Restoring Warped Document Images through 3D Shape Modeling

Chew Lim Tan, *Senior Member*, *IEEE*, Li Zhang, Zheng Zhang, and Tao Xia

Abstract—Scanning a document page from a thick bound volume often results in two kinds of distortions in the scanned image, i.e., shade along the "spine" of the book and warping in the shade area. In this paper, we propose an efficient restoration method based on the discovery of the 3D shape of a book surface from the shading information in a scanned document image. From a technical point of view, this shape from shading (SFS) problem in real-world environments is characterized by 1) a proximal and moving light source, 2) Lambertian reflection, 3) nonuniform albedo distribution, and 4) document skew. Taking all these factors into account, we first build practical models (consisting of a 3D geometric model and a 3D optical model) for the practical scanning conditions to reconstruct the 3D shape of the book surface. We next restore the scanned document image using this shape based on deshading and dewarping models. Finally, we evaluate the restoration results by comparing our estimated surface shape with the real shape as well as the OCR performance on original and restored document images. The results show that the geometric and photometric distortions are mostly removed and the OCR results are improved markedly.

Index Terms—Document image restoration, document image analysis, shape from shading, image warping, image distortion, OCR improvement.

1 Introduction

Scanning a document page from a thick bound volume often results in two kinds of distortions in the scanned image as in Fig. 1, i.e., shade along the "spine" of the book (photometric distortion) and warping in the shade area (geometric distortion). The distortion introduces problems in image restoration of archived documents and in word spotting for document retrieval from the digital library. Given these problems and the large number of books to be scanned, our objective in this research is to design an efficient content-free algorithm to restore the document images by correcting the above two types of distortion.

There have been related techniques reported in the literature. We classify them into two categories: 1) restoration approaches based on 2D document image processing and 2) restoration approaches based on 3D document shape discovery.

1.1 Restoration Approaches Based on 2D Document Image Processing

These approaches basically make use of 2D information in the document image to find distortion parameters in order to do restoration. Baird discusses a 10-parameter document image defect model [1] with further refinements [2], [3]. These models approximate some document printing and imaging parameters, including symbol size, digitizing resolution, affine spatial deformations, jitter, blurring, etc. Tang and Suen [4] adopt 2D geometrical transformations based on the finite element theory to approximate 3D distortions. Lavialle et al. [5] introduce a new text line straightening method based on an active contour model with the use of Bezier curves followed by cubic B-splines for more accurate results. Weng and Zhu [6] propose the use of two-dimensional spline functions in bicubic, biquadratic, and/or bilinear models to approximate three-dimensional nonlinear distortion curves. Tsoi and Brown [7] introduce an approach to restore warped document images using 2D image boundaries. We have also proposed a regression model of curved text lines [8] with further improvements in [9], [10], [11] to permit subsequent straightening of text lines.

1.2 Restoration Approaches Based on 3D Document Shape Discovery

Approaches in this category can be further divided into two subcategories. The first involves image capture with special setup, while the second works directly on camera or scanner images without special setup. We first examine image capture with special setup. First, Pilu [12] presents a novel method that models the paper deformations of a curled document using an applicable surface and then unrolls it to a plane to obtain the undistorted image. Brown and Seales [13] propose an image restoration approach through the acquisition of the document 3D geometry using a structured light 3D acquisition system and restoring it using a particle-based system and a mass spring model. The method does not correct the shading in the images. Doncescu et al. [14] propose a similar method in which a laser projector is used to project a 2D light network on the document surface to capture the 3D shape and then 2D distortions of the surface are corrected with a two-pass mesh dewarping algorithm. Yamashita et al. [15] propose a shape reconstruction method by using a two-camera stereo

- *C.L. Tan, L. Zhang, and Z. Zhang are with the School of Computing, National University of Singapore 3, Science Drive 2, Singapore 117543. E-mail: {tancl, zhangli, zhangz}@comp.nus.edu.sg.*
- *T. Xia is with the Department of Mathematics, National University of Singapore 2, Science Drive 2, Singapore 117543. E-mail: iesxiat@nus.edu.sg.*

Manuscript received 1 Oct. 2004; revised 18 Apr. 2005; accepted 14 June 2005; published online 13 Dec. 2005.
Recommended for acceptance by P. Torr.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0523-1004.

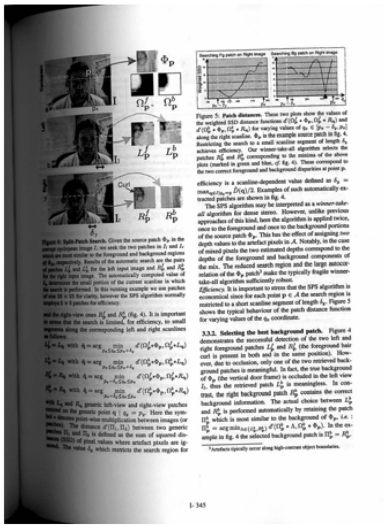

Fig. 1. A gray-scale image scanned from a bound document.

vision system. They reconstruct the 3D document surface shape by using a Non-Uniform Rational B-Spline curve representation and transform the two original images of curved surfaces to those of flat surfaces.

The other subcategory of the 3D shape discovery approach pertains to a reconstruction of the 3D model from images without special setup. Cao et al. [16], [17] build a general cylindrical model from a camera image without special setup and use the skeleton of horizontal text lines in the image to help estimate the model parameters. Kanungo et al. [18] introduce a global degradation model for perspective distortion simulation in which the perspective geometry of the photocopier/scanner's optical system is studied with the assumptions that the warped surface is circular cylindrical and the lighting direction is vertical. Wada et al. [19] develop a complicated model to reconstruct the 3D book surface, incorporating interreflections (increased illumination on one part of the book caused by secondary reflections from another) with eight-parameters that are estimated a priori by using images of white flat slopes with known slants. This method requires the book spine to be strictly parallel to the scanning light and assumes the book surface to be cylindrical. Another limitation is its high computational cost in dealing with interreflections even with the tessellation method they propose. We [20], [21] have proposed a much more efficient system than Wada's model, which also assumes parallel alignment of the scanned documents with the light source. But, we have done away with the complicated computation of interreflection.

Comparing the two subcategories, the first can handle arbitrary image distortions including page crumple and folding, but it requires special setup to assist in 3D shape discovery. On the other hand, methods in the second subcategory obtain 3D information directly from the images. However, they can only handle the smooth distortion of image surface, such as cylindrical surfaces. The method reported in this paper belongs to the second subcategory.

1.3 Present Work

This paper reports new results from further improvements on our previous model [20], [21]. We now do not require the book spine to be parallel with the scanning light axis, which is unrealistic under normal scanning conditions. Furthermore, we propose a new method of computing the depth map to replace the estimation approach in [20] and the iterative algorithm in [21]. Basically, our model tries to uncover the geometrical and lighting information with respect to the scanner's structure and the book surface. This enables us to transform the warped book surface image into its flattened rendition. As a measure of success, we compare our reconstructed surface shape with the real surface shape using quantitative measures. We also show the shading recovery in terms of the intensity restoration along each column. Furthermore, we reimplement Wada's method and compare it with our method in terms of efficiency and OCR performance. A great speedup is achieved for our method with a good restoration result which is shown by the OCR testing conducted on various types of document images. The OCR results here serve as a secondary measure of the effectiveness of our method. Last but not least, we discuss several breakdown points of our method in terms of image contrast, surface shape, and material.

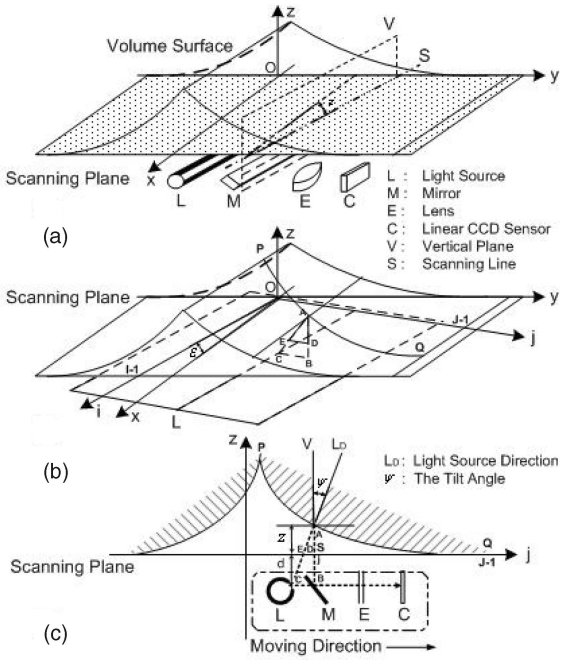

Fig. 2. The practical scanning conditions.

2 Practical Models

Fig. 1 shows a typical scanned gray-scale document image. Fig. 2a, adapted from [19], shows the structure of the image scanner and the coordinate system of the book surface. The scanner consists of a light source L, a linear CCD sensor C, a mirror M, and a lens E. The sensor C takes a 1D image $P(i)$ $(0 \leq i < image\ length\ I)$ along the scanning line S and moves with L, M, and E. Thus, a 2D image $P(i,j)$ $(0 \leq j < image\ width\ J)$ is formed from the sequence $P(i)$. Note that, in real scanning conditions, the book to be scanned may not be aligned strictly parallel with the

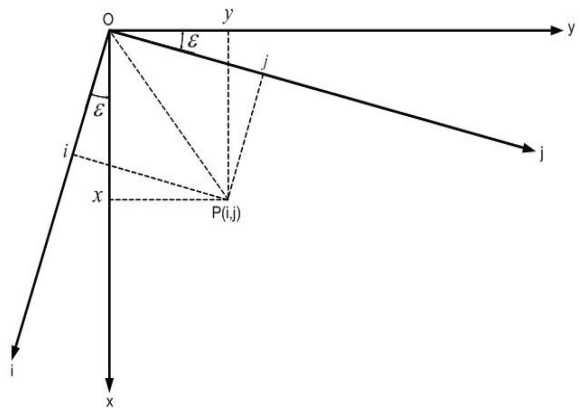

Fig. 3. Transformation between the i-j image indices and the x-y coordinates.

scanning light, i.e., the scanning line S may not be parallel to the book spine. Let ε be the angle between the scanning line S and the projection of book spine on the scanning plane. As shown in Fig. 2a, we define the 3D coordinate system of the book surface as:

- The origin O is located at the point on the scanning plane, i.e., $P(0,0)$ of the 2D image.
- The *x*-axis is located on the scanning plane and parallel to the book spine. It has positive values for the book spine.
- The *y*-axis is located on the scanning plane and parallel to the horizontal book boundary. It has positive values for the book surface captured by the scanner.
- The *z*-axis is perpendicular to the x-y plane and has positive values for the book surface.

Thus, the image coordinate system and the book surface coordinate system share the same vertical axis, i.e., *z* axis. The transformation between the image indices i-j and book surface coordinates x-y is shown in Fig. 3. As defined earlier, ε is the angle between the *i-axis* and *x-axis*, and ε has positive values for counterclockwise document skew and negative values for clockwise skew. We then define the function $x(i,j)$ and $y(i,j)$, which return the corresponding x and y coordinates for a point $P(i,j)$ as follows:

$$x(i,j) = \sqrt{i^2+j^2} \cdot \cos\left(\arctan\frac{j}{i} - \varepsilon\right), \quad (1)$$

$$y(i,j) = \sqrt{i^2+j^2} \cdot \sin\left(\arctan\frac{j}{i} - \varepsilon\right), \quad (2)$$

where $0 \le i < I$ and $0 \le j < J$. Similarly, the function $i(x,y)$ and $j(x,y)$, which return the corresponding image index (i,j) for given coordinate (x,y), are defined as:

$$i(x,y) = \sqrt{x^2+y^2} \cdot \cos\left(\arctan\frac{y}{x} + \varepsilon\right), \quad (3)$$

$$j(x,y) = \sqrt{x^2+y^2} \cdot \sin\left(\arctan\frac{y}{x} + \varepsilon\right). \quad (4)$$

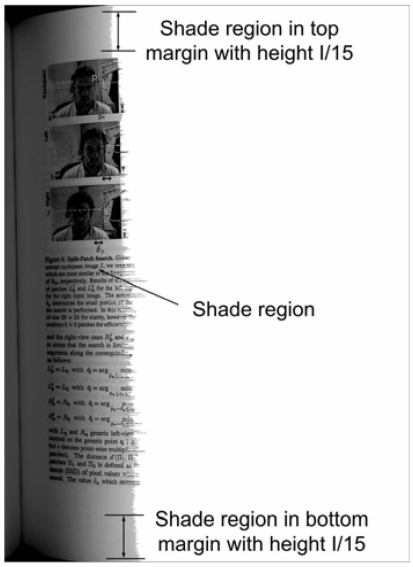

Fig. 4. The shade boundary detected for Fig. 1.

The document skew ε is detected by the following two steps:

Step 1. Detect the shade region using the run-length method in our previous paper [10]. The shade region detected for Fig. 1 is shown in Fig. 4.

Step 2. Extract the partial shade regions within the top and bottom margins of height $I/15$.

Step 3. Apply the Hough Transform on the extracted shade regions to find the straight line L that fits the maximum number of pixels in the shade region. If the number of pixels on L exceeds 50 percent of the pixels on the extracted partial shade regions, we set ε to be the angle between L and the vertical image boundary; otherwise, set $\varepsilon = 0$.

Fig. 2b shows a condition where a point C on the directional linear light source L casts a beam of light on a point A on the book surface. E is the intersection point of line AC and the scanning plane. B and D are the projection points of A on the scanning plane and the imaging plane (where the light source and CCD array lie on), respectively. Thus, plane ABC is perpendicular to L, and curve PAQ is the intersection of plane ABC and the book surface. Note that curve PAQ is the cross-section shape of the book surface on the j-z plane instead of the y-z plane since the plane ABC is perpendicular to the light source L (also the *i-axis*) instead of *x-axis* (also the book spine). Fig. 2c shows the cross-section shape of the book surface in the j-z plane, where V is the vertical to the scanning plane, L_D is the light source direction, ψ is the tilt angle between the light source and the normal to the scanning plane, z is the distance from A to the scanning plane, and d is the distance from scanning plane to light source.

We specify our problem using two practical models: a 3D geometric model and a 3D optical model. We introduce the following assumptions in our practical models:

Assumption 1. *The discussion of our method in this paper is based on generalized cylindrical book surfaces. By cylindrical,*

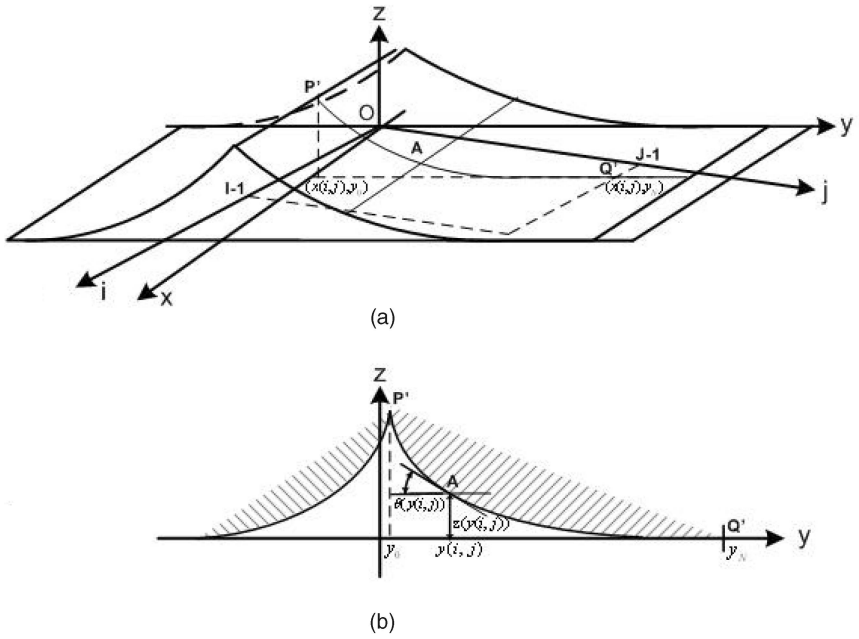

Fig. 5. The cross-section shape of the book surface in the (a) x-y-z space and the (b) y-z plane.

we mean a uniform height along the book spine, while by generalized, we mean a smooth cross-sectional shape including circular, elliptical, hyperbolic, and parabolic. This is reasonable as the book pages are usually uncreased and the book is normally leveled because of its uniform weight.

Assumption 2. *The book surface is Lambertian, i.e., no specular reflections and uniform brightness of the reflected light in all directions. This is reasonable as most papers are nearly Lambertian. Even with non-Lambertian material, its gray-scale image obtained from scanning does not affect much of the restoration result.*

2.1 The 3D Geometric Model

Suppose the 3D point A in Fig. 2b corresponds to the pixel $P(i,j)$ in the scanned document image. By Assumption 1, the z values are uniform along the x-axis. Thus, the 3D coordinates of A in the x-y-z space can be represented as $(x(i,j), y(i,j), z(y(i,j)))$, where $x(i,j)$ and $y(i,j)$ are calculated from (1) and (2), respectively. Fig. 5a shows a cross-section shape P'AQ' of the book surface at $x(i,j)$ in the y-z plane. Note that plane PAQ in Fig. 2b is perpendicular to the i-axis, while plane $P'AQ'$ here is perpendicular to the x-axis. I and J denote the image length and width, respectively. y_0 denotes the y value of the book spine and y_n denotes the largest y value of the points on curve $P'AQ'$.

Fig. 5b shows the cross-section shape in the y-z plane. Our Assumption 1 reduces the 3D book surface in the x-y-z space to the 2D cross-section shape in the y-z plane. Mathematically, we have $z(y_N) - z(y(i,j)) = \int_{y(i,j)}^{y_N} z'(y)dy$. Since $y(i,j)$ and y_N are integers and $z(y_N) = 0$, in our discrete y-z coordinate system, the depth map $z(y(i,j))$ is represented as:

$$z(y(i,j)) = \sum_{y_k=y_N}^{y(i,j)+1} \int_{y_k-1}^{y_k} -z'(y) = \sum_{y_k=y_N}^{y(i,j)+1} \tan\theta(y_k) \quad \left(y_0 \le y(i,j) \le y_N, \theta \in \left[0, \frac{\pi}{2}\right]\right), \tag{5}$$

where $y_N = y(l, J-1)$ for $x(l, J-1) = x(i,j)$ and $0 \le l < I$, y_0 is the y coordinate of the point separating the book pages in the cross-section shape, and $\theta(y_k)$ the slant angle at y_k. Note that y_N is not a constant, but varies with $x(i,j)$. The image indices corresponding to y_k can be calculated by substituting $(x(i,j), y_k)$ to (3) and (4), respectively.

2.2 The 3D Optical Model

For the real-world scanning conditions shown in Fig. 2, we formulate our shape-from-shading (SFS) problem by considering the following four factors:

- **Factor 1:** Proximal and moving light source. The linear directional light source of the scanner moves during the scanning process and is located very close to the book surface. This implies that the illuminant intensity varies with respect to the location on the book surface.
- **Factor 2:** Lambertian reflection. The book surface is Lambertian, i.e., the reflected light intensity observed by the CCD sensor is proportional only to the cosine of angle of incidence.
- **Factor 3:** Nonuniform albedo distribution. This is due to the printed surface of the book.
- **Factor 4:** Document skew ε. The book to be scanned is not necessarily aligned parallel to the linear directional light source of the scanner.

Let us first consider an ideal shape-from-shading problem, which satisfies:

1. A distant and fixed light source: This implies that the illuminant intensity and the light source direction are constant all over the object surface.
2. Lambertian reflection.
3. Uniform albedo distribution: The albedo is constant all over the book surface.
4. No document skew: The book spine is parallel to the linear directional light source of the scanner, i.e., $\varepsilon = 0$.

The problem under these ideal conditions can be formulated as:

$$I_o(p) = I_s \cdot k \cdot \cos\varphi(p), \tag{6}$$

where $I_o(p)$ denotes the reflected light intensity observed at p, which is a 2D point in the scanned image. I_s is the illuminant intensity, k is the albedo on the surface, and $\varphi(p)$ is the angle between the direction of the light source and the surface normal at the 3D point on the book surface corresponding to p.

2.2.1 Proximal and Moving Light Source (Factor 1)

With a proximal and moving light source, the illuminant intensity is no longer constant over the object surface. The illuminant intensity on one point p is now a function of the location of p and that of the light source corresponding to p. We can formulate the problem as follows:

$$I_o(p) = I_s(s(p), l(p)) \cdot k \cdot \cos\varphi(p), \tag{7}$$

where $s(p)$ and $l(p)$ denote the 3D point on the book surface and light source location corresponding to p, respectively.

2.2.2 Lambertian Reflection and Nonuniform Albedo Distribution (Factors 2 and 3)

Next, we can formulate our problem by incorporating Lambertian reflection and nonuniform albedo distribution characteristics into (7):

$$I_o(p) = I_s(s(p), l(p)) \cdot k(s(p)) \cdot \cos\varphi(p), \tag{8}$$

where $k(s(p))$ denotes the albedo at $s(p)$, and $\cos\varphi(p)$ denotes the reflectance property.

2.2.3 Document Skew ε (Factor 4)

Finally, as shown in Fig. 2a and Fig. 2b, we take the document skew ε into account. By the coordinate system in Fig. 2 and (8), the relationship between the image intensity $P(i,j)$ (pixel value) and the reflected light intensity is represented as follows:

$$\begin{aligned} P(i,j) &= \alpha \cdot I_o(i,j) + \beta \\ &= \alpha \cdot I_s(z(y(i,j))) \cdot k(x(i,j), y(i,j)) \cdot \cos\varphi(y(i,j)) + \beta. \end{aligned} \tag{9}$$

- $P(i,j)$: The image intensity at (i,j) in the observed image, i.e., the scanned image.
- α, β: The gain and bias of the photoelectric transformation in the scanner, respectively.
- $x(i,j)$, $y(i,j)$: The x and y coordinates of the 3D point corresponding to $P(i,j)$, which are calculated from (1) and (2), respectively.
- $z(y(i,j))$: The z coordinate of the 3D point on the book surface corresponding to $P(i,j)$ in the observed image. By Assumption 1, the z values are constant along x-axis. $z(y(i,j))$ is the distance between the scanning plane and the book surface, which is represented as (5).
- $I_s(z(y(i,j)))$: The illuminant intensity distribution on the y-z plane when taking the 1D image at $(x(i,j), y(i,j))$, i.e., $I_s(z(y(i,j)))$ is the practical representation of $I_s(s(p), l(p))$. Due to the constant tilt angle ψ of the directional linear light source of the scanner, as shown in Fig. 2b and Fig. 2c, the 3D points with the same z value have the same distance $(AD + BD)/\cos\psi$, i.e., $(z(y(i,j)) + d)/(z(y(i,j)) + d)/\cos\psi$, from the light source. Note that the document skew ε has no effect on this distance, and the book surface depth $z(y(i,j))$ is the only effective factor. Based on the directional linear light source model, $I_s(z(y(i,j)))$ can be represented as follows:

$$\begin{aligned} I_s(z(y(i,j))) &= \frac{I_D(\psi)}{((z(y(i,j)) + d)/\cos\psi)} \\ &= \frac{I_D(\psi) \cdot \cos\psi}{z(y(i,j)) + d}, \end{aligned} \tag{10}$$

  where ψ is the tilt angle between the light source direction and the normal to scanning plane and $I_D(\psi)$ is the directional distribution of the illuminant intensity.
- $k(x(i,j), y(i,j))$: The albedo of the 3D point corresponding to $P(i,j)$.
- $\cos\varphi(y(i,j))$: The cosine of the angle φ between the light source direction and the surface normal at the 3D point $(x(i,j), y(i,j), z(y(i,j)))$. By Assumption 2, the book surface is Lambertian. The reflected light intensity is proportional only to $\cos\varphi$. Since the directional linear light source of the scanner has a constant tilt angle ψ and, by Assumption 1, both the light source direction and the surface normal are constant along x-axis, we can represent φ as $\varphi(y(i,j))$. Note that $\varphi(y(i,j))$ is neither in the plane PAQ nor the plane $P'AQ'$. We shall discuss the relation between $\varphi(y(i,j))$ and $\theta(y(i,j))$ (the slant angle in the cross-section shape $P'AQ'$) in the next section.

In the experiments, parameters α, β, and $I_D(\psi)$ are estimated a priori using calibration images attached with the scanner. Parameters d and ψ are estimated by measurement of the geometry of the projected scanner light.

3 Reducing the 3D Practical Models to a Unique 2D Cross-Section Shape

Note that, after we built the geometric model and optical model, we have reduced the 3D book surface reconstruction problem to the 2D cross-section shape reconstruction problem. However, we are facing the following two problems:

Problem 1. *In the geometric model, the 2D cross-section shape in the y-z plane of the 3D book surface is not unique since y_N varies with $x(i,j)$ in (5). That means the number of $z(y(i,j))$ and $\theta(y(i,j))$ pairs is different for different $x(i,j)$.*

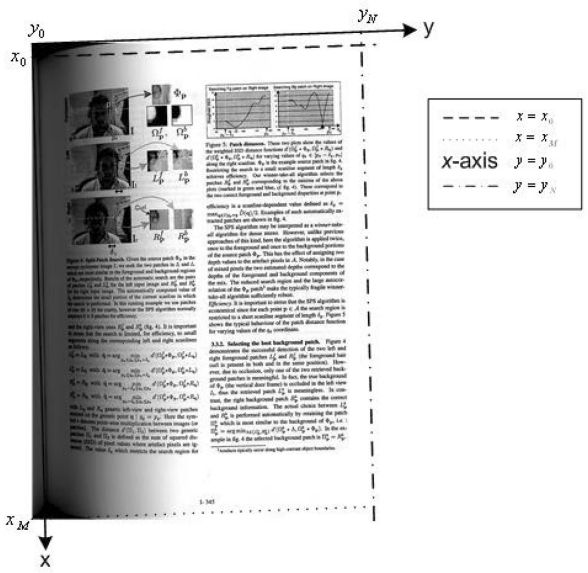

Fig. 6. The processing area of the document image in Fig. 1.

Problem 2. *Both $\theta(y(i,j))$ in the geometric model (5) and $\varphi(y(i,j))$ in the optical model (9) are unknown, which leads to a greater number of unknown variables than the number of equations in our practical models.*

The solution of the depth map $z(y(i,j))$ (2D cross-section shape on the y-z plane) cannot be found under these conditions. In this section, we describe a method to reduce the 3D book surface to a unique 2D cross-section shape by solving Problems 1 and 2.

3.1 The Processing Area of the Document Image (Solving Problem 1)

In order to reduce the 3D book surface to a unique 2D cross-section shape, we define four lines in the x-y plane: $x = x_0$, $x = x_M$, $y = y_0$, and $y = y_N$, where

$$x_0 = \max(0, x(0, J-1)), \tag{11}$$

$$x_M = \min(x(I-1, 0), x(I-1, J-1), \tag{12}$$

$$y_0 = \max(0, y(I-1, 0)), \tag{13}$$

$$y_N = \min(y(0, J-1), y(I-1, J-1)). \tag{14}$$

I and J are the image length and width, respectively, $x(0, J-1)$, $x(I-1, 0)$, $x(I-1, J-1)$, $y(I-1, 0)$, $y(0, J-1)$, and $y(I-1, J-1)$ are calculated from (1) and (2).

We define the processing area of the document to be the rectangle bounded by the four lines defined by (11)-(14). Fig. 6 shows the processing area of Fig. 1. The 3D surface is now reduced to a unique 2D cross-section shape between y_0 and y_N on the y-z plane.

3.2 The Relation between $\theta(y(i,j))$ and $\varphi(y(i,j))$ (Solving Problem 2)

By knowing the document skew ε, we may derive the relation between $\theta(y(i,j))$ and $\varphi(y(i,j))$. Thus, we can reduce the number of unknown variables by one. Fig. 7a shows the book surface and its two coordinate systems. The plane ABC and ABC′ are perpendicular to the light source (also the i-axis) and the book spine (also the x-axis), respectively, and intersect the book surface at curve PAQ (the cross-section shape of the book surface in the j-z plane) as shown in Fig. 2b and curve P′AQ′ (the cross-section shape of the book surface in y-z plane) as shown in Fig. 5a, respectively.

Fig. 7b shows the geometric structure of $\theta(y(i,j))$ and $\varphi(y(i,j))$ at $y(i,j)$ ($y_0 \le y(i,j) \le y_N$, where y_0 and y_N are defined by (13) and (14), respectively). $\theta(y(i,j))$ denotes the slant angle of the 2D cross-section shape P′AQ′ at $y(i,j)$, $\varphi(y(i,j))$ denotes the angle between the light source direction L_D and the surface normal $N(y(i,j))$ at $y(i,j)$ (due to Assumption 1, the surface normal of the points along x-axis are constant, which can be expressed as $N(y(i,j))$), denotes ψ the constant tilt angle of the scanner light source, and denotes V the vertical to the scanning plane.

In Fig. 7c, we define a local coordinate system for point A on the book surface, whose origin is offset from the origin O in Fig. 7a to the point A itself, and the orientations of all the axes remain unchanged. We rename the axes as x', y', z', i', and j', corresponding to the x, y, z, i, and j axes in Fig. 7a, respectively. Thus, the angle between x'-axis and i'-axis is equal to ε. We define L' and N' to be the unit vector for L_D and $N(y(i,j))$ in Fig. 7b, respectively, i.e., $||L'|| = ||N'|| = 1$. In Fig. 2b, since $\theta(y(i,j))$, $N(y(i,j))$, and V are coplanar, and $N(y(i,j))$ and V are perpendicular to the two edges of $\theta(y(i,j))$, the angle between $N(y(i,j))$ and V is equal to $\theta(y(i,j))$, i.e., the angle between the unit vectors $||N'||$ and $||V'||$ in Fig. 7c is equal to $\theta(y(i,j))$.

Let $L'_{i'j'z'} = [i_{L'}, j_{L'}, z_{L'}]$ and $N'_{x'y'z'} = [x_{N'}, y_{N'}, z_{N'}]$ denote the unit vector representation for L' and V' in the i'-j'-z' and x'-y'-z' coordinate systems, respectively. According to Fig. 7c, $N'_{x'y'z'} = [0, \sin\theta(y(i,j)), \cos\theta(y(i,j))]$ and $L'_{i'j'z'} = [0, \sin\psi, \cos\psi]$. Let $L'_{x'y'z'} = [x_{L'}, y_{L'}, z_{L'}]$ denote the vector representation for L' in the x'-y'-z' coordinate system. The relation between $L'_{x'y'z'} = [x_{L'}, y_{L'}, z_{L'}]$ and $L'_{i'j'z'} = [i_{L'}, j_{L'}, z_{L'}] = [0, \sin\psi, \cos\psi]$ in matrix form are shown as follows:

$$L'_{x'y'z'} = \begin{bmatrix} \cos\varepsilon & \sin\varepsilon & 0 \\ -\sin\varepsilon & \cos\varepsilon & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \sin\psi \\ \cos\psi \end{bmatrix} = \begin{bmatrix} \sin\varepsilon \cdot \sin\psi \\ \cos\varepsilon \cdot \sin\psi \\ \cos\psi \end{bmatrix}. \tag{15}$$

Since $\varphi(y(i,j))$ is the angle between L' and N', we have:

$$\begin{aligned} &\cos\varphi(y(i,j)) \\ &= < L'_{x'y'z'}, N'_{x'y'z'} > \\ &= \cos\varepsilon \cdot \sin\psi \cdot \sin\theta(y(i,j)) + \cos\psi \cdot \cos\theta(y(i,j)). \end{aligned} \tag{16}$$

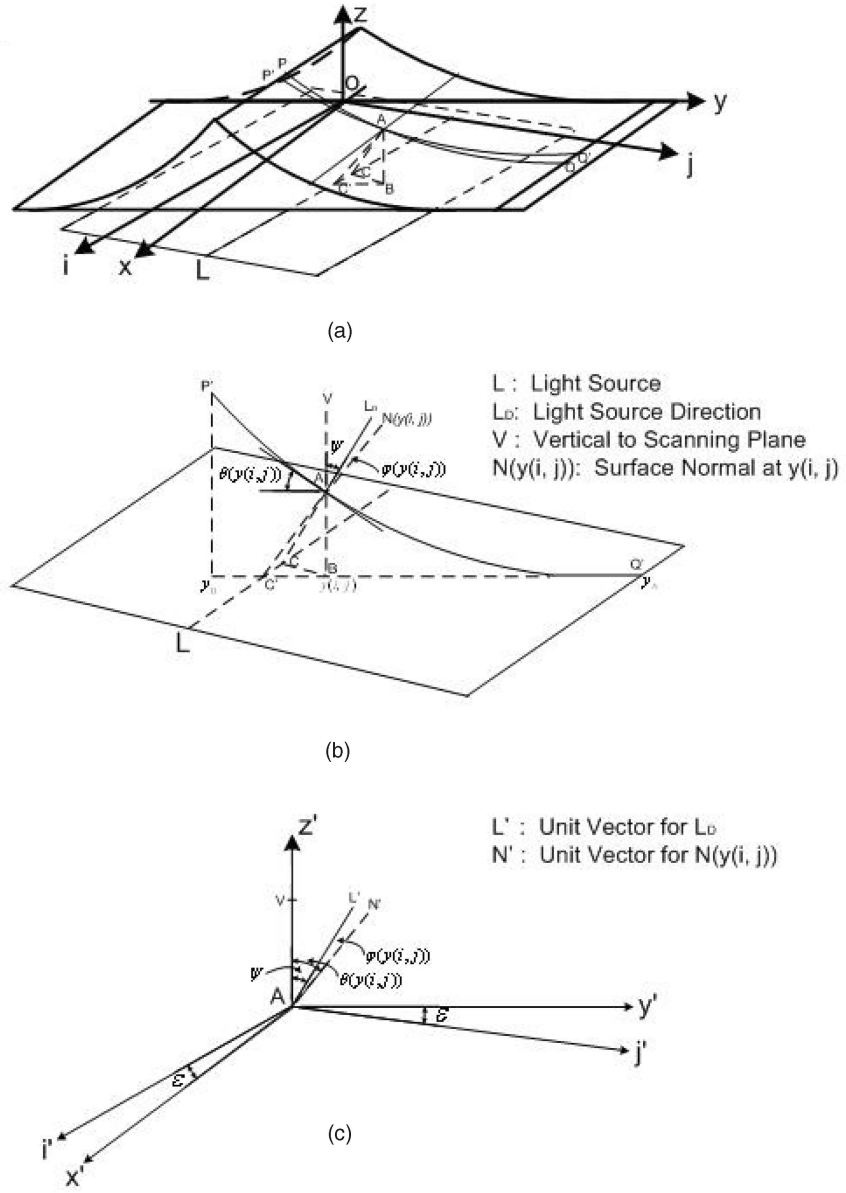

Fig. 7. The schematic drawing of the relation between $\theta(y(i,j))$ and $\varphi(y(i,j))$.

Let $a = \cos\varepsilon \cdot \tan\psi$ and we have:

$$\begin{aligned}
&\cos\varphi(y(i,j)) \\
&= \cos\psi \cdot [\cos\varepsilon \cdot \tan\psi \cdot \sin\theta(y(i,j)) + \cos\theta(y(i,j))] \\
&= \cos\psi \cdot [a \cdot \sin\theta(y(i,j)) + \cos\theta(y(i,j))] \\
&= \cos\psi \cdot \sqrt{a^2+1} \cdot \left[\frac{a}{\sqrt{a^2+1}} \cdot \sin\theta(y(i,j)) + \frac{1}{\sqrt{a^2+1}}\right. \\
&\quad \left. \cdot \cos\theta(y(i,j))\right].
\end{aligned} \tag{17}$$

Since $\left(\frac{a}{\sqrt{a^2+1}}\right)^2 + \left(\frac{1}{\sqrt{a^2+1}}\right)^2 = 1$, we introduce an angle γ, such that

$$\sin\gamma = \frac{a}{\sqrt{a^2+1}}, \tag{18}$$

$$\cos\gamma = \frac{1}{\sqrt{a^2+1}}, \tag{19}$$

$$\tan\gamma = \frac{\sin\gamma}{\cos\gamma} = a = \cos\varepsilon \cdot \tan\psi. \tag{20}$$

By substituting (18), (19), and $a = \cos\varepsilon \cdot \tan\psi$ to (17), we have

$$\begin{aligned}
&\cos\varphi(y(i,j)) \\
&= \cos\psi \cdot \sqrt{a^2+1} \cdot [\sin\gamma \cdot \sin\theta(y(i,j)) + \cos\gamma \cdot \cos\theta(y(i,j))] \\
&= \sqrt{\cos^2\varepsilon \cdot \sin^2\psi + \cos^2\psi} \cdot \left[\sin\gamma \cdot \sin\theta(y(i,j))\right. \\
&\quad \left. + \cos\gamma \cdot \cos\theta(y(i,j))\right] \\
&= \sqrt{\cos^2\varepsilon \cdot \sin^2\psi + \cos^2\psi} \cdot \cos(\theta(y(i,j)) - \gamma).
\end{aligned} \tag{21}$$

From (20), we derive that $\gamma = \arctan(\cos\varepsilon \cdot \tan\psi)$. Substituting γ to (21), we have the relation between $\varphi(y(i,j))$ and $\theta(y(i,j))$ as follows:

$$\begin{aligned}&\cos\varphi(y(i,j))\\&=\sqrt{\cos^2\varepsilon\cdot\sin^2\psi+\cos^2\psi}\cdot\cos(\theta(y(i,j))\\&-\arctan(\cos\varepsilon\cdot\tan\psi)).\end{aligned}\tag{22}$$

Note that, when the document skew $\varepsilon = 0$, i.e., N', L', and V' are coplanar, (22) is simplified to

$$\cos\varphi(y(i,j)) = \cos(\theta(y(i,j)) - \psi).$$

4 Reconstruction of the Book Surface Shape and Albedo Distribution

Remember that our problem is how to restore the document image by removing the shade (deshading) and correcting the warping (dewarping) in the shade. For deshading, we need the albedo distribution $k(x(i,j), y(i,j))$ and for dewarping we need the shape of the book surface shape $z(y(i,j))$. In this section, we will discuss how to solve $z(y(i,j))$ and $k(x(i,j), y(i,j))$ by adopting the two practical models.

4.1 Reconstruction of Book Surface Shape

Since most book surfaces have a uniformly colored background (typically, an unprinted white background), we can assume that pixel values $P_w(y_n)$ $(y_0 \le y_n \le y_N)$ corresponding to the background can be obtained as the maximum for each y value in the scanned document image, i.e., we scan the image pixels along the direction of the x-axis (in the x-y-z 3D coordinate system) instead of the i-axis (in the i-j 2D image index system), and set $P_w(y_n)$ as the maximum intensity value of each scan:

$$\begin{aligned}&P_w(y_n) = \max_{x_m} P(i(x_m, y_n), j(x_m, y_n))\\&(x_0 \le x_m \le x_M, y_0 \le y_n \le y_N),\end{aligned}\tag{23}$$

where x_0, x_M, y_0, and y_N are calculated from (11), (12), (13) and (14), respectively, and $(i(x_m, y_n), j(x_m, y_n))$ is the corresponding image index for (x_m, y_n), which are calculated from (3) and (4). We then extract the global maximum pixel value $P_w^{\max}$ over all y_n columns by:

$$P_w^{\max} = \max_{y_n} P_w(y_n).\tag{24}$$

By (9), the constant albedo k_w of the white background is calculated as follows:

$$k_w = \frac{P_w^{\max} - \beta}{\alpha\cdot I_s(0)\cdot\cos\psi}.\tag{25}$$

The optical model of the white background with the constant albedo k_w is represented as:

$$P_w(y_n) = \alpha\cdot k_w\cdot I_s(z(y_n))\cdot\cos\varphi(y_n) + \beta,\tag{26}$$

where $P_w(y_n)$ is from (23) and k_w from (25).

By substituting (10) into (25) and (26), we derive the following representation of $z(y_n)$ from the optical model:

$$z(y_n) = \left[\frac{(P_w^{\max} - \beta)\cdot\cos\varphi(y_n)}{(P_w(y_n) - \beta)\cdot\cos\psi} - 1\right]\cdot d.\tag{27}$$

Replacing $\cos\varphi(y_n)$ by (22), we have:

$$z(y_n) = \left[\rho\cdot\frac{\cos(\theta(y_n) - \arctan(\cos\varepsilon\cdot\tan\psi))}{P_w(y_n) - \beta} - 1\right]\cdot d,\tag{28}$$

where ρ is a constant represented as:

$$\rho = \frac{(P_w^{\max} - \beta)\cdot\sqrt{\cos^2\varepsilon\cdot\sin^2\psi + \cos^2\psi}}{\cos\psi}.\tag{29}$$

By substituting y_n into (5) (geometric model), we derive another representation of $z(y_n)$:

$$z(y_n) = \sum_{y_k = y_N}^{y_n+1}\tan\theta(y_k)\quad\left(y_0 \le y_n \le y_N, \theta\in\left[0, \frac{\pi}{2}\right]\right),\tag{30}$$

where y_0 and y_N are from (13) and (14), respectively.

Note that, we now have two representations of $z(y_n)$, i.e., (28) from the optical model and (30) from the geometric model, and there are only two unknowns $z(y_n)$ and $\theta(y_n)$. Furthermore, the initial values of $z(y_n)$ and $\theta(y_n)$ are known: $z(y_N) = 0$ and $\theta(y_N) = 0$. Thus, in theory, the solution of the depth map $z(y_n)$ can be found. We propose the following algorithm to compute the numerical solutions of $z(y_n)$ and $\theta(y_n)$:

Step 1. Rewrite (28) in the following form:

$$\begin{aligned}&\cos(\theta(y_n) - \arctan(\cos\varepsilon\cdot\tan\psi))\\&= \left[\frac{P_w(y_n) - \beta}{\rho}\right]\cdot\left[\frac{z(y_n) + d}{d}\right].\end{aligned}\tag{31}$$

Thus, if $\theta(y_n) < \arctan(\cos\varepsilon\cdot\tan\psi)$,

$$\theta(y_n) = \arctan(\cos\varepsilon\cdot\tan\psi) - \arccos\left[\frac{P_w(y_n) - \beta}{\rho}\cdot\frac{z(y_n) + d}{d}\right],\tag{32}$$

else

$$\theta(y_n) = \arctan(\cos\varepsilon\cdot\tan\psi) + \arccos\left[\frac{P_w(y_n) - \beta}{\rho}\cdot\frac{z(y_n) + d}{d}\right].\tag{33}$$

Step 2. Set $z(y_N) = 0$, $\theta(y_N) = 0$, and Exceed = 0 (since $\theta(y_n)$ is monotonously increasing starting with value 0 from y_N to y_0, this control flag "Exceed" indicates whether the θ value has exceeded $\arctan(\cos\varepsilon\cdot\tan\psi)$).

Step 3. For $(y_n = y_N - 1, y_n \ge y_0, y_n - -)$

- 3.1 Compute $z(y_n)$ by (30).
- 3.2 If $\theta(y_n + 1) \ge \arctan(\cos\varepsilon\cdot\tan\psi)$, set Exceed =1.
- 3.3 If Exceed =0, compute $\theta(y_n)$ by substituting $z(y_n)$ to (32), else compute $\theta(y_n)$ by substituting $z(y_n)$ to (33).

End of the Algorithm

4.2 Reconstruction of the Albedo Distribution

By (9), the albedo distribution is represented as:

$$k(x(i,j), y(i,j)) = \frac{P(i,j) - \beta}{\alpha\cdot I_s(z(y(i,j)))\cdot\cos\varphi(y(i,j))}.\tag{34}$$

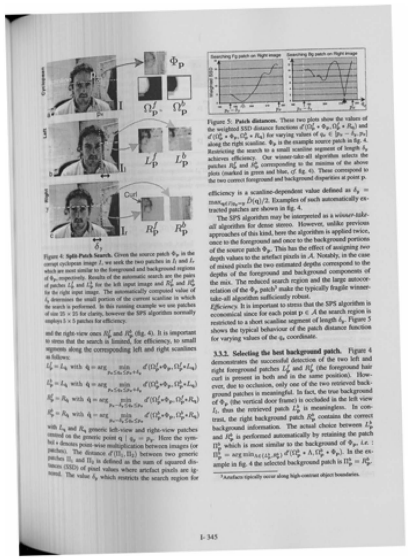

Fig. 8. Image generated by the deshading model for the processing area defined in Fig. 6.

Substituting (22) into (34), we have:

$$k(x(i,j),y(i,j)) = \frac{P(i,j)-\beta}{\alpha \cdot I_S(z(y(i,j))) \cdot \sqrt{\cos^2\varepsilon \cdot \sin^2\psi + \cos^2\psi} \cdot \cos(\theta(y(i,j)) - \arctan(\cos\varepsilon \cdot \tan\psi))}, \quad (35)$$

where $z(y(i,j))$ and $\theta(y(i,j))$ are looked up from the depth map $z(y_n)$ and $\theta(y_n)$, respectively, $I_s(z(y(i,j)))$ is calculated from (10), $P(i,j)$ is the pixel value in the observed image (scanned image), and α, β, ε, and ψ are known constants.

5 Restoration of Document Image

We propose two models, i.e., deshading and dewarping models, to restore the document image by using the book surface shape and albedo distribution obtained in the last section.

5.1 Deshading Model

If the physical document page is flat while scanning, all the z values are zero, i.e., $z(y(i,j)) = 0$, and the page surface normal is the vertical, i.e., $\varphi(y(i,j)) = \psi$. By (9), the optimal image intensity $P^*(i,j)$ for point $(x(i,j), y(i,j), 0)$ would be

$$P^*(i,j) = \alpha \cdot k(x(i,j),y(i,j)) \cdot I_s(0) \cdot \cos\psi + \beta, \quad (36)$$

where $k(x(i,j),y(i,j))$ is calculated by (35), $I_s(0)$ is calculated by (10), and α, β, and ψ are known constants. Therefore, our restoration system recalculates the image intensity for each pixel by (36). The deshading result is shown in Fig. 8.

5.2 Dewarping Model

Taking the image generated by the deshading model as input, we next build our dewarping model. The removal of distortion is done based on the converged shape $z(y_n)$ of the book surface. Note that distortion along the *y*-axis is only due to orthogonal projection, while perspective distortion along *x*-axis is due to perspective projection. Thus, the dewarping model consists of three parts: 1) restoration along the *x*-axis, 2) restoration along the *y*-axis, and 3) correction of document skew ε.

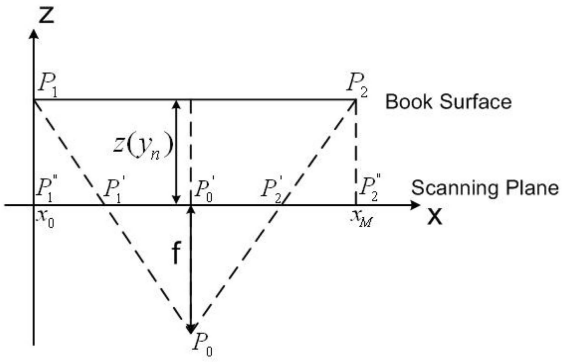

Fig. 9. Perspective projection on a slice of the x-z plane at y_n.

5.2.1 Restoration along x-axis

Fig. 9 shows a slice of the x-z plane at y_n ($y_0 \le y_n \le y_N$), where P_1P_2 denotes a line parallel to the *x*-axis on the book surface, $P_1'P_2'$ denotes the line on the scanning plane corresponding to P_1P_2 viewed by the CCD sensor (i.e., $P_1'P_2'$ is the project image of P_1P_2 in the image plane), $P_1P_1'' = z(y_n)$, $P_1''P_2'' = 2 \times P_1''P_0' = x_M - x_0$ (x_0 and x_M are constants from (11) and (12), respectively), and the focal length $f = P_0P_0'$.

We restore the *x*-axis distortion by regenerating the image intensity $P^{**}(i,j)$ according to (37) for each pixel on y_n (stretching $P_1'P_2'$ to the length of P_1P_2, from perspective equations, this stretching proves to be uniform for each y_n and proportional to $z(y_n)$),

$$P^{**}(i,j) = \begin{cases} P^*(i(x_r,y(i,j)),j(x_r,y(i,j))) & \text{if } x_r \text{ is an integer} \\ P^*(i(floor(x_r),y(i,j)),j(floor(x_r),y(i,j))) \cdot (ceil(x_r) - x_r) & \\ + P^*(i(ceil(x_r),y(i,j)),j(ceil(x_r),y(i,j))) \cdot (x_r - floor(x_r)) & otherwise, \end{cases} \quad (37)$$

where

- $P^*(i(x_r,y(i,j)),j(x_r,y(i,j)))$: Pixel value regenerated from (36), which is corresponding to the 3D point $(x_r, y(i,j), z(y(i,j)))$.
- x_r: The corresponding x value on P_1P_2 for $x(i,j)$ on $P_1'P_2'$. x_r is calculated as follows:

$$\begin{aligned} x_r &= x(i,j) \cdot \frac{P_1'P_2'}{P_1P_2} + P_1'P_1'' \\ &= x(i,j) \cdot \frac{f}{z(y(i,j)) + f} + \frac{z(y(i,j)) \cdot \frac{1}{2}(x_M - x_0)}{z(y(i,j)) + f} \\ &= \frac{2 \cdot x(i,j) \cdot f + z(y(i,j)) \cdot (x_M - x_0)}{2 \cdot (z(y(i,j)) + f)}, \end{aligned} \quad (38)$$

  where $x(i,j)$ and $y(i,j)$ are calculated according to (1) and (2) and f is a known constant.
- $floor(x_r)$: Returns the largest integer that is less than or equal to x_r.
- $ceil(x_r)$: Returns the smallest integer that is greater than or equal to x_r.

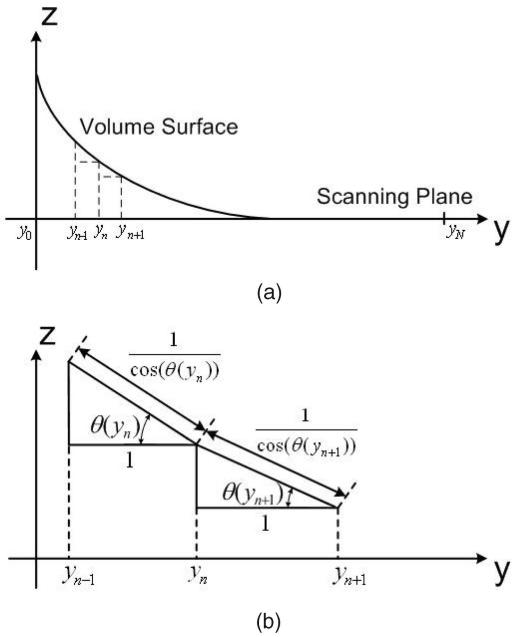

Fig. 10. Orthogonal projection on a slice of the y-z plane.

5.2.2 Restoration along y-axis

Distortion along the y-axis is only due to orthogonal projection as shown in Fig. 10a. Fig. 10b shows that the true width between two adjacent y value the y_{n-1} and y_n is equal to $\frac{1}{\cos\theta(y_n)}$. Thus, the correction along the y-axis is done by the following two steps: First, we add the estimated true width w calculated from (39) and, then, we stretch the observed image to its true width along the y-axis using similar method discussed in the last section.

$$w = \sum_{y_n=y_0}^{y_N} \left(\frac{1}{\cos(\theta(y_n))} - 1 \right). \tag{39}$$

The dewarping result is presented in Fig. 11.

5.2.3 Correction of Document Skew ε

As shown in Fig. 11, the document has a skew angle ε with the image boundary. We simply rotate the image at the center by ε degree, and the final restored image is shown in Fig. 12. The experiment shows that the geometric and photometric distortions are mostly removed, indicating that the shape is accurately estimated enough for image restoration.

6 Experimental Results

To test the effectiveness of our method, we conducted a series of experiments on about 20 pages of images from a thick bound book scanned using an EPSON Expression 1640XL with various resolutions and contrasts. Based on the results, we show how well the warped surface shape is reconstructed, how well the shading is recovered, how efficient our method is comparing to other method, and also how much our method improves the OCR performance. Finally, we also discuss some of the breakdown points of our method in terms of image contrast, surface shape, and material, etc.

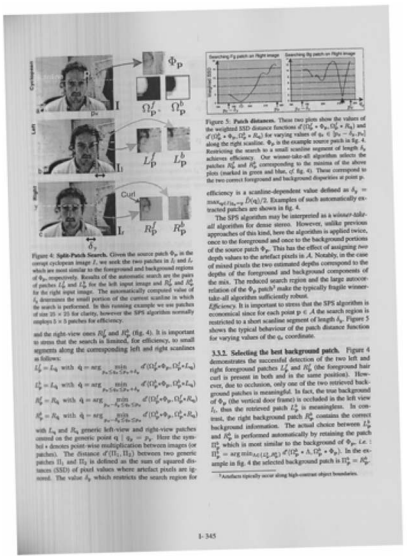

Fig. 11. Image generated by the dewarping model for the processing area defined in Fig. 6.

6.1 Reconstructed Shape versus Real Shape

For our cylindrical model, in order to ascertain that the shape has been correctly discovered, we need only to compare the cross-sectional shape. We try to capture the real cross-section shape of the warped surface using a small handheld scanner from the side of the thick bound book. The cross-section curve is then extracted from this scanned image and compared with the cross-section curve we obtained using our shape reconstruction method. Fig. 13 shows the comparison among the real cross-section shape, initial estimated shape (when z=0), and the final shape with an image of size $1,250 \times 1,970$. As we can see, the final shape is very close to the real cross-section shape.

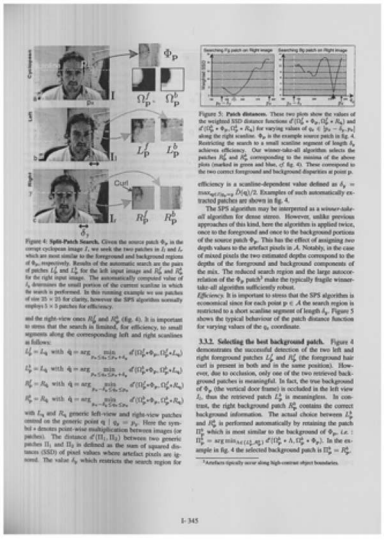

Fig. 12. Final restored document image for Fig. 1.

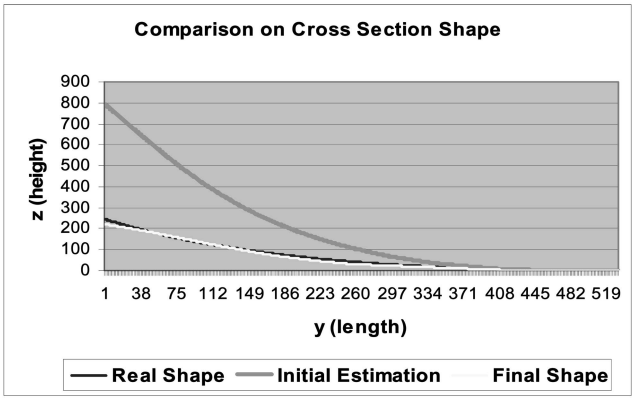

Fig. 13. Comparison on the cross-section shape.

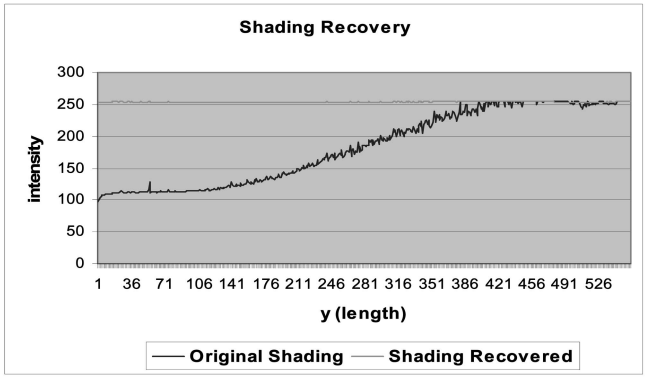

Fig. 14. Shading recovery.

6.2 Shading Recovery

We also illustrate the shading recovery on the same image above in terms of the maximum intensity along each column of the image. As we can see from Fig. 14, the shading is well restored near the spine region.

6.3 OCR Results Comparison

As a measure of success, we evaluate the restoration results by carrying out OCR testing on both original and restored images obtained using our method. We perform the OCR test using three well-known commercial OCR products, namely, Scansoft OmniPage Pro 14.0, IRIS Readiris Pro 9.0, and ABBYY FineReader Pro 7.0. Precision and recall [22] on characters and words defined below are used as the evaluation metrics for the OCR results:

$$\text{Character Precision} = \frac{\text{Number of characters correctly detected by OCR}}{\text{Total number of characters detected by OCR}},$$

$$\text{Character recall} = \frac{\text{Number of characters correctly detected by OCR}}{\text{Total number of characters in the document}},$$

$$\text{Word Precision} = \frac{\text{Number of words correctly detected by OCR}}{\text{Total number of words detected by OCR}},$$

$$\text{Word Recall} = \frac{\text{Number of words correctly detected by OCR}}{\text{Total number of words in the document}}.$$

We scanned 100 document pages with resolution 200, 300, and 400 ppi, respectively. Tables 1, 2, 3, and 4 show the average character and word precision/recall, respectively, for the original restored images. In Tables 5 and 6, we summarize the improvements on character and word precision/recall by our method. Our method significantly improves the character precision by 10.55 percent to 13.82 percent, character recall by 7.47 percent to 17.43 percent, word precision by 11.48 percent to 15.28 percent, and word recall by 9.93 percent to 17.94 percent for the three tested OCR products. These improvements demonstrate the success of our restoration method. Moreover, our method is very efficient while achieving good restoration results.

6.4 Comparison with Wada's Method

We chose Wada's method for comparison as it is closely related to ours in restoring scanner images. Wada's method assumes parallel alignment of the book spine with the scanning light. It also assumes the presence of interreflection, which adds on computational cost. As a comparison, we

TABLE 1
Average Character Precision and Recall for the Original Scanner Document Images

	Average Character Precision (%)			Average Character Recall (%)		
Resolution	OmniPage	Readiris	FineReader	OmniPage	Readiris	FineReader
200 ppi	83.96	80.34	82.93	90.49	79.13	79.01
300 ppi	84.14	83.31	84.13	90.94	80.92	80.16
400 ppi	84.96	85.14	85.01	91.43	81.22	80.92

TABLE 2
Average Word Precision and Recall for the Original Scanner Document Images

	Average Word Precision (%)			Average Word Recall (%)		
Resolution	OmniPage	Readiris	FineReader	OmniPage	Readiris	FineReader
200 ppi	81.87	78.39	80.48	87.11	76.13	77.64
300 ppi	82.54	82.15	81.11	87.93	78.74	78.60
400 ppi	82.59	84.19	82.60	88.01	78.89	79.29

TABLE 3
Average Character Precision and Recall for the Restored Images

	Average Character Precision (%)			Average Character Recall (%)		
Resolution	OmniPage	Readiris	FineReader	OmniPage	Readiris	FineReader
200 ppi	96.80	93.45	96.75	98.53	94.10	96.44
300 ppi	97.54	94.74	97.71	98.73	94.87	96.98
400 ppi	97.91	95.69	98.46	98.90	95.30	97.27

TABLE 4
Average Word Precision and Recall for the Restored Images

	Average Word Precision (%)			Average Word Recall (%)		
Resolution	OmniPage	Readiris	FineReader	OmniPage	Readiris	FineReader
200 ppi	96.02	93.27	95.75	97.54	93.22	95.43
300 ppi	96.41	94.63	96.39	97.89	94.10	96.54
400 ppi	96.43	95.67	96.76	97.94	94.85	96.61

TABLE 5
Improvement on Average Character Precision and Recall by Our Method

	Improvement on Average Character Precision (%)			Improvement on Average Character Recall (%)		
Resolution	OmniPage	Readiris	FineReader	OmniPage	Readiris	FineReader
200 ppi	12.84	13.11	13.82	8.04	14.79	17.43
300 ppi	13.40	11.43	13.58	7.79	13.95	16.82
400 ppi	12.95	10.55	13.45	7.47	14.08	16.35

TABLE 6
Improvement on Average Word Precision and Recall by Our Method

	Improvement on Average Word Precision (%)			Improvement on Average Word Recall (%)		
Resolution	OmniPage	Readiris	FineReader	OmniPage	Readiris	FineReader
200 ppi	14.15	14.88	15.27	10.43	17.09	17.78
300 ppi	13.87	12.48	15.28	9.96	15.36	17.94
400 ppi	13.84	11.48	14.16	9.93	15.96	17.32

TABLE 7
Time and OCR Results Comparison with Wada's Method

	Skewed image (200ppi)			Specular image (200ppi)		
	OCR precision	OCR recall	Time(s)	OCR precision	OCR recall	Time(s)
Wada's	72.36%	78.22%	214.52	80.15%	87.53%	246.34
Ours	83.96%	90.49%	2.12	81.26%	90.15%	2.81

reimplemented Wada's algorithm to include the above assumptions. We tested both Wada's algorithm and our method on several images involving skewed scanning and specular document page by comparing the time and OCR results. In Table 7, we show results of two images, one skewed and the other specular. We can see a higher computational cost of Wada's algorithm, while our method handles skewed scanning and specular pages very well. An example of restoration results on a skewed image is shown in Fig. 15.

6.5 Discussion on Breakdown Points

During our experimentation, we observe that the shape reconstruction process relies on the contrast of the input image particularly near the book spine. This is because the shape from the shading problem recovers the shape based on the shading information. Therefore, when the contrast is too high or too low, the shading information near the book spine will be indifferent regardless of the depth level. Fig. 16 shows an example of reconstructed shape with various image contrasts. We can see that the reconstructed shape near the spine region has a large deviation from the real shape for a low contrast image, although the further part merges well. Therefore, to achieve a good restoration result, a proper contrast needs to be adjusted.

Our Assumption 1 imposes a generalized cylindrical surface, which means that our method may fail when the

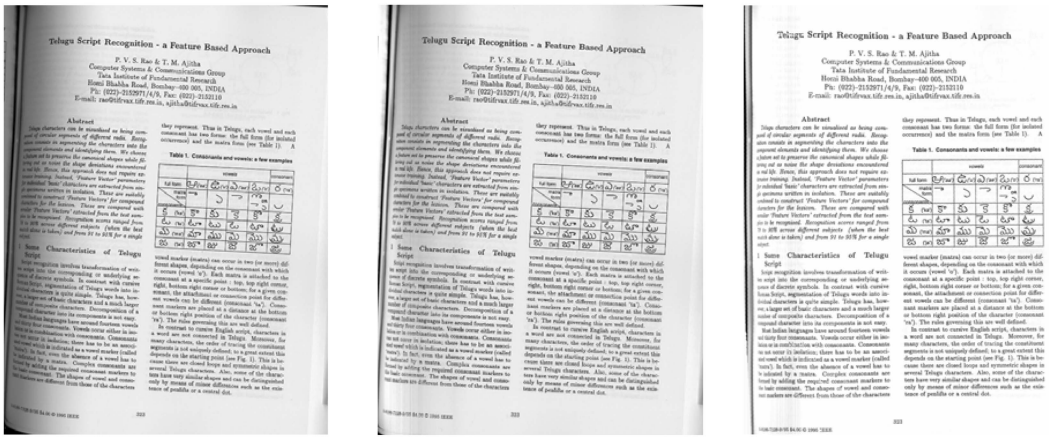

Fig. 15. Restored image comparison with Wada's method. (a) Original Image. (b) Wada's Method. (c) Our Method.

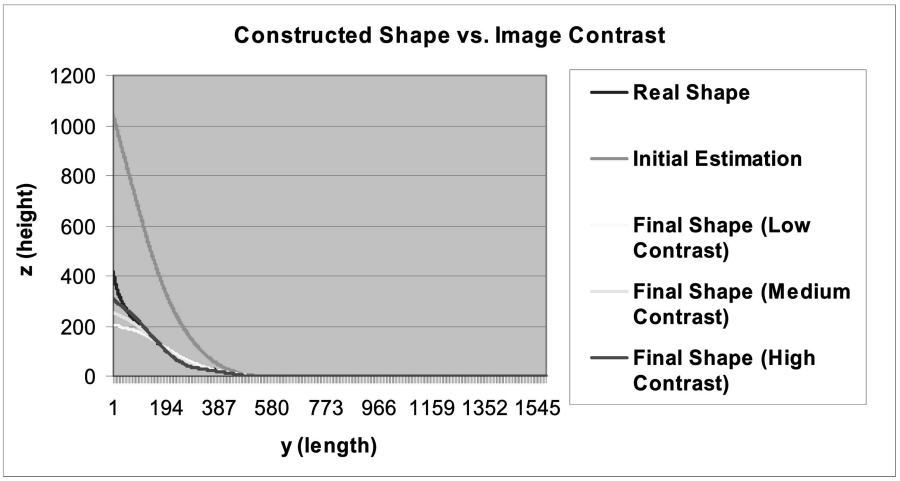

Fig. 16. Constructed shape versus image contrast.

height changes along the spine. However, our method can be adapted to include other smooth surfaces with nonuniform height, such as conical surfaces. What we need to do is just to calculate the height $z(x, y)$ at each different point (x, y) instead of considering a uniform cross-section shape along the *x-axis*.

We also assume in our Assumption 2 that the warped surfaces are Lambertian, as is the case for most of the book surfaces. Nevertheless, during our experiments, we also tried several images of non-Lambertian surfaces, such as those pages in magazines. The results show that after converting the images to grayscale and applying our method to it, the restoration result is still satisfactory, as shown in Table 7 about the OCR results on specular images. On the other hand, there are also techniques available that can separate specular component swith Lambertian components, such as one proposed by Ragheb and Hancock [23].

7 Conclusion and Future Works

In this paper, we described the image distortion problems while scanning thick, bound documents, and discussed the real-world shape-from-shading techniques for recovering the 3D shape of a book surface from its scanner image. We presented a restoration system that solves these problems using the shape obtained. This system can successfully remove the shade, adjust the warped surface, and deskew the images. The experiments showed well reconstructed surface shape and a marked improvement on the OCR test on the restored document image.

Our method is a new attempt that has proven superior to other methods based on the 3D document shape discovery. In comparison to the existing methods, our method requires neither textual contents nor document boundaries in the document image, nor special camera setups. Our method neither assumes the document to be scanned aligned strictly parallel to the scanning direction, nor does it assume the document surface to be strictly circular cylindrical. Moreover, it is much more efficient than Wada's method with satisfactory restoration results.

Our future work will look into the following areas:

- Extend the system to handle color document images, instead of only gray-scale ones.

- Though most papers are Lambertian, some documents may contain plates, which have specular reflection property. Specular models may be further studied to find the best one fitting this situation.

ACKNOWLEDGMENTS

This research is supported in part by NUS URC grant no. R252-000-202-112 and Agency for Science, Technology, and Research, Singapore, grant no. 042 101 0085.

REFERENCES

[1] H. Baird, "Document Image Defect Models," *Proc. Workshop Syntactic and Structural Pattern Recognition,* pp. 38-46, June 1990.
[2] H. Baird, "Document Image Defect Models and Their Uses," *Proc. Int'l Conf. Document Analysis and Recognition,* pp. 730-734, Oct. 1993.
[3] H. Baird, "Document Image Quality: Making Fine Discriminations," *Proc. Int'l Conf. Document Analysis and Recognition,* pp. 459-462, Sept. 1999.
[4] Y.Y. Tang and C.Y. Suen, "Image Transformation Approach to Nonlinear Shape Restoration," *IEEE Trans. Systems, Man, and Cybernetics,* vol. 23, no. 1, pp. 155-171, Jan./Feb. 1993.
[5] O. Lavaille, X. Molines, F. Angella, and P. Baylou, "Active Contours Network to Straighten Distorted Text Lines," *Proc. Int'l Conf. Image Processing,* pp. 1074-1077, Oct. 2001.
[6] Y. Weng and Q. Zhu, "Nonlinear Shape Restoration for Document Images," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* pp. 568-573, 1996.
[7] Y.C. Tsoi and M.S. Brown, "Geometric and Shading Correction for Images of Printed Materials—A Unified Approach Using Boundary," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 240-246, 2004.
[8] Z. Zhang and C.L. Tan, "Restoration of Document Images Scanned from Thick Bound Document," *Proc. Int'l Conf. Image Processing,* pp. 1074-1077, Oct. 2001.
[9] Z. Zhang and C.L. Tan, "Recovery of Distorted Document Image from Bound Volumes," *Proc. Int'l Conf. Document Analysis and Recognition,* pp. 429-433, 2001.
[10] Z. Zhang and C.L. Tan, "Straightening Warped Text Lines Using Polynomial Regression," *Proc. Int'l Conf. Image Processing,* pp. 977-980, Sept. 2002.
[11] Z. Zhang and C.L. Tan, "Correcting Document Image Warping Based on Regression of Curved Text Lines," *Proc. Int'l Conf. Document Analysis and Recognition,* pp. 589-593, Aug. 2003.
[12] M. Pilu, "Undoing Page Curl Distortion Using Applicable Surfaces," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 67-72, Dec. 2001.
[13] M.S. Brown and W.B. Seales, "Image Restoration of Arbitrarily Warped Documents," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, no. 10, pp. 1295-1306, 2004.
[14] A. Doncescu, A. Bouju, and V. Quillet, "Former Books Digital Processing: Image Warping," *Proc. Int'l Workshop Document Image Analysis,* pp. 5-9, 1997.
[15] A. Yamashita, A. Kawarago, T. Kaneko, and K.T. Miura, "Shape Reconstruction and Image Restoration for Non-Flat Surface of Document with a Stereo Vision System," *Proc. Int'l Conf. Pattern Recognitiontion,* 2004.
[16] H. Cao, X. Ding, and C. Liu, "A Cylindrical Model to Rectify the Bound Document Image," *Proc. Int'l Conf. Computer Vision,* vol. 2, pp. 228-233, Oct. 2003.
[17] H. Cao, X. Ding, and C. Liu, "Rectifying the Bound Document Image Captured by the Camera: A Model Based Approach," *Proc. Int'l Conf. Computer Vision,* vol. 1, pp. 71-74, Oct. 2003.
[18] T. Kanungo, R.M. Haralick, and I. Phillips, "Nonlinear Global and Local Document Degradation Models," *Int'l J. Imaging System and Technology,* pp. 220-230, Oct. 1994.
[19] T. Wada, H. Ukida, and T. Matsuyama, "Shape from Shading with Interreflections under a Proximal Light Source: Distortion-Free Copying of an Unfolded Book," *Int'l J. Computer Vision,* vol. 24, no. 2, pp. 125-135, 1997.
[20] Z. Zhang, C.L. Tan, and L.Y. Fan, "Estimation of 3D Shape of Warped Document Surface for Image Restoration," *Proc. Int'l Conf. Pattern Recognition,* 2004.
[21] Z. Zhang, C.L. Tan, and L.Y. Fan, "Restoration of Curved Document Images through 3D Shape Modeling," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 10-16, 2004.
[22] M. Junker, R. Hoch, and A. Dengle, "On the Evaluation of Document Analysis Components by Recall, Precision and Accuracy," *Proc. Int'l Conf. Document Analysis and Recognition,* pp. 713-716, 1999.
[23] H. Ragheb and E.R. Hancock, "Separating Lambertian and Specular Reflectance Components Using Iterated Conditional Modes," *Proc. British Machine Vision Conf.,* pp. 541-552, Sept. 2001.

Chew Lim Tan received the BSc (Hons) degree in physics in 1971 from the University of Singapore, the MSc degree in radiation studies in 1973 from the University of Surrey, U.K., and the PhD degree in computer science in 1986 from the University of Virginia. He is an associate professor in the Department of Computer Science, School of Computing, National University of Singapore. His research interests include document image and text processing, neural networks, and genetic programming. He has published more than 230 research publications in these areas. He is an associate editor of *Pattern Recognition* and has served on the program committees of several coneferences. He is the current president of the Pattern Recognition and Machine Intelligence Association (PREMIA) in Singapore. He is a senior member of the IEEE and the IEEE Computer Society.

Li Zhang received the BSc degree (Hons) and the MSc degree from the School of Computing at the National University of Singapore in 2003 and 2004, respectively. She is currently a PhD candidate at the National University of Singapore and also serves as a research assistant in the Center for Information Mining and Extraction (CHIME), School of Computing, National University of Singapore. She is a member of the Pattern Recognition and Machine Intelligence Association (PREMIA) in Singapore. Her research interests are in computer vision, pattern recognition, document image analysis and processing, and information extraction and retrieval.

Zheng Zhang received the BSc (Hons) and PhD degrees in computer science from the National University of Singapore University in 2000 and 2005, respectively. He is a senior consultant at Frontline Ltd. in Singapore. His research mainly focuses on document image restoration, analysis, understanding, and recognition.

Tao Xia received the BS, MS, and PhD degrees from Peking University, Beijing, China, in 1991, 1994, and 1997, respectively, all in applied mathematics. He is a research fellow in the Institute of Engineering Science at the National University of Singapore from 1999. His main research interests are in the areas of wavelet/frame theory and applications, image and video processing and compression, and pattern recognition and computer vision.