Geometric and Photometric Restoration of Distorted Documents

Mingxuan Sun[1], Ruigang Yang[1], Lin Yun[1], George Landon[1], Brent Seales[1], Michael S. Brown[2]

[1] Department of Computer Science; University of Kentucky; Lexington, KY 40506
[2] School of Computer Engineering; Nanyang Technological University; Singapore 639798

Abstract

We present a system to restore the 2D content printed on distorted documents. Our system works by acquiring a 3D scan of the document's surface together with a high-resolution image. Using the 3D surface information and the 2D image, we can ameliorate unwanted surface distortion and effects from non-uniform illumination. Our system can process *arbitrary* geometric distortions, not requiring any pre-assumed parametric models for the document's geometry. The illumination correction uses the 3D shape to distinguish content edges from illumination edges to recover the 2D content's reflectance image while making no assumptions about light sources and their positions. Results are shown for real objects, demonstrating a complete framework capable of restoring geometric and photometric artifacts on distorted documents.

1. Introduction

This paper focuses on the recovery of 2D content printed on documents and similar items whose shapes are distorted (i.e. not flat). Our technique aims to restore the 2D content of these items such that the geometry and illumination are rectified. This work has significant relevance to applications in digital libraries and museums, where printed items are routinely imaged for purposes of archival and scholarly study.

The challenge addressed in this paper is straight-forward: while printed content is 2D, the material itself after printing does not necessarily remain planar. This gives way to two types of distortion that must be corrected: (1) geometric distortion and (2) shading distortion. Printed material's may become distorted for several reasons, including folding and bending, damage and decay, or in many cases, by design. Figure 1 shows an example of a distorted document targeted by our application. The original content printed on the document is distorted by the document's geometry and by shading artifacts from non-uniform illumination attributed to the non-planar geometry. The goal of our system is to recover the 2D content void of geometric and shading artifacts.

Geometric distortion correction has received a great deal of attention. Techniques ranging from simple 2D "de-skewing" to general distortion correction have been devel-

Figure 1: The 2D content printed on the document suffers from geometric and shading distortion. The second image shows 2D content restored.

oped [3, 7, 27]. On the other hand, the problem of correcting shading in the context of document restoration, beyond the purpose of simple binarization, has received significantly less attention. Nevertheless, these are companion problems: shape distortion induces shading irregularities that are difficult to remove at image acquisition. Some of the few examples of work addressing both shape and shading restoration [24, 27] assume restrictive geometric models, which limits their utility in general cases.

Our Contribution In this paper, we present a system that rectifies both the geometry and illumination of 2D content printed on *arbitrarily* distorted documents. Our framework acquires a 3D scan of the document's surface together with a high-resolution image. The 3D shape is used to compute a mapping between the image of the 2D content distorted by the document's surface and the desired non-distorted 2D content on a plane. This 3D data is also used to classify intensity changes (image gradient) into two classes–illumination change and reflectance change–based on the depth variation. Illumination changes, which are caused by the distorted surface normal (e.g., variation in 3D shape) are removed from the image gradient field. Using the modified image gradient field, we can re-integrate to a corrected image, free of shading artifacts.

The approaches of our system are *cooperative* in nature,

where the 3D shape and input image are used to correct for both geometry and shading, and the results from these procedures are combined to generate the final restored 2D document void of geometric and shading distortion. Compared to simulation-based techniques that also utilize 3D shape, we do not need to explicitly compute the number of light sources in the scene, nor their positional information. This makes our approach much more practical. From a system standpoint, our contribution is a complete pipeline for geometric and photometric restoration of arbitrarily distorted documents.

The remainder of this paper is organized as follows: section 2 presents related work; section 3 discusses the overall system, including our techniques for geometric and photometric restoration; section 4 shows results on several real examples; sections 5 and 6 provide a closing discussion and conclusion.

2. Related Work

Techniques to correct geometric distortion in documents has been well studied, especially those targeting common geometric distortion that can be robustly modelled. For example, *deskewing* techniques correct for off-axis distortion by aligning the document's content (lines of text, for example) with the $x - y$ axis of the image (e.g., see [3, 15, 18]). Another class of algorithms corrects the very specific *binder-curl* effect found near a book's spine. Techniques such as [9, 10, 24, 27] exploit the constrained shape of bound materials: parameterized geometric models serve as a tool to fit the surface shape and undo the distortion. A number of researchers have developed methods for correcting distortion from *arbitrarily* distorted materials, where very limited assumptions about the document's geometry are made [7, 8, 22]. In these approaches, a 3D reconstruction of the material's surface shape is acquired. The 3D model is then mapped to a plane using various numerical techniques ranging from iterative-relaxation to conformal mapping. In a similar system, 3D shape information from shading cues [16] is used to guide a restoration process for printed material.

Correcting shading distortion in documents, outside simple image thresholding used on black and white text-based documents, has received significantly less attention. Outside the context of document processing, however, one finds a rich set of approaches to shading correction. The notion of intrinsic images [4], for example, introduces the idea of decomposing an image into two intrinsic images: a reflectance image and an illumination image. With this intrinsic decomposition, the reflectance image is decoupled from the scene's illumination. Recovering these two intrinsic images from a single image, however, is an ill-posed problem as the number of unknowns is twice the number of equations [25]. Nevertheless, progress has been made toward achieving this decomposition under certain conditions. Many techniques

(e.g., [6, 14, 17, 21]) are based on the *Retinex* theory [19], which assumes that the gradients along reflectance changes have much larger magnitudes than those caused by illumination. These techniques work well for slow-varying illumination changes, but cannot handle sharp edges (as seen in Figure 1). Recent algorithms have used machine learning techniques to separate illumination and reflectance [5, 23]. While they can deal with more complex situations, the result depends largely on the training data-set.

The above approaches typically do not require 3D information, but make some strong assumptions about the scene. In a different direction, computer graphics researchers use 3D information to recover arbitrary textures (reflectance) [11, 13, 20, 26]. These approaches use physically-based light transport simulation to remove the effect of existing lighting conditions. *A priori* information of the lighting conditions are typically required for good results.

For our approach, our shading correction exploits the presence of the 3D geometry, necessary for geometric correction. The 3D geometry is used to classify edges as illumination edges or reflectance edges. This will be used to remove the illumination edges to help rectify the overall illumination of the 2D content.

3. System Overview

A diagram of our system is shown in Figure 2. A 3D reconstruction and high-resolution image of the document are acquired via a structured light-system. This data is used to correct geometric and shading distortion separately. The two results are combined, using the mapping for the geometric correction with the adjusted image to produce the final output. Sections 3.1, 3.2, and 3.3 detail our acquisition setup and the algorithms used to correct the geometry and shading distortion.
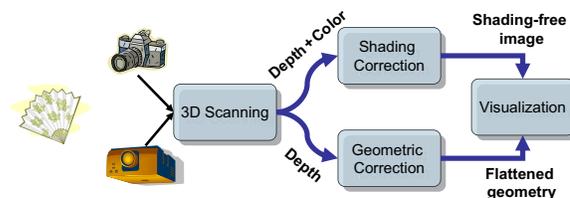

Figure 2: A system diagram of the overall document restoration process.

3.1. Acquisition System

Our 3D scanning system is shown in Figure 3. The setup uses two digital video cameras, a high-resolution still digital camera, and a laser. The laser and two video cameras acquire the 3D geometry using standard structure-lighting techniques(e.g. [1]), while the high-resolution still camera

captures the 2D image. The laser is mounted on a PC controlled pan/tilt unit to allow the light to be swept across the document. The cameras are calibrated with a calibration pattern, while the laser remains uncalibrated. The stripe is detected by the two video cameras with sub-pixel accuracy by fitting a Gaussian curve to the laser's intensity pattern. These detected points in the stereo cameras are triangulated to obtain a dense 3D point cloud on the document's surface. The 3D points are then projected into the high-resolution texture camera. The result is a depth image in the form:

$$I(u, v) = (x, y, z, r, g, b)$$

, where (u, v) are the coordinates for the texture camera's images, and x, y, z are the pixel's corresponding recovered 3D value, and r, g, b are the corresponding colors.

The laser stripe generator, with 5mW power output at 635nm wavelength, produces a 1mm wide stripe on the document surface. Two video cameras acquire images at a resolution of $640 \times 480 pixels$ and the digital camera acquires a high-resolution image of 3072×2048. Our system can acquire a 3D surface in approximately three minutes and typically produces around three hundred thousand 3D points. The 3D points have an average spacing of $0.6mm$ and an average depth accuracy of $0.5mm$ over a typical scan of $200 \times 280mm$ surface area.
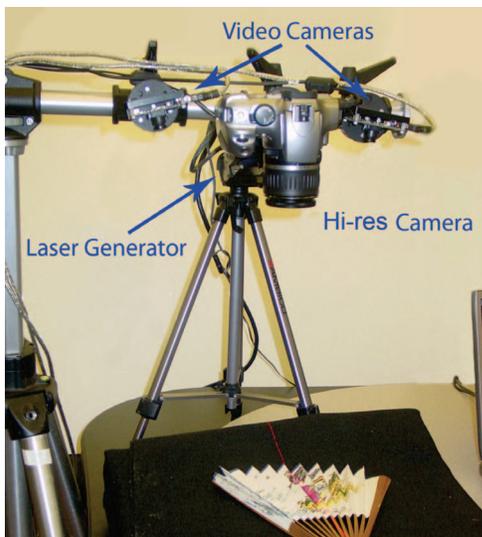

Figure 3: Our 3D scanner setup used to acquire the input data for our restoration pipeline.

3.2. Geometric Correction

We use a technique similar to [7] and [22] for geometric correction. These approaches push 3D surface points to a plane while maintaining distances between reconstructed 3D points. However, while [7] uses a finite-element structure of shear, structural, and bend springs, we find that can use the topology of a triangulated surface mesh directly with comparable results.

Our algorithm begins by constructing a triangulated mesh from reconstructed points on the material's surface. We resample the point cloud by sampling the depth map $I(u, v)$. For an A4 size(210 x 297 mm) paper, we get a mesh model with around 30,000 triangles.

Let $E = \{e_{i_1, j_1}, \ldots, e_{i_n, j_n}\}$ be the set of n edges in the triangled mesh. The edges are composed of two 3D points, denoted as $\mathbf{x}_i$ and $\mathbf{x}_j$, where $\mathbf{x} = (x, y, z)$. The Euclidean distance, d_{ij}, between these points is defined as $d_{ij} = ||\mathbf{x}_i - \mathbf{x}_j||_2$. Let the distance $\hat{d}_{ij}$ be the starting distance computed for all edges e_{ij} on the material original reconstructed 3D surface.

It is assumed that if the document was planar then all 3D points $\mathbf{x_i}$ would lie on the $z = 0$ plane. We find an optimal mapping from the 3D surface to a planar one (i.e. $z = 0$) subject to the following constraint:

$$min \sum_{e_{ij} \in E} ||d_{ij} - \hat{d}_{ij}||_2. \tag{1}$$

The above equation allows the $\mathbf{x}_i$ to move on the $z = 0$ plane while attempting to maintain overall edge length. Using an iterative relaxation method to vary the points on the $x - y$ plane subject to Eq. 1 we can obtain a mapping to the $x - y$ plane of the 3D mesh where edge distortion is minimized. Using this 3D to 2D mapping, 2D pixels in the original image, which are projections of the starting 3D points, are warped to their corresponding 2D location in the $x - y$ plane to undistort the image. Figure 4 shows an example of the flattening process. The input image, 3D structure, 3D structure pushed a plane, and resulting output image are shown. The corresponding triangulated mesh is also shown in the bottom row. Note that shading artifacts have not been corrected in this example. This can the illusion that the recovered image is still distorted.

3.3. Shading Correction

We use the intrinsic image model to correct for shading, which models an image f as the product of the scene's reflectance r and the illumination i, i.e.,

$$f = i \cdot r. \tag{2}$$

With this model, if we can obtain r, we will have the reflectance of the document which can then be uniformly illuminated. As mentioned in Section 2, the intrinsic decomposition is ill-posed given a single input image. The problem we try to address, is to correct shading from a single image with known geometry but unknown arbitrary reflectance and illumination. It is, however, still an ill-posed problem. Even with known geometry, the lighting effect and surface reflectance are still interchangeable. Therefore some further assumptions have to be made.
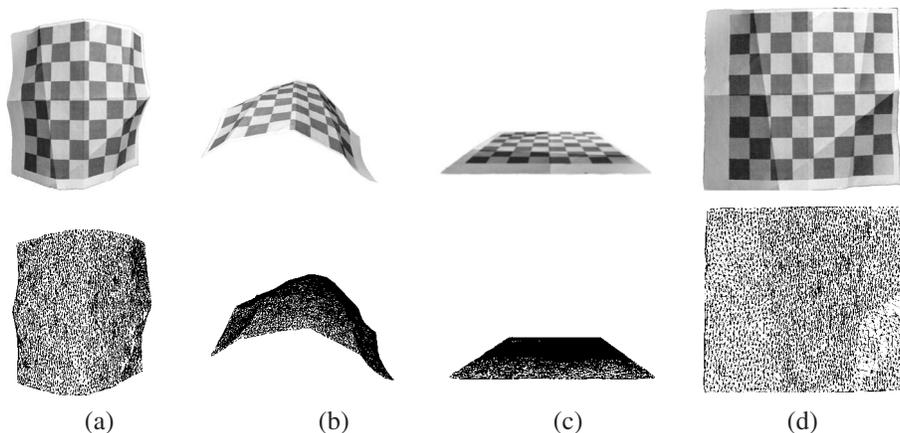
COMPUTER SOCIETY

Figure 4: Geometric correction: [Top row]: (a) input image; (b) view of 3D surface (c) 3D surface flattened to a plane;(d) output image with geometry corrected. [Bottom row]: Corresponding triangulated mesh for each image in the top row.

We assume that any noticeable visual discontinuity, i.e., an edge, in the image is caused by either the reflectance or the illumination. Under this assumption, if we can identify and remove the intensity change caused by the *illumination edges*, we can restore the original reflectance image. Thus, the problem of shading correction is reduced to a classification problem of illumination edges and reflectance (content) edges. Previous approaches have used machine learning techniques [5, 23] to perform this classification from a single image, but this problem is inherently ambiguous from only photometric information.

Since we know the document's geometry, we can identify illumination edges based on depth discontinuities on the document's surface. Based on this basic idea, we have developed a shading correction algorithm.

Before presenting the details, we first define the notion of illumination edges and reflectance (content) edge.

Illumination Edge An illumination edge is the union of connected *illumination pixels* in the input image. An illumination pixel corresponds to a 3D surface point that is at most C^0 continuous with respect to its surrounding points.

Reflectance Edge A reflectance edge is the union of connected *reflectance pixels* in the input image. A reflectance pixel corresponds to a 3D surface point that is at most C^0 continuous in reflectance (or color).

Note that both types of edges create the visually discontinuously edges in the input image.

3.3.1 Illumination edge detection

Given an image and its corresponding per-pixel depth map, we would like to detect illumination edges. By its definition, an illumination edge corresponds to 3D points that are at most C^0 continuous in geometry. We can find C^1 discontinuity by examining the *gradient image* of the depth map, $I(u, v)$, introduced in section 3.1. The edges detected in the gradient image, denoted as $\{E_i\}$, correspond to illumination edges. This simple and robust 2D image processing technique avoids processing the depth data in 3-space.

In practice, we have to deal with object boundaries, i.e. the edges of the document. We simply define a maximum depth gradient threshold to remove edges caused by object discontinuity.

3.3.2 Image restoration

Our image restoration procedure starts with the gradient field (ΔI) of the input image. After illumination edge detection, we set the gradient of the illumination pixel (i.e., $p \in \{E_i\}$) to zero to remove the illumination change at p in ΔI. We now wish to integrate the gradient field in order to reconstruct an image without illumination artifacts. This can be achieved by solving the Poisson equation [2] which is able to reconstruct a scalar function from a guidance vector field and a given boundary condition.

Let Ω stand for the image to be restored and w the guidance vector field, Poisson equation with Dirichlet boundary conditions is formulated as:

$$\nabla^2 f = \nabla \cdot w, f \mid_{\partial\Omega} = f^* \mid_{\partial\Omega}, \qquad (3)$$

where f is the unknown scalar function, f^* provides the desirable values on the boundary $\partial\Omega$, $\nabla \cdot w = \frac{\partial w_x}{\partial x} + \frac{\partial w_y}{\partial y}$ is the divergence of $w = (w_x, w_y)$, $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator.

Our expected solution is therefore the unique solution of the discrete form of the Poisson equation (3), where f is discretized naturally using the underlying discrete pixel grid of

IEEE
COMPUTER
SOCIETY

the image, Ω, and the vector guidance field w is set to be:

$$w(p) = \begin{cases} 0 & \text{if pixel } p(x,y) \in \{E_i\}, \\ \nabla I(p) & \text{otherwise.} \end{cases} \quad (4)$$

In this form, the discrete Poisson equation is a sparse linear system that can be solved numerically using techniques such as Gauss-Seidel iteration with successive over-relaxation.

To deal with color input, we independently solve the Poisson equation in different color channels. To minimize global color drift due to the scale ambiguity in Poisson equation, we have used the YUV color space.

3.3.3 Reflectance edge removal

The techniques presented above makes an assumption that the illumination edges and reflectance edges do not overlap. While this assumption is made in other shading correction algorithms (e.g., [19, 14]), it is rarely true in practice. When two types of edges intersect or overlap, the restored images usually exhibits some color-bleeding, i.e., the reflectance (content) edges are blurry near shading boundaries; an example is shown in Figure 5 top row.

We have developed a technique to ameliorate the bleeding problem. For every identified illumination pixel, we compute its gradient direction in both the depth image and the input image. If the two directions differ by a large amount, it means that it is likely to be an intersection of illumination change and reflectance change. In this case, we remove this pixel from the illumination pixel set by restoring its gradient value in the vector guidance field (Eq. 4). Figure 5 demonstrates its effectiveness with a synthetic checkerboard pattern.
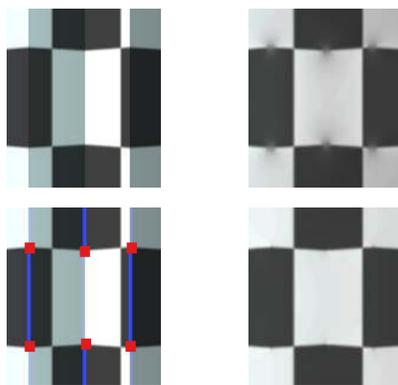

Figure 5: Effects with and without reflectance edge removal (RER). (Top-left) Input image; (top-right) restored image without RER; (bottom-left) Illumination mask with reflectance overlap removed (the overlap is expanded for illustration purpose); (bottom-right) image restored with RER.

It should be noted that if there are long overlaps of both types of edges the above approach will generate unsatisfactory results. In practice we found that it works best when the overlap is no longer than a few pixels, and the threshold is set to 30-40 degrees.

4. Results

Figure 6 shows an example of distorted document that has been processed by our system. The 2D content is of a checkerboard pattern. Figure reffig:all(a)-top shows the input image. The corresponding bottom image shows the 3D shape acquired by our system. Figure reffig:all(b)-top shows the resulting output image after geometric correction described in Section 3.2. The bottom image shows the 3D shape flattened by to the plane. Note that the 2D content may still appear distorted due to the non-uniform illumination still present from the input image. Figure reffig:all(c)-top shows the document's shading artifacts corrected using the technique outlined in Section 3.3. The bottom image shows a pixel mask denoted illumination edges as discussed in Section 3.3.1. Figure reffig:all(d) shows the final output combining the results from both restoration algorithms. The document content looks significantly better than the input image. The lines are straight and the illumination appears uniform.

Figure 7 shows another example. This is a typical example of a document that has been folded at one time and is not completely unfolded when imaged. This example has significant shading artifacts. The first image shows the input image. The second shows the geometry rectified with original input shading. The last image shows the final output with both shading and geometry correct. In this example, the flattened image alone would still be unusable due to the severe illumination effects. The additional shading correction provides a significant improvement in the restoration.

Our last example is an oriental fan (Figure 8). The design of the fan itself is problematic for imaging. The figure shows the original input, the 3D surface reconstruction, results from geometric rectification, and results from both geometric and shading rectification. While some bleeding is evident in the shading correction, the final image is nonetheless an improvement over the original input.

5. Discussion

Our system is successful in correcting both geometric and shading distortion. The nature of our approach does assume that the cameras can observe the entire surface, i.e. there are no self occlusions on the document surface. In addition, distortion from shape smaller than the accuracy of our scanner ($0.5mm$) cannot be addressed. Shape variations this small, however, are fairly rare in printed materials and would often be considered part of the texture (content).

In regards to correcting shading, our formulation can not deal with smooth shading changes or those caused by cast shadows. Smooth shading changes fit well with the *Retinex*
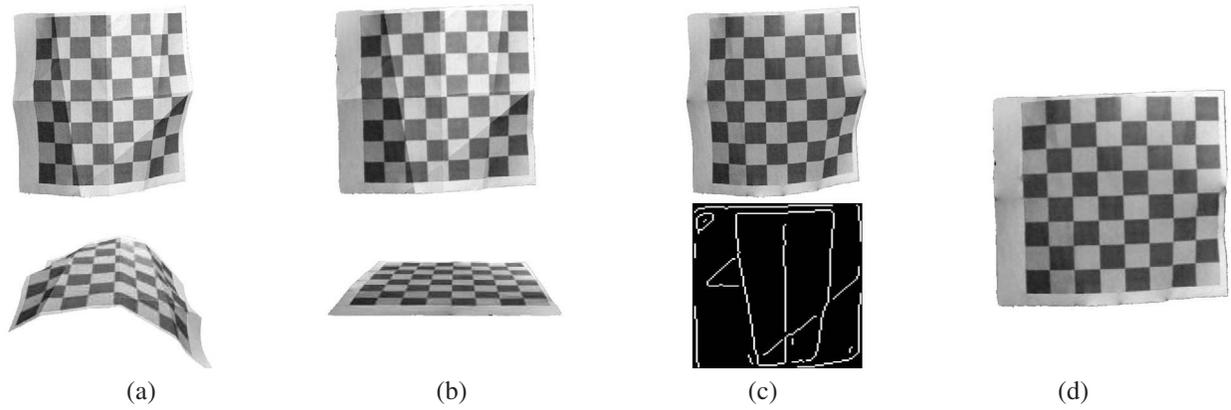
Figure 6: Our document restoration pipeline. (a) captured texture image (top) and geometry (bottom); (b) flattened document (top) and geometry (bottom); (c) input image with shading corrected (top) and the illumination edge mask (bottom); (d) results combined to generate the final output of the document without geometric and shading distortions.

Figure 7: (Left)Input image; (Middle) Flattened image without shading correction; (Right) Image with shading correction applied to the flattened image..
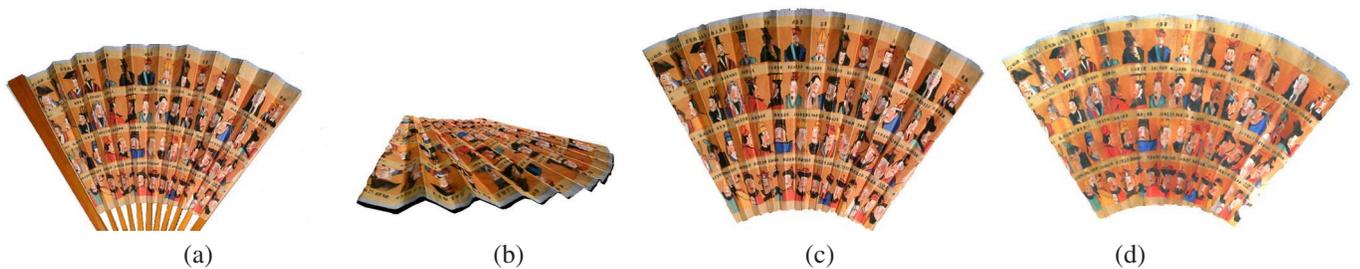

Figure 8: An oriental folding fan. (a) Input original image; (b) 3D view of the scanned data; (c) fan geometry flattened (using original input image). (d) final image with geometry and shading corrected image

theory [19], therefore they can be removed by many existing techniques, such as the *bilateral filtering* technique [21] which is relatively simple to implement. Sharp cast shadows are more challenging to remove. Under certain lighting assumptions it has been shown that hard shadows can be successfully removed by first finding a shadow invariant image [12]. While we have not implemented these techniques, we envision that the resulting image from our approach can be further processed to remove these shading artifacts we currently do not handle.

While 3D scanning may seem excessive for document imaging, for rare and valuable documents, e.g. those found in special collections and museums, such processing is often warranted. In addition, these types of documents are more likely to suffer from geometric distortion and require special imaging approaches. Very old documents, especially those made of animal skin, are often brittle and dry and are at risk for tearing and breaking when physically handled. Our system provides a viable alternative to physical restoration for such items.

6. Conclusion

We have demonstrated a system to restore the 2D content found on arbitrarily distorted documents. Our approach uses a structured-light setup to acquire the 3D surface of the document and a high-resolution image. The 3D data is used to correct geometric distortion and classify illumination and reflectance images to correct shading distortion. The results of these two processes are combined to compute the final output of an image with rectified geometry and illumination. Our approach can significantly improve the visual appearance of 2D content printed on distorted documents.

References

[1] M.D. Altschuler, B.R. Altschuler, and J. Taboada. Measuring Surfaces Space-Coded by a Laser-projected Dot Matrix. In *Imaging Application for Automated Industrial Inspection*, 1979.

[2] G. Arfken. Gauss's law, poisson's equation. In *Mathematical Methods for Physicists*, chapter 1.14, pages 74–78. Academic Press, 3rd edition, 1985.

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

[4] H.G. Barrow and J.M. Tenenbaum. *Recovering Intrinsic Scene Characteristics from Images*. Academic Press, 1978.

[5] M. Bell and W. T. Freeman. Learning Local Evidence for Shading and Reflection. In *International Conference on Computer Vision (ICCV)*, pages 670–677, 2001.

[6] A. Blake. Boundary Conditions for Lightness Computation in Mondrian world. *Computer Vision, Graphics, and Image Processing*, 32:314–327, 1985.

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

[8] M.S. Brown and C.J. Pisula. Conformal deskewing of warped documents. In *IEEE Computer Vision and Pattern Recognition*, 2005.

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

[10] 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.

[11] P. Debevec. Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography. In *Proceedings of SIGGRAPH*, pages 189–198, 1998.

[12] G. Finlayson, S. Hordley, and M. Drew. Removing shadows from images. In *ECCV'02*, page 823C836, 2002.

[13] A. Fournier, A. Gunawan, and C. Romanzin. Common Illumination between Real and Computer Generated Scenes. In *Proceedings of Graphics Interface*, 1993.

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

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

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

[17] B. K. P. Horn. Determining Lightness from an Image. *Computer Vision, Graphics, and Image Processing*, 3:277–299, 1974.

[18] 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.

[19] E. H. Land and J. J. McCann. Lightness and Retinex Theory. *Journal of the Optical Society of America*, 61:1–11, 1971.

[20] S. Marschner and D. Greenberg. Inverse Lighting for Photography. In *Proceedings of IS&T/SID Fifth Color Imaging Conference*, pages 262–265, 1997.

[21] B. M. Oh, M. Chen, J. Dorsey, and F. Durand. Image-based Modeling and Photo Editing. In *Proceedings of SIGGRAPH*, pages 433–442, 2001.

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

[23] M. F. Tappen, W. T. Freeman, and E. H. Adelson. Recovering Intrinsic Images from a Single Image. In *Advances in Neural Information Processing Systems (NIPS)*, 2003.

[24] 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*, 2004.

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

[26] Y. Yu, P. Debevec, J. Malik, and T. Hawkins. Inverse global illumination: recovering reflectance models of real scenes from photographs. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques (SIGGRAPH)*, pages 215–224, 1999.

[27] Z. Zhang, C.L. Tan, and L. Fan. Estimation of 3d shape of warped documents surface for image restoration. In *IEEE Computer Vision and Pattern Recognition*, 2004.