Jurnal Teknologi, 37(D) Dis. 2002: 117–126

SKEW DETECTION AND CORRECTION OF JAWI IMAGES USING GRADIENT DIRECTION

KHAIRUDDIN OMAR[1], ABD. RAHMAN RAMLI[2],
RAMLAN MAHMOD[3] & MD. NASIR SULAIMAN[4]

Abstract. This paper discusses an enhancement algorithm of skew and slant orientation of the Jawi images. A simple and quick algorithm that uses a gradient orientation histogram, similar to the one proposed by Sun and Si(1997) for the Latin text is suggested to overcome this skew and slant problem in Jawi text. This algorithm starts by carrying out the image gradient operation, followed by determining the orientation histogram and smoothing algorithm with a median filter. The histogram's maximum value is generated to obtain the skew angle followed by the correcting procedure of the early value obtained by using polynomial function. The angle is calculated analytically and finally converting the image accordingly. The slanted Jawi text uses the same approach as the above to get the slant angle. This image is sheared in order to get the right image. The outputs of this study show promising and convincing results.

Key Words: Gradient orientation histogram, the mask of sobel edge detector, skew, slant, median filter, and skew angle

Abstrak. Kertas kerja ini dikemukakan untuk menjelaskan tentang pembetulan pencongan dan erotan imej Jawi. Satu algoritma mudah dan pantas ini menggunakan histogram orientasi cerunan sama seperti yang telah dicadangkan oleh Sun dan Si(1997) terhadap teks Latin telah dicadangkan untuk mengatasi masalah pencongan dan erotan terhadap teks Jawi. Algoritma bermula dengan melaksanakan operasi cerunan imej; diikuti oleh penentuan histogram orientasi dan licinkannya menggunakan penapis median. Nilai maksimum histogram dicari untuk mendapatkan sudut pencong; diikuti oleh pembetulan nilai awal dengan menggunakan fungsi polinominal. Sudut dihitung secara analitik; dan akhir sekali imej diputarkan mengikut sudut pencong yang diperoleh. Eroten teks Jawi juga telah menggunakan kaedah yang sama untuk mendapatkan sudut erotan dan kemudian diricih supaya dapat menghasilkan imej yang betul. Hasil daripada algoritma ini menjanjikan keputusan yang baik.

Kata Kunci: Histogram orientasi cerunan, topeng penjejak sisi sobel, pencong, erotan, penapis median, dan sudut pencong

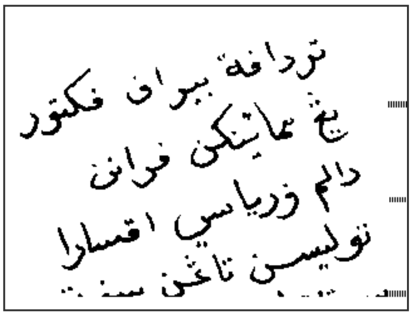
[1] Jabatan Sains dan Pengurusan Sistem, Fakulti Teknologi dan Sains Maklumat Universiti Kebangsaan Malaysia 43600 Bangi. Selangor Darul-Ehsan. e-mail:ko@ftsm.ukm.my

[2] Jabatan Komunikasi Sistem dan Komputer, Fakulti Kejuruteraan, Universiti Putra Malaysia 43400 Serdang. Selangor Darul-Ehsan. e-mail:arr@eng.upm.edu.my

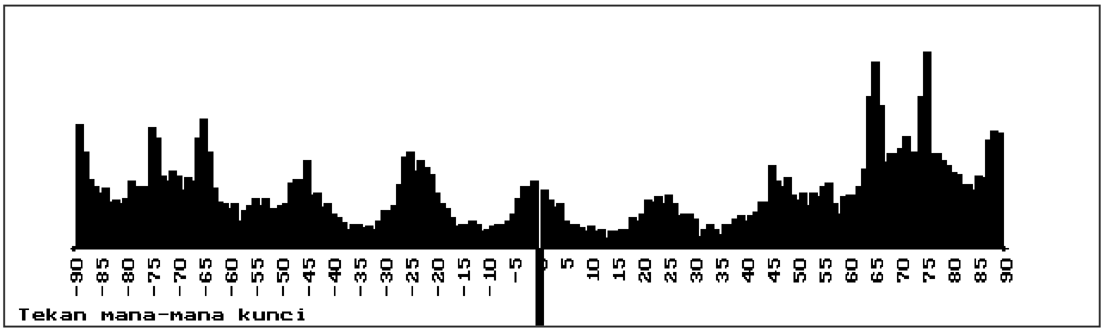
[3&4] Jabatan Sains Komputer, Fakulti Sains Komputer dan Teknologi Maklumat. Universiti Putra Malaysia. 43400 Serdang Selangor Darul-Ehsan. e-mail:ramlan@fsktm.upm.edu.my dan nasir@fsktm.upm.edu.my

1.0 INTRODUCTION

There are many application areas in document image processing, and these include selection of encoding methods for document archiving, retrieval, high quality facsimile, and digital reprographics as well as preprocessor for *Optical Character Recognition* (OCR). The first step in document analysis is to obtain a digitized raster image of the document using appropriate scanning system. It is followed by page layout analysis and character recognition. Before the structure of the text is obtained, a test is carried out to find out whether the document is skewed. If it is skewed, correction needs to be done. Uncompensated skew can cause serious performance deterioration. Figure 1(a) shows one of the skewed document images.



(a)



(b)

Figure 1 (a) Original skewed Jawi text image with 14o. (b) The gradient orientation histogram of Jawi image in (a).

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

Figure 2 Sobel Edge Detector Mask. (a) q operator. (b) p operator.

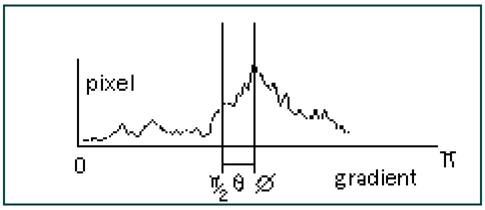


Figure 3 An illustration for skew detection of a document image using gradient orientation histogram in the range of [0, π]. The skew angle θ is the difference between $\pi/2$ and ϕ.

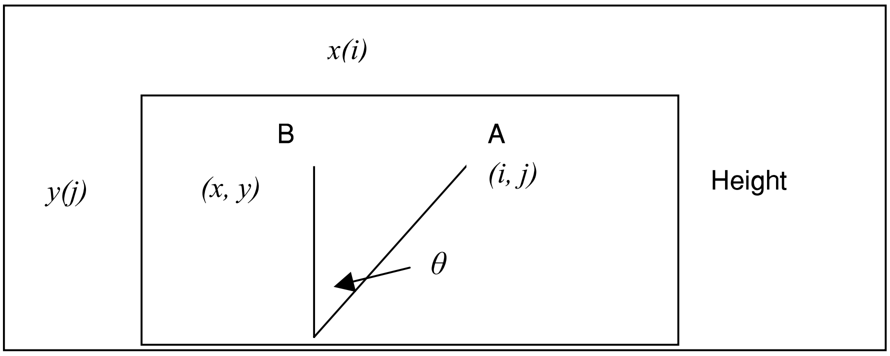


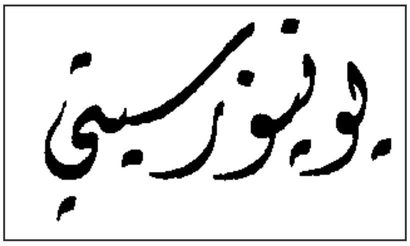
Figure 4 Shear Transformation on Jawi image.

Many techniques have been built for the correction of skewed document images. Basically, they can be described in five categories [1]:

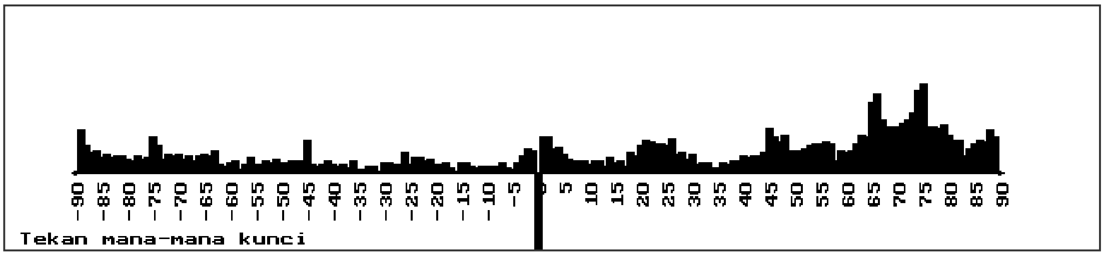
(1) Projection profile;
(2) Hough transform technique;
(3) Fourier method;
(4) Nearest neighbor clustering; and
(5) Correlation.

In technique (1), a series of projection profiles are obtained at a number of angles close to the expected orientation, and the variation is calculated for each of the profiles. The profile that gives the maximum variation corresponds to the projection with the best alignment to the text line. This projection angle is called the skew angle. Baird proposed modifications to the projection profile technique for very fast and accurate iterative convergence on the skew angle [2].

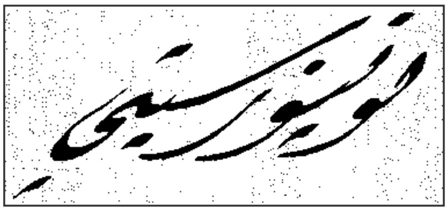
Spitz adapted Baird's alignment technique to fiducially point position data for detecting skew in compressed image [3]. Cardiello *et al.*, have proposed a technique using the projection histogram in which a sample region with the maximum average density of black pixels per row is rotated by pre-specified angles [4]. The horizontal



(a)



(b)



(c)

Figure 5 (a) A Jawi word with slant characters; (b) The gradient orientation histogram; (c) Slant Corrected word shear (slant angle obtained is 15°).

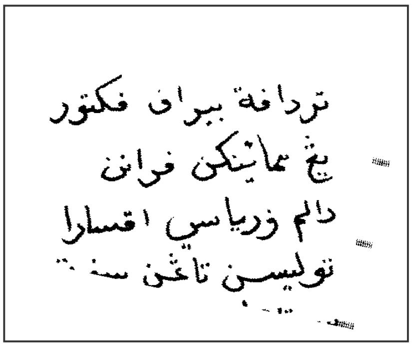


Figure 6 Slant corrected of Jawi word (slant angle is 14°)

projection histogram of the region is evaluated for each angle. The skew angle corresponds to a rotation for which the mean deviation of the histogram is maximized.

Nakano *et al.* [5], Srihari and Govindaraju [6], and Hinds *et al.*, [7] proposed skew detection techniques based on the Hough transform. The Hough transform maps each point in the original (x, y) plane to all points in the (ρ, θ) Hough space of lines through (x, y) with slope θ and distance ρ from origin. The dominant lines are obtained from peak in Hough space and thus the skew. One limitation of this technique is that if text is sparse, it may be difficult to select a peak in the Hough space.

Postl [8] has suggested two techniques for skew correction. One of them is called the simulated skew scan technique, and it can be classified in the projection profile categories. The other technique is based on the Fourier transform. In this technique, the direction in which the density of the Fourier space is the largest gives the skew angle. For large images, the Fourier technique can be computationally expensive, and very often for a document image, the largest density direction of the Fourier space is on a vertical line and the true density direction may not be the largest. This makes the detection for the skewness of a document image difficult using Fourier transforms.

A bottom-up technique to skew estimation based on nearest-neighbour clustering is described by Hashizume *et al.* [9]. In this work, 1-nearest-neighbors of all connected components are obtained. The direction vectors for all nearest-neighbor pairs are accumulated in a histogram, and the histogram peak is found to obtain the document skew angle. The advantage of this technique over the previous techniques mentioned before is that, it is not limited to any range of skew angle. But since only one nearest-neighbor connection is made for each component, connections with noise, subparts of characters, and connection between lines can reduce its accuracy.

Yan [10] has introduced a technique for determining the skew angle of an image by using cross-correlation between lines at a fixed distance. It is based on observations that the correlation between two vertical lines in an image of a skewed document and is maximized in general if one of the lines is shifted relatively to the other line such that the character base line levels for the two lines are coincident.

2.0 SKEW CORRECTION

In this paper, a simple and fast technique for determining the skew angle of an image and the slant angle of Jawi text by using the gradient orientation histogram as proposed by Sun and Si [11] for Latin text is presented. This technique is based on the observation that the gradient orientation should be mainly in direction perpendicular to the text line. After the skew angle has been determined, the image can then be rotated for correction. Slant characters can also be detected using the same gradient orientation technique.

2.1 Skew Correction Algorithm

For an image described by $f(x,y)$, the gradient vector $[p, q]^T$ at point (x,y) is defined as:

$$p = \partial f(x,y)/\partial x, \qquad q = \partial f(x,y)/\partial y. \tag{1}$$

The orientation for this gradient vector is:

$$\phi = \arctan(q/p), \text{ where } \phi = [-\pi, \pi]. \qquad (2)$$

The gradient can be obtained by using an $N \times N$ Sobel filtering operation. Figure 2 shows the Sobel edge detector masked with size 3×3 with N=5. According to [12], the bigger the neighbor's size, the easier the edge thickness and noise are obtained. Calculation on vector $[p, q]$ is done by splitting the kernel into 5×1 sub-kernel (compared to 2×1 by [11]) and iteratively calculates the responses. This splitting process has been described in Juslinda *et al.* [13]. The result is two responses for the *x*- and *y*-directions respectively. The gradient orientation at this point of the image can be obtained using Equation (2). For detecting the orientation of a skewed document, just half of the range of ϕ will be enough. The negative value of ϕ are offset by π to become positive.

For a skewed document, there will be more points in the image whose gradient orientation is perpendicular to the document text lines. It is expected that the statistical information about the gradient orientation of an image can be used for skew angles detection. Figure 1 (b) and Figure 3 show two examples of the statistical information of the gradient orientation of an image, in which ϕ is the skew angle. The orientation histogram for this gradient image can be obtained by using Equation (2). The angle (in the equation is a continuous function. Quantization is needed for the given range and resolution. The resolution of the obtained histogram depends upon the range considered and the number of points in the histogram. If we would like to have 360 points within the angle between 45° and 135°, then the angle resolution will be 0.25 degree. From the obtained histogram, which is $h(\phi)$, the orientation of the skewed document is defined as:

$$\theta = \begin{cases} \phi - \pi/2, & if \quad \phi = [0, \pi], \\ \pi/2 - \phi, & if \quad \phi = [-\pi/2, \pi/2]. \end{cases} \qquad (3)$$

where ϕ is the maximum value for $h(\phi)$. For an upright document image, ϕ will be $\pi/2$, hence the skew angle is zero. Figure 1(a) shows the original image of a skewed document, while Figure 1(b) gives the histogram of the gradient orientation of the original image.

After the initial document's skew angle has been obtained, the obtained initial value was refined by locally fitting a cubic polynomial function. A cubic polynomial function has the form $y = f(x) = ax^3 + bx^2 + cx + d$, where a, b, c, and d are constants. Note that the values of a, b, c, and d are needed to identify a specific polynomial. Basically, four non-collinear points determine a cubic polynomial. Consider the points (–1,0), (0,3), (1,0), and (3,0). If the x-values (in this case is the range considered) in list $L1$, and the y-values (in this case is the gradient orientation) in list $L2$, and then do a cubic regression computation (by using Maple's built-in statistical package), $a = 1$, $b = -3$, $c = -1$, and $d = 3$, with residual value (measure of fitness), $R^2 = 1$. The cubic function that contains the four given points is $y = x^3 - 3x^2 - x + 3$.

In our experiment, the value of a, b, c, and d are varied, depending the sample of image. After fitting a cubic polynomial function, the maximum value can be calculate analytically. The input image can be rotated for skew correction. During the rotation operation, bilinear interpolation of neighboring pixels is involved to reduce noise effect. The rotation can be done by using Equation (4), where (C_x, C_y) is the center of rotation and can be arbitrary, and γ_{ij}, $1 \leqslant i, j \geqslant 2$, are the elements of the rotation matrix.

$$\begin{bmatrix} X_{out} \\ Y_{out} \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} X_{in} \\ Y_{in} \end{bmatrix} - \begin{bmatrix} C_x \\ C_y \end{bmatrix} \tag{4}$$

The algorithm for the detection and correction of the skewed document image is given below:

The Skew Detection and Correction Algorithm (SDCA)

Input: Jawi text image

Output: Skewing free of Jawi image

Begin

Step 1. Perform quantization process by choosing the range of (and the number of points in the histogram.
Step 2. Perform gradient operation on the image by using Equation (1) and (2);
Step 3. Obtained orientation histogram and smooth it out with a median filter;
Step 4. Search for the maximum value of the histogram, $h(\phi)$ in order to obtain an initial skew angle;
Step 5. Refine the obtained initial value by locally fitting a cubic polynomial function, and calculate the maximum analytically; and
Step 6. Rotate image for skew correction by using Equation (4).
End.

2.2 Slant Correction Algorithm

Slant is the angle in degrees clockwise from vertical at which the characters were written. Slant correction is the process, which tries to normalize the slant of the characters in a line or paragraph to the vertical. The technique described in Section 2.1 can be used for character slant correction. After the slant angle has been obtained, a shear operation on this text line or paragraph can be used to correct the slant characters.

During the shear operation, for each pixel (i, j) in the original image is changed to the new coordinates (x, y) of this pixel in the slant-corrected images as follows:

$$y = i, \quad x = i - (\text{height} - j)*\tan(\theta).$$

where θ is the slant angle obtained. Point 'A' is transformed into point 'B' as shown in Figure 4. Point on the bottom row of the image will not be modified by the transformation. The height of the words (characters) is the same but the width of the image will probably change.

Figure 5(a) shows the original image with slanted word; Figure 5(b) shows the gradient orientation histogram; and Figure 5(c) shows the slant corrected image using this gradient technique. The output in Figure 5(c) looks more slanted than before (Figure 5(a)). However, according to Khairuddin [1], this type of slanted image can be much easier to segment by using his segmentation method.

3.0 EXPERIMENTAL RESULTS

The results of the algorithm explained earlier on some real images are given in this section. Samples of digitized Jawi text of size 640 × 480 pixels have been used for testing and scanned at 300 dpi. A set of 100 images has been used to test the algorithm. The experiments have confirmed the feasibility of the technique adopted in this paper. The performance of this technique is 78%. Figure 6 shows the image that has been corrected from Figure 1(a). The gradient is obtained by using a 5×5 Sobel filtering operation on all image points [11]. Then the gradient orientation at every point of the image can be obtained by using Equation (2). By taking a set of bins in the range of $90 - \phi_1$ and $90 + \phi_2$, the orientation histogram of the input image can therefore be obtained. No initial smoothing was applied to the original image before the gradient operation because the process of obtaining the histogram can cause canceling of the random noise contribution. Because of the quantization effect, the text letter stroke is mostly short zigzagged line segments. These line segments are usually in the vertical or horizontal condition. Therefore, in most of the gradient orientation histograms, the peak is at 45° and its multiples (such as 90°, 135°, 180°, 225°, and 315°). These distinct peaks have been removed by using a median filter over the histogram. Some other problems faced are some of the image pixels may be missing during slanting process (see Figure 5(c)) and creates too much noise. It must be noted that digital image coordinates, for example (k, l) are integers, whereas the arithmetic used in the algorithm is floating point arithmetic. Therefore, truncation or rounding operations are needed. These operations may force two image pixels, let say $a[i][j]$, and $b[i'][j']$ lying in adjacent position to be mapped on the same pixel $b[k][l]$. This fact creates pattern spikes in the sheared image, if the original image $a[i][j]$ is scanned and sheared on a pixel-by-pixel basis.

4.0 CONCLUSIONS

A simple and fast document image skew correction algorithm has been developed which employs only the gradient orientation histogram. This algorithm was originally implemented for Latin documents by Sun and Si [11]. Several experiments have been carried out. The results show that the gradient orientation technique is capable of obtaining the skew angle of a document image. The same algorithm can also be used for the detection of the slant angle of the letters in a text line. These slanted letters can be corrected using a shear operation (Equation (5)) to an upright position for the ease of recognition. Future work is needed to compare this technique with other techniques.

ACKNOWLEDGEMENT

We thank the reviewers for their valuable comments.

REFERENCES

[1] Khairuddin bin Omar, 2000. *Handwritten Jawi Text Recognition Using Multilevel Classifier. Ph.D. Thesis*. Department of Computer Science. Faculty of Computer Science and Information Technology. Universiti Putra Malaysia. Serdang.

[2] Baird, H. S., 1987. The Skew Angle of Printed Documents. *Proceeding of SPSE 40th Symp. Hybrid Imaging Systems*. Rochester, NY. 21-24.

[3] Spitz, A. L., 1992. Skew Determination in CCITT Group 4 Compressed Document Images. *Proc. Symp. Document Analysis and Information Retrieval*. Last Vegas, Nevada, USA. 11-25.

[4] Ciardiello, G., M. T. Degrandi, M. P. Roccotelli, G. Scafuro, and M. R. Spada. 1988. An Experimental System for Office Document Handling and Text Recognition. *Proceeding of International Conference on Pattern Recognition*. (2): 739-743.

[5] Nakano, Y., Y. Shima, H. Fujisawa, J. Higashino, and M. Fujinawa. 1990. An Algorithm for the Skew Normalization of Document Image. *Proceeding of International Conference on Pattern Recognition*. (2): 8-13.

[6] Srihari, S. N. and V. Govindaraju. 1989. Analysis of Textual Images using the Hough Transform. *Machine Vision Applications*, (2): 141-153.

[7] Hinds, S. S. J. L. Fisher, and D. P. D'Amato. 1990. A Document skew detection method using run-length encoding and the Hough Transform. *Proceeding of International Conference on Pattern Recognition*. (1): 464-468.

[8] Postl, W. 1986. Detection of Linear Oblique Structure and Skew Scan in Digitized Documents. *Proceeding of International Conference on Pattern Recognition*. 687-689.

[9] Hashizume, A., P. S. Yeh, and A. Rosenfeld. 1986. A Method of Detecting the Orientation of aligned Components. *Pattern Recognition Letters*. (4): 125-132.

[10] Yan, H., 1993. Skew Correction of Document Images using Interline Cross-Correlation. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*. 6(55): 538-543.

[11] Sun, C. and D. Si. 1997. Skew and Slant Correction for Document Images using Gradient Direction. *Proceeding of the 4th Inter. Conf. On Document Analysis and Recognition*, (1): 142-146.

[12] Pitas, I., 1993. *Digital Image Processing Algorithms*. New Jersey: Prentice Hall.

[13] Juzlinda binti Mohd Ghazali, Khairuddin bin Omar, Shahrul Azman bin Mohd Noh, and Md. Jan bin Nordin. 2002. "Analisis Tulisan Tangan Secara Automatik Memfokus Aspek Kecondongan". *Seminar Cabaran Alaf Baru Pengurusan Organisasi daripada Perspektif Teknologi Maklumat TEMPO2002*, Universiti Utara Malaysia.