Geometric and Shading Correction for Images of Printed Materials Using Boundary

Michael S. Brown and Yau-Chat Tsoi

Department of Computer Science

Hong Kong University of Science and Technology

Clear Water Bay, Hong Kong

{brown, desmond}@cs.ust.hk

Index Terms

Geometric and photometric correction, distortion removal, boundary interpolation.

ABSTRACT

A novel approach that uses boundary interpolation to correct geometric distortion and shading artifacts present in images of printed materials is presented. Unlike existing techniques, our algorithm can simultaneously correct a variety of geometric distortions, including skew, fold distortion, binder curl, and combinations of these. In addition, the same interpolation framework can be used to estimate the *intrinsic* illumination component of the distorted image to correct shading artifacts. We detail our algorithm for geometric and shading correction and demonstrate its usefulness on real-world and synthetic data typical of imaged print materials.

I. INTRODUCTION

Camera-based imaging of printed materials is becoming increasingly common [9]. While the printed content is inherently 2D, the physical media that the content is printed on is rarely 2D. As a result, two types of distortion are commonly present in images of these materials. The first is geometric distortion of the 2D content arising from material's non-planar shape. The second is shading artifacts also resulting from the non-planar shape. Examples of these effects can be seen in Figure 1.

Geometric and shading distortion is particularly troublesome for images of large materials, such as oversize books and art-like materials. Due to the nature of the objects, they are rarely flattened before imaging. Moreover, some materials simply cannot be made completely flat without risk of damage. This makes it difficult to avoid geometric distortion and shading artifacts. For these materials, distortion correction must be applied after imaging.

In this paper, we present a unified approach to correct both geometric and photometric (shading) distortion via the 2D boundary of the imaged material. Using boundary interpolation, we can compute a corrective mapping to simultaneously undo common geometric distortions, such as skew, binder curl, and fold distortion (and combinations of these). In addition, the same interpolant framework can be used to estimate the intrinsic illumination image. This estimated illumination image together with the original image can be used to remove shading artifacts. Our proposed approach is fast, works directly from a 2D image, and provides a general solution to correct both simple and complex geometric distortions and shading.

Shorter versions of this work have appeared in [19], [5], [4]. The work presented in [5], [4] discussed only geometric correction, while [19] introduced the technique to correct shading artifacts. In this paper, we expand on this previous work to provide greater detail to the algorithms and more thorough results, including experiments on synthetic data that provides ground truth for quantitative results.

The remainder of this paper is organized as follows: Section II discusses previous work, III and IV overviews our geometric and shading correction approach. Section V concludes our work.

II. RELATED WORK

Previous work to correct document distortion has focused primarily on two types of distortions: *skew* and *binder-curl*. Skew arises when the acquired image and the physical material's 2D content are not axis aligned. Approaches [1], [12], [14], [16] to correct this distortion compute planar transformations to rectify the imaged content. Binder-curl effect occurs in the region near a book's spine. Binder-curl results in an unsightly *curling* appearance in the 2D content. Approaches to undo this distortion [6], [7], [15], [20], [22] parameterize the distortion using cylindrical models which can be used to "unroll" the curved imagery.

These existing approaches are limited to planar documents or a single page of a bound volume. These approaches also assume the presence of only a single type of distortion. As a result, the skew and binder-curl correction are
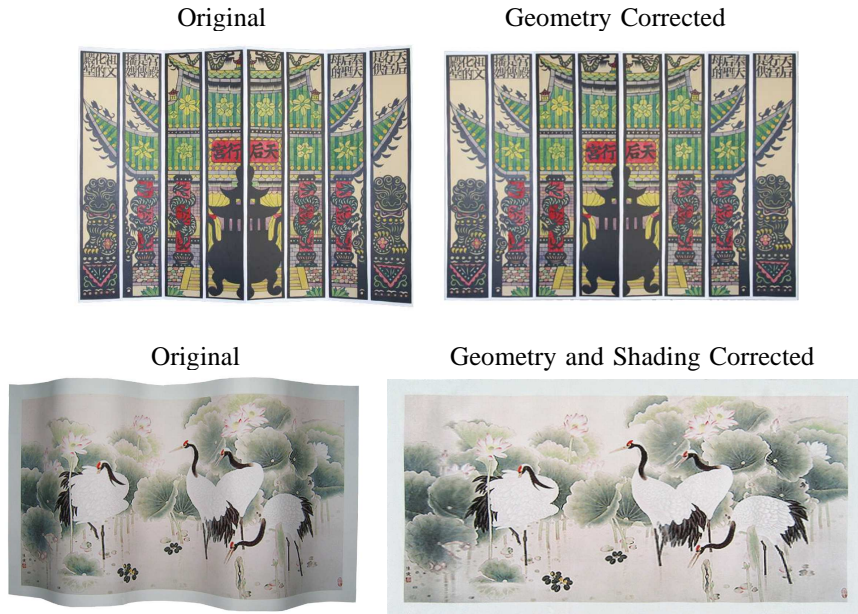
Fig. 1. (Top) An image of a large piece of art that suffers from fold distortion before and after being corrected with our algorithm. (Bottom) An image of a painting suffering for geometric distortion and shading artifacts before and after correction.

typically applied in tandem to remove all distortion. Furthermore, existing approaches assume that the targeted distortion is exhibited only once in the image. De-skewing algorithms cannot correct a material that exhibits multiple skewing (for example from folds), such as shown in Figure 1 (Top). Binder-curl approaches assume the image is of a single page and the curl is either on the left or right of the image. The image in Figure 1 (Bottom) with several curled regions cannot be corrected. None of the existing approaches can work for materials that exhibit both multiple skew (folds) and binder-curl.

Another drawback to existing approaches is that they do not address shading artifacts. Since these techniques target images of text-based documents and bound books, the acquired images are subsequently converted to bi-tonal representations for use in further processing, such as optical character recognition (OCR). Thus, shading artifacts are only considered in the binarization process, requiring the use of local vs. global thresholding to overcome variations in shading [22].

Recent approaches have addressed distortions from *arbitrarily* shaped materials [3], [17]. These approaches acquire a 3D reconstruction of the imaged material's surface in addition to a 2D image. Relaxation algorithms are used to flatten the 3D surface to a plane while minimizing spring energies between acquired 3D points. While these approaches work, the additional 3D shape information is rarely available and requires modifications to the imaging technology. In addition, these relaxation approaches are computationally slow [3], [17] and do not address shading artifacts. Recent work by Gumerov et al. [13] presented a system that can the 3D shape from a single image, which

can be subsequently used to restore the image, but again shading artifacts are not addressed.

The approach presented in this paper papers differs from previous work in several distinct ways. First, unlike traditional document correction techniques, our approach can handle all types of distortion in a single pass. Thus, skew and curl are undone all at one time. Also, we are not restricted to a single page at a time, as with the deskewing and curl removal algorithms. Moreover, we do not require a 3D model of the surface. While the 3D model approach can address more, we are able to address the most common distortions present in imaged images, namely, fold, curl, and skew. Lastly, we are the only technique that addresses the shading artifacts. Instead of simple binarization, our technique actually removes the shading to restore the image, which is particular useful for artwork and similar types of printed media.

III. Geometric Distortion Correction

A. Distortion Parameterization By Boundary

Our approach works under the assumption that the printed content will be rectilinear in its true planar format. Our goal is to find a parameterization between the desired rectilinear representation and the distorted input image. To do this, we consider how to model the material's 3D structure and how this model behaves under projection, i.e. imaging.

Several authors [6], [7], [20] have proposed cylindrical models to model binder-curl distortion of a single page. We use a more general model of a ruled surface composed of two *opposite-boundary* curves [10]. To visualize this model, consider a book distorted by binder curl represented in Figure 2. Given the top and bottom 3D boundary curves, $\mathbf{C}_1(u)$ and $\mathbf{C}_2(u)$, the entire 3D surface can be described as:

$$\mathbf{S}(u,v) = (1-v)\mathbf{C}_1(u) + v\mathbf{C}_2(u) \tag{1}$$

where $\mathbf{S}(u,v)$ is the equation for the 3D ruled surface, parameterized by u and v. Given $\mathbf{S}(u,v)$, every iso-parametric line $u = u_c$ is a straight line segment between $\mathbf{C}_1(u)$ and $\mathbf{C}_2(u)$ parameterized by v. Note, in the case of folds, if the fold is along the iso-parametric lines $u = u_c$, the surface can still be modeled using only the boundary curves. Not only does this model describe the 3D structure, it also fits our notion of how printed materials behavior. Ruled
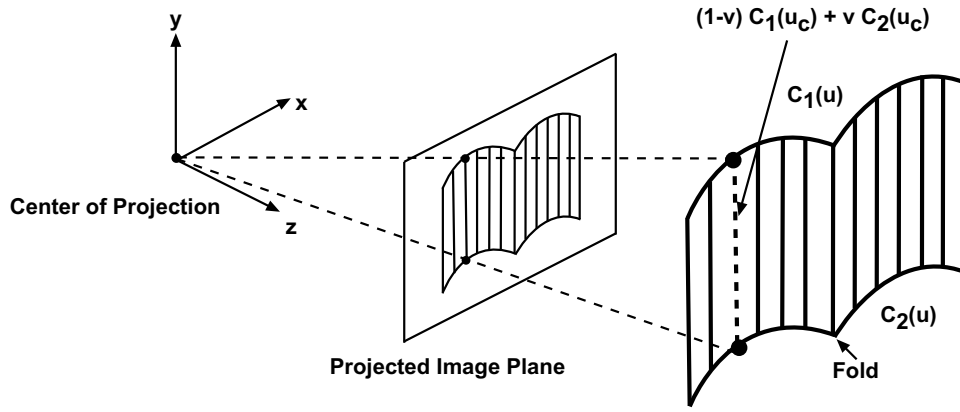
Fig. 2. Material represented by a ruled surface defined by two opposite curves, $\mathbf{C}_1(u)$ and $\mathbf{C}_2(u)$. The projection of this surface can also be parameterized by the resulting 2D boundary curves.

surfaces are developable surfaces and can be mapped to a plane without distortion. This is similar to our notion of a printed page; it can be flattened without introducing any distortion to the 2D content.

Although restrictive, this two-boundary ruled surface model is sufficient to model the *vast majority* of printed materials. Existing cylindrical models can be represented with our model. In the case of folded materials, the folds made are almost exclusively along iso-parametric lines to allow the material to be folded into equal pieces for storage. Thus, a document with several folds and curl effects can be modeled. We note that it is easy to violate this surface model; e.g. folding a page's corner over on itself introduces a fold against the iso-parametric lines, however such examples are not typical of the majority of imaged materials.

With the surface model in place, we examine its behavior under projection. The 2D projection of points on the two 3D boundary curves can be expressed as, $\mathbf{x}_1 = \tilde{\mathbf{P}}[\mathbf{C}_1(u_c)\ 1]^T$ and $\mathbf{x}_2 = \tilde{\mathbf{P}}[\mathbf{C}_2(u_c)\ 1]^T$, where $\tilde{\mathbf{P}}$ is the $3{\times}4$ projection matrix of the camera. Since lines are preserved under projection, the line segments passing through, $\mathbf{C}_1(u_c)$ and $\mathbf{C}_2(u_c)$ must pass through the 2D boundary points $\mathbf{x}_1$ and $\mathbf{x}_2$ as shown in Figure 2. Thus, the projection of this model can also be parameterized using the projected boundary curves in 2D.

While only the two opposite boundary curves are needed to parameterize the distortion, in practice we will use four boundary curves corresponding to the rectilinear edges of the imaged materials. With appropriate boundary interpolation, it is not necessary to explicitly denote which curves are the opposite boundary pair. In addition, small non-linearities in the imaging system, such as radial distortion, can be compensated for in the interpolation process.
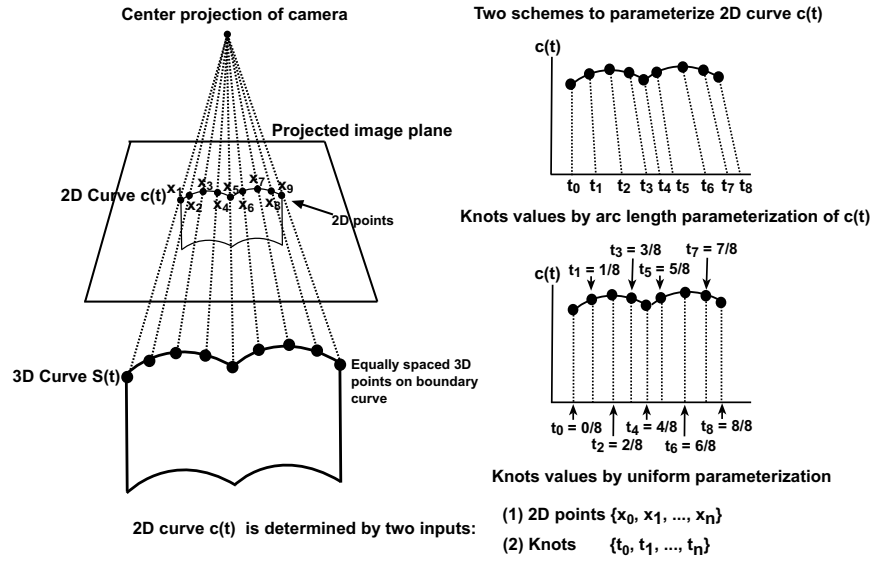
Fig. 3. The 2D curve $\mathbf{c}(t)$ is defined by $n + 1$ 2D feature points $\mathbf{x}_i$ *and* $n + 1$ knot values t_i. Two schemes are used to compute the knot's values. **Arc length** parameterization assigns knot values using their position along the arc length of $\mathbf{c}(t)$. **Uniform** parameterization assumes the knots are uniformly sampled in parameter space, i.e. $t_i = i/n$. If $\mathbf{x}_i$ are from equally sampled 3D points along the 3D curve, uniform parameterization can encode useful depth information.

B. Boundary Representation

Four boundary curves of the distorted material are represented by curves $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$, $\mathbf{c}_4$, representing the top, right, bottom and left sides of the material. Each of these curves is encoded using a natural cubic spline that defines continuous image coordinates along the curve. i.e. $\mathbf{c}_i = (x(t), y(t))$, where t is a parameter along the curve with range between 0 and 1.

Recall that 2D natural cubic splines (NCS) are defined by specifying a set of $n + 1$ 2D points, $\mathbf{x}_0$, $\mathbf{x}_1$, ..., $\mathbf{x}_n$ together with their corresponding parameter value (knots), t_0, t_1, ..., t_n. The 2D points $\mathbf{x}_i$ and knots t_i, uniquely define a set of piece-wise functions, $\mathbf{s}_i(i)$ such that:

$$\mathbf{s}(t) = \begin{cases} \mathbf{s}_0(t) & t \in [t_0, t_1] \\ \\ \mathbf{s}_1(t) & t \in [t_1, t_2] \\ \\ \vdots & \\ \\ \mathbf{s}_{n-1}(t) & t \in [t_{n-1}, t_n] \end{cases}, \tag{2}$$

where $\mathbf{s}_i(t) = (x_i(t), y_i(t))$, and $x_i(t)$ and $y_i(t)$ are cubic functions. The coefficients of these cubic pieces can be uniquely computed with the constraints that $\mathbf{s}_{i-1} = \mathbf{x}_i = \mathbf{s}_i$ for $(1 \leq i \leq n-1)$ and that $\mathbf{s}'$ and $\mathbf{s}''$ are continuous.

While the 2D points $\mathbf{x}_0$, $\mathbf{x}_1$, ..., $\mathbf{x}_n$ can be obtained by extracting image coordinates along the boundary, a corresponding set of knot values, t_i must also be specified. Two different parameterization approaches are used.

The first approach uses the approximated 2D arc-length, $|\mathbf{s}|$ of $\mathbf{s}(t)$ which is defined by the chord length of the points $\mathbf{x}_i$, such that each t_i is defined as:

$$|\mathbf{s}| = \sum_{0 < i < n-1} |\mathbf{x}_i - \mathbf{x}_{i+1}|_2$$

$$t_i = \begin{cases} 0 & \text{if } i = 0 \\[2ex] \frac{1}{|\mathbf{s}|} \sum_{0 < j < i} |\mathbf{x}_{j-1} - \mathbf{x}_j|_2 & \text{if } i > 0, \end{cases} \tag{3}$$

where $|\cdot|_2$ represents Euclidean distance. The idea of this parameterization is to move along the curve $\mathbf{s}(t)$ at a constant rate.

The second approach defines the parameterization *uniformly* with values $t_i = i/n$, where $n + 1$ is the number of control points. In this case, uniform parameterization encodes how to move along the 2D curve as it corresponds to a uniform sampling along the *arc-length* of the actual 3D curve. Consider a 3D curve called $\mathbf{S}(t)$, where t is parameterized by the 3D curve's *arc-length*. If we have the relationship:

$$\mathbf{x}_i = \tilde{\mathbf{P}}[\mathbf{S}(t = i/n) \ 1]^T, \tag{4}$$

where $\tilde{\mathbf{P}}$ is the projection matrix of the camera, i is the ith sample, and $n + 1$ is the number of samples, then $\mathbf{x}_i$ represents the projection of the 3D curves points at $t_i = i/n$ in $\mathbf{S}(t)$. This can be visualized in Figure 3. The features points $\mathbf{x}_i$ are obtained from the uniform *sampling* of the 3D curve and corresponds to a uniform *parameterization* of the projected 2D curve. Such parameterization can be used if we have prior knowledge that 2D image points are from 3D points samples by arc-length in 3D. This 2.5D information incorporates the depth change information of the 3D curve into the parameterization.

C. Boundary Interpolation

The restored image is defined over the parametric space u and v, where $u \in [0, 1]$ and $v \in [0, 1]$. Each curve $\mathbf{c}_i$ maps to its corresponding side of the rectilinear image (shown in Figure 4). For example, the (x, y)'s points
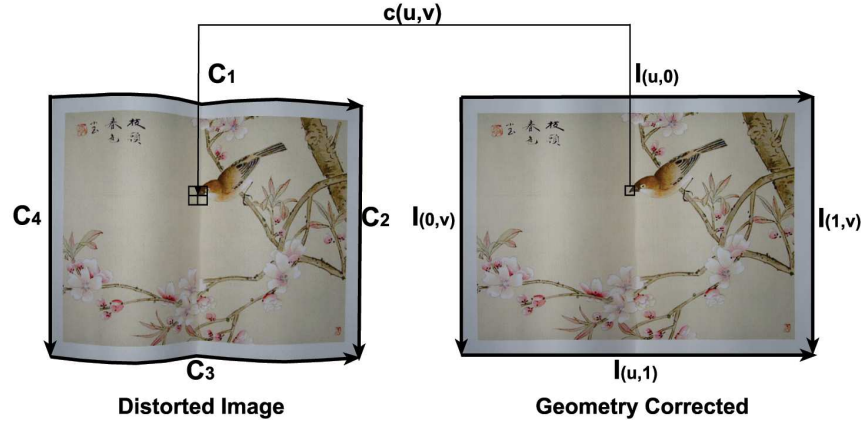
Fig. 4. Curves $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ correspond to the rectilinear edges of the corrected image $I(u,0), I(1,v), I(u,1), I(0,v)$ respectively. Coons patch interpolation using the $\mathbf{c}_i(\cdot)$ curves provides a mapping between restored (u,v) and the distorted (x,y) coordinates. Geometric correction can performed by pointwise bilinear samples of the (x,y) points. Note that shading correction has not been applied and the restored image may still appear distorted.

along the top boundary $\mathbf{c}_1(u)$ should map to $I(u,0)$, in the undistorted image. This curve-to-line mapping can be computed by simple sampling of 2D curve.

While the boundaries can be easily mapped to their correct locations, a 2D function to describe how to map (x,y) points inside the $\mathbf{c}_i$ curves is needed. This 2D function can be provided using a bi-linearly blended Coons patch [8], as follows:

$$
\begin{aligned}
\mathbf{c}(u,v) = \begin{bmatrix} 1-u & u \end{bmatrix} \begin{bmatrix} \mathbf{c}_4(v) \\ \mathbf{c}_2(v) \end{bmatrix} \\
+ \begin{bmatrix} \mathbf{c}_1(u) & \mathbf{c}_3(u) \end{bmatrix} \begin{bmatrix} 1-v \\ v \end{bmatrix} \\
- \begin{bmatrix} 1-u & u \end{bmatrix} \begin{bmatrix} \mathbf{c}_1(0) & \mathbf{c}_2(0) \\ \mathbf{c}_3(1) & \mathbf{c}_4(1) \end{bmatrix} \begin{bmatrix} 1-v \\ v \end{bmatrix}
\end{aligned}
\tag{5}
$$

Equation 5 is formed by a linear interpolation of two opposite-boundary curves (first two terms), with a corrective function based on the boundaries' corner points (third term). Because of the need for the corrective function in the interpolant, the function is not considered a bi-linear interpolation, but instead is a bi-linearly "blended" interpolant [10].

The advantage of using equation 5 with four curves, instead of two as defined in the original ruled surface

equation 2, is that it is not necessary to specify the true opposite boundary curves. One pair of opposite boundaries should be straight-lines, however, in practice they are not straight due to non-linearites in the image system. The non-linearities are encoded in the boundary splines and their effects will be compensated in the unwarping process.

D. Geometric Distortion Removal

Using the four boundary curves, the function $\mathbf{c}(u, v)$ provides a mapping between (u, v) coordinates in the rectilinear image space $I(u, v)$ to their corresponding (x, y) coordinates in the distorted image. Constructing the restored rectilinear image is performed by pointwise bilinear re-sampling of the distorted image using the relationship $(u, v) \rightarrow (x, y)$. The size of the restored image is specified either by the user or set according to pixel length of the horizontal and vertical c_i curves in the distorted image.

E. Experimental Results

In this section, we report the effectiveness of our geometric correction approach. For the first experiment, we perform the evaluation on synthetic data. Synthetic models are created using OpenGL to simulate the common distortion pattern of documents, where ground truths are available for direction verification. To gauge the quality, we compare projected 3D points on model's surface with the points calculated by Coons patch equation using boundary.

For the second experiment, the evaluation is performed on lettered document and OCR is used as a metric to gauge the quality. A lettered document is distorted in several different patterns and digitized using hand-held digital camera. OCR is then performed on these distorted images, and the rectified images using both arc-length and uniform parameterization. We compare the OCR results for the three documents configurations.

1) **Experiment 1 (Evaluation using synthetic data):** We evaluate our algorithm by using three synthetic models, (i) binder curl, (ii) folded tablet and (iii) 3 page fold-out, as shown in the first row of Figure 5. The mathematical description of these models are shown in the second row.

With the synthetic models, two sets of n 2D points (control point-set and Coons patch calculated point-set) are constructed (shown in Figure 6) and are used for evaluation. The control point-set, defined as $\Phi = \{\rho_1, \ \rho_2, \ \rho_3, \ ..., \ \rho_n\}$, is a set of screen coordinates constructed by projecting 3D points on model's surface

(i) Binder Curl	(ii) Folded Tablet	(iii) 3 Page Fold-Out
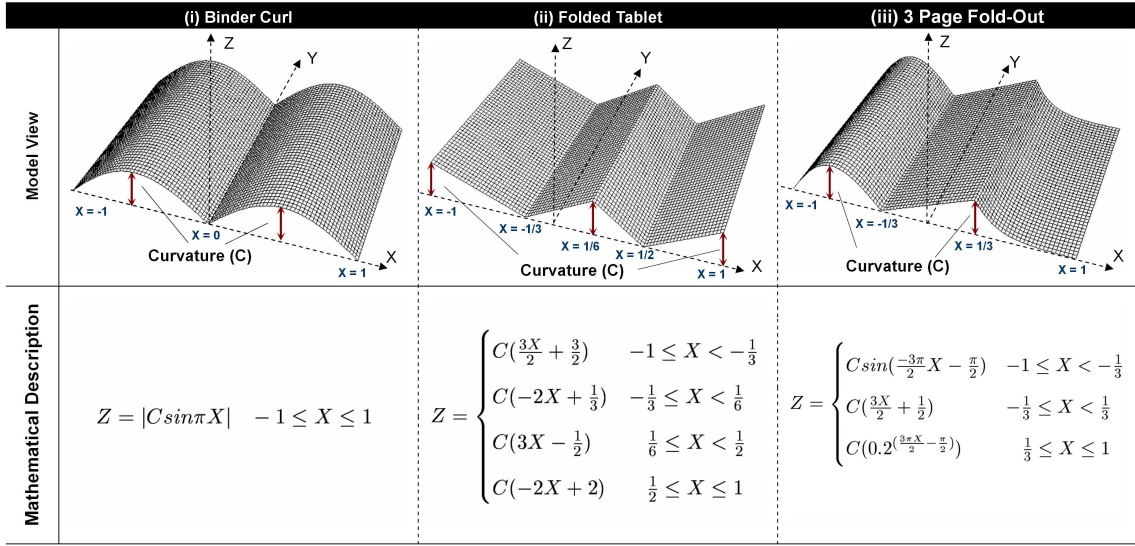
Fig. 5. Three synthetic models, (i) binder curl (ii) folded tablet (iii) 3 page fold-out, are created using OpenGL and provide ground truth to evaluate the quality of our geometric correction approach. The perspective view and the mathematical form of the models are shown in the first and second row respectively.

using OpenGL function *gluProject*, while those 3D surface points are selected by uniform sampling over the X-Y plane of the surface. The sampling rate of the 3D points can be depending on how many test samples we need for the analysis. For our experiment, we use n = 280×140 as the sampling rate in X and Y direction respectively.

The Coons patch calculated point-set, defined as, $\Psi = \{\zeta_1,\ \zeta_2,\ \zeta_3,\ ...,\ \zeta_n\}$, is constructed by computing the same set of points as the control point-set except using only boundary information. First, a set of control points are sampled along the 3D boundary's curves and projected on screen. Then, these projected control points are used to calculate n test points using Coons patch equation. Notice that the number of control points used for the calculation is generally less than the test samples generated along the boundary.

To evaluate our algorithm on different factors, we compute the max error E_{Max}, mean error E_{Mean} and the standard deviation E_{Std}, between the two data sets (Φ and Ψ) for each combination of curvature (C = 0.2, 0.3, 0.4 and 0.5), parameterization scheme (arc-length and uniform), and number of control points (12×4, 24×8 and 36×12) over each type of surface. More specifically, the error measures, E, are defined as follows.

$$E_{Max} = max(|\rho_i - \zeta_i|_2) \tag{6}$$

$$E_{Mean} = \frac{1}{n}\sum_{i=1}^{n}|\rho_i - \zeta_i|_2 \tag{7}$$
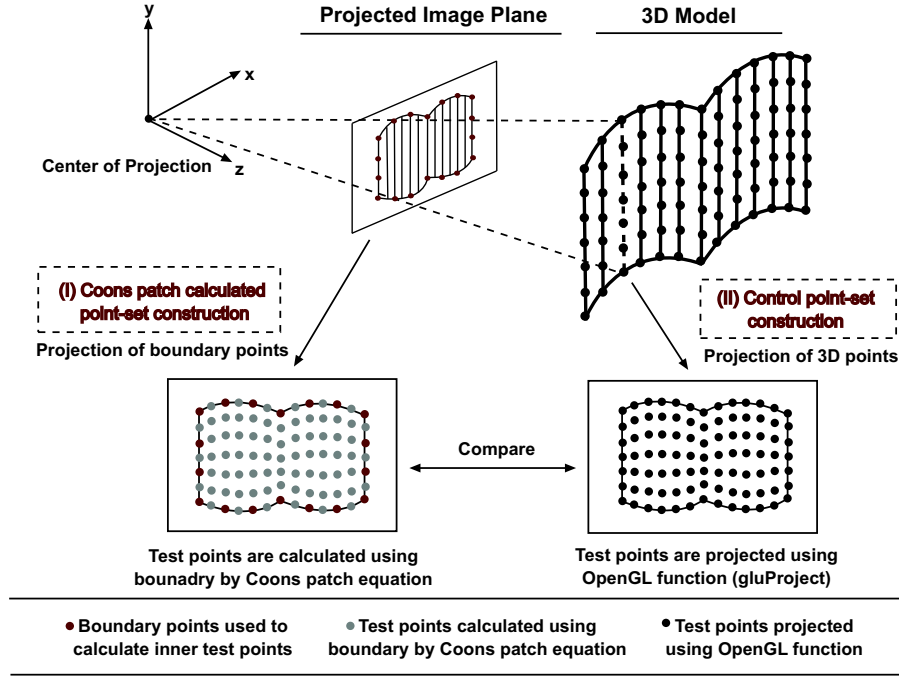
Fig. 6. The construction process of two data point sets: *control point-set* and *Coons patch calculated point-set* using synthetic models. Control point-set is a set of 2D screen coordinates constructed by projecting points on the 3D model's surface to the image plane, while Coons patch calculated point-set is constructed by computing the same set of test points using only boundary information.

$$E_{Std} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (|\rho_i - \zeta_i|_2 - E_{Mean})^2} \tag{8}$$

where ρ_i = (x,y) denotes the ith sample point in set Φ, ζ_i = (x,y) denotes the ith sample point in set Ψ and $|\cdot|_2$ represents Euclidean distance.

In Figure 7, we measure errors using binder curl model in four different curvatures arranged increasingly from left to right, where arc-length and uniform parameterization are used for the analysis. As the corrected results using different number of control points are visually similar, we only show the rectified image using 24×8 control points for clarity of display. Notice that the corrected image using arc-length parameterization has artifacts due to the depth distortion. The artifacts are more noticeable when the curvature of the surface is high. This can be verified by looking at the uneven black and white squares pattern. The depth distortion is efficiently removed when uniform parameterization is applied using same set of control points. This parameterization is justified because we know the 2D control points are from uniform samples in arc-length along the corresponding 3D curve of the surface. To evaluate the ability of our restoration process, we measure the errors between the calculated values using only boundary with the ground truth provided by the given model. The quantitative results are presented below the
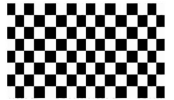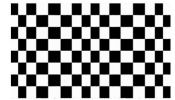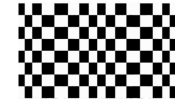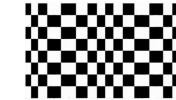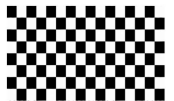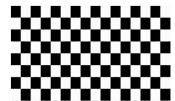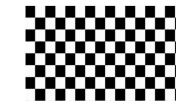
Binder Curl											
Curvature	**0.2**			**0.3**			**0.4**			**0.5**		
	Input		Model view	Input		Model view	Input		Model view	Input		Model view
Arc-length Result(24x8)												
Control Points	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12
Pixel Error Max	8.182	8.183	8.183	13.425	13.428	13.428	18.656	18.661	18.661	23.664	23.649	23.651
Mean	3.665	3.665	3.665	5.992	5.993	5.994	8.261	8.265	8.265	10.323	10.329	10.330
Std	2.789	2.789	2.788	4.438	4.437	4.437	6.080	6.079	6.079	7.661	7.658	7.657
Uniform Result (24x8)												
Control Points	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12
Pixel Error Max	0.183	0.105	0.095	0.416	0.157	0.125	0.827	0.240	0.160	1.374	0.409	0.241
Mean	0.067	0.055	0.055	0.150	0.060	0.057	0.351	0.075	0.056	0.646	0.114	0.072
Std	0.048	0.030	0.029	0.107	0.042	0.033	0.209	0.061	0.042	0.343	0.096	0.059

Fig. 7. The results of error measurement using binder curl synthetic model with different combination of parameters (curvature, parameterization scheme and number of control points). Input images (with model view), geometry corrected images and quantitative results are shown for comparison.

corrected image. From the results, we have the following observations:

1) Correction using uniform parameterization performs better than that of using arc-length. This is even more true for imaged document that suffered from significant depth distortion (more curvature). The error of the correction using uniform parameterization has less than a unit mean error deviation from the ground truth. This trend is valid over different curvature values.

2) Surface with higher curvature needs more control points to get a better approximation of boundary and thus a more accurate estimation of test sample points.

Same experiment is performed for synthetic models of folded tablet and 3 page fold-out inset. The results of the experiment are shown in Figure 8 and Figure 9. The organization of these results is same as Figure 7 and a similar trend of error can be observed.

2) **Experiment 2 (Evaluation using lettered documents with OCR):** This experiment uses OCR as a metric to evaluate the quality of the correction. A document with 390 letters is imaged. When the document is imaged completely flat, all characters can be successfully identify via OCR. The document is then distorted in five trials. OCR is performed on the unrestored image, and the restored images using both arc-length and uniform parameterization. We use a commercial OCR software, *Readiris Pro*, to perform the OCR. We compare the number
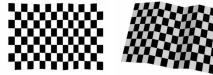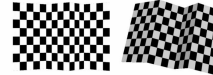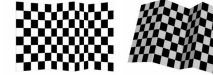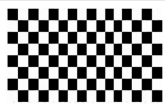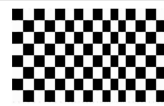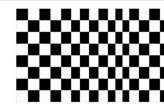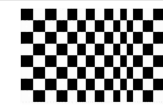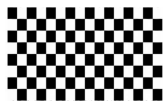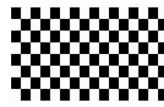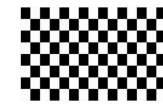
Folded Tablet												
Curvature	0.2			0.3			0.4			0.5		
	Input	Model view		Input	Model view		Input	Model view		Input	Model view	
Arc-length Result(2 4x8)												
Control Points	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12
Pixel Error Max	8.432	8.432	8.432	15.014	15.014	15.013	21.828	21.828	21.828	28.237	28.237	28.237
Mean	4.580	4.580	4.580	7.995	7.995	7.995	11.467	11.467	11.467	14.672	14.672	14.672
Std	2.268	2.268	2.268	4.168	4.168	4.168	6.133	6.133	6.133	7.928	7.928	7.928
Uniform Result (24x8)												
Control Points	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12
Pixel Error Max	0.132	0.132	0.132	0.172	0.172	0.172	0.136	0.136	0.136	0.177	0.177	0.178
Mean	0.057	0.057	0.057	0.055	0.054	0.054	0.063	0.062	0.062	0.062	0.062	0.062
Std	0.038	0.038	0.038	0.052	0.052	0.052	0.034	0.033	0.033	0.047	0.046	0.046

Fig. 8. The results of error measurement using folded tablet synthetic model with different combination of parameters (curvature, parameterization scheme and number of control points). Input images (with model view), geometry corrected images and quantitative results are shown for comparison.

of misses by the OCR algorithm for the three documents configurations. A miss is defined as any letter that is misclassified or any letters not identified that are present in the document. A hit is all letter correctly identified.

Figure 10 shows the results. The distorted image causes an error of 6-9%. The restored image, with arc-length parameterization reduces the error to 0-5%. When uniform parameterization is used, the error rate ranges between 0-2%, with the majority of trials having 100% accuracy.

IV. SHADING CORRECTION

Shading is a strong visual cue for shape. Correcting geometric distortion without addressing shading artifacts can produce restored images that still *appear* perceptually distorted (see Figure 4). For applications, such as text-based imaging with OCR processing, shading artifacts can be ignored, however, for images of materials where the original content is desired (for example, images of artwork), shading artifacts must be removed to produce a perceptually correct image.

One current approach is to manually adjust the lighting in the acquisition environment. This, however, can be a tedious process requiring time and expertise to find the correct orientation of the lights and materials. Moreover, this often requires additional lights which add cost to the overall imaging system. An image-based solution would
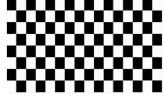
3 Page Fold-Out												
Curvature	0.2			0.3			0.4			0.5		
	Input	Model view		Input	Model view		Input	Model view		Input	Model view	
Arc-length Result(2 4x8)	(image)			(image)			(image)			(image)		
Control Points	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12
Pixel Error Max	13.550	13.542	13.542	22.517	22.508	22.503	30.938	30.923	30.905	38.440	38.339	38.310
Pixel Error Mean	6.922	6.915	6.914	12.314	12.300	12.297	17.382	17.351	17.343	21.702	21.633	21.623
Pixel Error Std	3.565	3.562	3.561	5.876	5.876	5.865	7.950	7.928	7.924	9.727	9.688	9.681
Uniform Result (24x8)	(image)			(image)			(image)			(image)		
Control Points	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12	12x4	24x8	36x12
Pixel Error Max	2.845	0.170	0.095	2.107	0.398	0.170	4.052	0.735	0.341	6.498	1.178	0.564
Pixel Error Mean	0.441	0.055	0.034	0.586	0.126	0.069	1.267	0.219	0.097	2.177	0.290	0.148
Pixel Error Std	0.626	0.048	0.022	0.500	0.115	0.044	0.948	0.203	0.075	1.619	0.278	0.129

Fig. 9. The results of error measurement using 3 page fold-out (curl + fold distortion) synthetic model with different combination of parameters (curvature, parameterization scheme and number of control points). Input images (with model view), geometry corrected images and quantitative results are shown for comparison.

undoubtedly be preferred.

A. Shading Artifacts Removal

For shading correction, we only consider the luminance component of the image. Our input image is represented in YUV colorspace and the Y-channel (luminance) of the input image, I_Y, can be expressed as the product of the *intrinsic* illumination image, L_Y, and the *intrinsic* reflectance image, R_Y, [2] as follows:

$$I_Y = L_Y \cdot R_Y, \tag{9}$$

where $\cdot$ is a pixel-wise multiple between the two images. If we can successfully compute L_Y, we can derive the intrinsic reflectance image R_Y as:

$$R_Y = e^{log I_Y - log L_Y}. \tag{10}$$

Once we compute the reflectance image R_Y, we can generate an image with uniform shading by multiplying the reflectance image by a constant c, such that $I_Y^{new} = cR_Y$. While the intrinsic image model is view-dependent, we only need to restore our image from a single viewpoint. As a result, this simple intrinsic image model can be used as a reasonable approximation of our imaged scene.

Given a single image, solving for the intrinsic illumination image is ill-posed as the number of unknowns (L

and R) is more than the given input I. Several authors have proposed approaches to estimate L from both single

and multiple temporal images [11], [18], [21]. For our algorithm, however, we can exploit the fact that almost

all printed materials have a uniformly colored (typically white) margin, or border, about the page, void of printed

content. The material's reflectance property can be assumed to be the same everywhere on the border. Intensity

variations along this uniform border are due to the amount of illumination present on the material's surface. These

intensities provide illumination samples of the material at the borders only. The illumination in the interior of the

material still needs to be computed. This again becomes a boundary interpolation problem which can be addressed

using an approach similar to equation 5.

B. Illumination Image Estimation

Let (u, v) represent 2D image coordinates aligned in the restored Y-channel image, I_Y. From Figure 11, we can

see that the intensity value at the 2D coordinates (x, y) along the boundary of $\mathbf{c}_1(u)$ should map to $L_1(u)$, i.e.

the top edge of the undistorted image. This correspondence holds for other curves, e.g. (x, y) coordinates along

the curve $\mathbf{c}_2(v)$ should map to $L_2(v)$. Given the intensity values along the boundaries, the pixel value of interior

points can be found using the following equation:

$$
\begin{aligned}
L_Y(u, v) = \begin{bmatrix} 1-u & u \end{bmatrix} & \begin{bmatrix} L_4(v) \\ L_2(v) \end{bmatrix} \\
+ \begin{bmatrix} L_1(u) & L_3(u) \end{bmatrix} & \begin{bmatrix} 1-v \\ v \end{bmatrix} \\
- \begin{bmatrix} 1-u & u \end{bmatrix} & \begin{bmatrix} L_1(0) & L_2(0) \\ L_3(1) & L_4(1) \end{bmatrix} \begin{bmatrix} 1-v \\ v \end{bmatrix},
\end{aligned}
\tag{11}
$$

where

$$L_i(u) = I_Y(\mathbf{c}_i(u)),$$

$$L_i(v) = I_Y(\mathbf{c}_i(v)),$$

where $\mathbf{c}_i(\cdot)$ returns the 2D image coordinate (x, y) along the appropriate boundary curve $\mathbf{c}_i$. The illumination image estimation is also based on Coons patch interpolation, however the data being blended is not (x, y) coordinates but intensity values along the border. Figure 11 shows the results of the illumination estimation. The desired intrinsic reflectance image and the subsequent image with shading artifacts removed can be computed using equation 10.

C. Experimental Results

In this section, we report evaluation results of our shading correction approach. The study is divided into two parts. We commence with a study of synthetic imagery aimed at evaluating the performance of the method on data with known ground truth. The second part of the study focuses on real world data and aims to demonstrate the utility of the method in real world situation.

1) **Experiment 1 (Evaluation using synthetic imagery):** In this experiment, we evaluate our shading correction approach by using synthetic imagery captured from the virtual scene. Three document models described in Section III-E are placed in the scene, mapped with texture and illuminated using directional light source (i.e. light source at infinity).

As we use OpenGL as the 3D modeling tool and it performs shading calculation in RGB colorspace, our shading correction approach follows the same track as the 3D tool in which we remove the shading artifacts for each RGB channels.

To quantify the quality of our algorithm, we compute the peak signal-to-ratio (PSNR) for image with shading correction and without shading correction. To ground truth the calculation, a reference image, I^{Ref}, is used. The reference image is generated by undistorting an imaged document with no shading. This can be achieved by turning off the shading function using OpenGL command *glDisable(GL_LIGHTING)*. The PSNR metric is defined over RGB colorspace and described as follows.

$$PSNR = 10log_{10}(\frac{S_{Max}}{MSE_{RGB}})$$

$$S_{Max} = 3 * (255)^2 \tag{12}$$

$$MSE_{RGB} = \frac{1}{w \times h} \sum_{0}^{w} \sum_{0}^{h} ((I_R^{Test} - I_R^{Ref})^2 + (I_G^{Test} - I_G^{Ref})^2 + (I_B^{Test} - I_B^{Ref})^2)$$

where I^{Test} denotes the testing image, I^{Ref} denotes the reference image, the subscript denotes the color channel of the colorspace, w and h denote the width and height of the testing image respectively.

Figure 12 shows the first synthetic example with binder curl distortion, where two different textures are mapped on to the model for evaluation. We use 24×8 control points to approximate the boundary curves, where arc-length and uniform parameterization are used for the correction. The geometric corrected results are shown in the first column. The second column shows the estimated illumination images using white border information from the input boundary. The third column shows the images with shading artifacts removed. The reference images are used as the ground truth and shown in the fourth column. To evaluate the quality of our photometric correction approach, quantitative results (PSNR) are presented in the right-most column. From the results, we observe that the shading corrected image has about three times higher PSNR than the one with shading artifact. This trend is valid over different parameterization schemes.

Same experiment is performed for the synthetic models of folded tablet and 3 page fold-out inset. Results of these experiments are shown in Figure 13 and Figure 14. The organization of the results is same as Figure 12 and a similar trend of improvement can be observed.

2) **Experiment 2 (Evaluation using real examples):** This experiment uses real examples to demonstrate the utility of our proposed approach. Results of the examples processed by our technique are presented. These examples use the two boundary parameterizations, (1) arc-length and (2) uniform, as mentioned in Section III-B. Using arc-length parameterization requires no knowledge about the depth change of the imaged material's boundary curves and acceptable results are obtained when the depth distortion is small compared to the size of the imaged material. Uniform parameterization is useful to correct item which has significant depth change. For

uniform parameterization, a paper checkerboard pattern is placed underneath the imaged material. This pattern undergoes same deformation as the printed media and is used to guide the uniform parameterization. We also show the computed illumination image for our input and show the results of the image with shading artifacts removed. Under controlled imaging environment, automatic border detection can be robustly performed using segmentation. For our experiments we use corner detection of the checkerboard pattern to supply the boundary points $\mathbf{x}_i$ since multiple parameterizations are used for comparison. For illumination estimation, the borders are adjusted slightly to be within the document's margin.

Example 1: Small Bird Example

Figure 15(a) shows the first example with binder curl distortion. A checkerboard pattern is placed underneath the page and uniform parameterization is used to correct geometric distortion. The result are shown in Figure 15(b). Figure 15(c) shows the estimated illumination image. Figure 15(d) shows the restored image with uniform illumination.

Example 2: Three Page Fold-Out Bird Example

Figure 16(a) shows an example of a three page fold-out inset from an oversized art book. Distortion is caused by both spline curl and folding. Arc-length parameterization, approximated by chord arc-length is used to correct the distortion. The result is shown in Figure 16(b). Figure 16(c) shows the estimated illumination image. Figure 16(d) shows the rectified image with shading corrected. This example appears rectified, but distortion due to the depth is present. This can be verified by comparing the page size of each segment in Figure 16(d) with that in the pressed flat control image shown in Figure 16(e).

Example 3: Buddhist Fold-Out Example

Figure 17(a) shows an example which is a large fold-out page from an oversized art book. This is an excellent representation of materials where existing techniques are not applicable. This book has many pages that fold-out to display very wide content. The page is $80cm \times 33cm$ in width and height and is very difficult to image completely flat and exhibits both fold and curl distortion.

Figure 17-I(b) shows the result using arc-length parameterization. The corners of the checkerboard pattern are used as the 2D feature points to define the boundary splines. Notice that the corrected image has artifacts due to the depth distortion. This can be verified by looking at the uneven black and white square pattern along the boundaries. Figure 17-II(b) shows the result which parameterized the boundary using the same 2D feature points but with uniform parameterization. This parameterization is justified because the 2D features are from uniform samples on the 3D curve obtained from the inserted pattern. The depth distortion has been corrected. Figure 17(c) shows the estimated illumination images using white border information from the input boundary. Figure 17(d) shows the new images under uniform illumination.

Example 4: Birds Example

Figure 18(a) shows another example of imaged art with significant "roll" distortion. Uniform parameterization is used to correct the geometric distortion, shown in 18(b). The contents of the imaged material have been rectified. Figure 18(c-d) shows the derived illumination image and resulting image with shading corrected.

V. Conclusion

This paper presents has presented two innovations for document imaging. First, we demonstrated how the printed materials can be modeled using a ruled surface composed of opposite boundary curves. This general surface model can model several common distortion found in printed materials, including: skew, folds, and binder-curl. We show that the projection of this model can be parameterized using Coons patch blending with four boundary curves. This 2D interpolation provides a mapping between the distorted and restored image allowing shape distortion to be corrected. In addition, 2.5D information can be easily incorporated into the interpolation process by changing only the curve's parameterization. This allows the correction of materials with significant depth distortion.

Secondly, we showed that the same interpolation framework can be used to estimate the intrinsic illumination image of the input image. From this intrinsic illumination image, we can compute the intrinsic reflectance image that exhibits no shading artifacts.

Currently, we use a physical pattern to guide the uniform parameterization. The use of such a pattern is reasonable in a restoration context. Our personal experience with digitization efforts, for preservation and archival purposes, is that a great deal of care and time is spent on each imaged item and additional props (such as rulers and color strips) are routinely incorporated into the imaging environment. Using a paper pattern under the imaged material would not present a great burden and could be reasonably adopted, especially if it is to facilitate restoration. We are currently working on techniques to remove the need for this pattern. Using the illumination image obtained from arc-length parameterization, we can apply shape from shading techniques to get an estimation of the 3D arc-length near the boundary to parameterize the 2D curve without the need of a pattern.

In summary, we have presented a novel approach based on boundary interpolation that can correct geometric distortion and shading artifacts present in images of printed materials. Our algorithm can simultaneously correct a variety of geometric distortions, including skew, folding, binder curl. In addition, the same interpolation framework can be used to estimate the illumination component of the input image which is used to correct shading artifacts. Our approach is fast, simple, and provides a unified approach to correct geometric and photometric distortions in imaged print materials.
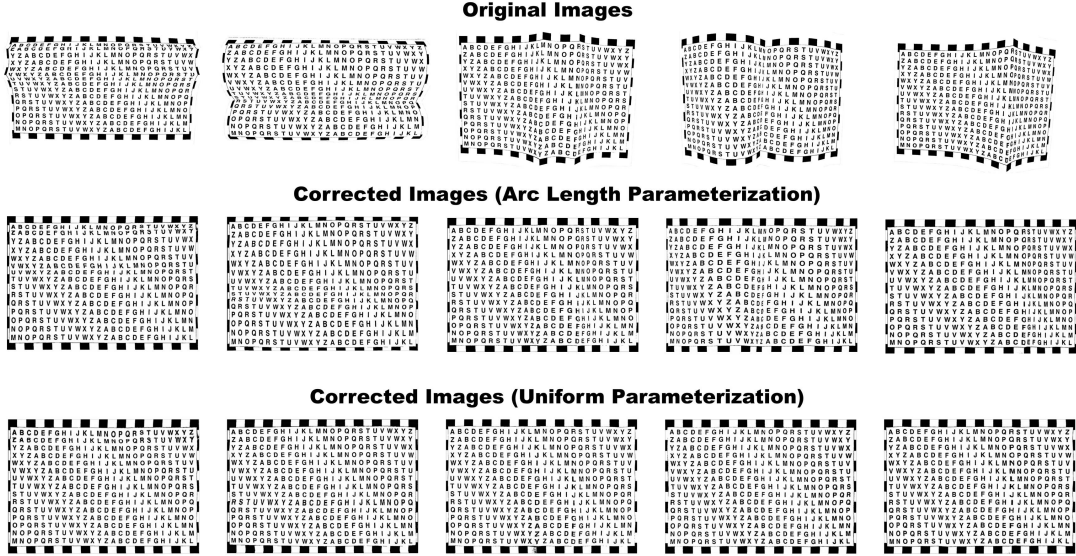
VI. ACKNOWLEDGMENTS

REFERENCES

[1] Avanidra and S. Chaudhuri. Robust detection of skew in document images. *IEEE Transactions on Image Processing*, 6(2):344–349, 1997.

[2] H.G. Barrow and J.M. Tenenbaum. *Recovering intrinic scene characterisitcs from images*. Academic Press, 1978.

[3] M. S. Brown and W. B. Seales. Document restoration using 3D shape. In *ICCV '01*, July 9-12 2001.

[4] M.S. Brown and Y.C. Tsoi. Correcting common distortions in camera-imaged library materials. In *IEEE/ACM Joint Conference on Digtial Libraries*, May 2003.

[5] M.S. Brown and Y.C. Tsoi. Undistorting imaged print materials using boundary information. In *Asian Conference on Computer Vision*, Feb 2004.

[6] H. Cao, X. Ding, and C. Liu. A cylindrical surface model to rectify the bound document image. In *ICCV'2003*, pages 228–233, 2003.

[7] H. Cao, X. Ding, and C. Liu. Rectifying the bound document image captured by the camera: A model based approach. In *Proc. 7th International Conference on Document Analsysis and Recognition*, 2003.

[8] S. Coons. Surfaces for computer aided design. Technical report, MIT, 1968. Available as AD 663 504 from the National Technical Information service, Springfield, VA 22161.

[9] D. Doermann, J. Liang, and H. Li. Progress in camera-based document image analysis. In *International Conference on Document Analysis and Recognition (ICDAR'03)*, pages p. 606–616, August 2003.

[10] G. Farin. *Curves and Surfaces for Computer Aided Geometric Design*. Academic Press, San Diego, CA, 1990.

[11] B.V. Funt, M.S. Drew, and M. Brockington. Recovering shading from color images. In *ECCV'92*, pages 123–132, 1992.

[12] B. Gatos, N. Papamarkos, and C. Chamzas. Skew detection in text line position determination in digitized documents. *Pattern Recognition*, 30(9):1505–1519, 1997.

[13] N. Gumerov, A. Zandifar, R. Duraiswami, and L. S. Davis. Structure of applicable surfaces from single views. In *European Conference on Computer Vision*, May 2004.

[14] H. Jiang, C. Han, and K. Fan. Automated page orientation and skew angle detection for binary document images. *Pattern Recognition Letters*, 18(7):125–133, 1997.

[15] T. Kanungo, R. Haralick, and I. Phillips. Global and local document degradation models. In *ICDAR-93*, pages 730–734, Tsukuba, Japan, October 20-22 1993.

[16] U. Pal and B. Chaudhuri. An improved document skew angle estimation technique. *Pattern Recognition Letters*, 8(17):899–904, July 1996.

[17] M. Pilu. Undoing paper curl distortion using applicable surfaces. In *CVPR '01*, Dec 11-13 2001.

[18] M. Tappen, W.T. Freeman, and E. Adelson. Recovering intrinsic images from a single image. In *Advances in Neural Information Processing Systems 15 (NIPS)*. MIT Press, 2003.

[19] Y.C. Tsoi and M.S. Brown. Geometric and shading correction of imaged printed materials: A unified approach using boundary. In *IEEE Computer Vision and Pattern Recognition*, June 2004.

[20] T. Wada, H. Ukida, and T. Matsuyama. Shape from shading with interreflections under proximal light source. In *ICCV '95*, pages 66–71, 1995.

[21] Y. Weiss. Deriving intrinsic images from image sequences. In *ICCV'01*, pages II: 68–75, July 2001.

[22] Z. Zhang and C. L. Tan. Restortation of images scanned from thick bound documents. In *ICIP '01*, Thessaloniki, Greece, Oct 2001.

Original Images

Corrected Images (Arc Length Parameterization)

Corrected Images (Uniform Parameterization)

Type	Trial 1			Trial 2			Trial 3			Trial 4			Trial 5		
	No. of Hits	No. of Misses	% of Error	No. of Hits	No. of Misses	% of Error	No. of Hits	No. of Misses	% of Error	No. of Hits	No. of Misses	% of Error	No. of Hits	No. of Misses	% of Error
Original	360/390	30/390	7.69%	365/390	25/390	6.41%	356/390	34/390	8.72%	358/390	32/390	8.21%	360/390	30/390	7.69%
Arc Length Parameterization	378/390	12/390	3.08%	390/390	0/390	0.00%	371/390	19/390	4.87%	375/390	15/390	3.85%	380/390	10/390	2.56%
Uniform Parameterization	385/390	5/390	1.28%	390/390	0/390	0.00%	387/390	3/390	0.77%	390/390	0/390	0.00%	390/390	0/390	0.00%

Fig. 10. The results of five trials using distortion correction with arc-length and uniform sampling on a controlled example with text. The table shows the results of using OCR to subsequently identify the letters in the restored image.
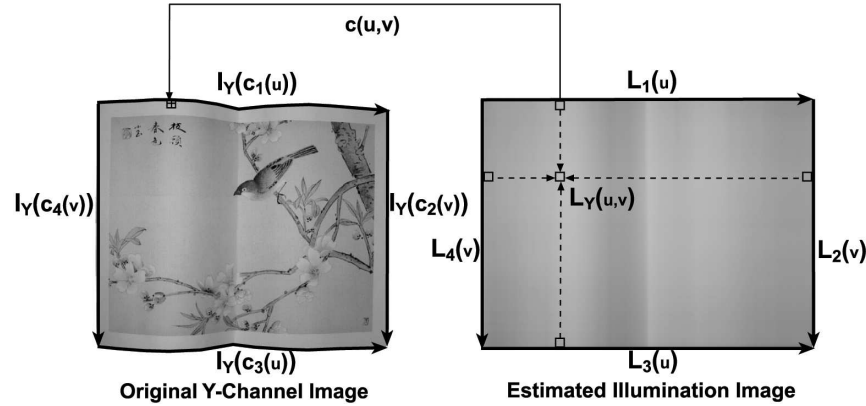
Fig. 11. Boundary intensity values along the curves c_1, c_2, c_3, c_4 correspond to the rectilinear edges of the shadow image $L_1(u), L_2(v), L_3(u), I_4(v)$ respectively. Each inner pixel value of coordinates (x, y) in the shadow image is found from boundary values using the function $L_Y(u, v)$ and pixel is re-sampled using bilinear interpolation.
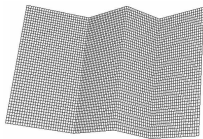
Binder Curl

Texture 1 (Fish example)

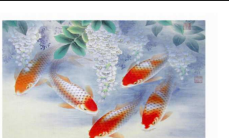|	Geometry Corrected Image with Shading (1)	Estimated Illumination Image	Geometry and Shading Corrected (2)	Reference Image	PSNR	
Arc-Length					(1)	15.2868
					(2)	43.1064
Uniform					(1)	14.9086
					(2)	45.7971

Texture 2 (Bird example)

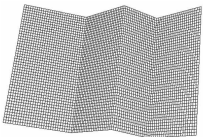|	Geometry Corrected Image with Shading (1)	Estimated Illumination Image	Geometry and Shading Corrected (2)	Reference Image	PSNR	
Arc-Length					(1)	15.2868
					(2)	43.1064
Uniform					(1)	14.9086
					(2)	45.7971

Fig. 12. Experimental results using synthetic example with binder curl distortion. Two textures are mapped on to the model and used for evaluation. Geometry corrected image with shading is shown in the first column. The second column shows the estimated illumination image. The third column shows images with geometry and shading corrected. The fourth column shows the reflectance image. Quantitative results using signal-to-noise ratio (PSNR) are presented in the right-most column.

Folded Tablet					
Texture 1 (Fish example)					
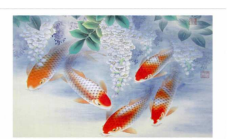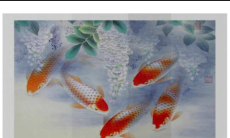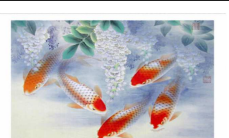
	Geometry Corrected Image with Shading (1)	Estimated Illumination Image	Geometry and Shading Corrected (2)	Reference Image	PSNR	
Arc-Length					(1)	15.6588
					(2)	41.3036
Uniform					(1)	15.6122
					(2)	41.2465

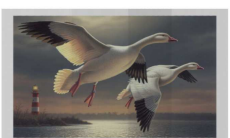| **Texture 2 (Bird example)**					
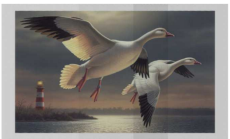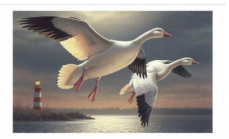
	Geometry Corrected Image with Shading (1)	Estimated Illumination Image	Geometry and Shading Corrected (2)	Reference Image	PSNR	
Arc-Length					(1)	17.4641
					(2)	41.7759
Uniform					(1)	17.4013
					(2)	41.7137

Fig. 13.   Experimental results of texture mapped synthetic model with fold distortion. Two different textures are used for evaluation. Geometry corrected image with shading is shown in the first column. The second column shows the estimated illumination image. The third column shows images with geometry and shading corrected. The fourth column shows the ground truth reflectance image. Quantitative results using peak signal-to-noise ratio signal-to-noise ratio (PSNR) are presented in the right-most column.
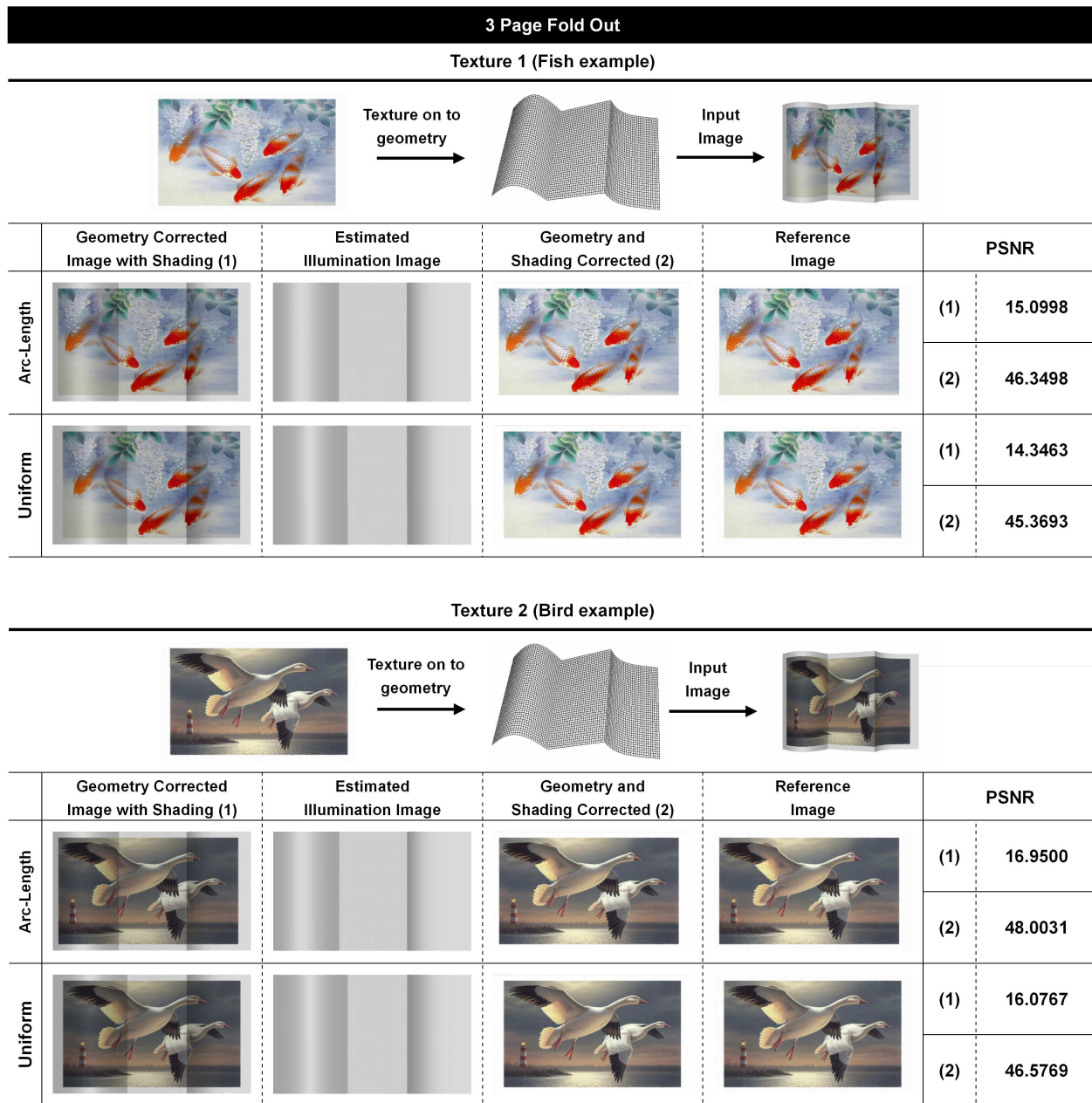
Fig. 14. Experimental results of synthetic model with a combination of fold and curl distortion. Two textures are used for evaluation. Geometry corrected image with shading is shown in the first column. The second column shows the estimated illumination image. The third column shows images with geometry and shading corrected. The fourth column shows the reflectance image. Quantitative results using signal-to-noise ratio (PSNR) are presented in the right-most column.

(a) Original Image

(b) Geometry Corrected

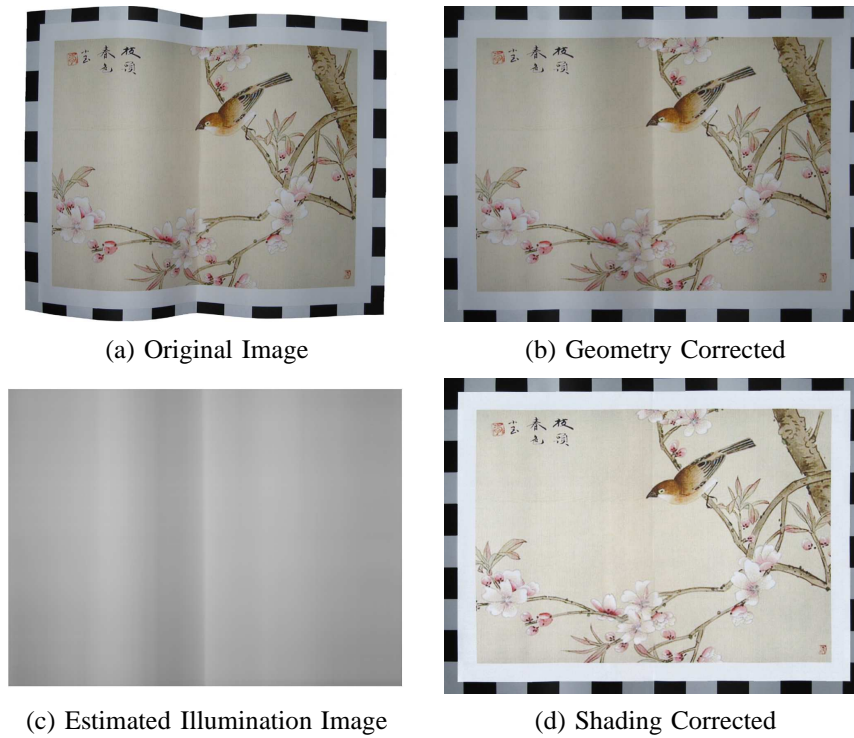(c) Estimated Illumination Image

(d) Shading Corrected

Fig. 15. Example of material with spine distortion. (a) Original image. (b) Corrected image with distortion removed. (c) Estimated intrinsic illumination image. (d) Shading corrected.

(a) Original Image

(b) Geometry Corrected Image

(c) Estimated Illumination Image

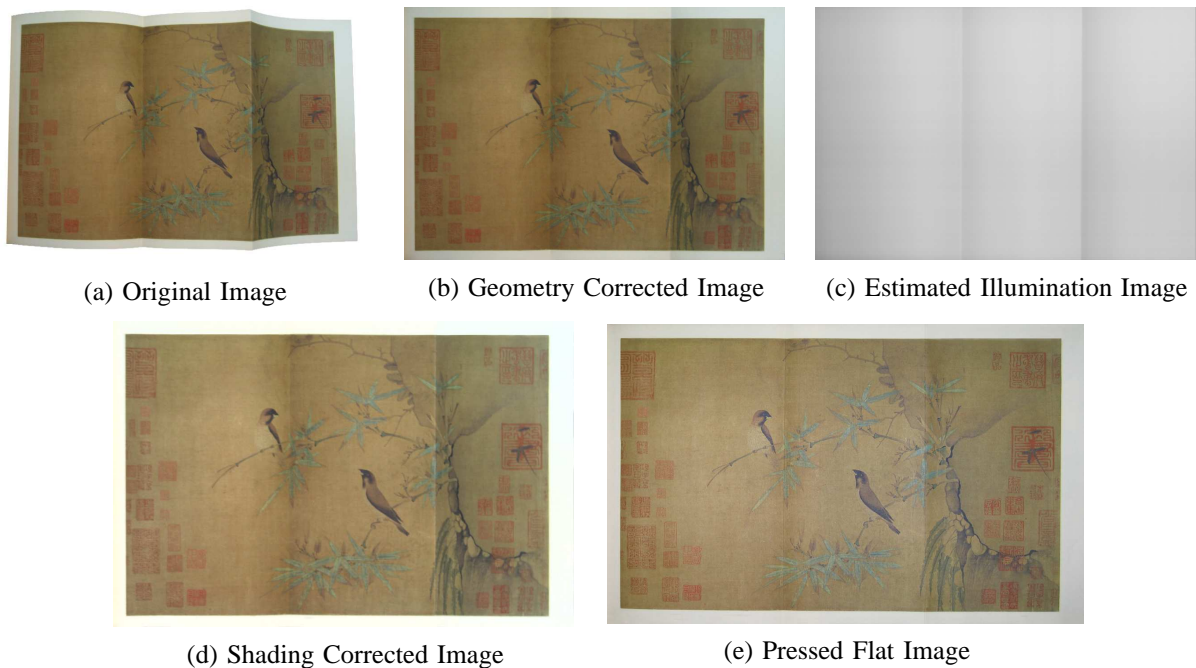(d) Shading Corrected Image

(e) Pressed Flat Image

Fig. 16. Example of material with both binder curl and fold distortion. (a) Original image (b) Corrected image with distortion removed (c) Estimated intrinsic illumination image (d) Shading corrected image (e) Pressed flat control image

(a) Original Image

(b) Geometry Corrected

(I) Arc Length Parameterization   **(II) Uniform Parameterization**

(c) Estimated Illumination Images

(I) Arc Length Parameterization   **(II) Uniform Parameterization**

(d) Shading Corrected

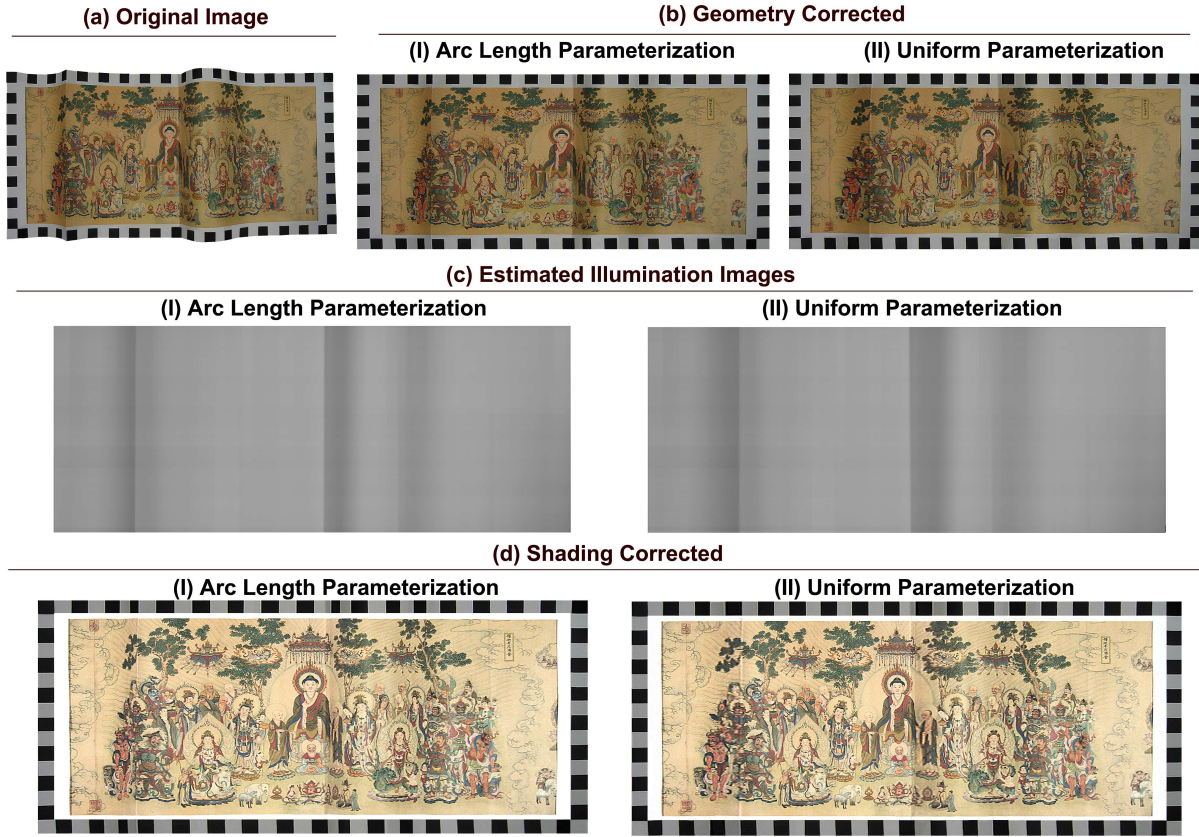**(I) Arc Length Parameterization**   **(II) Uniform Parameterization**

Fig. 17. Example of wide fold-out page from an art book, exhibiting both binder-curl and fold distortion. Corrected images using arc-length parameterization are shown on column I, those using uniform parameterization is shown in column II. The images are labeled: (a) Original image. (b) Geometry corrected. (c) Estimated illumination images. (d) Shading corrected.

(a) Original Image   (b) Geometry Corrected

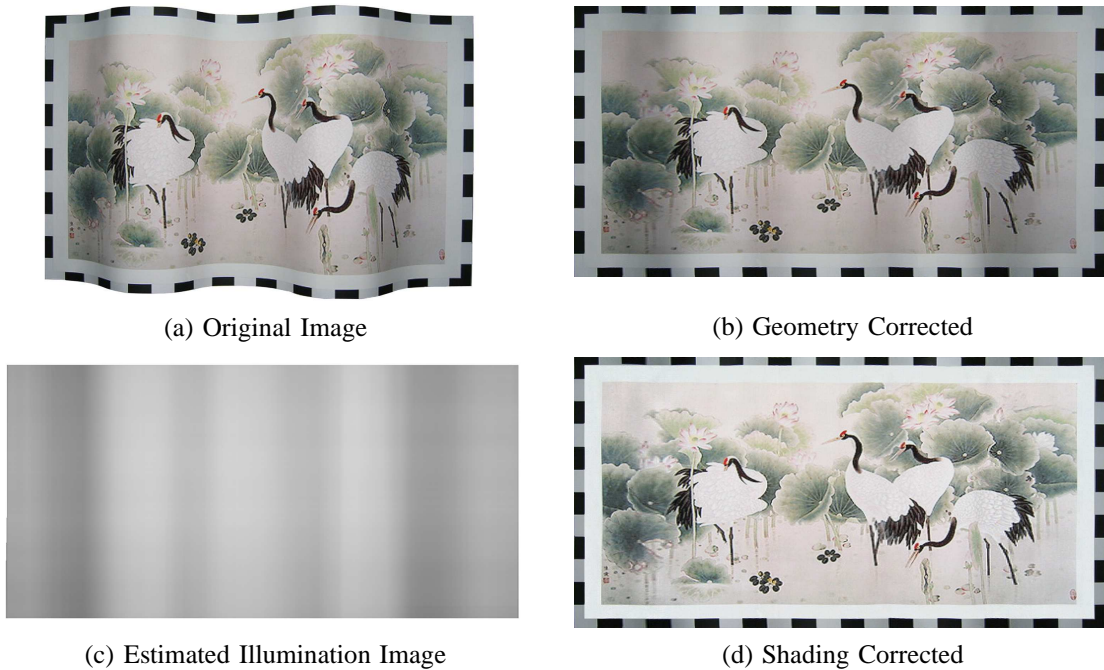(c) Estimated Illumination Image   (d) Shading Corrected

Fig. 18. Example of material with roll (curl) distortion. 2.5D information is incorporated to guide the uniform parameterization. (a) Original image. (b) Geometry corrected using uniform boundary sampling. (c) Estimated illumination image. (d) Shading corrected.