead of an ad hoc under sampling method.

The Radon Transform [15]

In 1917 radon published a paper describing a technique dealing with parallel beam project geometry. Later this method became known as the radon transform. Its inverse is commonly used for tomography applications. The current research has utilized this idea to detect the angle a text line makes with the horizontal direction and calculate the skewed angle of the text. The radon transform computes projection of an image matrix along specified directions. A projection of a two-dimensional function $f(x,y)$ is a line integral in a certain direction. For example, the line integral of $f(x, y)$ in the vertical direction is the projection of $f(x, y)$ onto the x-axis; the line integral in the horizontal direction is the projection of $f(x, y)$ on to the y-axis.

Projections can be computed along any angle. In general, the radon transform of $f(x, y)$ is the line integral of f parallel to the y axis

$$R = \int f(x' \cos q - y' \sin q, x' \sin q + y' \cos q) dy' \dots \dots (1)$$

Where:

$$\dots \dots (2)$$

R is a matrix containing the projection for

each angle. As one varies the angle, a different value is obtained for the projection. Taking the position of the maximum value of the matrix R gives the index which points to the angle that gave that maximum. This angle is the direction of the text lines with respect to the horizontal axis. The following is an example to illustrate the mechanics of the radon projection.

Example:

$$\begin{matrix} \text{Image} = & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ \text{number of projection} = & 7 \\ \text{angle vector} = & -45 & 45 \\ \mathbf{R} = & 0 & 0 \\ & 0.4142 & 0 \\ & 0.6742 & 0.2652 \\ & 0.4697 \\ & 0.2652 \\ & 0 \end{matrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos q & \sin q \\ -\sin q & \cos q \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Maxim um projection = 2.4697 at column 2 of the matrix R this corresponds to the second value in the angle vector which is the angle 45.

The number of points at which the projection is computed is as follows:

$$2 \cdot (\text{maxdim} - \text{floor}((\text{maxdim} - 1)/2)) + 3$$

Where:

maxdim = maximum dimension of image this number is sufficient to compute the projection at unit intervals.

$$\text{For the } 3 \times 3 \text{ image the number of projection} = 2 \cdot (3 - (3-1)/2) + 3 = 2 \cdot (3-1) + 3 = 7$$

Skew Angle Detection and Correction

Casual use of the scanner may lead to skew in the document image. Skew angle is the angle that the text lines of the document image make with the horizontal direction. Skew correction can be achieved in two steps namely,

- (a) Estimations of the skew angle and
- (b) Rotation of the image by the angle in (a) in the opposite direction.

The skew of the document image can come from the following:

- i- The document is skewed when it was scanned.
- ii- The printing of the text is skewed in the original image.
- iii- The skew could be a combination of the above two.

The current research detected and correct for the above three types. The page skew is detected through horizontal and vertical projections by detecting the extremities of the

document, calculating the skew angle and then correcting the page skew. The text line skew is detected through the use of the radon transform. The range of angle considered here is -25 to +25 with accuracy of .5 degree.

To speed up the process of the radon transform, the wavelet transform is used to reduce the dimension of the image at the same time reduce the amount of pixels to be processed.

For the wavelet expansion of a function $f(t)$, a two parameter system can be constructed as follows:(16)

$$f(t) = \sum_k \sum_j a_{j,k} \psi_{j,k}(t) \quad \dots\dots(3)$$

this wavelet expansion is in terms of two indices, the time translation index k and the scaling index j , where both j and k are integer indices and the $\psi_{j,k}(t)$ are the wavelet expansion (basis) transform. We took the Low-Low part of wavelet transform. This component contains all the important relevant features of the image. After the angle detection, the original image is rotated in the opposite direction.

The main steps of the algorithm is as follows:

The Document Skew Detection and correction algorithm

- Step 1: read gray image.*
- Step 2: Perform a two dimensional first level wavelet transform.*
- Step 3: take the Low-Low component of the transform.*
- Step 4: binarize this component.*
- Step 5: for theta = 65: step 0.5 TO 115 (i.e. +25 to -25)*
- Step 6: apply the Radon transform to the image in step 4.*
- Step 7: Find the maximum of the Radon transform matrix.*
- Step 8: Find the angle which corresponds to the maximum.*
- Step 9: Text line skew angle = line-angle = 90 – angle in step 8.*
- Step 10: Rotate original image by angle in step9.*
- Step11: Crop the rotated image in step 10 to original image size.*

Figure 1 and figure 2 are two examples of using the algorithm. Figure 1 is for a document which contains text only and the skew angle of the text is 5 degrees. Figure 2 is for a document of mixed text and picture with skew angle of -7degrees.

Conclusions

The following points are found to be of interest:

1. The reduction in image size and image pixel density by using the wavelet transform was found to be more robust since it relies on sound mathematical bases and not on ad hoc down sampling technique.

2. The angle accuracy of .5° was found to be the upper limit so as not to affect the machine-hand written classification. Anything higher than this and the algorithm will not work correctly. One could take a lower value but this will slow down the algorithm. No advantage is gained from lowering this value, as far as classification of machine hand written text is concerned.

3. The range of skew angle chosen which is -25 and 25 degrees is a choice of convince, one could have chosen a range of 0-180 degrees, but this has no practical value for document analysis.

The 180 degrees mean the document is upside down. The present research is not concerned with the orientation of the document whether it is upside down or not.

4. Any skew greater than the chosen range (-25 to 25) is sufficient to once visually whether document is skewed or not and then corrected manually, may be through rescanning. One can easily increase the range to -30+30 or -45 to +45 etc..... but increasing the range will affect adversely the processing speed.

5. One could have achieved speed in processing by considering a subset of the image (for example taking the central part of the image) and applying the algorithm to this sub image. It is true one can achieve a reduction in speed but might miss the main part of the text region of the document, and hence the angle detected will not be the angle of the text.

6. It was found that the presence of object other than text region in the document did not affect the detection of the skew angle as long as the text region is the dominant region, as depicted in figure 2. This was found by testing the algorithm on documents with picture regions and without this picture region and the result was the same.

References

- [1] W. postal. **Detection of linear oblique structures and skew scan in digitized documents.** Proc. Of the 8th international conference on pattern. Recognition, pp 687 689, Paris, France, 1986.
- [2] H.S. Baird. **The skew angle of printed documents.** Proc. Of the conference

- society of photographic scientists and engineers, volume 40, pages 21–24, Rochester, NY, may, 20-21 1987.
- [3] G. Ciardiello, C. Scafuro, M.T. Degrandi, M.R. Spada, and M.P. Hoccotelli. An experimental system for office document handling and text recognition. Proc. Of the 9th international conference on Pattern Recognition. Volume 2, pages 739–743, Roma, Italy, November 14–17 1988.
- [4] Y. Ishitani. **Document skew detection** based on local Region complexity. Proc. Of the 2nd international conference on document analysis and recognition, pages 49–52, Tsukuba, Japan, October 1993. IEEE Computer society.
- [5] A. Bagdanov and J. Kanai. Projection Profile Based Skew estimation algorithm for ,JPEG compressed images. Proc. Of the 4th international conference on document analysis and recognition, pages 401-405. UIm. Germany. August 1997.
- [6] S.N. Srihari and V. Govindaraju. Analysis of textual images using the Hough transform. Machine vision and applications. 2(3): 141-153. 1989.
- [7] S. Hinds, J. Fisher, and D. D'Amato. A document skew detection method using run—length encoding and the Hough transform. Proc of 10 th international conference on pattern recognition, pages 464 — 468. Atlantic city, N.J. June. 17-21 1990.
- [8] A.L. Spitz. Skew Determination in CCITT group 4 compressed document images. Proc. Of the symposium on document analysis and information retrieval, pages 11– 25, Las Vegas. 1992.
- [9] D.S. Le, G.R. Thoma, and H. Wechsler. Automated pages Orientation and skew Angle Detection for Binary Document Images. Pattern Recognition. 27(10):1325–1344, 1994.
- [10] Y. Min, S.B. Cho, and Y. Lee. A data Reduction Method for Efficient document skew Estimation based on Hough transformation. Proc of the 13th international conference on pattern, Recognition, pages 732-736. Vienna, Austria. August 1996. IEEE press.
- [11] U. Pal and H.P. Claudihuri. An improved document skew angle estimation technique. Pattern Recognition letters, 17(8) 899–904-1996.
- [12] B. Yu and A.K. Jain. A Robust and Fast Skew Detection Algorithm for Generic Documents. Pattern recognition. 29(10): 1599-1629, 1996.
- [13] F. Al-adhadh. Computer vision for machine and hand written documents. PhD. Thesis, informatics institute for postgraduate studies, 2003.
- [14] B.B Chaudhuri and U. Pal. A Complete printed Bangia OCR System. Pattern recognition Vol. 3 1. No. 5 pp. 531-549, 1998.
- [15] Matlab user's Guide ver2, Mathworks Inc. 2000.
- [16] Dia alzubaidy. Classification and Reconstruction of old Iraqi Civilization Images. Phd thesis, institute for postgraduate studies in informatics,Iraq,2003.



(a) one-level wavelet transform, (the top-left component is taken).

استناداً الى الكشف المقدم من قبل السيد مدير النظم في المكتب والمرفق
طياً والذي يبدأ بالتسلسل (١) وينتهي بالتسلسل (٨) فقد تم احتساب
الساعات الكلية للفرق التي نفذت الأنشطة اعلاه ((الجدول رقم ٣))
المرفق طياً فقد تم احتساب الساعات الإجمالية لمنفذي الأنشطة التي تم
استيفاء مبالغها وقيدت ايراداً لحساب المكتب .

(b) original document with 5 degree skew.

استناداً الى الكشف المقدم من قبل السيد مدير النظم في المكتب والمرفق
طياً والذي يبدأ بالتسلسل (١) وينتهي بالتسلسل (٨) فقد تم احتساب
الساعات الكلية للفرق التي نفذت الأنشطة اعلاه ((الجدول رقم ٣))
المرفق طياً فقد تم احتساب الساعات الإجمالية لمنفذي الأنشطة التي تم
استيفاء مبالغها وقيدت ايراداً لحساب المكتب .

(c) document rotated 5 degrees anti-clockwise as a result of applying the algorithm.

استناداً الى الكشف المقدم من قبل السيد مدير النظم في المكتب والمرفق
طياً والذي يبدأ بالتسلسل (١) وينتهي بالتسلسل (٨) فقد تم احتساب
الساعات الكلية للفرق التي نفذت الأنشطة اعلاه ((الجدول رقم ٣))
المرفق طياً فقد تم احتساب الساعات الإجمالية لمنفذي الأنشطة التي تم
استيفاء مبالغها وقيدت ايراداً لحساب المكتب .

(d) document cropped to original size.

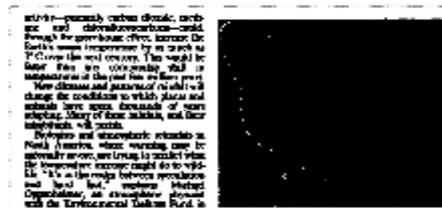
Figure 1 Document with 5 degree skew.



(a) Original document



(b) Binarized image



(c) image rotate 7 degrees clockwise detected by the algorithm



(d) Image cropped to original size.

Figure 2: image with mixed text and picture with -7 degrees skew.

خوارزمية لتحديد و تصحيح الانحراف في الوثائق باستخدام محولات المويجة و رادون

E.mail: monemrahma@yahoo.com

الخلاصة:

كجزء من البحوث الجارية في عملية تحليل صور الوثائق , تم تطوير خوارزمية لإيجاد الانحراف في الوثائق عن الأتجاه الأفقي . الأجزاء الرئيسية للخوارزمية تتركز على تحويل "الويجت" لتقليل عدد النقاط التي يتم التعامل معها , و كذلك تحويل "رادون" لاحتساب ميل الأسطر عن الأتجاه الأفقي . فحصت الخوارزمية على عدة نماذج من الوثائق بنجاح.