Equivalence of Different Methods for Slant and Skew Corrections in Word Recognition Applications

Petr Slavík
and Venu Govindaraju, *Senior Member, IEEE*

Abstract—Normalization of slant and skew is often used in processing a word image before recognition. In this paper, we prove the theoretical equivalence of different methods for slant and skew corrections. In particular, we show that correcting first for skew by rotation and then for slant by a shear transformation in the horizontal direction is equivalent to first correcting for slant by a shear transformation in the horizontal direction and then for skew by a shear transformation in the vertical direction. Our proof can be easily modified to prove equivalence of other methods for correcting the slant and skew.

Index Terms—Image preprocessing, slant normalization, skew normalization, handwriting recognition.

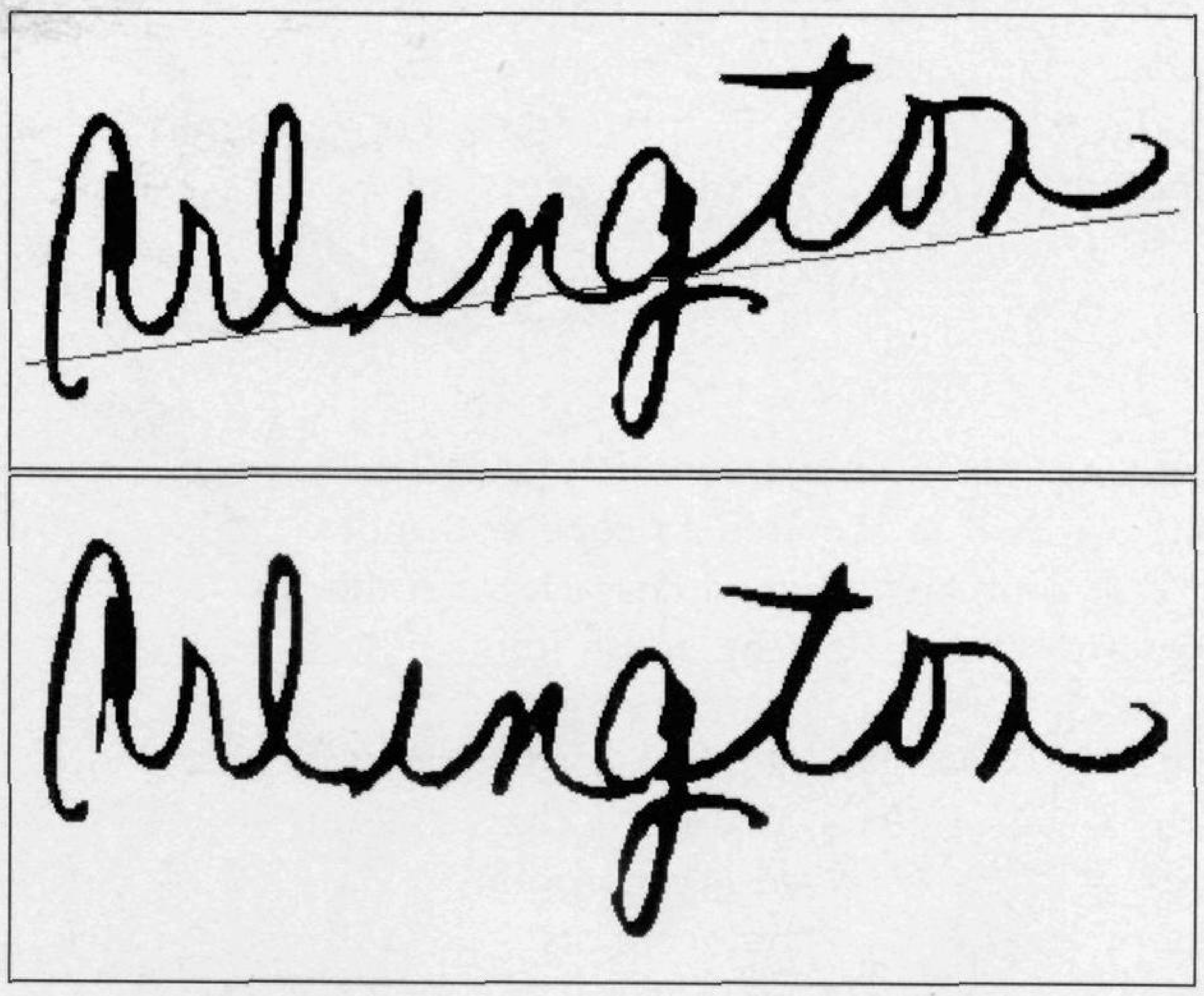

Fig. 1. Handwritten word image before and after skew correction.

1 Introduction

In most handwritten word applications, correcting the skew (deviation of the baseline from the horizontal direction—Fig. 1) and the slant (deviation of average near-vertical strokes from the vertical direction—Fig. 2) is an important preprocessing step. Typically, first the skew angle and then the slant angle are corrected [1], [3], [7]. In some cases, however, it might be more convenient to first correct the slant and then the skew [8]. In this paper, we prove the theoretical equivalence of these approaches showing that the order and the method in which the slant and skew angles are corrected is not important for word recognition applications.

Specifically, we consider the following two methods of correcting both the slant and skew angles: 1) the first standard method corrects the skew angle by rotation and then corrects the slant angle by a shear transformation in the horizontal direction [1], [3] 2) the second method first corrects the slant angle by shear-transforming the entire image in the horizontal direction and then corrects the skew angle by shear-transforming the entire image in the vertical direction [8].

Assuming that the pixel coordinates are real numbers, we will prove that the two images produced 1) by rotation followed by a shear transformation in x-direction and 2) by a shear transformation in x-direction followed by a shear transformation in y-direction differ only in their aspect ratios. And if reasonable limits are imposed on the possible values of skew and slant angles, the changes in the aspect ratio are negligible. For example, if the skew angle is 10^o and the slant angle is 30^o, the two aspect ratios would differ by only a multiplicative factor of 0.85.

Rounding of coordinates will ultimately produce slightly different images. This, however, should not influence a word recognizer in any significant way. High-level features [1], [2] (loops, ascenders, descenders, etc.) will remain unchanged while low-level features [4], [5], [6], [9] (various moments, distribution of slopes, or pixels in subregions of each segment, etc.) will change only slightly.

2 Mathematical Formulation

To justify the above claims, let us consider a general image of a handwritten word where the deviation of the baseline from the horizontal axis is α and the deviation of the average direction of near-vertical strokes from the line perpendicular to the baseline is β—see the top image in Fig. 3. Assume that the baseline intersects the origin of the coordinate system. This assumption will only simplify our formulas and has certainly no effect on the generality of our results.

2.1 First Method

Images in Fig. 3 show the individual steps of the first (traditional) method of correcting first the skew angle by rotation and then the slant angle by shear transformation in the horizontal direction. Given a point (x, y), its new coordinates (x', y') after rotating the entire image around the origin by angle $-\alpha$ (negative α) satisfy

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

After shear-transforming the point in the horizontal direction by angle $-\beta$ (negative β), its new coordinates (x'', y'') will be

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 1 & -\tan\beta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix},$$

where

$$T = \begin{pmatrix} \cos\alpha + \sin\alpha\tan\beta & sin\alpha - \cos\alpha\tan\beta \\ -\sin\alpha & \cos\alpha \end{pmatrix} \qquad (1)$$

is the matrix of the composition of rotation and shear transformation.

Define $\gamma = \beta - \alpha$; that is, γ is the deviation of the average near-vertical stroke from vertical direction—see Fig. 4. Using the fact that

$$\begin{aligned} \tan\beta - \tan\alpha &= (1 + \tan\alpha\tan\beta)\tan(\beta - \alpha) \\ &= (1 + \tan\alpha\tan\beta)\tan\gamma, \end{aligned}$$

- *The authors are with the Center of Excellence for Document Analysis and Recognition (CEDAR), Department of Computer Science and Engineering, State Unversity of New York at Buffalo, 226 Bell Hall, Amherst, NY 14260. E-mail: {slavik, govind}@cedar.buffalo.edu.*

Manuscript received 2 Sept. 1999; revised 7 Sept. 2000; accepted 28 Sept. 2000.
Recommended for acceptance by D. Dori.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 110536.