Proceedings of 2018 IEEE
International Conference on Mechatronics and Automation
August 5 - 8, Changchun, China

Research on Deskew Algorithm of Scanned Image

Yonggang Zhang
High voltage technique
Yunnan electric power technology Co.,LTD
Kunming, 650217,China
yonggangzhang@163.com

Jianming Zhong,Hong Yu,Liuyang Kong
High voltage technique
Electrical Research Institute of Yunnan Electric Power Grid Co,Ltd.
Kunming,650217, China
yuhong2388245@163.com

***Abstract* - It is difficult to ensure that the scanned document will not be skewed due to the high speed scanning of the paper document by the mechanical feed or the document itself. This will not only affect the visual effects, but also seriously affect the accuracy and speed of subsequent processing. Therefore, this paper presents a Hough transform algorithm based on automatic scanning image correction and cropping, the first to determine whether the scanned image is a text or non-text image, the image preprocessing based on different types of scanned images of the image binarization, Compared with the traditional algorithm has a good anti-interference ability, can accurately detect the scan image skew angle.**

Index Terms - binarization, automatic correction, Hough transform algorithm.

I. INTRODUCTION

With the rapid development of information technology, more and more applications have realized paperless offices such as online marking, questionnaire surveys, government documents, bank bills and digital libraries[1-4]. In practice, a large amount of information is recorded on paper documents, and paper documents are inconvenient to save, retrieve, modify and so on[5]. In order to make full use of the advantages of electronic documents, a large number of paper documents Convert to high-speed document scanner electronic document. Compared with the traditional paper documents, electronic documents have many advantages such as small storage space, convenient retrieval, easy to transfer and update, etc[6-8]. Meanwhile, in order to increase the copyright protection of electronic documents, the necessary confidential coding can be carried out to improve reliability Sex. And, if needed, electronic documents can easily be converted into various forms of paper documents at any time.

In the area of document image analysis and comprehension, the document image correction and clipping technology is an important part of it, and it has a significant impact on the subsequent image segmentation, layout analysis, character recognition and other work[9]. In the image recognition system, the quality of the image data directly affects the accuracy of recognition, and the actual scanned image tends to be inclined and produce redundant edges. In order to improve the recognition rate of the system and increase the stability of the system, it is necessary to quickly and accurately detect the tilt angle of the image and crop the redundant edges of the scanned image, and correct them. At present, most scanners do not have the functions of automatic image skew correction and cropping[10]. However, in practice, inclinations of printing and printing often occur, which makes it impossible to correctly correct tilt. Therefore, it is of great significance and value to research on the correction and cropping of scanned document images in the processing of document images such as digital library construction and online marking, and has a broad space for development.

Currently used to convert large quantities of paper documents into electronic documents commonly used high-speed scanner automatic feeding mode, which is different from the flatbed scanner, you can greatly speed up the document scanning speed, but due to mechanical paper or the document itself The reason is difficult to ensure that the scanned document will not be skewed, and high-speed scanners scan the paper size is generally larger than paper, so the scanned image will have a black edge or gray edge, as shown in Figure 1:

Fig.1 Scanning image

This image not only affects the visual effects, but also will seriously affect the accuracy and speed of subsequent processing, such as: layout analysis, character segmentation and so on. Skew angle is slightly larger will seriously affect the recognition efficiency of OCR (cursor character recognition) system, so the image tilt angle detection and correction in the document analysis is a very frequent operation, improve

document image tilt angle correction and cutting speed, you can To ensure the visual effect of the entire document image, but also can improve the follow-up work processing speed and reduce the complexity of the work. Therefore, the use of a certain degree of automatic correction and cropping algorithm to correct the scanned document image tilt and redundant edge cutting is to ensure the quality of scanned images is a key step.

Due to the complexity of the document image layout, it is very difficult to establish a common algorithm for image correction of oblique images. How to detect the tilt angle of the document image efficiently and accurately is the key to successful tilt correction. Therefore, in view of the above problems, this article is of great significance to the research of automatic correction and cropping algorithms.

II. THE TRADITIONAL ALGORITHM

At present, many researchers are devoted to the correction research of document scanning images. One of the most crucial problems is the detection of the inclination of scanned images. After years of exploration and research, many researchers have developed a variety of document tilt correction algorithms, can be divided into the following categories: cross-correlation based on the method of cross-section projection and Fourier transform based methods.

A. Based on Profile Projection Algorithm

Projection method is currently the most common method of tilt correction of images. The basic idea is that because there is a gap between the lines of text, irradiating text from multiple directions with a bundle of parallel rays, the text will block the light and leave behind Shadow, and the gap can pass through the light, through the direction of the maximum amount of light is the text of the image oblique direction, which can be the text of the image angle of deflection. Its operation process is to scan the different angles of the image projection to obtain its corresponding projection, and then select the profile of a certain eigenvalue, such as the gradient, the first eigenvector and variance, calculated according to these eigenvalues of the image slope.

Taking the projection in only two directions as an example, the images are projected in two directions, that is, the vertical direction and the horizontal direction respectively, and the corresponding vertical projection and horizontal projection are obtained. If the scanned text image is not tilted, and the textual behavior Is the horizontal direction of the premise, based on this analysis, its horizontal projection will have the maximum amplitude and frequency, the peak and peak distance should be equal to the height of the word, the trough and trough the distance between the text should be text Line spacing. Let "I (x, y)" be the value of the two-dimensional pixel, the image is H × W size image, there is:

Vertical projection:

$$I(x,y)=\begin{cases}1,(x,y)\\0,(x,y)\end{cases} \quad (1)$$

The projection of I (x, y) in the vertical and horizontal directions is:

Horizontal projection:

$$I(x)=\sum_{y=0}^{w-1} I(x,y) \quad (2)$$

Figure 2 is a horizontal and vertical scanning image projection schematic:

Fig.2 Schematic diagram of the image projection

Then some projection eigenvalues are analyzed to get the tilt angle. This method is very easy to implement, but if the projection of the entire image, the need for the projection direction too much, greatly increasing the amount of computation, the entire algorithm becomes complicated. And this method cannot be used to have the graphic and table scanned document images, and can only get good results in plain text images with small inclination angle. If the skew angle is large, the calculation result will be inaccurate, the error. The method of cross-correlation is based on the method of calculating the cross-correlation of pixel points on two row lines at equal distances in the image. If the document image is tilted, the pixels in the same row on two parallel scan lines will be translated. The correlation matrix is established by the pixels in the parallel lines and the projection is performed in the vertical direction. The tilt angle of the image corresponds to the global maximum in the obtained projection image. The accuracy of this method for estimating the tilt angle is high, but the correlation matrix is computationally intensive. Moreover, for some special cases, the accuracy of the image is reduced, for example, there are more illustrations in the image, and the accuracy is greatly reduced.

B. Methods Based on Cross-correlation

The cross-correlation based method is based on a special case of projection. Its basic principle is that when the scanned document image is skewed, the pixels on the same line will occur on two parallel scan lines by translating, we set up the matrix and make the corresponding vertical projection. By obtaining the global maximum of the projection image, we can obtain the tilt angle of the detected image.

Let I (x, y), where (0<x<X-1,0<y<Y-1), is the pixel value of each point of the grayscale image, making two vertical lines on the scanned image 2.2): and define the vertical cross correlation function Rl:

$$(x_0,s)=\sum\nolimits_{y} I(x_0,y)*I(x_0+d,y+s) \quad (3)$$

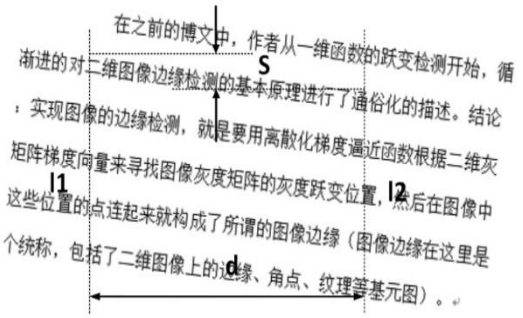

Fig.3 R(s)Construction image

According to Figure 3, the two vertical lines for the two parallel scan lines, pixels in the text line after moving in the distance d, resulting in a size of s gap, then Rl have the maximum value. Next, we take all the lines of distance d and add all the results to get the following formula:

$$R(s)=\sum_{y_0=0}^{Y-d} R_l(y_0,s) \qquad (4)$$

According to the range of s, we calculate R (s) under different values of s, find the corresponding peak, calculate the correlation matrix and make the corresponding vertical projection, and get the global maximum of the corresponding projection, and then get the scan image tilt angle. The calculation method of this method is undoubtedly very huge, but the estimation of the inclination is quite accurate, but the limitation of this algorithm is still better for the text with more texts. However, if the text contains more Picture, it will make the accuracy of tilt angle estimation reduced.

C. Based on Fourier Transform Method

Fourier transform is widely used in image processing, such as image denoising, image enhancement, image edge detection, image extraction, eigenvalue and image compression. In the image processing, the image of the mutation corresponds to the high-frequency component of the image, which is consistent with the edge of the image; so the use of Fourier transform. Image in the frequency-domain space for different components of the filter, you can complete the corresponding image processing. Fourier transform to detect tilt angle, we must first talk about the image. All pixel Fourier transform, the density of the largest. The direction corresponds to the tilt angle of the scanned image. But this method needs to be done for all pixels. Line Fourier transform, if the image is larger, the amount of computation required will be very large, so it is seldom used to detect the tilt angle.

When the above algorithm is used to process general document scanning images, the algorithm has high detection precision. However, due to mechanical paper feeding or document itself, various kinds of interference factors are inevitably generated during document scanning, for example, illumination, shaking of the machine itself, damage to the edges of paper documents, and the like. These factors lead to different colored lines or background images in the scanned image. The existing tilt angle estimation algorithms have poor detection accuracy for the tilt angle of such images, and also have great errors.

Based on the advantages and disadvantages of the above algorithms and the anti-jamming ability of the scanned image, a new algorithm for detecting the tilt angle and the clipping value of the scanned image is proposed in this paper. The experimental results show that the algorithm has strong anti-jamming ability, Can accurately detect the tilt angle of the scanned image, to achieve automatic correction of scanned images.

III. ALGORITHM IMPLEMENTATION

The method of skew correction of the scanned image in this paper is generally divided into two steps: (1) Detect and obtain the skew angle of the scanned image. (2) The scanned image is rotated to achieve the purpose of correction. From these two steps, we can see that it is very important to detect the angle of inclination of the scanned image, which is the core of the whole technology. Only after the angle of the image is obtained, the subsequent rectifying work can be carried out smoothly.

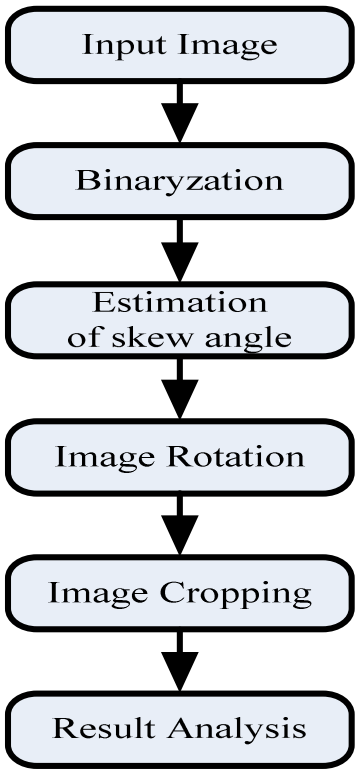

Fig.4 Specific flow of the algorithm

A. Image Preprocessing

Due to the different interference of the scanner in scanning the image, a variety of noise and interfering factors that affect the algorithm are produced in the scanning result, especially in high-volume scanning by mechanical feeding. So in the scanned image correction and cropping before the image to be a certain pre-treatment. Figure5 shows the binarized scanned image.

Fig.5 Binary scanned images

B. Skew Angle Detection

This paper detects the image tilt angle based on the new Hough transform algorithm. The idea of Hough transform algorithm is that the image space is transformed into the parameter space so that the straight line in the image space is mapped to a point in the parameter space and the point in the image space corresponds to a sinusoid in the parameter space. The sinusoid corresponding to the point intersects at a point in the parameter space, and the sinusoid corresponding to the point where the straight line passes through the above two points intersects at the intersection in the parameter space. Using the above characteristics, the problem of line detection in image space can be transformed into the problem of finding the local maximum in parameter space.

This paper detects the tilt angle algorithm is based on the probability of statistical methods to measure the value of tilt angle. First, select a point at regular intervals. For example, take a point (x, y) at every 10 pixels, for example, take three points (x_{n-2}, y_{n-2}) 、 (x_{n-1}, y_{n-1}) 、 (x_n, y_n) , and then calculate the approximate integer value k_r of the slope between two points.

$$k_r \approx \sum_{r=0}^{n} (y_r - y_{r-1}) / (x_r - x_{r-1}) \; r = 0,1,2\cdots \tag{5}$$

If they are the same straight line on the slope and the number of occurrences plus 1, the most frequent occurrence of the scan image is the final slope value. After obtaining the slope value, the tilt angle of the scanned image is obtained by using an inverse trigonometric function.

$$\theta = \arctan(k) \tag{6}$$

C. Image Rotation

After obtaining the tilt angle of the image, this article will re-arrange each pixel of the source image in the new image according to the rotation formula of the image. The following will be the introduction of the principle of image rotation, shown in Figure 6 as the image rotation schematic:

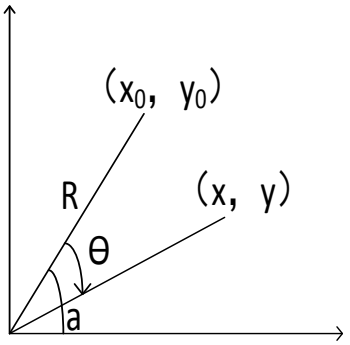

Fig. 6 image rotation schematic

If the rotation angle from point (x_0, y_0) rotation to point (x, y) rotation is the distance from point (x_0, y_0) to point, according to the trigonometric function:

$$\begin{aligned} x_0 &= R\cos\partial \\ y_0 &= R\sin\partial \end{aligned} \tag{7}$$

Then rotate the coordinates:

$$\begin{aligned} x &= R\cos(\partial - \theta) \\ y &= R\sin(\partial - \theta) \end{aligned} \tag{8}$$

From (11) and (12) shows, according to the sine and cosine theorem can be obtained:

$$\begin{aligned} \cos(\partial - \theta) &= \cos\partial\cos\theta + \sin\partial\sin\theta \\ \sin(\partial - \theta) &= \sin\partial\cos\theta - \cos\partial\sin\theta \end{aligned} \tag{9}$$

$$\begin{aligned} x &= R*\cos\partial\cos\theta + R*\sin\partial\sin\theta \\ &= x_0 * \cos\theta + y_0 * \sin\theta \end{aligned} \tag{10}$$

$$\begin{aligned} y &= R*\sin\partial\cos\theta - R*\cos\partial\sin\theta \\ &= y_0 * \cos\theta - x_0 * \sin\theta \end{aligned} \tag{11}$$

According to the tilt angle of the image to calculate the corresponding sine $d_{\sin}$ and cosine $d_{\cos}$ values, the formula as (12) below:

$$\begin{aligned} d_{\sin} &= \sin(\theta * \pi / 180) \\ d_{\cos} &= \cos(\theta * \pi / 180) \end{aligned} \tag{12}$$

Assuming that the pixel coordinates (x_0, y_0) of a certain point in the source image are, and the rotation of the image is calculated after the rotation of the new coordinates (x_{new}, y_{new}) , from (10), (11) and (12) shows a new coordinate calculation formula such as (13) as shown:

$$\begin{aligned} x_{new} &= x * d_{\cos} + y * d_{\sin} \\ y_{new} &= y * d_{\cos} - x * d_{\sin} \end{aligned} \tag{13}$$

After obtaining the new coordinates, we first create an image with no content and the same size as the source image, and then assign the image to the new image based on the

obtained gray value. This operation is repeated for the pixel value of each image of the entire image, that is, the rotation of the image is completed.

IV. IMAGE CROPPING

After the rotation correction and sharpness, we get a processed image with a black background while the latter is out of the display window to further crop the image.

For images that exceed the display window range, automatic cropping of the scanned image becomes easier to understand after you understand the cropping method of the image. After correction, we can regard the frame shape of the scanned image as a polygon, process the scanned image with the method of processing polygons, cut the scanned image, and cut out the part outside the displayed window, as shown in Figure 7 Cropped image:

Fig.7 to be cropped image

For a black background image, the part of the black background outside the valid part of the image is that we do not need, and this part is cut off to obtain the final ideal image. The clipping method used in this paper is to detect the frequency of the black pixel to detect, Scanning pixel value of the probability of statistics to estimate the size of the black border, when the black pixel value exceeds a certain frequency, that is here for the black part of the corresponding cutting operation, the completion of the scanned image processing, cropping The effect is shown in Figure 8. For textual images, strings can also be cropped. Characters can be composed of individual line segments or strokes, or they can be represented as dots. Different from the precision required for cropping, character cropping can be done in different ways. We need to treat each character as an indivisible whole, compare the character box with the crop window, and display the corresponding character if the entire box is inside the window.

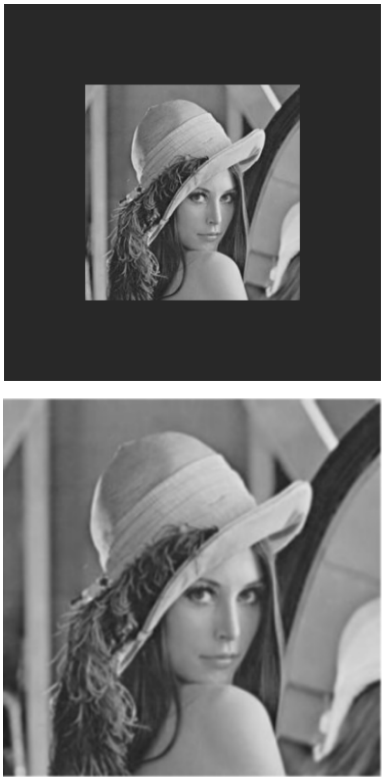

Fig.8 image cutting effect diagram

V. EXPERIMENTS AND RESULTS ANALYSIS

A. Experimental Results

After the actual test, the scanned images of the conventional image tilt correction test, text image tilt correction test, image cutting test, the test results of the scanned image as shown below:

From Figure 9, we completed the correction of the tilt of the normal image. The results show that this algorithm can correct the tilt of the non-document image.

(a) The source image (b) Corrected image

Fig.9 conventional image correction effect

The algorithm has good correction effect on both scanned text and scanned normal images. The deviation of the

rectified images from horizontal is less than ± 2 °, which can meet the practical application requirements. As shown in Figure10, the algorithm cuts the effect of the corrected image.

Fig.10 image cutting effect diagram

B. Experimental Results Analysis

According to the experimental results, we can see that the algorithm in this paper not only can automatically cut the scanned image, but also can detect and correct the skewed image, and emphasize the sharpness of the processed image. The algorithm has the effect of normalizing the scanned images, which is different from other scanning textual image rectifying algorithms, which can identify the image types and perform corrective actions according to different image types. The algorithm is fast, accurate and suitable for most situations under the application, you can face the actual completion of the corresponding correction work.

VI. CONCLUSION

this paper presents a Hough transform algorithm based on automatic scanning image correction and cropping, the first to determine whether the scanned image is a text or non-text image, the image preprocessing based on different types of scanned images of the image binarization, Compared with the traditional algorithm has a good anti-interference ability, can accurately detect the scan image skew angle.And the experiment shows the results of the algorithm.

REFERENCES

[1] A. D. Bagdanov and J. Kanai. Evaluation of Document Image Skew Estimation Techniques. *Proc. SPIE, San Jose*, CA, 1996, 2660: 343-353.

[2] J. J. Hull. Document image skew detection: Survey and annotated bibliography. Document Analysis Systems II, J. J. Hull and S. L. Taylor, Eds. *Singapore: World Scientific*, 1998: 40-4.

[3] P.V C.Hough. Method and means for recognizing complex pattern [P]. *U. S. Patten 3069654*, 1962.

[4] D. S. Le, G. R. Thoma, and H. Wechsler. Automated page orientation and skew angle detection for binary document images. *Pattern Recognition*, 1994, 27(10): 1325-1344.

[5] A. Amin and S. Fischer. A document skew detection method using the hough transform. *Pattern Anal*. Appl., 2000, 3(3): 243-253.

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

[7] Dance C R. Perspective estimation for document images [C]. Electronic Imaging 2002. *International Society for Optics and Photonics*, 2001: 244-254.

[8] Gatos B, Pratikakis I, Perantonis S J. Adaptive degraded document image binarization [J].*Pattern recognition*, 2006, 39(3): 317-327.

[9] Kavallieratou E, Fakotakis N, Kokkinakis G. Skew angle estimation for printed andhandwritten documents using the Wigner-Ville distribution [J]. *Image and Vision Computing*,2002, 20(11):813 – 824.

[10] Lu Y, Tan C L. A nearest-neighbor chain based approach to skew estimation in document images [J]. *PATTERN RECOGNITION LETTERS*, 2003, 24(3):2315-2323.