Journal of Network Intelligence ©2021 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **6**, Number **2**, May **2021**

Correction of Slanted Text Pictures Based on Modified Opencv

Ran Lu, Yongjin Xu, Zhanhe Wu

State Grid Zhejiang Marketing Service Center
Hangzhou 311100 , China
451928256@qq.com, xvyongj@126.com, zhanhe.wu@holley.cn

Yijie Zheng, Lingxin Ren, Feifei Ding, Jianjun Li*

School of Coomputer Science and Technology
Hangzhou Dianzi University
No.1 of the 2nd street, XiaSha, Hangzhou, 310018, China
936177378@qq.com, 1137026854@qq.com, dingfeifei7@gmail.com, jianjun.li@hdu.edu.cn

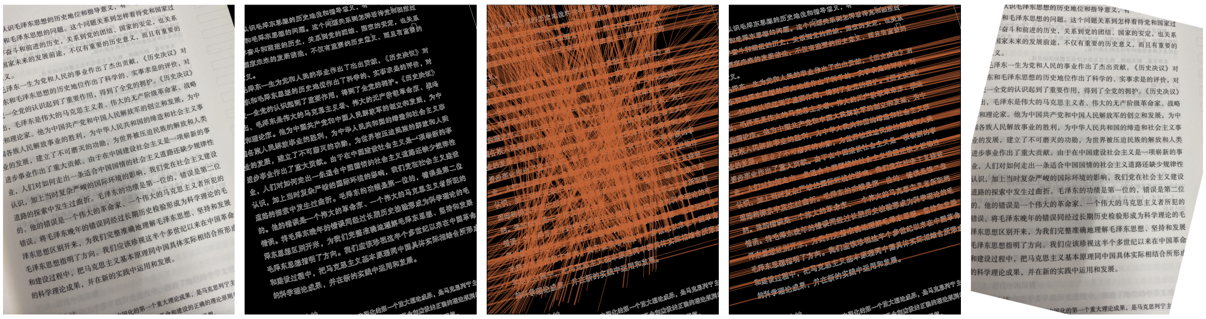
Chin-Chen Chang*

Department of Information and Engineering Computer Science
Feng Chia University
No. 100, Wenhwa Rd., Seatwen, Taichung, Taiwan 40724, R.O.C.
alan3c@gmail.com

Received August 2020; revised December 2020

Abstract. *Correction of slanted or tilted text is very important in many applications, such as optical character recognition (OCR), text detection and recognition. It is not difficult to find that the recognition accuracy of OCR has a great relationship with the tilt angle of the recognized picture, resulting in reduced OCR adaptability and great limitations. As for the problem, we proposed an innovative optimization to estimate the angle of tilting texts in a picture based on Hough transform and modified OpenCV, which reduces the negative effect of background noises. Our algorithm outperforms the state-of-the-art methods in both performances and robustness.*
Keywords: Image recognition; Hough transform; Tilt correction; OCR

1. **Introduction.** Correction of tilted text is an important problem in the field of image processing. It has a wide application in the real world such as the traffic sign automatic recognition, geographic information system, optical character recognition (OCR), etc. Take OCR recognition as an example, OCR is widely used everywhere. Today's students used to use the apps, such as homework help and APE search, to download the homework, complete them and then upload to the teacher. It is not difficult to find that when uploading photos the user prompts to take a picture along the horizontal line because the recognition error rate will greatly increase once the tilt angle of the picture is too large. That is to say, the tilt of the picture seriously affects the recognition accuracy of OCR. However, it is hard to ensure that the input picture is fixed without any tilt in reality, for example manufactories process pictures or in a rush time, the picture is always placed randomly. So the direction correction has to be performed before OCR recognition, so that the accuracy of text recognition can still be guaranteed under various conditions, and its robustness is improved [1].



(a) Input image (b) Binary image (c) Lines detected by Hough transform (d) Lines detected after optimized (e) Final corrected image

FIGURE 1. The overview of our proposed method. Each steps are concluded in the subtitle and shown the correction result.

Although the application scenarios of text image tilt correction and the processing methods involved are different, the basic principles are roughly the same. In this paper, we propose a ruler optimization selection method based-on Hough transform to further reduce the influence of noise and enhance robustness. As shown in Figure 1, we firstly grayscale the input image and obtain the binary image through edge detection and preprocessing. Then, the straight line detection is carried out based on Hough transform. Different with the method of directly setting threshold, we adopt the novel way of optimizing parameters in the process of straight line detection. On the other hand, the straight lines detected by the initial Hough transform contain a large number of noisy lines, which directly affects the final correction result. Therefore, our approach exploits a ruler optimization method to suppress the noise straight lines and calculate the final rotation angle. Finally, we obtained the rotation angle through the straight lines detection and completed the correction of the tilted text image. Our approach is also able to be applied to other scenarios such as license plate, ID card and graphic text.

2. **Related work.**

2.1. **Comparison of various methods for text image correction.** Nowadays, there are many excellent methods in the field of skewed text image correction, such as line-based detection, k-nearest neighbor cluster-based, projection-based methods, and contour-based methods [2–8].

The projection-based correction method detects the tilt of the picture by using the horizontal and vertical projection features of the picture in which the statistical characteristics such us the variance and mean square deviation of the projection, the projection feature vector, and the gradient direction field are main concerns [9, 10]. However, this method is only suitable for the detection of the inclination of the plain text image. As the image becomes larger and the complexity of the layout increases, the probability of error greatly increases. The method based on straight line detection is to detect the straight line in the image through various transformation algorithms, calculate the inclination angle, and then rotate through the angle. The K-nearest neighbor rule is to find the K nearest neighbors of all connected centers, calculating the vector direction of each pair of neighbors and statistics to generate a histogram and the peak of the histogram corresponds to the inclination of the entire image. The contour-based method is to calculate the inclination angle from the contour of the image itself, but this method has the greatest limitation and cannot be detected without a regular boundary [11].

An image may contain many lines, drawing, texts and so on, but the principal direction of the image is always along with the text inside. Therefore, detecting the text direction is very important and useful to determine the direction of the image [12]. Among them, straight line detection includes Hough transform method, least square method and two-point methods [13–16]. Recently, Bera et al. [17] use two novel core-region detection techniques to estimate both the slope and slant angles of offline handwritten word images. It improves the performance of word and character recognition systems. To improve the robustness of complex scenario such as multi-skewed text lines or overlapping words. Kundu et al. [18] applied Generative Adversarial Networks (GANs) to perform Text-line extraction, which produces impressive results.

The Hough transform method is widely used because of its good anti-interference ability and extremely high recognition accuracy. It can also recognize non-continuous straight lines that cannot be recognized by the least square method and the two-point method, so we chose to implement the straight line based on the Hough transform method.

2.2. Related work of hough transform.

2.2.1. *The basic principle of Hough transform.*

A straight line $y = k * x + b$ can be determined by two points $A = (X1, Y1)$ and $B = (X2, Y2)$;

At the same time, $y = k * x + b$ can be written as an expression about (k, b);

$$\begin{cases} b = -k * x_1 + y_1 \\ b = -k * x_2 + y_2 \end{cases} \tag{1}$$

Since the parameter k can be taken to be infinite, we use polar coordinates to represent straight lines. For any point in the image space, we have an expression: $\rho = x * cos\alpha + y * sin\alpha$. Therefore, a point under the Cartesian coordinate system can be converted into a straight line under the polar coordinate system; a sinusoidal curve under the Cartesian coordinate system can be converted into an intersection point of many curves under the polar coordinate system; Therefore, the Hough transform is to convert the problem of a straight line in the Cartesian coordinate system to the problem of finding the intersection point in the polar coordinate system.

2.2.2. *The main steps of the Hough transform.*

Create a discretized grid space, corresponding to ρ and α, respectively;

Perform edge extraction to obtain the edge information of the image;

To traverse all points in the space, we take the vote method to, that is, accumulate and add 1 to the counter of this point;

Traverse the $\rho - \alpha$ space to find the local extreme point, the coordinates of these points (ρ, α) are the possible straight line in the original image.

The basic strategy is Hough transform: The edge data points in image space are calculated to update accumulation and then selects local extreme values zoomed. In general, a straight line can be detected by finding the number of curves that intersect at a point in the plane. The more curves intersecting at a point means that the lines represented by this intersection consists of more points. We can define the number of curves intersecting at a point by setting the threshold of points on the line.The Hough transform tracks the intersection between the corresponding curves of each point in the image. If the number of curves intersecting at a point exceeds the threshold, it is regarded as a straight line.

2.2.3. *Advantages and disadvantages of Hough transform.*

Table 1. Given a fixed threshold, the numbers of detected straight lines of different images.

	Length	Width	Detected straight lines
Picture1	3024	4032	233
Picture2	3024	3930	178
Picture3	3024	4032	94119
Picture4	968	910	4

Advantages: Hough transfer in line detection can convert the more difficult global detection problem into solve the local peak problem in a linear parameter space and it is relatively easy to be solved. This method has a strong capability of anti-interference and it is insensitive to some phenomenons, such as incompleted image lines, noise image, unstructured image and so on.
Disadvantages: The efficiency of the Hough transform is not very high since its time complexity is the level of $O(n^2m)$ in the detection process. Only the straight line can be detected when using Hough transform since it ignores the length information of the line. Also, the detection accuracy is constrained by the discrete parameter intervals since the detection process with Hough transform is discretized.

3. **Method.**

3.1. **Optimization of manual tuning.**

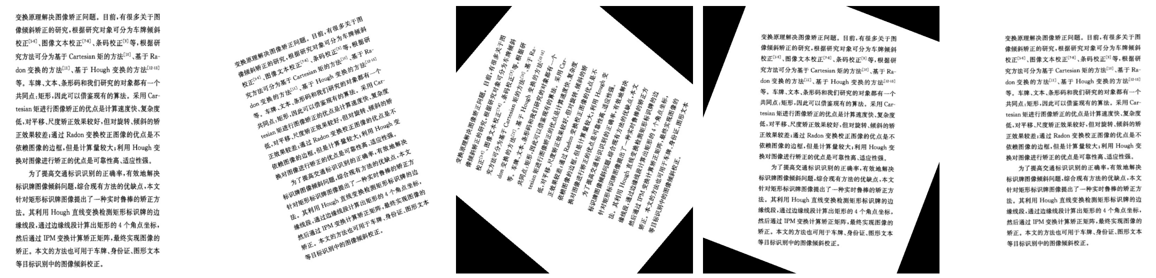
We selected a large number of pictures containing text to test the processing efficiency of the process, and found that between different pictures. There is a big difference in management efficiency.

As shown in Table 1, when the threshold is 370, the processing times of different images are different, and the numbers of detected straight lines are also very different.

The above data shows that the main factor that determines the processing efficiency is the threshold parameter of the accumulated count value, that is, the fifth parameter in the function. The larger of the value and the stricter of the requirement for straight line judgment. Also the less of the number of noise straight lines extracted and the less of the number of operations; On the contrary, the smaller of the value and the greater of the number of noise straight lines extracted and the more of the number of operations. After a large amount of calculations, we initially obtained an experimental conclusion and the result is shown in Table 2. When the extracted feature lines are about 700, the comprehensive performance of the operation efficiency and recognition accuracy of the process is the best.

Obviously, this threshold parameter is affected by factors such as content, picture quality, light, etc., and manual tuning is not realistic, so we thought of designing an effective algorithm that can automatically determine the appropriate threshold parameter according to different pictures.

It is easy to find that the threshold size and the number of detected lines are positively correlated, so we used the idea of dichotomy to determine the upper and lower bounds of a threshold, and the dichotomy makes the final threshold parameters more satisfactory in the conclusion. The time complexity of the subdivision is $O(log_2 n)$, and the exponential efficiency calculation will only cause a very small increase in the amount of calculations



(a) Input image (b) The input image rotated by 20 degrees (c) Corrected image by Hough transform (d) Corrected image by projection-based method (e) our method

FIGURE 2. The Comparison results of different methods after correction.

TABLE 2. Given a threshold of the best performance, the numbers of detected straight lines of different images.

	Length	Width	Detected straight lines
Picture1	3024	4032	691
Picture2	3024	3930	700
Picture3	3024	4032	729
Picture4	968	910	693

on the original process, but it can get accurate threshold parameters, which greatly optimizes the overall process. It also solves the problem of manual tuning and enhances the adaptability and robustness of the process.

3.2. **Optimization for eliminating the effect of disturbance noise.** We selected a large number of images containing text to test the accuracy of the process. The test results show that when the test image contains only text or contains large amounts of text, it has a better performance, but when the picture contains legends, charts, etc. with clear straight line content, or when the picture has a lot of noise, the rotation angle is not accurate enough. This is due to the influence of disturbance factors, which makes some of the extracted noise straight lines fit the distribution of disturbance factors. Our initial implementation was to record the average of the tilt angles of all the noise straight lines to become the final rotation angle. In this case, it has a great error effect on the final result.

After many experiments, it was found that the extracted feature line is still full of the correct rotation angle of the figure, but the wrong feature line is much different from this angle. Once the error amount is slightly larger, the result will be produced. Very deviating. So we thought of choosing the mode or the most inclined angle interval as the final result, excluding the influence of the disturbance line from the amplitude.

The specific method is: the idea of ruler selection is mainly used, that is, all feature lines are sorted into extremes, divided into equal-length small intervals, and the median of the interval containing the largest number of feature lines is taken as the final rotation angle. Since the complexity of sorting is only $O(nlog_2(n))$, and the number of characteristic

TABLE 3. The quantitative experimental results of our method on different rotated images, which tilts from $-80°$ to $+80°$.

Actual inclination	Average detection degree	Average error	Accuracy (%)	Average running time/s
-80	-80.23	0.23	99.71	1.76
-70	-70.31	0.31	99.55	1.81
-60	-59.86	0.14	99.76	1.81
-50	-50.14	0.14	99.72	1.72
-40	-40.24	0.24	99.4	1.77
-30	-29.92	0.08	99.73	1.83
-20	-19.63	0.37	98.15	1.82
-10	-10.10	0.1	99	1.80
10	10.12	0.12	98.81	1.85
20	20.21	0.21	98.96	1.89
30	29.75	0.25	99.16	1.88
40	40.11	0.11	99.72	1.78
50	49.99	0.01	99.98	1.75
60	60.12	0.12	99.8	1.81
70	69.87	0.13	99.81	1.79
80	80.35	0.35	99.56	1.77

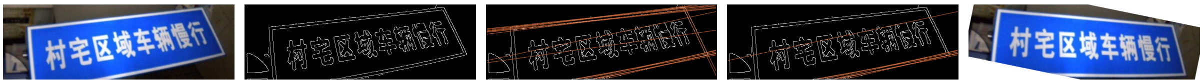
TABLE 4. Comparison results of detection degree with different methods.

method	Actual inclination	detection degree	Accuracy (%)
Projection-based [9]	20°	21°	95.2
Hough transform [10]	20°	40°	50
CTPN [19]	20°	20.8°	96.4
Ours	20°	19.5°	97.5

straight lines n is optimized around 700, this operation will not have substantial impact on the running time. The details of the proposed method are described in Algorithm 1.

4. **Experiment.** This algorithm is implemented in C++ language. The test environment is Intel Core i5 2.9GHz CPU, 8G memory, OSX operating system. We randomly intercepted 100 images with different typesetting content from the Baidu library, and rotated them to varying degrees, tilting from $-80°$ to $+80°$, intercepting every 10°, finally 1600 images were obtained, and then the algorithm described above was used to correct the tilt of the images.

The results are shown in Table 3, it is seen that the average accuracy of the algorithm reaches 99.43%, which shows that the accuracy of the algorithm is higher. In addition, in the time test, the experimental results show that the average operation time of this algorithm is about 1.8s under this environment, and the operation time of the charts of various sizes and degrees is extremely different, which verifies that it is on different charts.



(a) Input image (b) Binary image (c) Lines detected by Hough transform (d) Lines detected after optimization (e) Final corrected image

FIGURE 3. The correction result on oblique text images of road traffic signs.

TABLE 5. The ablation experiment result.

Actual inclination	Hough transform [10]	Hough transform with optimization	Accuracy improved (%)
20°	5.66°	19.5°	39.3

Algorithm 1 : The overview of our proposed Algorithm.

Require:
Slat image I_n;
The initial search threshold of characteristic straight lines $l = 1$ and $r = 700$;
Ensure:
The corrected image O_n;

```
Slat image I_n transform to binary image B_n;
while (l ≤ r) do
    mid ← (l + r)/2
    lines ← HoughLines(mid)
    (l, r) ← update(l, r)
end while
lines ← store(lines)
lines ← sorted(lines)
degree ← average(array(lines))
rotatedmage ← rotateImage(degree)
O_n ← Rect(srcWidth, srcHight)
return Corrected image O_n;
```

In addition, we also conducted comparative experiments on different methods. Among of them, we choose two classical methods for comparison. One is projection-based, the other is Hough transform. As shown in Table 4, our method obtain the accuracy of 97.5%, which performance better than the other two. Furthermore, the visualization result in Figure 2 also shows that our method is able to correct the tilted image effectively.

Since our slant correction method is based on Hough transform, the noise straight line has a great error effect on final correction result. Therefore, we conduct ablation experiment to validate the proposed optimization method is able to eliminate the effect of disturbance noise. As shown in Table 5, our optimization method achieves 39.3% accuracy improved. It is also consistent to the results of (c) (d) in Figure 1. When the optimization method apply in Hough transform, the noise straight line is eliminated sharply. In addition, we also conduct experiments on oblique text images of road traffic signs. The experimental results are shown in Figure 3, it can be seen that our method also achieves good performance in the case of non-plain text, which further indicates that our approach has a broad application prospect.

5. **Conclusion.** After discovering the limitations of the tilt angle of the image on the OCR technology, we proposed and designed a set of programs based on Opencv, using the Hough transform line detection method as the core, to implement an efficient algorithm for the rotation of the tilted text image. In the experiment, we found the inadequacies and shortcomings of the initial version. Combining the knowledge learned in the algorithm, we gave a personal solution or optimization method, such as determining the threshold value by a minute and determining the rotation angle by a ruler, This makes the final version of the program effective, accurate, and robust in various environments. Experimental results show that the algorithm realizes the detection and correction of large-scale text image tilt, and has good performance in time and space, and has strong practicality.

Acknowledgements. This work was supported in part by National Science Fund of China no.61871170; Key Research and Development Plan of Zhejiang: No .2021C03131; The Basic Research Program of KY2017210A001; Key Laboratory of Brain Machine Collaborative Intelligence of Zhejiang Province.

REFERENCES

[1] B. B. Fast and D. R. Allen. Ocr image pre-processor for detecting and reducing skew of the image of textual matter of a scanned document, US 1997.

[2] C. A. Boiangiu, A. Topliceanu, and I. Bucur. Efficient solutions for ocr text remote correction in content conversion systems. Control Engineering and Applied Informatics, vol. 15, no. 1, pp. 22–32, 2013.

[3] R. Islam, M. R. Islam, and K. H. Talukder. An approach to extract text regions from scene image. In 2016 International Conference on Computing, Analytics and Security Trends (CAST), vol. 1, no. 1, pp. 138–143, 2016.

[4] F. Y. Bu, C. S. Liu, and X. Q. Ding. A fast algorithm of skew detection and correction on gray business card image. Journal of Chinese Information Processing, vol. 18, no. 1, pp. 62–69, 2004.

[5] F. Y. Xie, Z. G. Jiang, and L. Wang. Skew detection for complex document image based on blank bar directional fitting. Journal of Computer Applications, vol. 26, no. 7, pp. 1587–1596, 2006.

[6] L. R. Dueire and A. B. Tenorio. A New Algorithm for Skew Detection in Images of Documents. In 2004 International Conference Image Analysis and Recognition (ICIAR), vol. 3212, pp. 234–240, 2004.

[7] F. F. Zeng, F. F. Wu, X. Ke, and X. Wang. A fast correction method for text image tilt. Computer Applications and Software, vol. 32, no. 4, pp. 181–185, 2015.

[8] T. Li and Y. Zhang. Correction of perspective text image based on gradient method. In 2010 International Conference on Information, Networking and Automation (ICINA), vol. 2, pp. 312–316, 2010.

[9] S. L. Zhang, W. B. LI, and J. JI. Skew correction method for document image based on projection. Computer Engineering and Applications, vol. 46, no. 3, pp. 166–168, 2010.

[10] J. T. Li, H. Q. Zhao, Q. Y. Peng, G. L. Zhan, and Y. F. Li. Document image skew correction method based on characteristic sample point detection and hough transform. Journal of Convergence Information Technology, vol. 7, pp. 576–584, 2012.

[11] W. Long. Research on partial differential equation text image layout detection algorithm based on projection histogram method. Journal of Jiamusi Vocational Institute, vol. 0, no. 2, pp. 271–273, 2018.

[12] A. Herout, M. Dubska, and J. Havel. Variants of the hough transform for straight line detection. In Real-Time Detection of Lines and Grids; Springer: London, UK, vol. 22, no. 6, pp. 3–16, 2013.

[13] H. F. Li, D. Pao, and R. Jayakumar. Improvements and systolic implementation of the hough transformation for straight line detection. Pattern Recognition, vol. 22, no. 6, pp. 697–706, 1989.

[14] S. M. Qiu and Y. R. Xia. A new fast hough transform algorithm. Computer Engineering, vol. 30, no. 2, pp. 148–150, 2004.

[15] Y. H. Zhang. Nameplate ocr image rotate correction method based on hough transform. Electrical Measurement and Instrumentation, vol. 8, no. 6, pp. 125–128, 2015.
[16] Y. QU and L. P. Yang. Hough transform ocr image slant correction method. Journal of Image and Graphics, vol. 6, no. 2, pp. 178–181, 2001.
[17] S. K. Bera, A. Chakrabarti, S. Lahiri, E. H. B. Smith, and R. Sarkar. Normalization of unconstrained handwritten words in terms of slope and slant correction. Pattern Recognition Letters, vol. 128, pp. 488–495, 2019.
[18] S. Kundu, S. Paul, S. K. Bera, A. Abraham, and R. Sarkar. Text-line extraction from handwritten document images using gan. Expert Systems with Applications, vol. 140, no. 112916, pp. 1–12, 2020.
[19] T. L. Zheng, X. F. Wang, and X. Xu. A novel method of detecting chinese rendered text on tilted screen of mobile devices. IEEE Access, vol. 99, pp. 1–8, 2020.