Fast and Robust Perspective Rectification of Document Images on a Smartphone

Williem† williem_060689@inha.edu
Christian Simon† cssen@inha.edu
Sungdae Cho‡ s-d.cho@samsung.com
In Kyu Park† pik@inha.ac.kr

†Dept. of Information and Communication Eng., Inha University, Incheon 402-751, Korea
‡Mobile Communication Division, Samsung Electronics Co. Ltd., Suwon 443-742, Korea

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 is too heavy to be implemented on a mobile device. In addition, they fails to reject the false case of the perspective rectification. To ameliorate such problems, the proposed framework is designed to pursue a fast and robust algorithm to detect horizontal and vertical vanishing points efficiently and robustly. Then, perspective rectification is performed using both horizontal and vertical vanishing points. In addition, the proposed framework has an adaptive scheme to detect the false case of the perspective rectification and skip the procedure without using the vertical vanishing point. Note that, the proposed framework is designed for consumer application so that bad results are rejected before they are shown for users. 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, optical character recognition (OCR) has been developed to read the characters in a document image captured from mobile device. However, existing OCR algorithms do not work correctly when the input image has serious distortion which destroys the text structure. More specifically, the existing OCR applications cannot detect the text regions accurately because of the geometric distortion caused by perspective projection and document rotation. 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. In addition, an adaptive algorithm is required to reject the false case of the perspective rectification. Note that the adaptive algorithm is important for consumer application because we need to reject failure case and show only the good look result to the users. The proposed framework is implemented and tested on an off-the-shelf smartphone. 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.

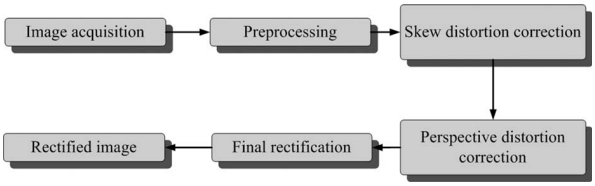

Figure 1. Pipeline of perspective rectification framework

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 [3]. 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. In the perspective distortion correction step, an adaptive scheme, to decide whether perspective rectification is used or not, is proposed. The proposed framework is designed for consumer application so that it is important to adaptively choose the selected result to be shown. The algorithm sequence is illustrated in Fig. 1.

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. Then, an adaptive thresholding method is utilized to binarize the text regions.

2.2. Skew Rectification

To rectify the skew distortion, the proposed framework extracts the candidate horizontal lines using Hough line detection algorithm [2]. 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 [1]. 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

To rectify the perspective distortion, the vertical vanishing point is detected and sent to infinity. The idea of the vertical vanishing point detection is to extract the reliable paragraphs and their left and right boundaries. To detected the paragraph, paragraph blob detection is performed by employing the morphological dilation and the connected component analysis. 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. For each intersection point of boundary pair, the perspective projective histogram is calculated and the maximum of histogram difference is computed. Then, 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. If there is no reliable vanishing point, the proposed framework is designed to reject the perspective rectification result.

3. Experimental Result

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. We test the proposed framework on the input document images with aligned paragraphs which are usual case in the published documents. The total 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). Based on the experimental results, the proposed framework is proven to obtain fast and accurate performance to rectify perspective distorted documents. The proposed framework is practicable for consumer usage because it contains adaptive scheme to reject failure cases.

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

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

4. Conclusion

In this paper, a fast and robust method to rectify image with geometric distortion was proposed. In addition, an adaptive algorithm to reject the failure case of the perspective rectification was also added in the proposed framework. Note that the adaptive scheme is important to select the best results to be shown for users. 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. Note that the proposed framework can obtain the rectified results for less than a second and is practicable for being used by consumers.

Acknowledgement

This work was supported by Samsung Electronics and Inha University.

References

[1] P. Clark and M. Mirmehdi. Rectifying perspective views of text in 3D scenes using vanishing points. *Pattern Recognition*, 36(11):2673–2686, November 2003.

[2] R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, January 1972.

[3] X. C. Yin, H. W. Hao, J. Sun, and S. Naoi. Robust vanishing point detection for mobilecam-based documents. In *Proc. of International Conference on Document Analysis and Recognition*, pages 136–140, September 2011.