

Research and Realization of Perspective Correction Technology for Document Image

Liu Shouliang Li Yiming Gao Yupeng Yang Kai

Physics and Electronic Information Engineering College Wenzhou University Wenzhou 325000
lsl198776@gmail.com

Abstract

In this paper, we set up a system which can correct the perspective distortion of document image. This system includes the operation of image preprocessing. This paper takes the information of document frame or document basing on the number of document frame. This paper uses the information obtained to correct the perspective distortion of document image. Experimental results show that this method can effectively correct perspective distortion image and the calibration accuracy and efficiency can be significantly increased.

1. Introduction

With the development of the information technology in today's society, the needs for digital processing of document information are increased. Digital storage with its advantages of easy storage, easy handling and taking up less space, has been used in various fields of social life. The most common methods of digital processing of document information are through the scanner. However, the scanner is too heavy and not easy to carry anytime, anywhere. so it's not conducive for the input of document information. With the popularity of digital cameras, phones and other equipments, document recognition technology is gradually going to the areas of the visual document images. Now we can get the document information by taking photos, then we can recognise the information and get the Electronic document. However, due to the tilt of the document itself and the influences of human factors, it's a common problem that the captured image will be perspective transformation. If the image has a perspective transformation, it's will create great difficulties for the identification of the document information, perspective distortion correction will be necessary.

The technology of image geometric distortion

correction includes image tilt, perspective and distortion correction. The technology of image tilt correction is mature. Relative to the image tilt correction, there are many difficulties and places to be improved in image perspective distortion correction. This paper studies the perspective correction of the document image. The literature [1] proposed a perspective distortion correction method for the small-scale document image. However, due to the special nature of small document, Document is smaller and the text content information is also relatively simple and arranged in the specification, so the information extraction and document perspective correction is relatively simple. the method has less good correction effect for larger document images, such as the images of book document. The method [2] uses the extended two-dimensional projection to locate the horizontal vanishing point, and then uses the document alignment and spacing between lines of text information to locate the vertical vanishing point, while it's not so efficient for small document containing the edges. Technology [3] proposed a new document image correction method, but because of the method involved in lines of text tends to computing and morning surface estimates, so the computing speed is slow. For small documents that contains the edge information can be extracted by the edges of a straight line to determine the corrected information. For large documents which has no border or the border information is incomplete, we can also extract the text content of a straight line to determine the corrected information. This method can correct small and large documents Effectively.

2. Image Processing

Document images taken by digital cameras and mobile phones, due to the complexity of the shooting environment and equipment, will inevitably produce image noise and uneven illumination. In

such cases, the image correction will produce a great error in the correction results .The correction results will not be satisfactory .The introduction of the methods below is in order to solves these problems , and the image preprocessing for image preprocessing in this article consists of two main aspects: the image denoising and the uneven illumination of the processing .

2.1 Image Denoising

The text image has own noise characteristic, it's noise usually refers to the spots and empty.The so-called spot is refers to the image burr, the stain and the content which is not relative with document. Median filter with the advantages of better inhibition in the area of point noise and interference pulse and so on , and compared with other denoising methods can effectively preserve the edge of the image contour information .Also, because the text image just consists of text and background , under normal circumstances the text is black and the background is white ,so the gray is relatively simple.2D median filtering principle is to be discussed to use pixel value to instead the gray value of the pixels with the gray value of adjacent pixels Median filtering the image gray level is the more simple and effective . The image shown below is the text image by adding Gaussian , salt and pepper noise and median filter image :



Figure 1

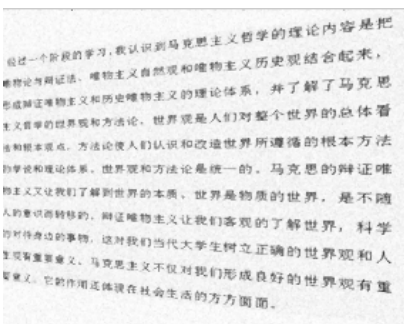


Figure 2

2.2 Uneven Illumination Processing

About Uneven processing on the image light, the classical method is the top-hat method in the article[4]. Method[5] proposes the improved algorithm which has its own merit .In above methods the most critical thing is to select a global threshold .I will use the adaptive threshold method of threshold selection,basing on the characteristics of the text images .Adaptive threshold selection is based on the different light conditions of different parts of image , And then we split the image into several areas which have different background pixel gray-scale.Then we select the best threshold and set the pixels which is higher than the threshold to be prospect,otherwise to be the background .This method is more flexible and accurate than using a global threshold , the correction effect is very ideal .Figure3 is for uneven lighting binarization image . Figure 4 is for after correction of uneven illumination binarization image.

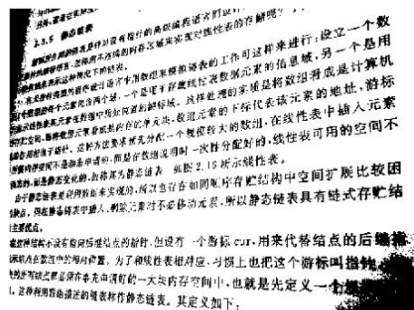


Figure 3

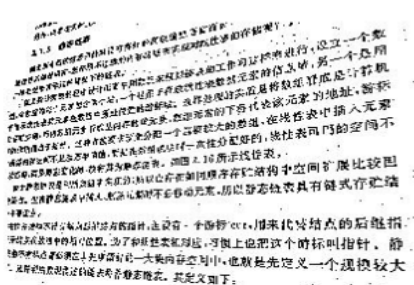


Figure 4

3. Image Perspective Correction

3.1 Principles of Perspective Correction

If the document is small and the distance of photography is short,we may get the full edge information in the image.While, Otherwise we can't get enough edge information.In both cases, we can not always regard them as the same, because

different case has the different Correction difficulty and efficiency. The former is relatively simple, The latter is relatively difficult. Therefore, in the perspective correction to make judgments :

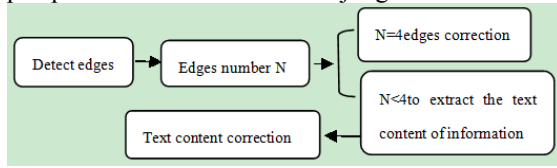


Figure 5

Firstly we can use hough transform method to extra-ct the geometric distortion of image edges for images which contain complete edge information after the judgment. Then we use the contour vertex coordinates to computer rectangular aspect ratio in order to obtain the matrix of distortion correction. Then the distorted image can be to be corrected.

As to the document image which don't have the complete edge information, we need to To extract the text content of the information. Perspective deformation of the potential text box in the text image is a convex quadrilateral. It corresponds to a rectangle in the original text. Correspondence between convex quadrilateral and rectangular can determine the corresponding perspective transformation matrix. we can correct the perspective distortion with the the corresponding matrix. The corresponding point - to - point relationship between horizontal document image and its perspective image can be expressed by perspective transformation matrix:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & h_8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (1)$$

x'_i flat text image coordinates, the coordinates of the perspective of text images. The above equation can be abbreviated as: $X' = HX$ (2),

H is non-singular matrix, we use Non - homogeneous coordinates of a point instead of point of homogeneous coordinates. Corresponding points x and x' have non- homogeneous coordinates (x, y) and (x', y') , Perspective transformation can be written in the form of non-homogeneous :

$$\begin{aligned} x' &= \frac{h_0x + h_1y + h_2}{h_6x + h_7y + h_5} \\ y' &= \frac{h_3x + h_4y + h_5}{h_6x + h_7y + h_5} \end{aligned} \quad (3)$$

Make $h=(h_0, h_1, h_2, h_3, h_4, h_5, h_6, h_7)$ T, if we know the four points of horizontal document image and its

perspective image, we can start work below:

$$\begin{cases} (x_1, y_1, 1, 0, 0, 0 - x_1x'_1 - y_1y'_1)h = x'_1 \\ (0, 0, 0, x_1, y_1, 1 - x_1x'_1 - y_1y'_1)h = y'_1 \end{cases} \quad (4)$$

you can get a h_i in the corresponding matrix when you get the solution of equations. Then we use Perspective transform solutions of the corresponding matrix H to correct perspective transformation of each pixel in the perspective of the text image. Finally we correct the perspective document image to be original front view.

3.2 Perspective Information Extraction

Hough transform can be extracted directly to the image that contains the complete edge information , and therefore we don't repeat it here .It is necessary to extract the text content information, if the document don't have the full edge information. First of all , we want to start the work of image binarization, this is because the gray image data contains too much information, which is not so necessary in perspective correction. Image binarization , greatly reduces the image data and saves processing time .Figure 6 is the picture after the binarization.:

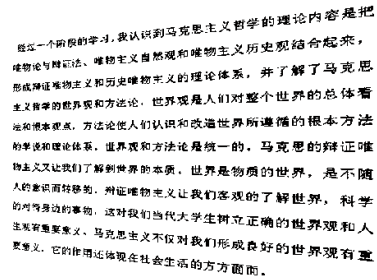


Figure 6

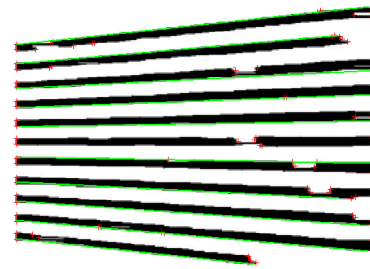


Figure 7

After image binarization processing, The data we are dealing with is simple. However , we need the straight lines of text, So we need to extract lines of text line. We use the run-length smoothing algorithm to smooth the image, And then we use smoothed image to analysis connected component , later to fit

the straight line of the line of text.As is shown in Figure 7.

When run chart is extracted ,we will begin to extract the straight line of run chart .You can use the top , middle, and bottom endpoint coordinates of each line of black connectivity. We use least squares method and the three groups of points to fit a line three straight lines of document,the three lines are top,middle,and bottom line. we take a minimum of a linear fitting a straight line as this line of text from the minimum error of a three lines basing on least squares method.As is shown in figure 7,we calculate the length of average line basing on straight line parameters.Trying to find the line greater than 0.7L as the top straight line of convex quadrilateral from top to bottom.Then Trying to find the line greater than 0.7L as the bottom straight line of convex quadrilateral from bottom to top.

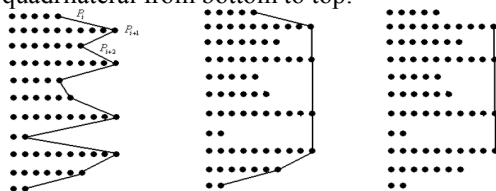


Figure 8

It's more difficult to extract the vertical line, because there is no straight line to extract in the document which don't contain the edge information.As is shown in the left image of figure 8 ,we get left concave point of the edge by repeated the search on the left broken line. Delete these points and get line which has no concave point ,as is shown in Figure 8. Finally Intercept the straight line and get vertical straight line of the right border,as is shown in the right image of figure 8.Similarly it's available to get vertical straight line of the left border.Calculate the point of intersection of the left line with top line and the point of intersection of the left line with bottom line. Similarly calculate the point of intersection of the right line with top line and bottom line.The four-point is just the four vertices of the convex quadrilateral.Finally, by the above methods we can correct perspective document image as shown below:



Figure 9



Figure 10

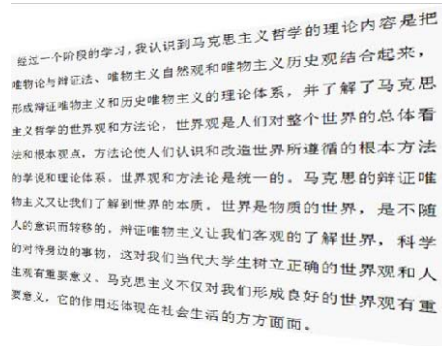


Figure 11

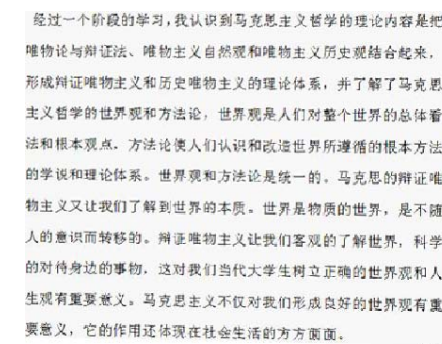


Figure 12

4. Conclusion

When we correct the perspective document images which contains the edge information ,the result is satisfactory,as to others we can also achieve a certain effect, while it's not so satisfactory.The algorithm can be developed in future study,specially to find a better solution for the extraction of the vertical line.

References

- [1] MaYichao, Dai Ruwei, Wang Chun—Heng “An Integrated Approach for Perspective Distortion Correction of Small Square Document Images” *Pattern Recognition*

and Artificial Intelligence 2006(08).

- [2] Clark P, Mirmhedi M. “Rectifying Perspective Views of Document in3D Scenes Using Vanishing Points”. *Pattern Recognition*, 2003, 36(11): 2673—2686.
- [3] Lu s J, Chen B M, Ko C C. “Perspective Rectification of Document Images Using Fuzzy Set and Morphological Operations”. *Image and Vision Computing*, 2005, 23(5): 541—553.
- [4] P.T.Jackway , Improved morphological Top- Hat [J]. *IEEE Electronics Letters*, 2000, 36 (14):1194—1195.
- [5] Sun Zhonggui, Reduce “Non-uniform Illumination Achieved by Matlab” *Image processing* 2008.
- [6] Lu s J, Chen B M, Ko C C. “Perspective Rectification of Document Images Using Fuzzy Set and Morphological Operations”. *Image and Vision Computing*, 2005, 23(5): 541—553.
- [7] Chang Jun. Miao Ligang, “Correcting radial distortion of document images by morphology”, *Journal of Computer Applications* 2010(04).
- [8] Gao Yupeng, Li Yiming, Hu Zhongyi. “Skewed text correction based on the improved Hough Transform”[C]. *Image Analysis and Signal Processing*, 2011: 368-372.
- [9] Chen Deyun, Yin Fang, Wu Rui, “ Perspective Distortion Correction Method of Scene Text Based on Vanishing Point”, *Journal of Nanjing University of Science and Technology*. 2011(8).
- [10] Li Zheng, Yang Yang, Xie Bin, Wang Hong, “Skew angle detection and correction of document images based on Hough transform”. *Computer Applications* 2005(3).