

Active Flattening of Curved Document Images via Two Structured Beams

Gaofeng MENG Ying WANG Shenquan QU Shiming XIANG Chunhong PAN

National Laboratory of Pattern Recognition (NLPR)

Institute of Automation, Chinese Academy of Sciences (CASIA)

No.95, Zhongguancun East Road, Beijing 100190, P.R. China

{gfmeng, ywang, shenquan.qu, smxiang, chpan}@nlpr.ia.ac.cn

Abstract—Document images captured by a digital camera often suffer from serious geometric distortions. In this paper, we propose an active method to correct geometric distortions in a camera-captured document image. Unlike many passive rectification methods that rely on text-lines or features extracted from images, our method uses two structured beams illuminating upon the document page to recover two spatial curves. A developable surface is then interpolated to the curves by finding the correspondence between them. The developable surface is finally flattened onto a plane by solving a system of ordinary differential equations. Our method is a content independent approach and can restore a corrected document image of high accuracy with undistorted contents. Experimental results on a variety of real-captured document images demonstrate the effectiveness and efficiency of the proposed method.

Keywords-document image processing; geometric rectification; structured beams; developable surface interpolation;

I. INTRODUCTION

As a competitive substitute, digital cameras exhibit great advantages against the flatbed scanners in the digitization of documents [7]. For instance, cameras are portable, fast responsive and allow non-contact imaging of materials that are not suitable for a flatbed scanner. However, camera-based imaging often introduces unwanted geometric distortions. This is especially true when one manages to capture images of an opened thick and bound book (see Figure 1). These unwanted distortions will cause serious problems to the subsequent processing steps and thus need to be first rectified.

In the past several years, geometric distortion rectification of document images has received great attention and many methods are proposed. Representative works include [9], [3], [12], [13], [1], [11], [2], [6], [5], [10], [8]. These methods can be roughly classified into several categories, including the methods assuming a restricted page shape [3], [11], [8], the methods relying on local or global image transformation [6], [12], [10], the methods based on range data [9], [1], [2] and the methods using multiple images of different views [13], [5].

To simplify and make the estimation process stable, a cylindrical surface is often used to model the document shape. Tan *et al.*[11] develop a shape-from-shading technique for rectifying the geometric distortion in a scanned document image. They use a cylindrical surface to represent

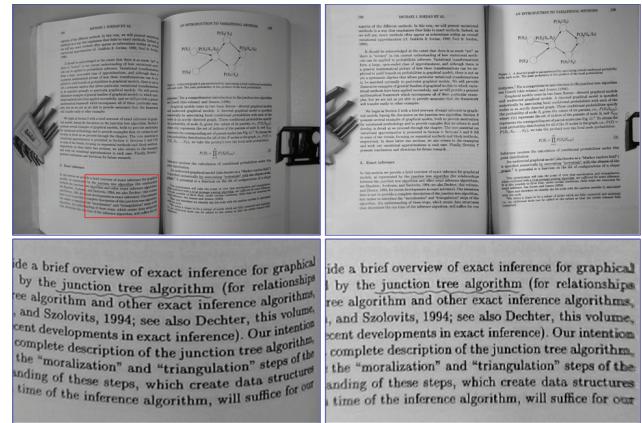


Figure 1. Document images captured by a digital camera often suffer from serious geometric distortions. (left) a curved document image captured from an opened book page, (right) geometric correction result by our method.

the page shape of an opened book. By doing so, they successfully reduce the problem of 3D shape estimation to a problem of 2D directrix recovery. Cao *et al.*[3] propose a cylindrical surface model for rectifying the distortions in a camera-captured document image. To recover the 3D directrix, their method requires the image plane of camera to be parallel to the rulings of the cylindrical page.

In our previous work [8], we noticed that the horizontal text lines are line convergence symmetric after perspective projection. Based upon this observation, we can perform a metric rectification by constructing an isometric mesh-grid from only two curved text lines. Our previous approach also extend Cao *et al.*'s work. It allows the camera to capture documents in an arbitrary view angle and position.

Although cylindrical surface, in most cases, provides a satisfactory approximation to the curved document pages, it fails to documents with a non-cylindrical shape. Observing that most document pages curve into a developable shape, Liang *et al.*[6] use a developable surface to model the curved page. To flatten the document page, a finite number of planar strips are employed to approximate the developable surface. These strips are then rectified piece by piece and finally merged together. However, due to the estimation errors, gaps and overlapping often appear in the neighboring

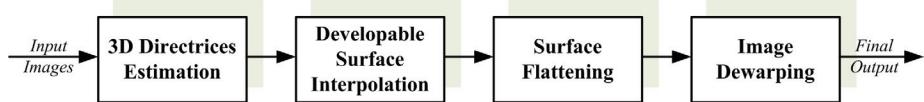


Figure 2. An overview of the proposed document restoration process.

strips. Tsoi and Brown [12] present a boundary interpolation based method to estimate a global warping function. Their method is applicable to a variety of geometric distortions, including skew, folding, binder curl and their combinations. Stamatopoulos *et al.*[10] propose a method that uses a coarse-to-fine strategy to remove the undesirable document image distortions. In their method, a global transformation is first implemented to project a curved surface to a rectangle area. Then pose normalization is performed on word level to restore the local distortions.

The key problem of geometric correction of a document image is to recover its 3D shape. A direct way to this is to use the range data obtained from a 3D scanner [9], [1], [2]. Pilu [9] uses a developable surface to fit the range data and then flatten the surface to yield an undistorted image. Brown *et al.*[1], [2] build a 3D scanning system to obtain the 3D information of a document page. The system consists of two digital video cameras, a high-resolution still digital camera and a laser light source mounted on a PC-controlled pan/tilt unit. The scanned 3D point cloud is first triangulated to obtain a 3D mesh, then a conformal mapping is used to map the 3D mesh onto a plane to rectify the geometric distortions. The methods using 3D range data generally assume no parametric model of the page shape and thus is very suitable for correcting arbitrary distortions. However, the requirements for expensive setups, as well as the limitations in speed (several minutes for scanning one page) and stability, often make them less attractive in applications.

Another way to estimate the 3D shape is to use multiple images captured at different viewpoints [13], [5]. Koo *et al.*[5] describe an interesting method using two images taken from different views. They first register an image pair using SIFT feature points, then epipolar constraints are used to recover the 3D surface. Theoretically, the performance of multi-view based methods depends on image contents. A sufficient number of feature points must be extracted so as to obtain an accurate estimation of the 3D page surface. However, this is often not satisfied to documents with sparse contents. Moreover, SIFT matching often fails to points with large distortions and image blurring. Thus, page shapes near these points are usually underestimated, leading to poor corrections in these areas.

In this paper, we present a method to correct geometric distortions in a document image captured by a digital camera. Our method uses two structured beams illuminating upon the page surface to recover two 3D curves. A devel-

opable surface is then interpolated to the two curves. The surface is flattened later onto a planar surface by solving a system of differential equations to obtain a global dewarping transformation. This transformation is finally used to correct the geometric distortions in the image. Figure 1 illustrates an example of the correction results of our method.

Our method is distinguished by three main highlights. The first is that our method is a content independent approach. It uses active beams rather than image contents, *e.g.*, text lines or image features, to recover the 3D document shape, and can thus be applied to documents with sparse contents. The second is that our method adopts a more general developable assumption on the curved page shape. It can address most types of distortions in documents that fail to methods using the cylindrical assumptions. Furthermore, instead of using 3D points cloud scanned from document pages, only two spatial curves are used to recover the 3D page surface. Thus, our method has a much lower computational complexity and can be implemented very efficiently. Our method is also a fully automatic approach. It does not involve any user interactions or parameters fine-tuning. This makes it very appealing to nonexpert users in applications.

II. THE PROPOSED METHOD

To begin with, we assume that document pages curve into a developable surface. Since real paper is approximately unstretchable and once being bent can be rolled out onto a plane without stretching and tearing, this assumption holds for most documents. Figure 2 shows the overview of the proposed document rectification method. The method consists of four main steps, *i.e.*, 3D directrices estimation using structured beams, developable surface interpolation, surface flattening and image dewarping transformation. More details of these steps will be given in the following subsections.

A. Directrix Estimation via Structured Beams

We designed a system to acquire two spatial curves on the page surface. The system is illustrated in Figure 3(a), which consists of a high-resolution digital camera and two laser light sources. The lasers are fixed on the top bracket and yield two beams of 1-mm width on the document page surface. The digital camera has a fixed focus and is used to capture high-resolution images of the beams and the document page.

To automatically extract the beams from the document image, we capture two images in a continuous shooting mode by toggling on and off the switch of lasers. Figure

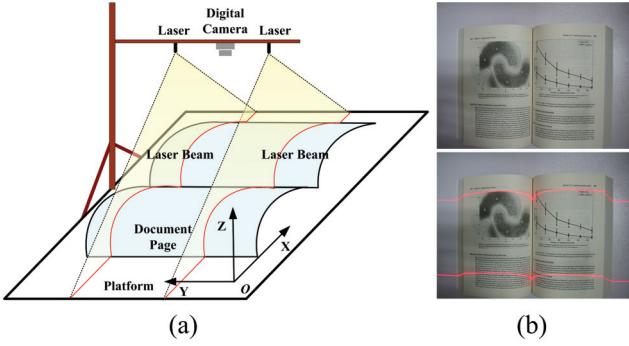


Figure 3. The system used to acquire the 3D directrix of page surface. (a) the designed system, (b) the captured image pair with and without beams.

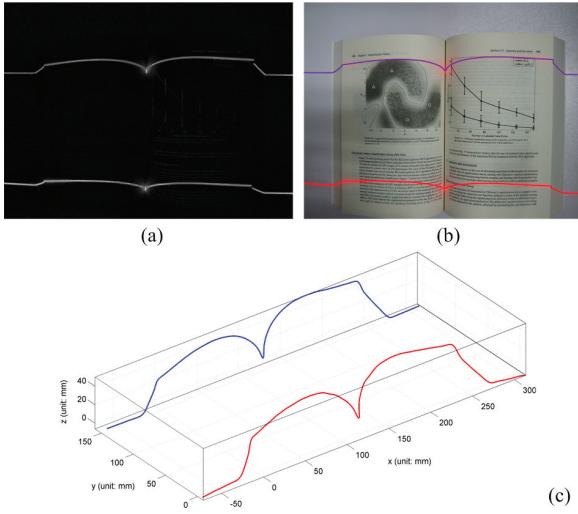


Figure 4. The extracted beams and the recovered directrices of the page surface in Figure 3(b). (a) the difference map of beam images, (b) the extracted beams, (c) the recovered spatial directrices.

3(b) shows the images with and without beams captured by our system. The difference map of the two images is then computed and binarized to locate the beams. The width of the beams is also reduced to a sub-pixel accuracy by fitting a Gaussian curve to its intensity profile.

The reconstruction of the spatial directrices from the beams is straightforward. To begin with, we need to calibrate the digital camera and the laser beams. Once the camera and the beams are calibrated, the estimation of the spatial directrix reduces to the computation of intersections of rays of beam points with the beam plane. The spatial directrix is also fitted by a cubic B-spline curve to connect the broken segments and reduce the fluctuations due to image quantization errors. Figure 4 gives the extracted beams and the recovered spatial directrices of the page surface.

B. Developable Surface Interpolation

A developable surface is a special case of the ruled surface. Denote $\mathbf{C}_0(t)$ and $\mathbf{C}_1(s)$ ($t, s \in [0, 1]$) as two differentiable space curves, respectively. Also define a correspondence $s = \phi(t)$. The correspondence is generally a monotonically increasing function that assigns each point t of $\mathbf{C}_0(t)$ to a point s of $\mathbf{C}_1(s)$. The following equation

$$\mathbf{X}(t, v) = (1 - v)\mathbf{C}_0(t) + v\mathbf{C}_1(\phi(t)), t \in [0, 1], v \in \mathbb{R}, \quad (1)$$

defines a ruled surface, where $\mathbf{C}_0(t)$ and $\mathbf{C}_1(s)$ are called the directrices of the surface \mathbf{X} , and the lines connecting points on $\mathbf{C}_0(t)$ and the corresponding points on $\mathbf{C}_1(s)$ are called the rulings. Note that a directrix of \mathbf{X} will intersect with all the rulings.

A ruled surface is developable if

$$\left\langle \frac{d\mathbf{C}_0(t)}{dt} \times \frac{d\mathbf{C}_1(s)}{ds}, \mathbf{C}_1(s) - \mathbf{C}_0(t) \right\rangle = 0. \quad (2)$$

That is to say, the vectors $d\mathbf{C}_0/dt$, $d\mathbf{C}_1/ds$ and $\mathbf{C}_1(s) - \mathbf{C}_0(t)$ should be coplanar for all points on the developable surface. The notation $\langle \cdot, \cdot \rangle$ above represents the inner product operator.

A developable surface is uniquely determined by two different directrices on it. Given two directrices of a developable surface, namely, $\mathbf{C}_0(t)$ and $\mathbf{C}_1(s)$, the problem of constructing a developable surface that interpolates to the directrices is to determine a proper correspondence $s = \phi(t)$ ($t \in [0, 1]$) between the directrices.

The correspondence function $\phi(\cdot)$ should satisfy the boundary conditions and be monotonically increasing, i.e., $\phi(0) = 0$, $\phi(1) = 1$ and $\phi'(t) \geq 0$. For a smooth document page surface, since the rulings do not intersect within the boundary of the page, we further require $\phi(\cdot)$ to be strictly monotonically increasing. A bound is thus imposed to $\phi'(t)$, i.e.,

$$0 < a \leq \phi'(t) \leq b, \quad t \in [0, 1] \quad (3)$$

where a, b are the lower and upper bounds of $\phi'(t)$, respectively.

An optimal correspondence function $\phi(\cdot)$ can be obtained by solving the following variational problem:

$$\begin{aligned} \min_{\phi(\cdot)} \int_0^1 & \left| \left\langle \frac{d\mathbf{C}_0}{dt} \times \frac{d\mathbf{C}_1}{ds}, \mathbf{C}_1(\phi(t)) - \mathbf{C}_0(t) \right\rangle \right| dt \\ \text{s.t. } & \phi(0) = 0, \phi(1) = 1, a \leq \phi'(t) \leq b \end{aligned} \quad (4)$$

However, it is generally difficult to obtain a closed-form solution of the above problem. Here, we propose a much efficient method that can yields a numerical solution of Eq.(4).

Discretizing t and s between $[0, 1]$ by using a fixed sampling interval Δ_t and Δ_s , respectively, yields two sequences of samples:

$$\begin{aligned} 0 &= t_0 < t_1 < \dots < t_{n-1} < t_n = 1, \\ 0 &= s_0 < s_1 < \dots < s_{m-1} < s_m = 1. \end{aligned} \quad (5)$$

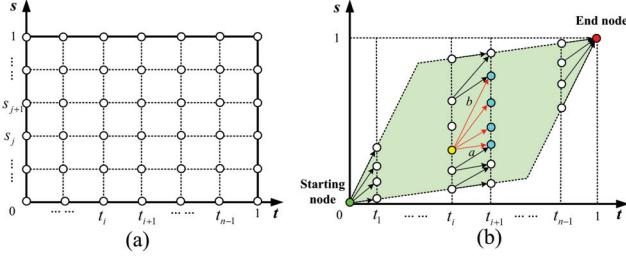


Figure 5. The directed graph G for computing the correspondence function. (a) discretization of the $t-s$ plane, (b) the constructed graph. All the vertices of the graph locate in a parallelogram. The slopes of its edges are a and b , respectively.

We then construct a weighted directed graph G with its vertices located at (t_i, s_j) of the $t-s$ plane, as illustrated in Figure 5(a). In G , only points with adjacent t -coordinates are connected. Two vertices (t_i, s_j) and (t_{i+1}, s_{j+k}) are connected with a directed edge only if the slope of the line passing through the two points falls between $[a, b]$, i.e.,

$$a \leq \frac{s_{j+k} - s_j}{t_{i+1} - t_i} = \frac{k\Delta_s}{\Delta_t} \leq b. \quad (6)$$

A cost matrix of correspondence can also be computed based on the samples. Each of its elements is given as:

$$w_{ij} = \left| \left\langle \frac{d\mathbf{C}_0(t_i)}{dt} \times \frac{d\mathbf{C}_1(s_j)}{ds}, \mathbf{C}_1(s_j) - \mathbf{C}_0(t_i) \right\rangle \right|. \quad (7)$$

$$(0 \leq i \leq n, 0 \leq j \leq m)$$

The edge connecting (t_i, s_j) and (t_{i+1}, s_{j+k}) is assigned a weight $w_{i+1,j+k}$. This weight gives the cost of the correspondence between t_{i+1} and s_{j+k} .

The problem of solving an optimal correspondence function now turns to be finding a shortest path from the starting node $(0, 0)$ to the end node $(1, 1)$ on the graph G , as illustrated in Figure 5(b). This can be done by using the famous Dijkstra's algorithm [4]. Figure 6 illustrates the estimated correspondence function and the constructed developable page surface.

C. Surface Flattening and Image Rectification

Flattening a developable surface can be efficiently implemented by unfolding two directrices on it. This can be done by taking advantage of the invariance of geodesic curvature of curves on the surface.

Let $\mathbf{C}(s)$ be a curve lying on a developable surface S , where $s \in [0, \ell]$ is the arc-length parameter of the curve, and ℓ is the total length of the curve. The curvature vector of the curve at a point P is given by

$$\mathbf{K} = \kappa \mathbf{n}, \quad (8)$$

where κ is the curvature magnitude, and \mathbf{n} is the unit principal normal vector of the curve at point P . The curvature vector \mathbf{K} can be further divided into two orthogonal

components, i.e.,

$$\mathbf{K} = \mathbf{K}_n + \mathbf{K}_g, \quad (9)$$

where \mathbf{K}_n normal to the surface of S is the normal curvature vector, and \mathbf{K}_g tangential to the surface is the geodesic curvature vector. The magnitude of \mathbf{K}_g can be computed by

$$k_g(s) = \|\mathbf{K}_g(s)\| = \langle \mathbf{n}_s(s), \mathbf{C}'(s) \times \mathbf{C}''(s) \rangle, \quad (10)$$

where \mathbf{n}_s is the unit normal vector of the surface S at P .

Geodesic curvature plays an important role on the flattening of a curve on a developable surface. Denote the curve after flattening $\mathbf{C}(s)$ onto a plane by $\mathbf{r}(s)$, where $s \in [0, \ell]$ is the corresponding arc-length parameter. When the developable surface is flattened, the geodesic curvature of any point on \mathbf{C} will remain unchanged. That is,

$$\begin{aligned} k_g(s) &= \langle \mathbf{n}_s(s), \mathbf{C}'(s) \times \mathbf{C}''(s) \rangle \\ &= \langle \mathbf{n}_p, \mathbf{r}'(s) \times \mathbf{r}''(s) \rangle, s \in [0, \ell], \end{aligned} \quad (11)$$

where \mathbf{n}_p is the unit normal vector of the plane. According to this property, we can reconstruct the flattened curve on the plane by solving a system of differential equations.

Assume $\mathbf{r}(s)$ is on the plane $Z = 0$. Thus, we can rewrite $\mathbf{r}(s) = (x(s), y(s), 0)^T$ ($s \in [0, \ell]$). Substituting it into Eq.(11) and taking $\mathbf{n}_p = (0, 0, 1)^T$, we have

$$x'(s)y''(s) - x''(s)y'(s) = k_g(s). \quad (12)$$

On the other hand, recall that s is the arc-length parameter of $\mathbf{r}(s)$. Thus, we also have

$$x'^2(s) + y'^2(s) = \left| \frac{d\mathbf{r}}{ds} \right|^2 = \frac{dx^2 + dy^2 + dz^2}{dx^2 + dy^2 + dz^2} = 1. \quad (13)$$

Supplying some necessary boundary conditions, we can obtain the following system of differential equations:

$$\begin{cases} x'y'' - x''y' = k_g \\ x'^2 + y'^2 = 1 \\ x(0) = x_0, y(0) = y_0, x'(0) = \mu_0, y'(0) = \nu_0 \end{cases} \quad (14)$$

where $(x_0, y_0)^T$ is the starting position of $\mathbf{r}(s)$ on the plane, and $(\mu_0, \nu_0)^T$, satisfying $\mu_0^2 + \nu_0^2 = 1$, is the flattening direction of the curve.

Fortunately, Eq.(14) has the closed-form solutions. Differentiating the second equation in Eq.(14) with respect to s yields

$$x'x'' + y'y'' = 0. \quad (15)$$

Substituting it into the first equation and using the second equation, we have

$$x'' = -k_g y', \quad y'' = k_g x'. \quad (16)$$

Using these equations, we can easily eliminate the terms y' and y'' in Eq.(14), thus obtaining a second-order ordinary differential equation of x :

$$\begin{cases} x''^2 = k_g^2 (1 - x'^2) \\ x(0) = x_0, x'(0) = \mu_0 \end{cases}. \quad (17)$$

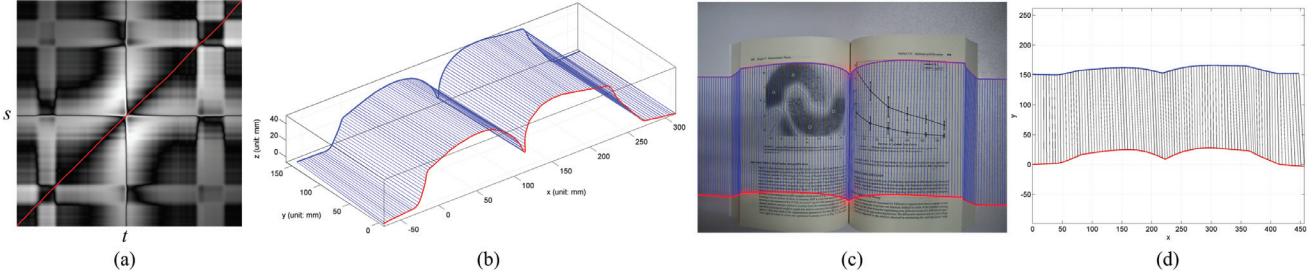


Figure 6. Developable page surface interpolation. (a) the cost matrix of correspondence and the estimated correspondence function (red curve), (b) the interpolated developable page surface, (c) all 3D rulings are also mapped back onto the input image, (d) the surface is flattened by unrolling its two directrices onto a plane.

Solving this equation and ignoring a trivial linear solution yields

$$x(s) = \pm \int_0^s \sin \varphi(t) dt + x_0, \quad (18)$$

where

$$\varphi(t) = \int_0^t k_g(s) ds \pm \sin^{-1} \mu_0. \quad (19)$$

Similarly, substituting Eq.(18) back to Eq.(16), we can solve for $y(s)$ as below:

$$y(s) = \mp \int_0^s \cos \varphi(t) dt + y_0. \quad (20)$$

Generally, the flattening of a curve onto a plane usually has two solutions. Which solution is selected depends on which side of the surface we choose to touch the plane in the flattening process.

A developable surface can be flattened onto a plane by unrolling its two directrices using Eq.(18) and Eq.(20). Then all the rulings can be reconstructed on the plane according to the correspondence between the two curves. Figure 6(d) gives the flattened surface and the reconstructed rulings on the plane. Once the surface is flattened, we actually establish a dewarping transformation. This transformation is then applied to the image to correct its geometric distortions. Figure 7 illustrates the rectification result of the image. A close-up image patch is also given to show the details. In the result, some words on the right side get blurred after image correction. This is mainly due to the insufficient resolution in image areas of large distortions.

III. EXPERIMENTAL RESULTS

A. Example Results

Figure 8 shows the results of geometric rectification on six document images with different types of distortions. To eliminate the effects of shading artifacts on the correction results, all images after geometric correction are also binarized by an adaptive threshold algorithm.

The first image in the figure is captured from a blank page of an opened notebook. From the correction result, we can see that the horizontal reference lines in the image

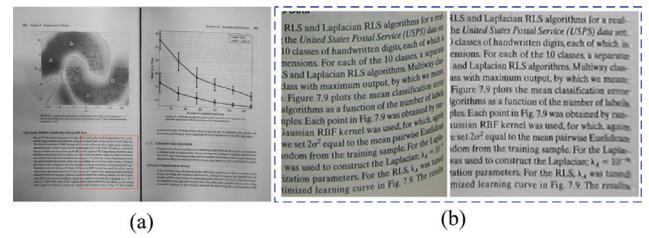


Figure 7. The geometric rectification result. (a) the rectified image, (b) the close-up patches of the original input (left) and the rectified image (right) in the red box.

are all straightened. This indicates that the distortions in the original image are satisfactorily corrected. The second and the third image are captured from pages of two opened books. The two pages have a cylindrical shape with large curvatures. The correction results show that our method can well remove such cylindrical distortions in the images.

To test the performance of our method on non-cylindrical page shapes, we take a sheet of paper and manually squeeze it into different non-cylindrical shapes. The fourth and the fifth image in the figure illustrate two such examples. From the results, we can see that our method can also well correct non-cylindrical distortions in the captured images.

The last image is captured from a sheet of paper with folding distortions. Since the page surface is not differentiable along the folding lines, the page is not a developable surface but an applicable surface. The correction result shows that our method can still yield good results for this case. However, due to a smooth approximation to the folding areas, small distortions also remain near the folding lines in the corrected results. This may be further improved if one can use a non-smooth approximation to the extracted spatial directrices of document surface.

B. Quantitative Evaluation

To make a quantitative evaluation on the rectification results of our method, we capture six images of a checkerboard pattern on a sheet of paper by manually squeezing it into different shapes. The checkerboard pattern consists of

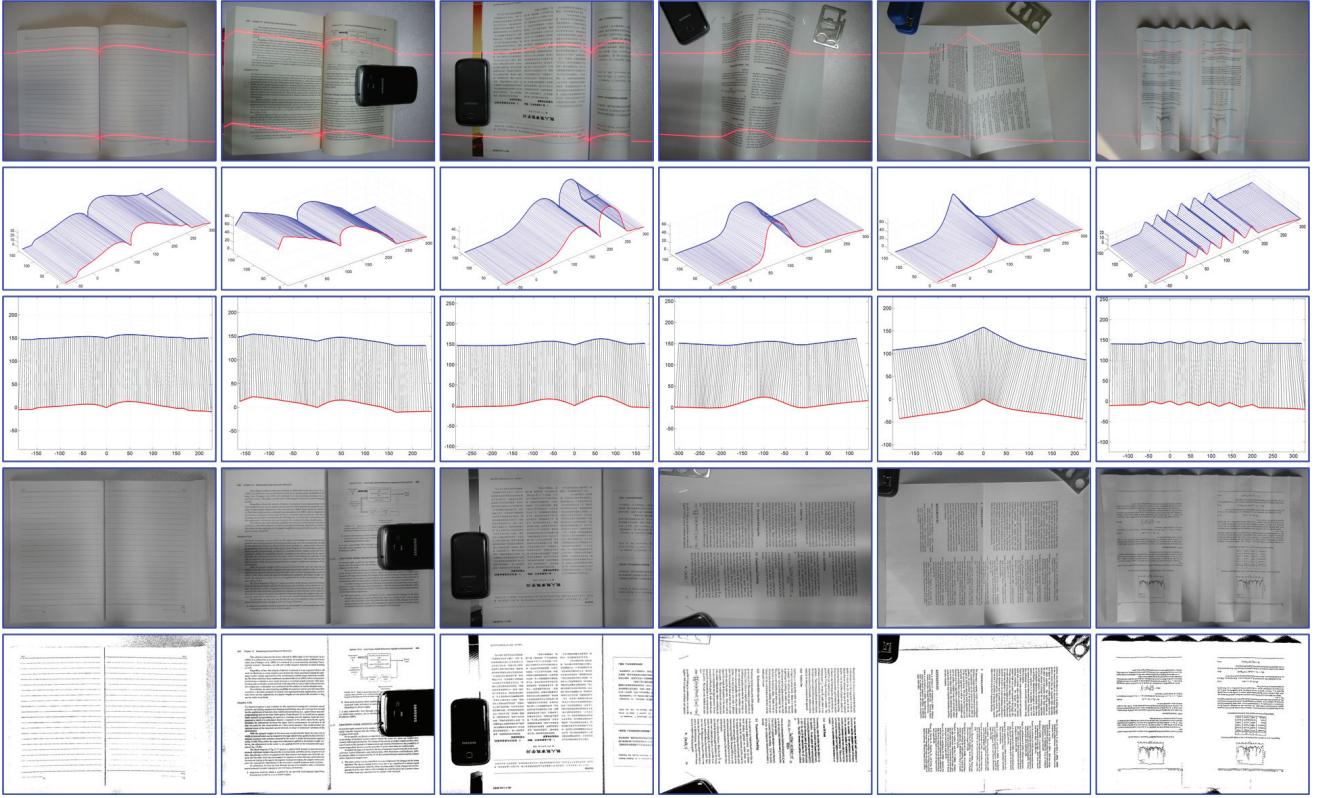


Figure 8. Some examples of the rectification results by our method on different types of geometric distortions. The first row: the captured images with beams. The second row: the recovered developable surfaces. The third row: the flattened surfaces on the plane. The fourth row: the geometric distortion corrected results. The fifth row: shading artifacts are removed by an adaptive binarization algorithm.

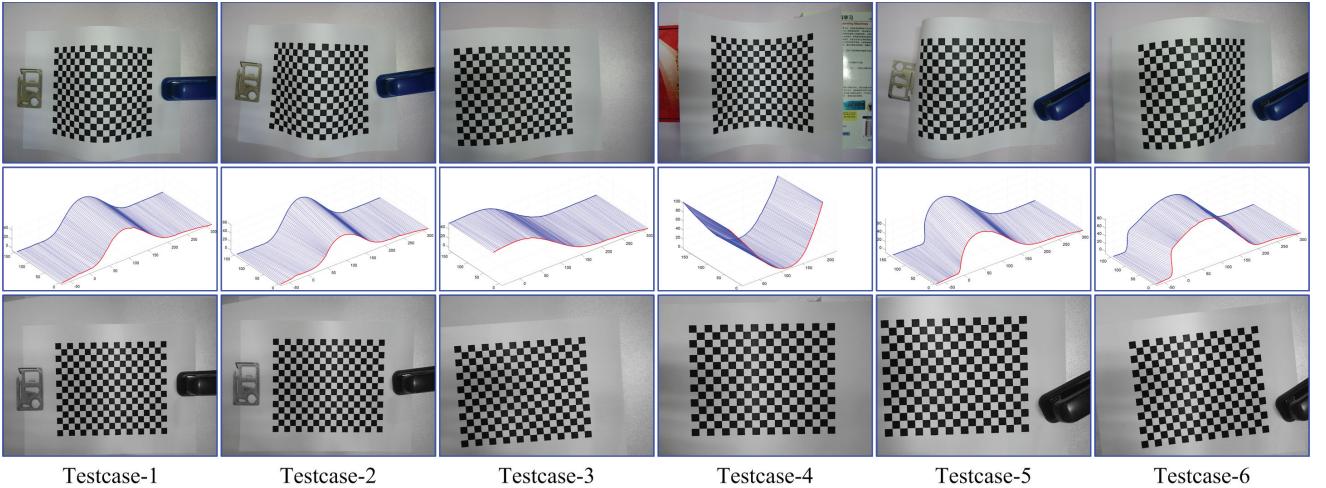


Figure 9. Distorted checkerboard patterns for quantitative evaluations of our method. (top) the images of distorted checkerboard patterns, (middle) the recovered page surfaces, (bottom) the geometric correction results.

15×19 squares. The size of each square is $10mm \times 10mm$. Figure 9 illustrates the images of the distorted checkerboard patterns and the geometric correction results.

To measure the accuracy of the correction results, we first

extract Harris corner points with sub-pixel precisions from each image. Then these points are matched to the ground-truth points in the undistorted checkerboard pattern by a distance-preserving transformation. Table I gives the dis-

Table I
RECTIFICATION ERRORS (UNIT: MM) OF THE DISTORTED CHECKERBOARD PATTERNS WITH THE GROUND-TRUTHS.

Distance Errors	Testcase-1	Testcase-2	Testcase-3	Testcase-4	Testcase-5	Testcase-6
Mean Distance	0.2404	0.3422	0.2107	0.2841	0.2059	0.3307
Max Distance	0.7766	1.3088	0.8812	0.9943	0.6875	1.0786
Standard Deviation	0.1465	0.2230	0.1328	0.1774	0.1275	0.1838

tance errors (unit: mm) of these points with the corresponding ground-truths. From the results, we see that our method can achieve an average rectification error about 0.3mm and the largest rectification error is below 1.4mm. This demonstrates the high correction accuracy of the proposed method.

C. Comparisons

We also compared our method with several state-of-the-arts, i.e., Tsoi and Brown’s method [12], Stamatopoulos *et al.*’s method [10] and our previous method published on TPAMI 2012 [8]. Tsoi and Brown’s method [12] interpolates a bi-linear blended Coons surface to four image boundaries to estimate a global warping function. Stamatopoulos *et al.*’s method [10] first conducts a coarse correction by extracting the top-most and the bottom-most text lines. A post-processing step is then followed to align each word to its baseline. Our previous method [8] assumes the document page shape is a cylindrical surface. Under this assumption, we can perform a metric rectification by constructing an isometric image mesh from only two curved horizontal text lines.

Figure 10 illustrates the comparisons of the four methods. To facilitate the extraction of image boundaries in Tsoi and Brown’s method [12], we deliberately add a rectangular box around the documents in our experiments. The horizontal edges of the box are also used as the estimates of text baselines in Stamatopoulos *et al.*’s method [10] and our previous method [8], respectively.

The first image (Image-1) in the top row of Figure 10 is captured from a document page with a cylindrical shape. As can be seen from the results, all methods can produce a desirable result with straightened text lines. Tsoi’s method and Stamatopoulos *et al.*’s method cannot restore a uniform font size in image areas with large distortions, while the proposed method and our previous method in [8] can address this problem much better. When the document page has a non-cylindrical shape, all the three compared methods fail, as illustrated in the bottom row of the figure (Image-2). In comparison, the proposed method works quite well in this case. It can still yield a high quality correction result, which is comparable to the given ground-truth.

Table II gives the OCR accuracy of the original images and the corrected results of the four methods in Figure 10. A famous commercial OCR product, ABBYY FineReader PE

10, is used for character recognition. The OCR accuracy in [8] is computed, which is based on the edit distance between the OCRed texts and the ground-truth texts. From the results, we can see that correcting the geometric distortions in document images can greatly improve the OCR performance.

IV. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a much efficient method to correct the geometric distortions in a document image captured by a digital camera. Unlike the text-lines based methods, which are often ruined by the challenging problems of text-lines extraction and high accurate baseline fitting, our method is a content independent approach. It can thus be applied to various types of documents, including documents with complex layouts, non-textual or sparse textual contents. Our method also adopts a more general developable assumption on the page shape. It can address more general types of geometric distortions that generally fail to methods using the cylindrical assumption. Finally, instead of using 3D range data, our method reconstructs a developable surface from only two structured beams. It thus has a much lower computational complexity and can be implemented very efficiently.

In our current work, non-uniform shading in the rectified images is not addressed. Such photometric artifact is mainly due to the non-planar page surface. It will be very interesting in the future to further correct this artifact by using the recovered 3D page surface. The proposed method can also be easily extended from two beams to multiple beams. This extension will provide a more accurate estimation of the curved page surface. Moreover, it will also help to locate the boundaries of images and thus enable the adaptivity of the setup to document pages of different sizes.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers and Area Chair for their valuable comments. This work was supported in part by the Projects of the National Natural Science Foundation of China (Grant No.61370039, 61305049, 61272331, 61175025).

REFERENCES

- [1] M. S. Brown and C. J. Pisula. Conformal deskewing of non-planar documents. In *CVPR’05*, pages 998–1004, Washington, DC, USA, 2005.

Original	Tsoi	Stamatopoulos	Meng2012	Our method	Ground-truth
[2] M. S. Brown, M. Sun, R. Yang, L. Yun, and W. B. Seales. Restoring 2d content from distorted documents. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 27(11):1904–1916, 2007.					
[3] H. Cao, X. Ding, and C. Liu. A cylindrical surface model to rectify the bound document image. In <i>ICCV'03</i> , pages 228–233, 2003.					
[4] E. W. Dijkstra. A note on two problems in connexion with graphs. <i>Numer. Math.</i> , 1(1):269–271, 1959.					
[5] H. I. Koo, J. Kim, and N. I. Cho. Composition of a dewarped and enhanced document image from two view images. <i>IEEE Transactions on Image Processing</i> , 18:1551–1562, 2009.					
[6] J. Liang, D. DeMenthon, and D. Doermann. Geometric rectification of camera-captured document images. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 30(4):591–605, 2008.					
[7] J. Liang, D. Doermann, and H. Li. Camera-based analysis of text and documents: a survey. <i>International Journal on Document Analysis and Recognition</i> , 7(2):84–104, 2005.					
[8] G. Meng, C. Pan, S. Xiang, J. Duan, and N. Zheng. Metric rectification of curved document images. <i>IEEE Transactions</i>					
[9] M. Pilu. Undoing paper curl distortion using applicable surfaces. In <i>CVPR'01</i> , volume 1, pages 67–72, 2001.					
[10] N. Stamatopoulos, B. Gatos, I. Pratikakis, and S. J. Perantonis. Goal-oriented rectification of camera-based document images. <i>IEEE Transactions on Image Processing</i> , 20(4):910–920, 2011.					
[11] C. L. Tan, L. Zhang, Z. Zhang, and T. Xia. Restoring warped document images through 3d shape modeling. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 28(2):195–208, Feb. 2006.					
[12] Y. C. Tsoi and M. S. Brown. Geometric and shading correction for images of printed materials: a unified approach using boundary. In <i>CVPR'04</i> , volume 1, pages 240–246, 2004.					
[13] Y. C. Tsoi and M. S. Brown. Multi-view document rectification using boundary. In <i>CVPR'07</i> , pages 1–8, 2007.					

Table II
OCR ACCURACY OF THE CURVED DOCUMENT IMAGES AND THE CORRECTION RESULTS IN FIGURE 10.

	Original	Tsoi [12]	Stamatopoulos [10]	Meng2012 [8]	Our method
Image-1 (cylindrical)	53.96%	90.66%	90.31%	94%	96.04%
Image-2 (non-cylindrical)	33.99%	68.26%	68.40%	68.26%	89.61%

- [2] M. S. Brown, M. Sun, R. Yang, L. Yun, and W. B. Seales. Restoring 2d content from distorted documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11):1904–1916, 2007.
- [3] H. Cao, X. Ding, and C. Liu. A cylindrical surface model to rectify the bound document image. In *ICCV'03*, pages 228–233, 2003.
- [4] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numer. Math.*, 1(1):269–271, 1959.
- [5] H. I. Koo, J. Kim, and N. I. Cho. Composition of a dewarped and enhanced document image from two view images. *IEEE Transactions on Image Processing*, 18:1551–1562, 2009.
- [6] J. Liang, D. DeMenthon, and D. Doermann. Geometric rectification of camera-captured document images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4):591–605, 2008.
- [7] J. Liang, D. Doermann, and H. Li. Camera-based analysis of text and documents: a survey. *International Journal on Document Analysis and Recognition*, 7(2):84–104, 2005.
- [8] G. Meng, C. Pan, S. Xiang, J. Duan, and N. Zheng. Metric rectification of curved document images. *IEEE Transactions*

on Pattern Analysis and Machine Intelligence, 34(4):707–722, Apr. 2012.

- [9] M. Pilu. Undoing paper curl distortion using applicable surfaces. In *CVPR'01*, volume 1, pages 67–72, 2001.
- [10] N. Stamatopoulos, B. Gatos, I. Pratikakis, and S. J. Perantonis. Goal-oriented rectification of camera-based document images. *IEEE Transactions on Image Processing*, 20(4):910–920, 2011.
- [11] C. L. Tan, L. Zhang, Z. Zhang, and T. Xia. Restoring warped document images through 3d shape modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(2):195–208, Feb. 2006.
- [12] Y. C. Tsoi and M. S. Brown. Geometric and shading correction for images of printed materials: a unified approach using boundary. In *CVPR'04*, volume 1, pages 240–246, 2004.
- [13] Y. C. Tsoi and M. S. Brown. Multi-view document rectification using boundary. In *CVPR'07*, pages 1–8, 2007.