Cursive word skew/slant corrections based on Radon transform

Jian-xiong Dong and Ponson Dominique
IMDS Software
75 rue Queen, Montréal, Québec, H3C 2N6
{jdong,dponson}@imds-world.com

Adam Krzyżak and Ching Y. Suen
CENPARMI, Concordia University
Montréal, Québec, Canada H3G 1M8
krzyzak@cs.concordia.ca, suen@cenparmi.concordia.ca

Abstract

This paper presents two fast and robust algorithms for word skew and slant corrections based on Radon transform. For the skew correction, we maximize a global measure which is defined by Radon transform of image and its gradient to estimate the slope. For the slant correction, Radon transform is used to estimate the long strokes and a word slant is measured by the average angle of these long strokes. Compared with the previous methods, these two algorithms do not require the setting of parameters heuristically. Moreover, the algorithms perform well on words of short length, where the traditional methods usually fail.

1. Introduction

In cursive word recognition, correcting the skew (deviation of the baseline from the horizontal direction–Fig. 1(a)) and the slant (deviation of average near-vertical strokes from the vertical direction –Fig. 1(b)) is an important preprocessing step. The slant and slope are introduced by writing styles. Both corrections can reduce handwritten word shape variability and help the later operations such as segmentation and feature extraction.

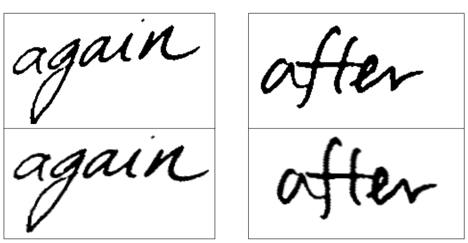

Figure 1: (a) The original handwritten word image and after skew correction, (b) the original word image and after slant correction.

For the skew and slant corrections, the crucial step is to detect the skew and slant angles correctly. In the literature, several methods have been proposed to deal with this problem. In [1], image contour is used to detect those minima. Morita et al. [2] proposed a method based on mathematical morphology to obtain a pseudo-convex hull image. Then minima are detected on the pseudo-convex image and a reference line is fit through those points. The primary challenge in these methods is the rejection of spurious minima. Also, the regression-based methods do not work well on short words because of the lack of a sufficient number of minima points. The other approaches for the detection of the slope angle is based on density distribution. In [3], several histograms are computed for different y projections. Then the entropy is calculated for each of them. The histogram with the lowest entropy will determine the slope angle. In [4], the Wigner-Ville distribution is calculated for several horizontal projection histograms. The slope angle is selected by Wigner-Ville distribution with the maximal intensity. The main problem for these distribution-based methods is high computational cost since an image has to be rotated for each angle. For the slant estimation, the most common method is the calculation of the average near-vertical strokes [5]. These methods use different criteria to select near-vertical strokes. The slopes of those selected strokes are estimated from the contours. The main disadvantage of these methods is that many heuristic parameters have to be specified. Vinciarelli et al. [6] proposed a technique based on a cost function which measures slant absence across the word image. The cost function is evaluated on multiple shear transformed word images. The angle with the maximal cost is taken as a slant estimate. Kavallieratou et al. [4] proposed a slant estimation algorithm based on the use of vertical projection profile of word images and the Wigner-Ville distribution. The approaches based on the optimization are relatively robust. However, the above two methods are computationally heavy since multiple shear transformed word images corresponding to different angles in an interval have to be calculated.

In this paper we propose a new algorithm for skew correction and a fast algorithm for slant correction. Both algorithms are based on Radon transform [7]. Compared with previous methods, they are not only robust but also computationally more efficient while the principle of computational efficiency usually is not taken into account in the

Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR'05)
1520-5263/05 $20.00 © 2005 IEEE

previous methods. In Section 2, a new algorithm for slope estimation is presented. In Section 3 we describe an efficient algorithm to calculate the average slant angle. The experimental results are shown in Section 4. Finally, some conclusions are drawn.

2. Slope estimation

Intuitively, it can be observed in Fig. 1 that a handwriting signal oscillates most frequently in the writing direction. When a handwritten signal is projected onto the axis orthogonal to the writing direction, its energy is also concentrated. For a binary word image[1], edge and foreground pixels characterize the shape variations and image energy, respectively. A global measure is designed to capture these two properties. The skew angle is selected, which corresponds to the maximal measure. First we introduce some notations. The 2D image is denoted as $f(x, y)$. The Radon transform [7] of f is the integral of $f(x, y)$ over a line in the xy-plane, which is illustrated in Fig. 2. Mathematically, the Radon transform of f is given by

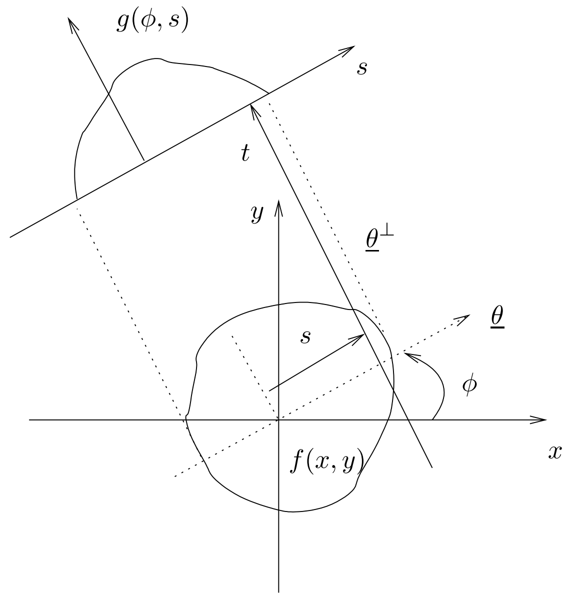

Figure 2: Illustration of parameters in the definition of Radon transform.

$$\begin{aligned} R\{f\} &\equiv g(\phi, s) \quad (2.1)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(s-(x\cos\phi + y\sin\phi))dxdy \end{aligned}$$

With the change of variables defined by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} \quad (2.2)$$

[1] A black/white image. The foreground is defined by black pixels.

(2.2) reduces to

$$\begin{aligned} g(\phi, s) &= \int_{-\infty}^{\infty} f(s\cos\phi - t\sin\phi, s\sin\phi + t\cos\phi)dt \\ &= \int_{-\infty}^{\infty} f(s\underline{\theta} + t\underline{\theta}^{\perp})dt, \quad (2.3) \end{aligned}$$

where $\phi \in [0, \pi)$, $s \in \mathbb{R}$, $\underline{\theta} = (\cos\phi, \sin\phi)$ and $\underline{\theta}^{\perp} = (-\sin\phi, \cos\phi)$. We define the Radon transform of gradient flow by

$$g'(\phi, s) = \int_{-\infty}^{\infty} |\,\mathrm{grad}\, f(s\underline{\theta} + t\underline{\theta}^{\perp}) \cdot \underline{\theta}^{\perp}|dt \quad (2.4)$$

This transform is depicted in Fig. 3. For a binary handwritten word image, the Radon transform of gradient flow is mainly used to estimate the average frequency in the projection direction.

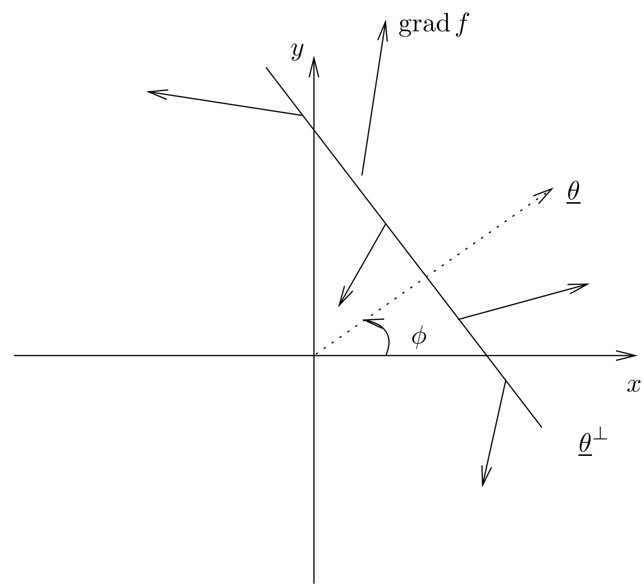

Figure 3: The Radon transform of gradient flow in $\mathbb{R}^2$. The gradient vector components parallel to the line of integration are used.

The angle ϕ^* can be obtained by maximizing the following function:

$$\phi^* = \arg\max_{\phi} \int_{-\infty}^{\infty} g(\phi, s)g'(\phi, s)ds \quad (2.5)$$

Finally, when a handwritten word image is deskewed, $\underline{\theta}^{*\perp}$ should correspond to the x axis. Therefore, the skew angle is $\frac{\pi}{2} + \phi^*$. In order to solve the optimization in (2.5), we propose an efficient algorithm. First let us define some notations. The image width and height of image $f(x, y)$ are denoted by w and h, respectively. $f(x, y)$ is a black/white(1/0) image. We map the input image from the Cartesian domain to the polar domain, where $x - \frac{w}{2} = r\cos\alpha$ and $y - \frac{h}{2} = r\sin\alpha$, $\alpha \in [0, 2\pi)$ and $r \le r_{\max}$. In order to reduce the computational cost, we use zero-crossing to replace $|\,\mathrm{grad}\, f(s\underline{\theta} + t\underline{\theta}^{\perp}) \cdot \underline{\theta}^{\perp}|$ in Eq.(2.4). In the binary image, zero-crossing occurs at the edge point. Then $g'(\phi, s)$ represents the number of crossing points parameterized by ϕ and s. Fig. 4 illustrates Radon transformation in the polar domain.

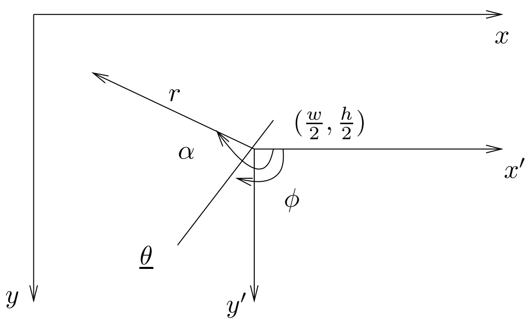

Figure 4: Radon transformation in the polar domain.

When the foreground pixel (r, α) is projected onto the unit vector $\underline{\theta}$, its value is given by

$$(r\cos\alpha, r\sin\alpha) \cdot (\cos\phi, \sin\phi) = r\cos(\alpha - \phi) \quad (2.6)$$

Now we propose the following efficient algorithm to detect the skew angle.

Fast algorithm for skew angle detection

Input: Input image $f(x, y)$, $K = \lceil r_{\max} \rceil$. matrix $A[K][360]$, matrix $B[2K][2K]$, vector $\mathrm{v}_1[2K]$, vector $\mathrm{v}_2[2K]$ and vector $\mathrm{v}_3[180]$. The foreground points $\{(x_i^f, y_i^f)\}$, $i = 1, \cdots, M$[2]; the edge points $\{(x_j^e, y_j^e)\}$, $j = 1, \cdots, N$.

Output: skew angle

Initialization: Pre-compute the caching tables

1.1 For i =1 to K step 1
 For j =1 to 360 step 1
  $A[i][j] = \lfloor i \times \cos(\frac{\pi}{180} \times j) + K \rfloor$
 End j
End i

1.2 For $j = 1$ to $2K$ step 1
 For i =1 to $2K$ step 1
  $l = \lfloor \sqrt{(i-K)^2 + (j-K)^2} \rfloor$
  $m = \lfloor (\mathrm{atan2}(j-K, i-K) + \pi) * \frac{180}{\pi} \rfloor$
  $B[j][i].r = l$
  $B[j][i].\alpha = m$
 End i
End j

[2]Note that the origin of the coordinate system is $(\frac{w}{2}, \frac{h}{2})$.

Radon transform in the polar domain

2.1 For $j = 1$ to 180 step 1
 $\mathrm{v}_1[i] \leftarrow 0$, $\mathrm{v}_2[i] \leftarrow 0$, $i = 1, \cdots, 2K$.
 For $i = 1$ to M step 1
  $x \leftarrow K + x_i^f$
  $y \leftarrow K + y_i^f$
  $r = B[y][x].r$
  $\alpha = B[y][x].\alpha$
  $k = A[r][(360 + \alpha - j) \bmod 360]$
  $\mathrm{v}_1[k] \leftarrow \mathrm{v}_1[k] + 1$
 End i
 For $i = 1$ to N step 1
  $x \leftarrow K + x_i^e$
  $y \leftarrow K + y_i^e$
  $r = B[y][x].r$
  $\alpha = B[y][x].\alpha$
  $k = A[r][(360 + \alpha - j) \bmod 360]$
  $\mathrm{v}_2[k] \leftarrow \mathrm{v}_2[k] + 1$
 End i
 Smooth v_1 and v_2 using moving window of size 5.
 $\mathrm{v}_3[j] \leftarrow 0$
 $\mathrm{v}_3[j] \leftarrow \mathrm{v}_3[j] + \mathrm{v}_1[i] \times \mathrm{v}_2[i]$, $i = 1, \cdots, 2K$
End j

Calculate maximal value

3.1 $i^* = arg\max_i \mathrm{v}_3[i]$, $i = 1, \cdots, 180$.

3.2 return $i^* + 90$.

In algorithm (2), most float operations are replaced with integer ones. Although the precision of Radon transforms is reduced, the detection accuracy of skew angle is not decreased. The reason is that in step (3.1) we need to obtain an angle which corresponds to the maximum response. High-precision Radon transform is not necessary in our algorithm. In addition, the most expensive operations can be pre-computed as look-up tables at the initial stage and the number of evaluated angles can be much smaller than 180.

3. Fast algorithm for slant estimation

In our algorithm, the basic idea is to estimate the average angular slant of long vertical strokes. Geometric justification is provided in Fig. 5. In this figure, line segment AB is a part of a long vertical stroke; CD and EF are parts of two short vertical strokes, respectively. In order to classify these two cases, we project AB, CD,EF onto the direction $\underline{\theta}^{\perp}$. Then we calculate the ratio of the projection length to the maximal range for a fixed s in (2.2), defined by

$$r(\phi, s) = \frac{g(\phi, s)U(g(\phi, s) - \tau_1)}{\sup_t\{t|f(s\underline{\theta} + t\underline{\theta}^{\perp}) > 0\} - \inf_t\{t|f(s\underline{\theta} + t\underline{\theta}^{\perp}) > 0\}} \quad (3.1)$$

where τ_1 is a tolerance parameter and sup and inf are "least upper bound" and "greatest lower bound" of a bounded non-empty set, and the function $U(.)$ is given by

$$U(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

If this ratio r is large, then probably there exists a long segment for a fixed s on the current projection direction. For example, for segment AB, the ratio is 1.0; for segments CD and EF, the ratio is (CD+EF)/CF. Let $I(\phi, s)$ be an indicator function, defined by

$$I(\phi, s) = \begin{cases} 1 & \text{if } r(\phi, s) > \tau_2 \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

where τ_2 is a tolerance parameter. Then the following formula is used to approximate the slant angle:

$$\phi^* = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(\phi, s) g(\phi, s) \phi d\phi ds \quad (3.4)$$

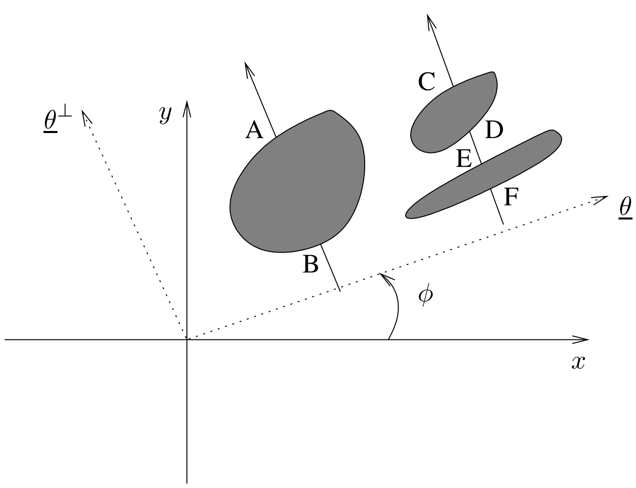

Figure 5: Illustration of slant angle detection.

Fast algorithm for slant angle detection

Input: deskewed image $f(x, y)$, $K = \lceil r_{\max} \rceil$. matrix $A[K][360]$, matrix $B[2K][2K]$, vector $v_1[2K]$, vector $v_2[2K]$, vector min_v$[2k]$ and vector max_v$[2k]$, and vector $v_3[180]$. The foreground points $\{(x_i^f, y_i^f)\}$, $i = 1, \cdots, M$[3]; the edge points $\{(x_j^e, y_j^e)\}$, $j = 1, \cdots, N$.

Output: Slant angle

Initialization: Pre-compute the caching tables

[3]Note that the origin of the coordinate system is $(\frac{w}{2}, \frac{h}{2})$.

1.1 For i =1 to K step 1
    For j =1 to 360 step 1
        $A[i][j] = \lfloor i \times \cos(\frac{\pi}{180} \times j) + K \rfloor$
    End j
End i

1.2 For $j = 1$ to $2K$ step 1
    For i =1 to $2K$ step 1
        $l = \lfloor \sqrt{(i-K)^2 + (j-K)^2} \rfloor$
        $m = \lfloor (\text{atan2}(j-K, i-K) + \pi) * \frac{180}{\pi} \rfloor$
        $B[j][i].r = l$
        $B[j][i].\alpha = m$
    End i
End j

Radon transform in the polar domain

2.1 totalSum← 0; AccumulatedAngle← 0

2.2 For $j = 1$ to 180 step 1
    $v_1[i] \leftarrow 0, v_2[i] \leftarrow 0, i = 1, \cdots, 2K$.
    For $i = 1$ to M step 1
        $x \leftarrow K + x_i^f$
        $y \leftarrow K + y_i^f$
        $r = B[y][x].r$
        $\alpha = B[y][x].\alpha$
        $k = A[r][(360 + \alpha - j) \text{ mod } 360]$
        $v_1[k] \leftarrow v_1[k] + 1$
    End i
    min_v$[i] \leftarrow 0$, max_v$[i] \leftarrow 0, i = 1, \cdots, 2K$.
    For $i = 1$ to N step 1
        $x \leftarrow K + x_i^e$
        $y \leftarrow K + y_i^e$
        $r = B[y][x].r$
        $\alpha = B[y][x].\alpha$
        $k_1 = A[r][(360 + \alpha - j) \text{ mod } 360]$
        $k_2 = A[r][(450 + j - \alpha) \text{ mod } 360]$
        if $k_2 <$ min_v$[k_1]$ then min_v$[k_1] \leftarrow k_2$
        if $k_2 >$ max_v$[k_1]$ then max_v$[k_1] \leftarrow k_2$
    End i
    $v_2[i] \leftarrow$ max_v$[i]$ − min_v$[i], i = 1, \cdots, 2K$.
    Smooth v_1 using moving window of size 5.
    sum← 0
    For $i = 1$ to $2K$ step 1
        if $v_1[i] > t_1 \times v_2[i]$ and $v_2[i] > t_2$ then
            sum ← sum + $v_1[i]$
        End
    End i
    totalSum←totalSum + sum;
    AccumulatedAngle←AccumulatedAngle * j
End j

Calculate average slant angle

3.1 return 90 - AccumulatedAngles/totalSum.

Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR'05)
1520-5263/05 $20.00 © 2005 IEEE

4 Experimental results

The experiments were conducted on IMDS cursive word database. Presently the vocabulary comes from the words of Collins Frequency Band 5, in which these words are most frequently used in daily life. The size of this lexicon is 670. Our samples are written by 78 people of different backgrounds including Arabic, Asian, Canadian, French, students and university professors, and company employees. No constraints are imposed on the writers in order to get most natural handwritten samples. Each writer writes samples in blank boxes on the form containing 670 words. This means there are no two samples from the same writer for each word. The training and testing sets consist of 38795 and 13733 samples, respectively.

In order to investigate how the skew and slant correction affect the performance, we consider the recognition task in two cases: with or without skew/slant corrections in the phase of preprocessing. The experimental results are shown in Table 1. It can be observed that skew and slant corrections can improve the accuracy of cursive word recognition system. We also visually inspected skew/slant corrections in all images in the training set and have found that the proposed methods perform very well, even on short words. Fig. 6 shows some examples.

Table 1: Recognition accuracy with/without skew and slant correction.

Skew and slant correction	Test error rate
Yes	44%
No	47%

(a)
(b)
(c)

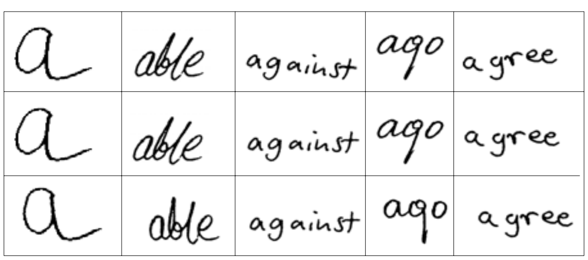

Figure 6: Some examples of skew & slant corrections. (a) Original images, (b) after skew correction, (c) after skew and slant corrections.

5. Conclusions

This paper presents two new algorithms for cursive word skew and slant corrections based on Radon transform. The two algorithms are not only robust but also much more efficient than the previous ones. In both algorithms no heuristic parameters require to be set. Moreover, they perform well on short words, where the traditional methods usually fails. Experimental results on IMDS cursive word database have shown that the proposed skew and slant algorithms can enhance the recognition accuracy.

Acknowledgments

This research has been support by Industrial Research Fellowship of Natural Sciences and Engineering Research Council of Canada and IMDS Software.

References

[1] S. Madhvanath, G. Kim, and V. Govindaraju, “Chaincode contour processing for handwritten word recognition,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, no. 9, pp. 928–932, 1999.

[2] M. Morita, J. Facon, F. Bortolozzi, S. Garnés, and R. Sabourin, “Mathematical morphology and weighted least squares to correct handwriting baseline skew,” in *Proc. IEEE International Conference on Document Analysis and Recognition*, Bangalore, India, Sept. 1999, pp. 430–433.

[3] M. Côté, E. Lecolinet, M. Cheriet, and C. Suen, “Automatic reading of cursive scripts using a reading model and perceptual concepts,” *International Journal on Document Analysis and Recognition*, vol. 1, no. 1, pp. 3–17, 1998.

[4] E. Kavallieratou, N. Fakotakis, and G. Kokkinakis, “New algorithms for skewing correction and slant removal on word-level,” in *Proc. IEEE 6th International Conference on Electronics, Circuits and Systems*, Pafos, Cyprus, Sept. 1999, pp. 1159–1162.

[5] R. Bozinovic and S. Srihari, “Off-line cursive script word recognition,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 8, pp. 63–84, 1989.

[6] A. Vinciarelli and J. Luettin, “A new normalization technique for cursive handwritten words,” *Pattern Recognition Letters*, vol. 22, no. 9, pp. 1043–1050, 2001.

[7] J. Radon, “Über die bestimmungen von funktionen durch ihre integralwerte längs gewisser mannigfaltkeiten,” *Ber. Vehr. Sächs. Akad. Wiss.*, vol. 69, pp. 262–277, 1917.