Restoring Warped Document Image Based on Text Line Correction

[*]Ching-Tang Hsieh, Sheng-Chao Lee, Cheng-Hsiang Yeh
Dep. of Electrical Engineering Tamkang University, New Taipei, Taiwan, R.O.C
**Correspondending Author: hsieh@ee.tku.edu.tw*

Abstract

Document images captured by camera often suffer from warping and distortions because of the bounded volumes and complex environment light source. These effects not only reduce the document readability but also the OCR recognition performance. In this paper, we propose a method to combine non-linear and linear compensation for correcting distortions of document images. First, due to the broken text result of Otsu binarization, an image preprocessing is used to remove the effect of background light. Second, the dewarping method using the cubic polynomial fitting equation is proposed to find out the optimal approximate text line for vertical direction rectification. Finally, we use linear compensation for horizontal direction rectification. Experimental results demonstrate the robustness of the proposed methodology and improve the accuracy rate of OCR recognition.

Keywords*: Document image, Warping, Morphology, Text line.*

1. Introduction

Document digitization with cameras or scanners that becomes more and more popular. Document image acquired by scanners or cameras which often suffer from variety different distortions because of the objects volume or environment. Document image will affect not only the text recognition accuracy rate but also the readability of documents. This effect will cause that difficult to analyze the contents of the file and will reduce the recognition accuracy rate of the OCR approaches.

At last decade, many techniques for document image rectification have been proposed [1] that can be classified into two categories. First category is based on 2-D distorted document image processing [2-4] and second one is 3-D document shape of surface reconstruction [5-7]. The former techniques do not depend on special equipment or prior information but rely on 2-D information. The other techniques in latter category obtain 3-D information by using special equipment or 3-D model reconstruction.

In 2-D image processing category, rectification techniques only use available information in distorted document images. Zhang and Tan [2] divide the distorted document images into shaded and non-shaded region. Their method uses polynomial to regress the warped text line with quadratic reference curves. They find out text curves by clustering connected components and move the connected components to restore straight horizontal baseline. But this method only fits for shaded region. Masalovitch and Mestetskiy [3] use outer skeleton of text for the distorted document as a rectification method. And they use continuous branches to define interlinear spaces of document.

First, the branches are approximated by cubic Bezier curve to find out specific distortion of each interlinear space. Then, they build a whole approximation document as rectification result. This method is sensitive to the deformation of text block vertical borders. Nikolaos Stamatopoulos et al. [4] propose a coarse-to-fine strategy as two stage rectification method for correcting warping document. In their method, coarse rectification step aims to restore large distortions for whole document and the fine rectification step aims to restore local distortions for each word. This method needs a lot of time to calculate parameters.

On the other hand, the 3-D document rectification techniques rely upon 3-D information extraction. Li Zhang et al. [5] use dedicated laser range scanner to capture the 3-D structure of warped document. Their method uses an irregular triangular mesh to represent the 3D geometry of the document. In correction step, aiming to different external forces they use the different rectification formula to correct warped document. Chew Lim Tan et al. [6] combine with the shape-from-shading and the book cross-section to reconstruct the 3-D surface of deformation document. A self-similarity measure is utilized for line tracing in [7]. They use local stroke statistics for text orientation estimation, and utilize 3-D coordinate grid in correction step to model geometric surface of document and perform photometric rectification.

doi:10.4156/rnis.vol14.82

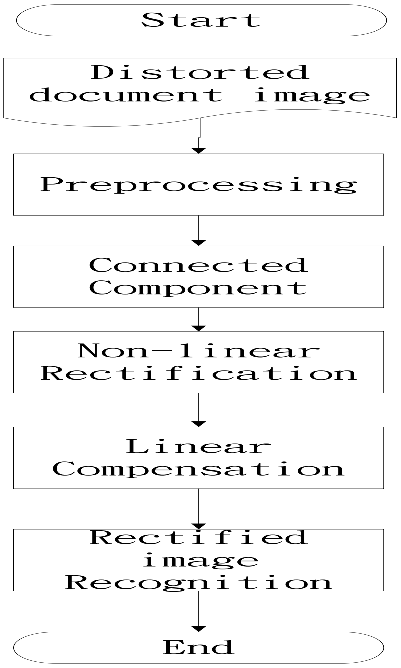


Figure 1. System Flowchart

The remainder of this paper is organized as follows. We focus on the propose method which is detailed described in section 2. First, our preprocessing step a method for removing light source is applied. Second, describe the natural cubic curves calculation for non-linear rectification step. Finally, we based on the results of second step performing linear compensation. In section 3 we demonstrate the experimental platform and the performance of our proposed method. Then show the performance of our propose method with figures and OCR recognition accuracy rate tables. In section 4 we discuss conclusions and future work.

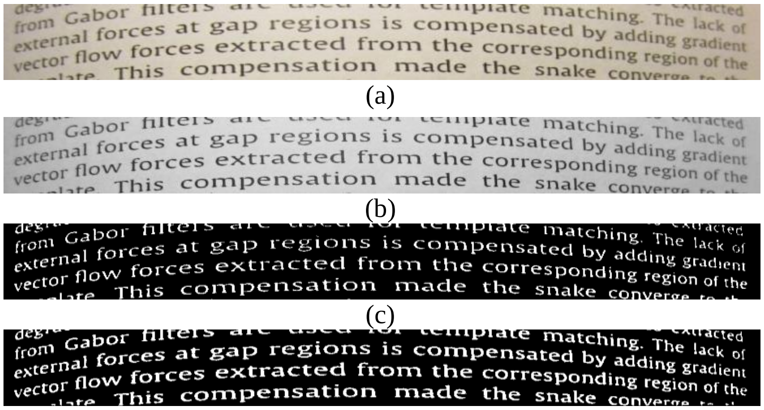
2. Proposed method

Without special equipment to obtain 3D depth information, we combine with nonlinear and linear rectification strategy using only 2D information in this paper. In the preprocessing step, we use the method for contrast enhancement and remove the impact of the light source. In the non-linear rectification step, we correct large horizontal distortion by the least squares method. Finally, we aim at word level using linear compensation for each word. The system flowchart of proposed method is shown in Figure 1.

A. Preprocessing

Before we proceed with major rectification process we apply a preprocessing step to extract the binarization text from original distorted image with different light source. First, we transform the original image into grayscale image by using formula (1).

$$I_{gray} = \frac{(I_R + I_G + I_B)}{3} \quad (1)$$


from Gabor filters are used for template matching. The lack of
external forces at gap regions is compensated by adding gradient
vector flow forces extracted from the corresponding region of the
This compensation made the snake converge

(a)

(b)

(c)

(d)

Figure 2. (a)Original image (b)Gray image (c)Direct using Otsu method (d)After [8] method using Otsu method

Using J. Moraleda [8] proposed method can overcome the different light source document images. Furthermore, it can reduce the probability of text destroyed. So, the recognition accuracy rate can be improved. We display preprocessing results of distorted document image in Figure 2.

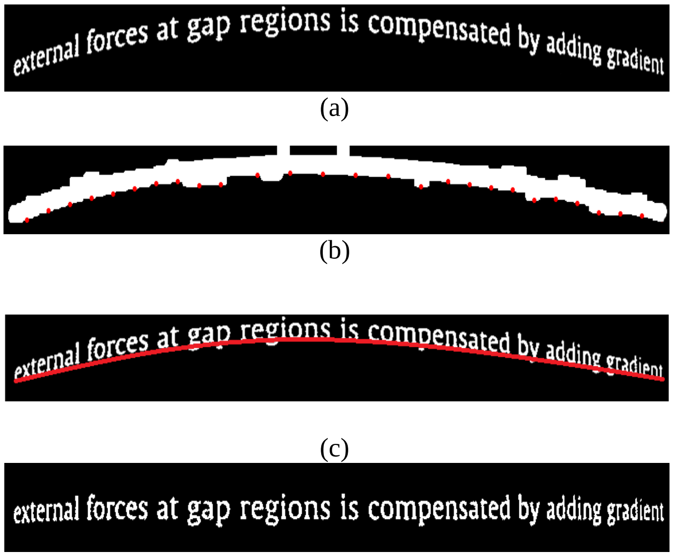
B. Non-Linear Rectification

In this step, we first perform the dilated operation to make the text block. Then the distorted text content uses the region growing [9] to label connected components. We cut the connected components by equidistance and find out the point of the text line. And then the natural cubic curve of baseline is obtained by using the least squares method as follows:

$$y_{im} = a_i x_m^3 + b_i x_m^2 + c_i x_m + d_i \tag{2}$$

$$\Delta y = y_{im} - y_{i0} \tag{3}$$

$$y_{im}' = y_{im} + \Delta y \tag{4}$$


external forces at gap regions is compensated by adding gradient

(a)

(b)

(c)

(d)

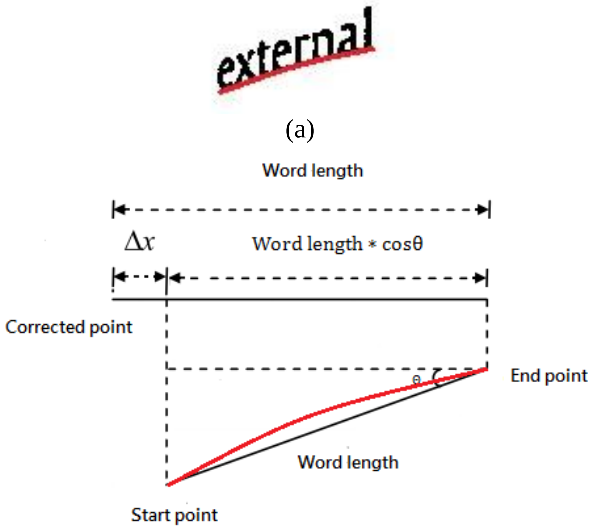
Figure 3. (a)binarized result (b)after dilating and cutting operation (c)calculate cubic curved line of baseline points (d)result of horizontal rectification where the y_i represents each line in document and $i = 1,2,3 \dots n$. The a_i, b_i, c_i and d_i in formula (2) respectively represents the same text line equation coefficients, and is obtained by the least squares method.

The length of text baselines are calculated according to each connected component which will be corrected. We represent connected component x-axis coordinates as $x_m = x_0, x_1, x_2 \dots x_{L-1}$. Then we use formula (2) to calculate the height position y of the x_0 coordinate of each baseline. Text horizontal rectification is based on connected component position y_{i0} of each baseline. The corrected position y' is obtained by formula (3) and (4). The correction process result is shown in Figure 3.

C. Linear Compensation

After extracting the respective word, we use the baseline cubic equation to approximate triangular straight as shown in Figure 4. According to compensation pixel approach we vertically extend the corrected document image to the target area. In addition, the vertical linear compensation can remove respective word distorted to increase the document readability and OCR recognition rate.

In this step, we perform linear pixel offset compensation of each text line with the word level to overcome offset of document image. The proposed method extracts single text line with connected components label, and then use horizontal projection method to segment words respectively. Finally, we perform the vertical compensation after non-linear rectification.



(a)

(b)

Figure 4. (a)extracted respective word segments with baseline (b) approximate length of the text.

Assuming the $f_{goal}(x', y')$ and the $f_{src}(x, y)$ represent the image of the target and source. The dx_{ij} and dy_{ij} are pixel position offset of a word. The corresponding relationship is described as formula (5).

$$\begin{cases} y'_{ij} = y_{ij} + dy_{ij} \\ x'_{ij} = x_{ij} + dx_{ij} \end{cases} \tag{5}$$

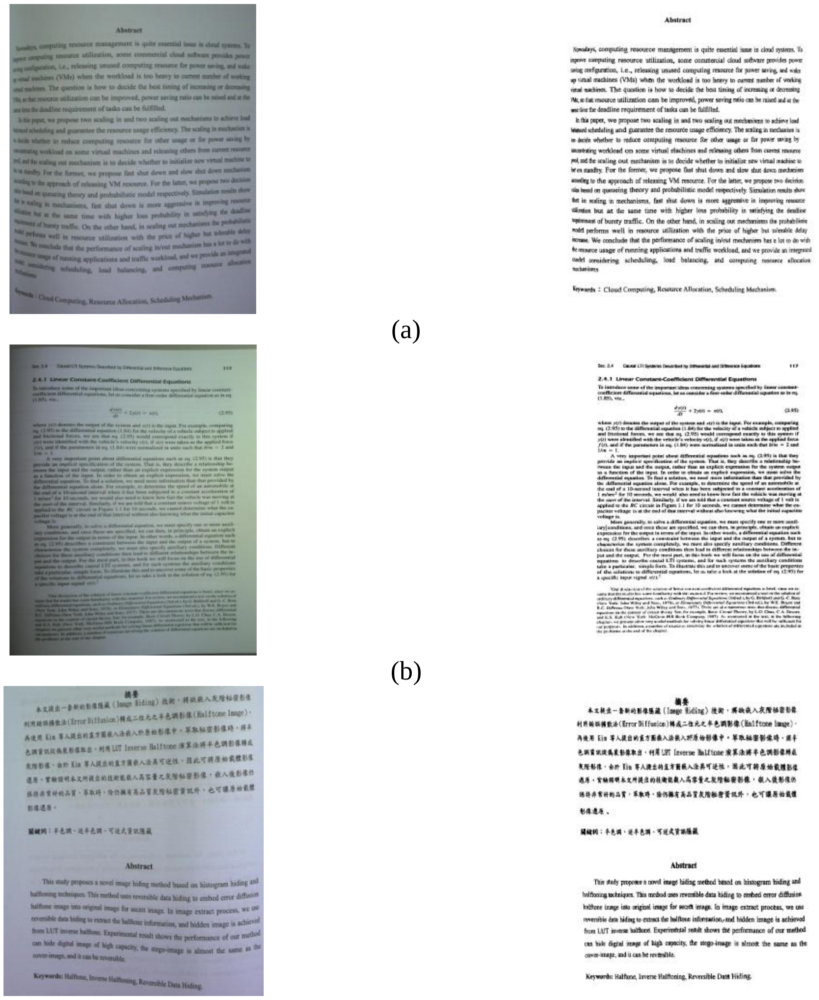
$$dx_{ij} = x_{ij} \times \frac{\Delta x}{\text{Word length} \times \cos\theta} \tag{6}$$

where the (x_{ij}, y_{ij}) is the i_{th} row of the j_{th} word of the text line. The main purpose of this step is to determine the x-axis direction pixel offset. The horizontal correction has been dealt with in step B. The word offset of text line in the i_{th} row j_{th} word calculated by using formula (6). The results of step C correct the text vertical distortions for each word. Therefore, the complete document distortions will be corrected through the non-linear rectification and linear compensation.

3. Experimental result

Experimental platform is Win 7 with professional Visual C++ 2010, and document camera is AF DocExpress 300. In order to confirm the validity of our proposed method, we use commercial software ABBYY FineReader [10] and Tesseract [11] with character recognition processing. Furthermore, the recognize languages is limited by the commercial Software. The results of our experiments, we use variety of different types of distorted document images 40. Average correction time-consuming for each distorted image is 5.54 seconds, and the resolution of the images is 2048 × 1536.

The correction results are illustrated in Figure 5, in which (a) the distorted image of the book pages, (b) distorted document images containing mathematical formula, (c) the document images containing different language. The corrected results show that the proposed method can achieve effective rectification with continuous text line. Experimental results of character and word accuracy rate, respectively, as shown in table I and table II.



Abstract

Nowadays, computing resource management is quite essential issue in cloud systems. To improve computing resource utilization, some commercial cloud software provides power saving configuration, i.e., releasing unused computing resource for power saving, and wake up virtual machines (VMs) when the workload is too heavy to current number of working virtual machines. The question is how to decide the best timing of increasing or decreasing VMs, so that resource utilization can be improved, power saving ratio can be raised and at the same time the deadline requirement of tasks can be fulfilled.

In this paper, we propose two scaling in and two scaling out mechanisms to achieve load balanced scheduling and guarantee the resource usage efficiency. The scaling in mechanism is to decide whether to reduce computing resource for other usage or for power saving by concentrating workload on some virtual machines and releasing others from current resource pool, and the scaling out mechanism is to decide whether to initialize new virtual machine to be on standby. For the former, we propose fast shut down and slow shut down mechanism according to the approach of releasing VM resource. For the latter, we propose two decision rules based on queueing theory and probabilistic model respectively. Simulation results show that in scaling in mechanisms, fast shut down is more aggressive in improving resource utilization but at the same time with higher loss probability in satisfying the deadline requirement of bursty traffic. On the other hand, in scaling out mechanisms the probabilistic model performs well in resource utilization with the price of higher but tolerable delay increase. We conclude that the performance of scaling in/out mechanisms has a lot to do with the resource usage of running applications and traffic workload, and we provide an integrated model considering scheduling, load balancing, and computing resource allocation mechanisms.

Keywords：Cloud Computing, Resource Allocation, Scheduling Mechanism.

(a)

Sec. 2.4 Causal LTI Systems Described by Differential and Difference Equations 117

2.4.1 Linear Constant-Coefficient Differential Equations

To introduce some of the important ideas concerning systems specified by linear constant-coefficient differential equations, let us consider a first-order differential equation as in eq. (3.85), viz.,

[illegible]

(b)

摘要

本文提出一套新的影像隱藏(Image Hiding)技術，將欲嵌入灰階秘密影像利用錯誤擴散法(Error Diffusion)轉成二位元之半色調影像(Halftone Image)，再使用 Kim 等人提出的直方圖嵌入法嵌入於原始影像中。萃取秘密影像時，將半色調資訊提取出來，利用 LUT Inverse Halftone 演算法將半色調影像轉成灰階影像，由於 Kim 等人提出的直方圖嵌入法具可逆性，因此可將原始載體影像還原。實驗證明本文所提出的技術能嵌入高容量之灰階秘密影像，嵌入後影像仍保持非常好的品質，萃取時，除仍擁有高品質灰階秘密資訊外，也可讓原始載體影像還原。

關鍵詞：半色調、逆半色調、可逆式資訊隱藏

Abstract

This study proposes a novel image hiding method based on histogram hiding and halftoning techniques. This method uses reversible data hiding to embed error diffusion halftone image into original image for secret image. In image extract process, we use reversible data hiding to extract the halftone information, and hidden image is achieved from LUT inverse halftone. Experimental result shows the performance of our method can hide digital image of high capacity, the stego-image is almost the same as the cover-image, and it can be reversible.

Keywords: Halftone, Inverse Halftoning, Reversible Data Hiding.

(c)

Figure 5. The correction results of document images

Table 1. Accuracy Rate of Characters

	Total number of characters	**Total number of wrong characters**	**Accuracy Rate**
Direct using ABBYY	69201	7211	89.58%
After rectification using ABBYY	69201	879	98.73%
Direct using Tesseract	69201	12422	82.05%
After rectification using Tesseract	69201	678	99.02%

Table 2. Accuracy Rate of Characters

	Total number of words	**Total number of wrong words**	**Accuracy Rate**
Direct using ABBYY	12319	1307	89.39%
After rectification using ABBYY	12319	473	96.16%
Direct using Tesseract	12319	2496	79.74%
After rectification using Tesseract	12319	515	95.82%

4. Conclusions

We proposed a method which combined with non-linear rectification and linear compensation to correct distortions of document images. Experimental results on distortion document images not only demonstrate the robustness of proposed methodology but also improve the effects of OCR recognition. Our future work will focus on extension of proposed method application in handwritten documents or other research.

5. References

[1] J. Liang, D. Doermann, and H. Li, “Camera-based analysis of text and documents: A survey,” Int. J.Document Analysis and Recognition, vol.7, no. 2-3, pp. 84–104, 2005.

[2] Z. Zhang and C. L. Tan, “Correcting document image warping based on regression of curved text lines” International Conference on Document Analysis and Recognition, Edinburgh, Scotland, pp. 589–593, 2003.

[3] A. Masalovitch and L. Mestetskiy, “Usage of continuous skeletal image representation for document images dewarping” In 2nd Int. Workshop on Camera-Based Document Analysis and Recognition, Curitiba, Brazil, pp. 45-53, 2007.

[4] N. Stamatopoulos, B. Gatos, I. Pratikakis, and S. J. Perantonis, “Goal-Oriented Rectification of Camera-Based Document Images,” IEEE Trans. Image processing, vol. 20, no. 4, pp. 910–920, Apr. 2011.

[5] L. Zhang, Y. Zhang, and C. L. Tan, “An improved physically-based method for geometric restoration of distorted document images,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 4, pp. 728–734, Apr. 2008.

[6] C. L. Tan, L. Zhang, Z. Zhang, and T. Xia, “Restoring warped document images through 3-D shape modeling,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 2, pp. 195–208, Feb. 2006.

[7] Y. Tian and S. G. Narasimhan, “Rectification and 3D Reconstruction of Curved Document Images,” IEEE Computer Vision and Pattern Recognition (CVPR), pp. 377-384, Jun. 2011.

[8] J. Moraleda, “Large scalability in document image matching using text retrieval,” Pattern Recognit. Letters J., vol. 33, no. 7, pp. 863–871, May. 2012.

[9] R. Adams, and L. Bischof, “Seeded region growing,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 16, no. 6, pp. 641-647, June, 1994.

[10] ABBYY FineReader OCR: http://finereader.abbyy.com/

[11] Tesseract: http://code.google.com/p/tesseract-ocr/S.