스마트폰에서 문서 영상의 기울어짐 및 원근왜곡 보정 기법

Williem*, Christian Simon*, In Kyu Park*, and Sungdae Cho**
*School of Information and Communication Engineering, Inha University, **Samsung Electronics
{williem_060689@inha.edu, cssen@inha.edu, pik@inha.ac.kr, s-d.cho@samsung.com}

Abstract

This paper presents a perspective rectification framework for mobile device that is fast and robust to recovering the fronto-parallel view of perspectively distorted document images. The conventional approaches suffer from high computational complexity which is too heavy to be implemented on a mobile device. To ameliorate such problem, the proposed framework is designed to pursue a fast and robust algorithm to detect horizontal and vertical vanishing points efficiently and robustly. While Hough line detection and skew projective histogram algorithm are utilized to detect the horizontal vanishing point, a paragraph boundary detection and perspective projective histogram are used to detect the vertical vanishing point. Finally, perspective rectification is performed using both horizontal and vertical vanishing points. We demonstrate the performance of the proposed framework on various challenging examples to confirm that the proposed system is fast and robust in rectifying the perspectively distorted images.

1. Introduction

During the last decade, mobile camera is used to replace the usage of flatbed scanner to recapture a document. Optical Character Recognition (OCR) has become an important technology to read characters in a document image captured from mobile device. However, existing OCR application do not work correctly when the input image has serious distortion which destroys the text structure.

In this paper, we proposed a perspective rectification framework to recover the text structure so that the readability of OCR algorithm could be improved significantly. The proposed framework is implemented and tested on an off-the-shelf smartphone. To the best knowledge of the authors, the proposed framework is the first autonomous perspective rectification framework which is implemented on a mobile device. The experimental results show that the proposed framework can obtain highly accurate results with fast running time on the smartphone. The OCR readability is improved significantly.

This paper is organized as follows. Section 2 describes the proposed perspective rectification framework in detail. The experimental result is presented in Section 3. Finally, we give a conclusive remark in Section 4.

2. Proposed Algorithm

The main idea of the proposed algorithm is to send the horizontal and vertical vanishing points to infinity, which is similar to the conventional approaches. However, the novel part of the proposed approach is the way to find the horizontal and vertical vanishing points in fast and robust way. The proposed approach consists of the preprocessing step, skew rectification step, and perspective rectification step. In preprocessing step, the input image is converted to the downsampled binary text image. Then, the horizontal vanishing point is detected and sent to infinity to rectify the skew distortion. Finally, the vertical vanishing point is also detected and sent to infinity to rectify the perspective distortion. The algorithm sequence is illustrated in Fig. 1.

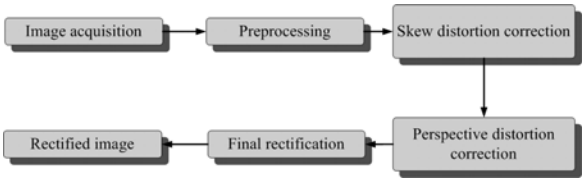

Fig 1. Pipeline of perspective rectification framework

2.1 Preprocessing

To achieve fast computational speed, the input image is downscaled so that proposed algorithm is run on the normalized image with small resolution. Instead of working on color space, the proposed framework utilizes the grayscale image for the algorithm development. In the binarization step, both adaptive thresholding method [1] and the text selection algorithm [2] are used to binarize text regions only.

2.2 Skew Rectification

To rectify the skew distortion, the proposed framework extracts the candidate horizontal lines using Hough line detection algorithm [3]. Then, all intersection points from all candidate lines are computed. All intersection points are clustered using K-means clustering algorithm and then each cluster is considered as a candidate of horizontal

vanishing point. Given a candidate, skew projective histogram is computed to find the best horizontal vanishing point. The detail of how to compute the projective histogram is described in [4] for the perspective rectification. To compute the best candidate of the horizontal vanishing point, the number of peaks in the derivative histogram larger than a threshold. Then, the skew rectification is performed by warping the input image using homography matrix which is modelled by manipulating the detected horizontal vanishing point.

2.3 Perspective Rectification

The idea of the vertical vanishing point detection is to extract the reliable paragraphs and their left and right boundaries. To detect the paragraph, paragraph blob detection is performed by employing the morphological dilation and the connected component analysis. Each paragraph blob is considered as candidate paragraph. For each detected paragraph, the left and right vertical boundaries are detected by computing the first and the last binary pixels on the top and bottom part of the paragraph. Several criteria are utilized to reject unreliable boundaries, such as too short lines and lines near image boundary. After the boundary detection, the intersection point of each boundary pair is computed, such that the intersection point is considered as the candidate for vertical vanishing point. For each candidate, the perspective projective histogram is calculated and the maximum of histogram difference is computed. To find the best candidate of vertical vanishing point, the point with maximum number of peaks in the derivative histogram value is selected as vertical vanishing point. Using the detected horizontal and vertical vanishing points, perspective rectification is performed to restore the fronto-parallel view of distorted image. The extended homography matrix for perspective rectification is modelled using horizontal and vanishing points information.

3. Experimental Results

The proposed algorithm is implemented and tested on an off-the-shelf smartphone (Samsung Galaxy S4) which has an embedded OCR application called *Optical Reader* on it. The robustness of the proposed framework can be viewed from rectified images, which are shown in Fig 2. For every pair of the input and the rectified images, we measure the readability improvement of the OCR algorithm by utilizing *Optical Reader* program. It is shown that the proposed method increases the readability of the OCR algorithm significantly. The average computational time for the whole process (RGB to grayscale conversion to the final perspective transformation) is around 700~800 ms on the mobile hardware (using CPU only). Note that it is faster than Yin's algorithm [5] which takes more than one second for CPU computation. Based on the experimental results, the proposed framework is proven to obtain fast and accurate performance to rectify perspective distorted documents.

(a) (b) (c) (d)

Fig 2. Results of perspective rectification. (a) Perspectively distorted input document images; (b) OCR results of the (b); (c) Perspective rectified images; (d) OCR results of (c).

4. Conclusion

In this paper, a fast and robust method to rectify image with geometric distortion was proposed. While Hough line detection and skew projective histogram were utilized to rectify the skew distortion, adaptive paragraph boundary detection and perspective histogram were used to rectify the perspective distortion. The experimental results proved that the proposed method recovered the fronto-parallel view from the perspective distorted document image. The proposed framework was implemented and tested on the off-the-shelf smartphone.

Acknowledgment

This work was supported by Samsung Electronics and Inha University.

References

[1] D. Bradley and G. Roth, “Adaptive thresholding using the integral image,” Journal of Graphics Tools, vol. 12, no. 2, pp. 13-21, 2007.

[2] P. Clark and M. Mirmehdi, “Finding text regions using localized statistical measures,” Proc. of BMVC, pp. 675-684, 2000.

[3] R. O. Duda and P. E. Hart, “Use of the Hough transformation to detect lines and curves in pictures,” Communications of the ACM, vol. 15, no. 1, pp. 11-15, January 1972.

[4] P. Clark and M. Mirmehdi, “Rectifying perspective views of text in 3D scenes using vanishing points,” Pattern Recognition, vol. 36, no. 11, pp. 2673-2686, November 2003.

[5] X. C. Yin, H. W. Hao, J. Sun and S. Naoi, “Robust vanishing point detection for mobilecam-based documents,” Proc. of ICDAR, pp. 136-40, 2011.