

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.

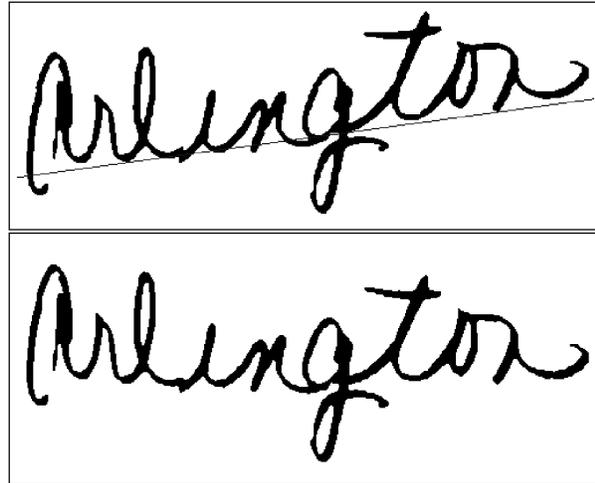


Fig. 1. Handwritten word image before and after skew correction.

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° and the slant angle is 30° , 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

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

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 University 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.

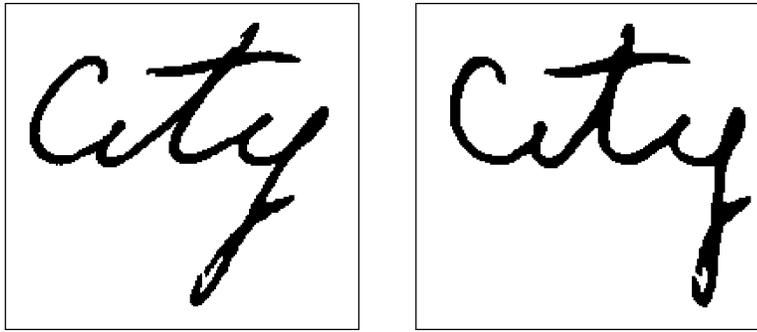


Fig. 2. Handwritten word image before and after slant correction.

we can rewrite (1) as

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

$$= \cos \alpha \begin{pmatrix} 1 + \tan \alpha \tan \beta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\tan \gamma \\ -\tan \alpha & 1 \end{pmatrix} = D_T \cdot R,$$

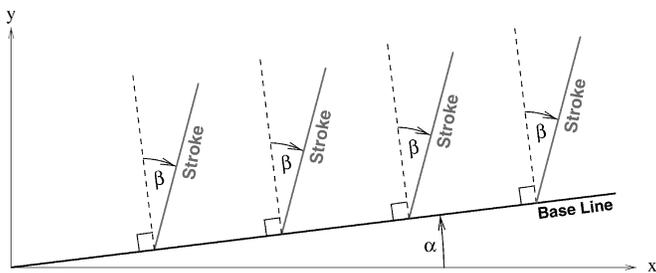
where

$$R = \begin{pmatrix} 1 & -\tan \gamma \\ -\tan \alpha & 1 \end{pmatrix} \quad (3)$$

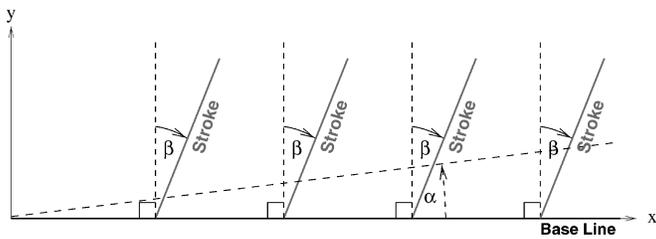
is a matrix for simultaneously shear-transforming in the horizontal direction by angle $-\gamma$ and in the vertical direction by angle $-\alpha$, and

$$D_T = \begin{pmatrix} \cos \alpha + \sin \alpha \tan \beta & 0 \\ 0 & \cos \alpha \end{pmatrix}$$

(2) is a simple diagonal matrix that squeezes the image in the y -direction and stretches (or squeezes—depending on the value of $\cos \alpha + \sin \alpha \tan \beta$) the image in the x -direction. It is to be noted that the determinant of T is equal to 1, hence, the area of the image is preserved.



↓ Rotate by angle $-\alpha$



↓ Shear in the x -direction by angle $-\beta$

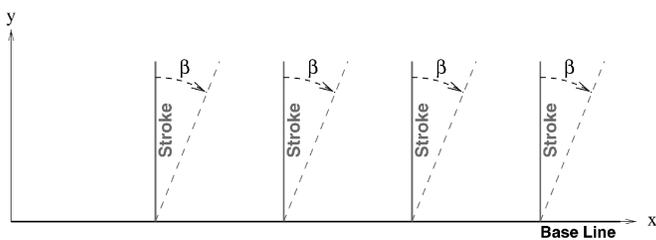
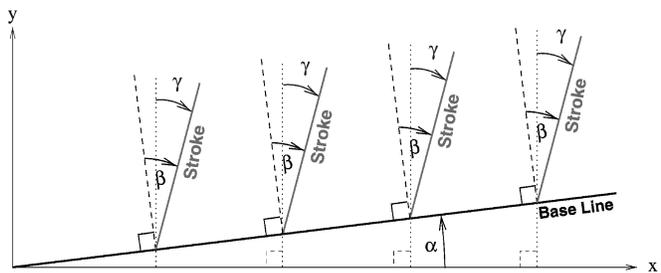
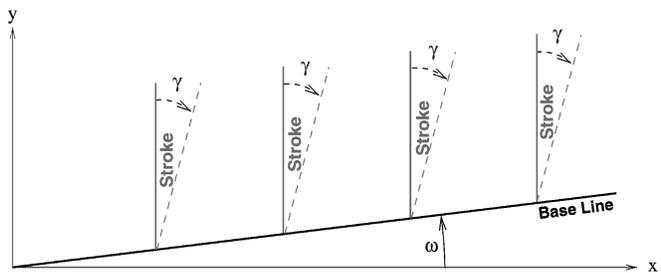


Fig. 3. First method of correcting the skew and slant angles of a handwritten word image.



↓ Shear in the x -direction by angle $-\gamma$



↓ Shear in the y -direction by angle $-\omega$

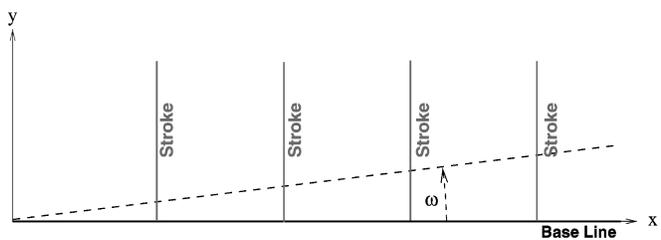


Fig. 4. Second method of correcting the skew and slant angles of a handwritten word image.

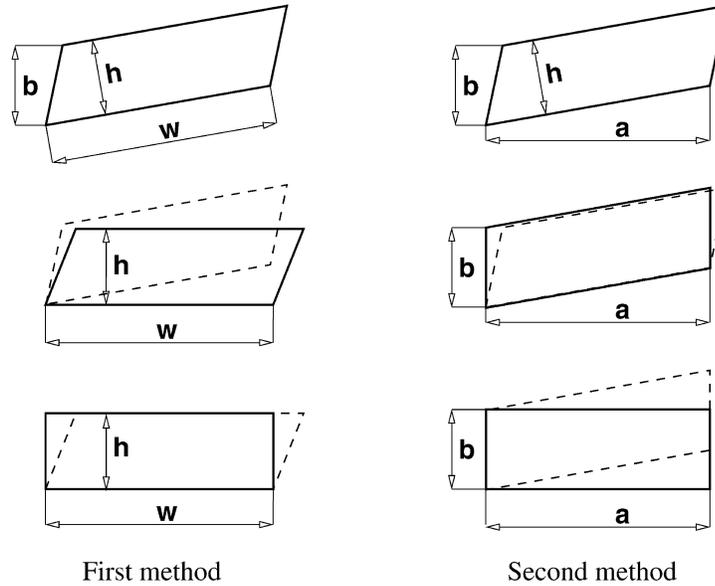


Fig. 5. Comparison of factors preserved by first method (rotate to correct skew then shear-transform to correct slant) and the second method (shear-transform to correct slant then shear-transform to correct skew). Note that $w * h = a * b$, hence the area of the “bounding box” is preserved.

2.2 Second Method

The second approach to slant and skew correction is shown in Fig. 4. In order to remove the slant angle, one first shear-transforms all the points in horizontal direction by angle $-\gamma$. This way, the average near-vertical strokes become vertical. The new coordinates (x', y') of a point (x, y) then become

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & -\tan \gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

We denote the skew angle after the shear transformation by ω . One can quickly realize that, in general, $\omega \neq \alpha$. By using the inverse of the above transformation, namely, $x = x' + y' \tan \gamma$, $y = y'$, and the baseline equation $y = x \tan \alpha$, one can show that the shear transformed baseline has an equation

$$y' = \frac{\tan \alpha}{1 - \tan \alpha \tan \gamma} x'.$$

Therefore,

$$\omega = \arctan \left(\frac{\tan \alpha}{1 - \tan \alpha \tan \gamma} \right).$$

It is clear from the discussion above that in order to correct the skew angle, one now has to shear-transform in the vertical direction by angle $-\omega$. The equation for this transformation is

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\tan \omega & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Combining the two transformations gives

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\tan \omega & 1 \end{pmatrix} \begin{pmatrix} 1 & -\tan \gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = N \begin{pmatrix} x \\ y \end{pmatrix}.$$

Here,

$$\begin{aligned} N &= \begin{pmatrix} 1 & -\tan \gamma \\ -\frac{\tan \alpha}{1 - \tan \alpha \tan \gamma} & 1 + \frac{\tan \alpha \tan \gamma}{1 - \tan \alpha \tan \gamma} \end{pmatrix} \\ &= \begin{pmatrix} 1 & -\tan \gamma \\ -\frac{\tan \alpha}{1 - \tan \alpha \tan \gamma} & \frac{1}{1 - \tan \alpha \tan \gamma} \end{pmatrix} = D_N \cdot R, \end{aligned}$$

where R was previously defined in (3) and

$$D_N = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{1 - \tan \alpha \tan \gamma} \end{pmatrix}$$

is a diagonal matrix that essentially stretches (or squeezes—depending on the signs of α and γ) the entire image in the y direction. Since the determinant of the entire transformation is again 1, the area of the image remains unchanged.

2.3 Comparing the Methods

Equations (2) and (4) clearly show the similarity between transformations T and N . The transformations differ only by the diagonal matrices D_T and D_N whose only effect is to squeeze (or stretch) the entire image in the x and y directions.

Let us now be more specific about the difference between transformations T and N . For simplicity, let us call “traditional” the image that would result from transforming the original image in the traditional way—by first rotating it and then shear-transforming it in the horizontal direction. Let “new” refer to the image resulting from transforming the original image by shear-transforming it first in horizontal and then in vertical direction.

Given the “new” image, one can easily obtain the “traditional” image by composing the transformations N^{-1} and T . Let us call this transformation S . Thus, we have

$$\begin{aligned} S &= N^{-1}T = D_T R R^{-1} D_N^{-1} = D_T D_N^{-1} \\ &= \begin{pmatrix} \cos \alpha + \sin \alpha \tan \beta & 0 \\ 0 & \cos \alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 - \tan \alpha \tan \gamma \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha + \sin \alpha \tan \beta & 0 \\ 0 & \cos \alpha - \sin \alpha \tan \gamma \end{pmatrix}. \end{aligned}$$

(4) Since S is a product of two matrices (N^{-1} and T) with unit determinants, the determinant of S has to equal 1, hence, $(\cos \alpha + \sin \alpha \tan \beta)(\cos \alpha - \sin \alpha \tan \gamma) = 1$. Therefore, to transform the “new” image to the “traditional” image, one has to multiply the x -coordinates of all points by $\cos \alpha + \sin \alpha \tan \beta$ and divide the

y -coordinates of all points by the same value. In particular, if $\alpha = 10^\circ$ and $\beta = 30^\circ$, one would have to multiply the x -coordinates by 1.085 and the y -coordinates by 0.922 while the area of the image would remain the same and the aspect ratio would increase by a multiplicative factor of 1.177. Fig. 5 shows what is and is not preserved by each transformation.

3 CONCLUSION

One could similarly show that shear-transforming first in the vertical direction and then in the horizontal direction [7] or shear-transforming simultaneously in both horizontal and vertical directions by appropriate angles are equivalent in the above sense to transformations T and N . However, when one shear-transforms simultaneously in both horizontal and vertical directions, the image area in general is not preserved.

REFERENCES

- [1] M. Brown and S. Ganapathy, "Preprocessing Techniques for Cursive Word Recognition," *Pattern Recognition*, vol. 16, no. 5, pp. 447-458, 1983.
- [2] P. Gader, M. Whalen, M. Ganzberger, and D. Hepp, "Handprinted Word Recognition on a NIST Data Set," *Machine Vision and Applications*, vol. 8, pp. 31-40, 1995.
- [3] W. Guerfali and R. Plamondon, "Normalizing and Restoring Online Handwriting," *Pattern Recognition*, vol. 26, no. 3, pp. 419-431, 1993.
- [4] G. Kim and V. Govindaraju, "Efficient Chain Code-Based Image Manipulation for Handwritten Word Recognition," *Proc. SPIE Symp. Electronic Imaging Science and Technology*, pp. 262-272, 1996.
- [5] G. Kim and V. Govindaraju, "A Lexicon Driven Approach to Handwritten Word Recognition for Real-Time Applications," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 4, pp. 366-379, Apr. 1997.
- [6] F. Kimura, M. Sridhar, and Z. Chen, "Improvements of a Lexicon Directed Algorithm for Recognition of Unconstrained Handwritten Words," *Proc. Second Int'l Conf. Document Analysis and Recognition*, pp. 18-22, 1993.
- [7] A. Senior, "Off-Line Handwriting Recognition: A Review and Experiments," Technical Report CUED/F-INFENG/TR 105, Eng. Dept., Cambridge Univ., Cambridge, England, Dec. 1992. available via FTP or from <http://espresso.ee.sun.ac.za/~schwardt/papers/pimage.html>.
- [8] P. Slavík and V. Govindaraju, "An Overview of Run-Length Encoding of Handwritten Word Images," Technical Report 2000-09, Dept. of Computer Science, SUNY at Buffalo, Aug. 1999. <http://www.cse.buffalo.edu/tech-reports/>.
- [9] Ø.D. Trier, A.K. Jain, and T. Taxt, "Feature Extraction Methods for Character Recognition—A Survey," *Pattern Recognition*, vol. 29, no. 4, pp. 641-662, 1996.