Circle Text Expansion as Low-rank Textures

Xin Zhang[*] and Fuchun Sun[†]

[*]*Department of Computer Science and Technology,Tsinghua University*
State Key Lab of Intelligent Technology and Systems
Tsinghua National Laboratory for Information Science and Technology, Tsinghua University
Beijing, China,100084
Email: xinzhang1111@gmail.com
[†]*Tsinghua University, Beijing, China*
Email: fcsun@mail.tsinghua.edu.cn

Abstract—Circle ring aligned text is very common in our daily life, such as university logo, advertisement, there are quite few methods to expand the circle text which will greatly improve the working range of optical character recognition (OCR) product and the accuracy of text segmentation. In this paper, a new method is proposed to handle this circumstance, called curve Transform Invariant Low-rank Textures(TILT). By change the Cartesian system into polar system, the transformed image matrix D can be decomposed into low-rank matrix A and a sparse error matrix E. Matrix A represent the text expansion image and E is the noises and other non-regular component of text image. All this consist of an optimized convex problem and can be solved by alternating direction method (ADM) method. The proposed method also provides a frame work for curve text expansion. Extensive experiments show the robustness of proposed method in expanding artificial and real text image, which contain English or Chinese texts.

Keywords-circle expansion; curve TILT; low-rank matrix; sparse error;

I. INTRODUCTION

Circle distribution text is very common in our daily life, such as the university logo, the advertisement, and the product name. Although the text looks beautiful, it makes the character recognition and segmentation harder. Thus, expanding the circle text into horizontal way can solve the above two problem. Because expanding circle ring is also optical character recognition (OCR)[1], [8] free and OCR can only deal with small rotation, so this function will definitely improve the OCR efficiency.

While many ways of text rectification has been proposed for skew and rotate correction, as far as we know, few methods works on circle distribution text. As the online translation becomes more popular, the picture download from internet, is hard to OCR to recognize, let alone translation. Tradition methods such as hough transform method [4], [2]does not works well because hough circle detection method[3] can find many circle and it is hard for hough transform works for most image with a single threshold. In this paper, we focus on changing the distribution of text from circle to line. Instead of using edge detection to estimate the parameter of circle image, we use the rank information to estimate the parameter. Why rank? Because rank represents the regularity of image texture and can integrate both the local and global information. Inspired by recent work on transform invariant low-rank textures (TILT)[13]our method is the application of TILT for circle distribution text correction. However, using TILT directly cannot solve the problem because TILT meant to solve the skew and rotate problem. We first propose to change the coordinate system from Cartesian system to polar system and then using different map transform to correct the circle distribution text.

Contributions.: In this paper, Because the proposed method based on the instinct symmetry structure and text regularity , it is less sensitive to the input points so it is more robust compared with the traditional hough method.

The main contributions of this paper are as follows:

- A new way for circle image expansion provides a frame work for curve text image correction. We can extend it to other curve text, such as ellipse text and quadratic curve. Only by changing the map transform matrix so long as the text image has some regularity or some symmetry structure.
- Compared with traditional hough method, our method do not need an accurate input and can still recovery the correct position of circle aligned text .This is because we integrate the low rank and sparse error to text structure and solve the convex optimization in every iteration. This model can reflect both the global structure and local symmetry of text and this is the reason why it outperforms the traditional method.
- The proposed method can applied to different language, such as English and Chinese. The proposed method can also handle the real image in our daily life which has much noise. The proposed method can greatly enlarge the work range for Optical Character Recognition product and increase the accuracy of text segmentation for circle image.

In next part , we will explain how to choose the feature for curve text so that it can meet the low-rank and sparse error requirement. In section 3, we describe the flow chart

of our algorithm. Section 4 explained the three main part of curve TILT algorithm in detail. Compare experiments and examples are shown in section 5.

II. Curve text Features

The circle ring text image can be represent by four parameters, the external circle radius, inner circle radius, the center of two circle. Here for simplicity, we consider two circles share the same center and it is always common especially for artificial image. Then we need some information to estimate those four parameters. Because user has already provide four points, we can calculate the initial parameters. Now we need the cue to detect wether the update parameters is correct. And the cues we used here are the low rank part and spares error of the input image. However, for the circle ring text image, because of its symmetry , general TILT may 'think' this position is always the correct position for text. So using general TILT can not solve this problem. The problem here is how to choose a good feature of circle image. A good TILT feature should have some properties: 1) it can be represented by a low-rank matrix and a sparse error; 2)it must distinguish the correct position and wrong position using the object function. 3) it is easy to calculate and have a sharp difference between right and wrong position. In order to meet the above requirement, we make some changes to the general TITL and get the Curve TILT algorithm. In TILT paper, it has already shown that TILT is sensitive to the horizontal and vertical edges. If the those edges become slant, then the TILT can correct them to the correct position. So it may be a good way to change the circle ring text distribution and circle expansion can do this job. Many work has already deal with circle expansion[10],[9], both in computer vision and biology[11]. In the biology, it is commonly used in Iris normalization and recognition. In computer vision, circle expansion is used for symmetry detection[6]. So circle ring expansion is very useful in many areas. In all,the simplest way is change the coordinate system from Cartesian system to polar system and using image interpolation method to generate a rectangle text image. This calculation is simple and efficiency.

III. Curve TILT algorithm

Figure 1 shows the flowchart of Curve TILT. The input image is a curved text image and the output is the circle expansion result and the circle estimated by Circle TILT. The top right image consist two circles, the 'red one' is the external estimated circle and the 'green one' is the inner circle estimated by Circle TILT. The bottom line shows the algorithm process. First, using the four input points and circle function, we estimate the initial $\tau = [R, r, x_0, y_0]$. Expanding the circle ring image and decompose the $D \circ \tau$ image into low rank texture A and the sparse error E, iterate update tau using alternating direction method (ADM) until the A arrives the lowest rank and E become the sparsest

error. (The initial rank (A)=27 and the $\|E\|_0 = 7.1867$ and the result initial rank(A)=26 and the $\|E\|_0 = 5.9566$). To illustrate the improvement for OCR product engine[5], we use a popular OCR engine-hanwang 5.0 version to test the useful of the method.At first, OCR can not recognize a single character, after the circle expansion ,the hanwang OCR product can recognize most of them.
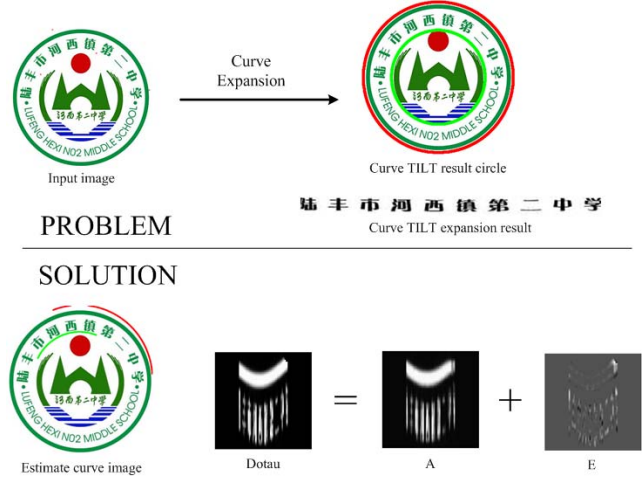

Figure 1.   Flowchart of Curve TILT algorithm

A. Estimate initial circle center and radius

In this paper, we need the user to plot 4 points, two points (x_0, y_0), (x_1, y_1) on external circle and two points (x_3, y_3), (x_4, y_4)on inner circle. Then use circle function(1) to calculate initial parameter (R, r, x0 ,y0).

$$
\begin{bmatrix}
(x_1 - x_0)^2 + (y_1 - y_0)^2 = (x_2 - x_0)^2 + (y_2 - y_0)^2 \\
(x_3 - x_0)^2 + (y_3 - y_0)^2 = (x_4 - x_0)^2 + (y_4 - y_0)^2 \\
R = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \\
r = \sqrt{(x_3 - x_0)^2 + (y_3 - y_0)^2}
\end{bmatrix} \quad (1)
$$

B. Change Coordinate System

In order to expanding the circle ring into a rectangle, one way to solve this problem is using polar system rather than Cartesian system, suppose that both the Cartesian system and polar system has the top-left points as the initial points. And the input image is a circle ring with center (x_0, y_0), and R is the radius of outer circle and r is the radius of inner circle. The output is a rectangle image with M as its height and N as its width. If we look the image as a matrix, then the output image is $I^0 \in R^{M \times N}$ and a pixel in has a coordinates as (x', y'); Thus, the map transform is as follows:

$$
\begin{bmatrix} x' \\ y' \end{bmatrix} \xrightarrow{\tau} \begin{bmatrix} (R-r)\frac{M+x-2}{M-1} \cos(\frac{(y-1)2\pi}{N}) + y_0 \\ (R-r)\frac{M+x-2}{M-1} \sin(\frac{(y-1)2\pi}{N}) + x_0 \end{bmatrix} \quad (2)
$$

Then if the center (x_0, y_0) and the two radius can be estimated correctly, then the output image should be a low rank one and the text are aligned in a horizontal way.

C. Circle Text Expansion Using Low-rank and Sparse Error

However, in practice, all the parameters $[R, r, x_0, y_0]$ are not known to us. How to estimate all the parameter has become a big problem. Inspired by work of Robust Principal Component Analysis(RPCA)[12], we solve this problem using character low rank textures. Both the English and Chinese characters have regular structures. Take Chinese character for example, most Chinese characters have many vertical and horizontal edges while most English characters have some symmetry structures, such as top-bottom symmetry and left-right symmetry. In this way, if we take the output image as a matrix, then lots of its columns and rows will have linear correlation and the rank of matrix will be low if the output image is correct rectified? However, in real world, it is not easy to find a totally symmetry character and often it contain some stroke that are neither horizon nor vertical. In addition to that, we cannot avoid noise in image which will also affect the symmetry of text image. In order to solve this problem, it is better to think the image is roughly 'low rank'. The correct line aligned text image matrix D can be decomposed into low rank texture matrix A and a sparse error matrix E:

$$D = A + E \tag{3}$$

Since we know the instinct feature of text image, what we do next is to find the map transform to meet the above requirement. Inspired by work TILT [13], then the circle ring text rectification problem can be generally formulated as below:

$$\begin{aligned} \underset{A,E,\tau}{Min} \quad & \|A\|_* + \lambda\|E\|_1 \\ s.t. \quad & D \circ \tau = A + E \end{aligned} \tag{4}$$

Where $\|\cdot\|_*$ denotes a matrix's nuclear norm (sum of all singular values) and $\|\cdot\|_1$ denotes a matrix's $\ell 1$ norm (sum of all absolute values). λ is the weight parameter having a positive value. In our work, D is the input circle ring image, A is the low-rank matrix of rectangle image recovered from the circle ring image, E is the sparse error. In this paper, since we already change the coordinate from Cartesian system to polar system, then the map transform τ contains four parameter: $[R, r, x_0, y_0]$.

In order to solve the problem (4), we need to linearize constrains so that the problem become a convex problem, then we get:

$$\begin{aligned} \underset{A,E,\tau}{Min} \quad & \|A\|_* + \lambda\|E\|_1 \\ s.t. \quad & D \circ \tau + J\Delta\tau = A + E \end{aligned} \tag{5}$$

Using alternating direction method (ADM)[7], problem can be solved by formulating augmented Lagrangian below:

$$\begin{aligned} L(A, E, \Delta\tau, Y, \mu) \quad & = \|A\|_* + \lambda\|E\|_1 \\ & + \langle Y, \, D \circ \tau + J\Delta\tau - A - E \rangle \\ & + \frac{\mu}{2}\|D \circ \tau + J\Delta\tau - A - E\|_F^2 \end{aligned} \tag{6}$$

Using (6),update the A,E and $\Delta\tau$ alternately. Because our map transform is different from only affine or projective transform matrix, the Jacobi matrix is totally different from TILT[13], (see the function(2))in our algorithm, the Jacobi is calculated as follows:

$$\nabla I = \frac{\partial}{\partial\varsigma}\left(\frac{I \circ \varsigma}{\|I \circ \varsigma\|_F}\right)|_{\varsigma=\tau} \tag{7}$$

In our experiment, the Jacobi matrix is:

$$\begin{bmatrix} \varphi_x \frac{M+X-2}{M-1} \cos\left[\frac{(y-1)2\pi}{N}\right] + \varphi_y \frac{M+X-2}{M-1} \sin\left[\frac{(y-1)2\pi}{N}\right] \\ -\varphi_x \frac{M+X-2}{M-1} \cos\left[\frac{(y-1)2\pi}{N}\right] - \varphi_y \frac{M+X-2}{M-1} \sin\left[\frac{(y-1)2\pi}{N}\right] \\ \varphi_y \\ \varphi_x \end{bmatrix}$$

Where is $\varphi_x = \frac{\partial}{\partial x}\left(\frac{I \circ \varsigma}{\|I \circ \varsigma\|_F}\right)|_{\varsigma=\tau} = \nabla I_x$,and $\varphi_y = \frac{\partial}{\partial x}\left(\frac{I \circ \varsigma}{\|I \circ \varsigma\|_F}\right)|_{\varsigma=\tau} = \nabla I_y$,and $\tau = \begin{bmatrix} R & r & x_0 & y_0 \end{bmatrix}^T$. For detail of the algorithm, interest reader may refer to TILT paper[13].

IV. EXPERIMENTAL RESULT FOR CIRCLE TEXT EXPANSION

In this section, we do extensive experiments to evaluate the performance of our algorithm and see how well our algorithm can recover the circle distribution text. As hough transform is one of the most popular method in detecting the circle in image, we also compare our result with hough result. For the dataset, we divided the dataset into two parts: one is the standard image, such as the school badge, the other part is the real image taken by camera. Both the dataset contains different size image, different font characters and different languages. The evaluation here is different for different database. For the standard image, we calculate the difference between the calculated $[R, r, x_0, y_0]$ and the ground truth value. As for the real image, there is no ground truth value for image so we see the correct horizontal line of each image. Because of our assumption that the text pixel is distrusted in circle ring. That the text appear in the circle ring should be wider near the external circle compared with the pixel near the inner circle. However, in our daily life, maybe all the ring text share a same width. So that is the reason why our result image has some scale distortion for text even when our parameter result is quite near the ground truth parameter.
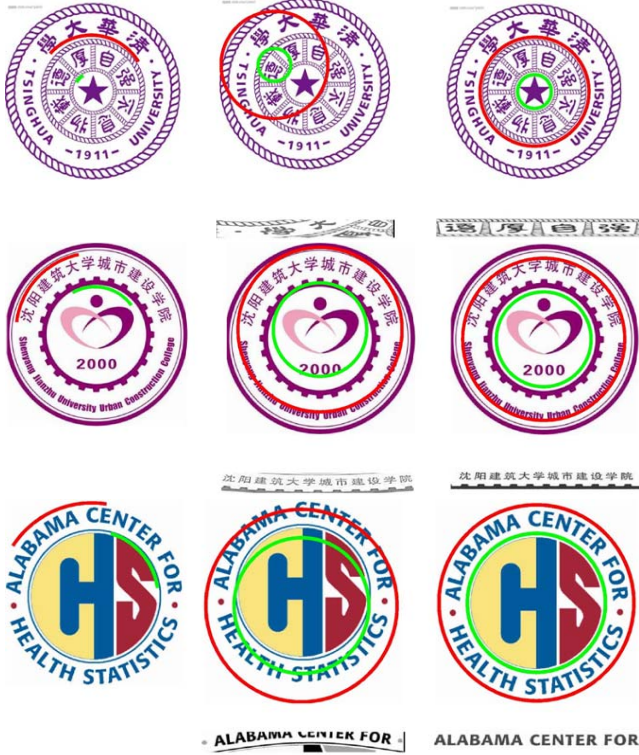
Figure 2. Hough vs Curve TILT method. The first row is the input image, the second row the result of hough method (top is the estimate circle and bottom is the expansion result). The third row is the result of circle TITL method. (top is the estimate circle and bottom is the expansion result).

Figure 3. Hough vs Curve TILT method. The first row is the input image, the second row the result of hough method (top is the estimate circle and bottom is the expansion result). The third row is the result of circle TITL method. (top is the estimate circle and bottom is the expansion result).

A. Expanding Circle Artificial Images

We first test on standard image because it contains less noise so that both hough can work at its best. For standard image, we calculate the difference between ground truth parameter and the rectified solution. If the tau is estimate correctly, then the circle ring is near the ground truth. From the expansion result, we can see how well the algorithm rectifies the circle ring image. The results are as follows:

For the first image, the ground truth of parameter is (R, r, x_0, y_0)=(125 40 281 228) and hough result is (124.6826 38.3187 210.5401 185.4993) and curve TILT result is (124.3286 38.6727 281.4232 229.7196) and for the second image ,the ground truth of parameter is (130.3085 76.9938 150 155.0000) ,the hough result is (132.2546 75.2732 150.7880 142.7361) and curve TILT result is (130.9433 76.5845 150.5178 156.9377). For the third image, the ground truth of parameter is (195 138 225 225), using hough circle detection method ,the result is (195.5411 137.9162 222.8547 252.7713) and using curve TILT method, the result is (195.1260 138.3314 225.6541 225.0884). The estimate result shows that most often TILT can calculate the parameter accurately compared with hough circle detection method. The reason for that is hough circle detection
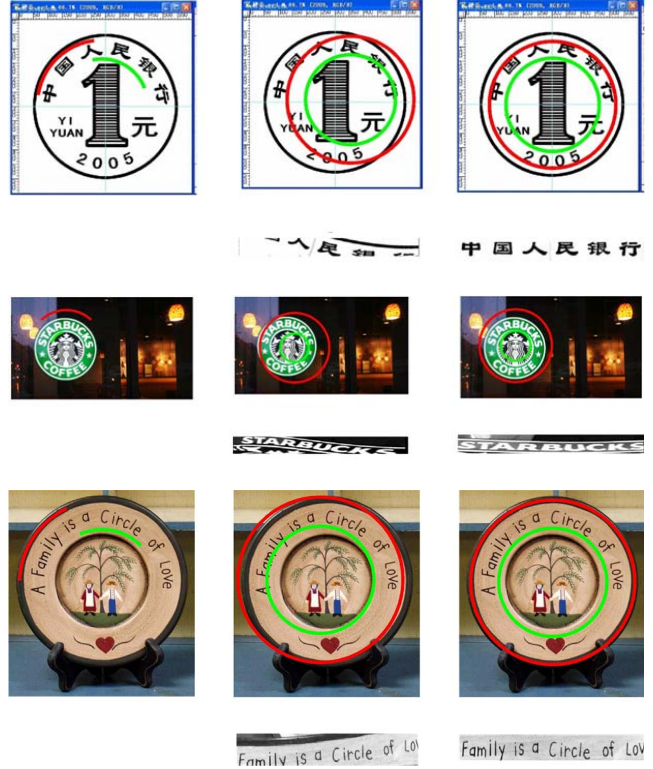
algorithm heavily relied on the first estimate the center of circle as well as internal and external circle radius. If the text image has complex background, say has lots of 'likely' circle edges, it is hard for hough decide which circle is the right text position.In other way,if the initial input is good and the image has small number of circle edges, it can achieve approximately same result with curve TILT method.

B. Expanding Circle Real Images

For real image, because we don't have the ground truth value, we can compare the result subjectively, that is from the aligned rectangle image. If the aligned rectangle text aligns horizontally, then the estimated parameter is close to the ground truth and vice versa. The results are as follows:

In fact, it is much harder to detect circle ring and calculate the parameter in the real image, because of vary illumination, noises and low contrast. All these factor effects the performance of Hough circle detection and Curve TILT algorithm. However, TILT is more robust than Hough because what Curve TILT using is the text regularity and symmetry structure. By harness both global and local structure of text image, TILT still works well even in very complex situation.

Figure 4. Curve TILT method. The top left image is the input image, the top right image is the result of curve TILT method. The bottom image is the circle ring expansion result.

Figure (4)show more results for curve TILT expansion for circle text.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed an efficient algorithm to rectify the circle aligned text into rectangle image using the low rank texture. As we can see that, by change the coordinate system from Cartesian to Polar system and then using the curve map transform is a good way to rectify the circle text image. And the rank is also a good indicator for the right position of text. The rectify image can dramatically improve the accuracy of text segmentation as well as the Optical character recognition product's performance. The current method can only deal with circle ring text. We will find a more general way to rectify the curved text, such as quadratic curve or cubic curve distribution text. In addition to that, we will also consider how to enhance the performance of TILT so that the coverage range becomes larger. Last, now the input is four points which is not so convince for user. Finding a better initial will also be explored in the future.

REFERENCES

[1] M. Cheriet, N. Kharma, C.-L. Liu, and C. Suen. *Character Recognition Systems: A Guide for Students and Practitioners*. Wiley-Interscience, 2007.

[2] S.-H. Chiu, J.-J. Liaw, and K.-H. Lin. A fast randomized hough transform for circle/circular arc recognition. *IJPRAI*, 24(3):457–474, 2010.

[3] T. D'Orazio, C. Guaragnella, M. Leo, and A. Distante. A new algorithm for ball recognition using circle hough transform and neural classifier. *Pattern Recognition*, 37(3):393–408, 2004.

[4] O. Ecabert and J.-P. Thiran. Adaptive hough transform for the detection of natural shapes under weak affine transformations. *Pattern Recognition Letters*, 25(12):1411–1419, 2004.

[5] Hanwang Corp. Hanwang Chinese OCR 5.0.

[6] S. Lee and Y. Liu. Skewed rotation symmetry group detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(9):1659–1672, 2010.

[7] Z. Lin, M. Chen, L. Wu, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices, 2009. UIUC Technical Report UILU-ENG-09-2215.

[8] S. Mori, C. Y. Suen, and K. Yamamoto. Historical review of OCR research and developmeng. *Proceedings of the IEEE*, 80(7):1029–1058, 1992.

[9] C.-H. Park, J.-J. Lee, M. J. T. Smith, and K.-H. Park. Iris-based personal authentication using a normalized directional energy feature. pages 224–232, 2003.

[10] Z. Sun and T. Tan. Ordinal measures for iris recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(12):2211–2226, 2009.

[11] A. Torii and A. Imiya. The randomized-hough-transform-based method for great-circle detection on sphere. *Pattern Recognition Letters*, 28(10):1186–1192, 2007.

[12] J. Wright, A. Ganesh, S. Rao, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices. *CoRR*, abs/0905.0233, 2009.

[13] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma. TILT: Transform invariant low-rank textures. In *Asian Conference on Computer Vision*, 2010.