2008

Innovative Techniques for Digitizing and Restoring Deteriorated Historical Documents

George V. Landon, Jr.
University of Kentucky

ABSTRACT OF DISSERTATION

George V. Landon, Jr.

The Graduate School
University of Kentucky
2008

Innovative Techniques for Digitizing and Restoring Deteriorated Historical Documents

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Engineering
at the University of Kentucky

By

George V. Landon, Jr.

Lexington, Kentucky

Director: Dr. W. Brent Seales, Professor of Computer Science

Lexington, Kentucky

2008

ABSTRACT OF DISSERTATION

Innovative Techniques for Digitizing and Restoring Deteriorated Historical Documents

Recent large-scale document digitization initiatives have created new modes of access to modern library collections with the development of new hardware and software technologies. Most commonly, these digitization projects focus on accurately scanning bound texts, some reaching an efficiency of more than one million volumes per year. While vast digital collections are changing the way users access texts, current scanning paradigms can not handle many non-standard materials. Documentation forms such as manuscripts, scrolls, codices, deteriorated film, epigraphy, and rock art all hold a wealth of human knowledge in physical forms not accessible by standard book scanning technologies. This great omission motivates the development of new technology, presented by this thesis, that is not-only effective with deteriorated bound works, damaged manuscripts, and disintegrating photo-negatives but also easily utilized by non-technical staff.

First, a novel point light source calibration technique is presented that can be performed by library staff. Then, a photometric correction technique which uses known illumination and surface properties to remove shading distortions in deteriorated document images can be automatically applied. To complete the restoration process, a geometric correction is applied.

Also unique to this work is the development of an image-based uncalibrated document scanner that utilizes the transmissivity of document substrates. This scanner extracts intrinsic document color information from one or both sides of a document. Simultaneously, the document shape is estimated to obtain distortion information.

Lastly, this thesis provides a restoration framework for damaged photographic negatives that corrects photometric and geometric distortions. Current restoration techniques for the discussed form of negatives require physical manipulation to the photograph. The novel acquisition and restoration system presented here provides the first known solution to digitize and restore deteriorated photographic negatives without damaging the original negative in any way.

This thesis work develops new methods of document scanning and restoration suitable for wide-scale deployment. By creating easy to access technologies, library staff can implement their own scanning initiatives and large-scale scanning projects can expand their current document-sets.

Keywords: Document Digitization, Document Restoration, Translucent Shape Reconstruction, Photometric Correction, Geometric Correction

George V. Landon, Jr.

April 23, 2008

Innovative Techniques for Digitizing and Restoring Deteriorated Historical Documents

By

George V. Landon, Jr.

Dr. W. Brent Seales

Director of Dissertation

Dr. Raphael Finkel

Director of Graduate Studies

April 23, 2008

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

George V. Landon, Jr.

The Graduate School
University of Kentucky
2008

Innovative Techniques for Digitizing and Restoring Deteriorated Historical Documents

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Engineering
at the University of Kentucky

By

George V. Landon, Jr.

Lexington, Kentucky

Director: Dr. W. Brent Seales, Professor of Computer Science

Lexington, Kentucky

2008

DEDICATION

To my wife Karen.

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor, Dr. Brent Seales, for providing advice, support, and example throughout my entire graduate career. I would also like to give a special thanks to my doctoral committee: Dr. Fuhua Cheng for his always constructive comments, Dr. Ruiguang Yang for his technical insights and concerns, and Dr. Mark Lauersdorf for providing additional motivation for my work with new applications. I would also like to thank the late Dr. Ross Scaife for his strong conviction to building bridges between the humanities and technology.

My wife, Karen, has been a constant source of encouragement, and, for that, I thank her. Also, I will thank my son, George III, who gave me added motivation to finish my graduate work. I would also like to thank my family for their love and support throughout the entire graduate process.

Table of Contents

List of Tables

List of Figures

List of Files

1. Dissertation_Landon_GV_08.pdf 13,437 KB

Chapter 1

Introduction

This thesis demonstrates new methods to digitize and restore various forms of historic deteriorated documentation.

Thesis

Recent large-scale document digitization initiatives have created new modes of access to modern library collections with the development of new hardware and software technologies. Most commonly, these digitization projects focus on accurately scanning bound texts, some reaching an efficiency of more than one million volumes per year. While vast digital collections are changing the way users access texts, current scanning paradigms can not handle many non-standard materials. Documentation forms such as manuscripts, scrolls, codices, deteriorated film, epigraphy, and rock art all hold a wealth of human knowledge in physical forms are not accessible by standard book scanning technologies. This great omission motivates the development of new technology, presented by this thesis, that is effective with deteriorated bound works, damaged manuscripts, and disintegrating photo-negatives but also easily utilized by non-technical staff.

It is extremely difficult to create uniform illumination across an entire surface when digitizing a non-planar document. This work solves this challenge first developing a novel light source calibration technique and then introducing a photometric correction develops an invertible intensity model using document shape and light source information. To expand the number and type of documents that can be digitized and restored, this thesis introduces a new transmissive document scanning framework. The photometric and geometric digitizer extracts intrinsic color information and shape distortion information from documents that are semi-transmissive in nature, which includes most forms of paper. This thesis also provides a breakthrough application in restoring deteriorated photographic negatives that have remained unprintable till now.

To begin this venture, one must first accept that documents, in all their forms, play a dominant role in mankind's cultural heritage. The significance of this cultural heritage is stated well by the United Nations Educational Scientific and Cultural Organization (UNESCO) Division of Cultural Heritage Programme 2004-2005 [89]:

> The cultural heritage, with its multifarious origins and as a receptacle of memory, embodies the symbolic value of cultural identities and constitutes a fundamental reference for structuring society. The various components of the physical heritage are powerful symbols and are the tangible manifestation of the diversity of peoples, cultures and beliefs that coexist or have existed successively in a single region.

Thus, an undeniable link exists between cultural heritage and all forms of documentation.

Originating from the first cave paintings to modern day web pages, humans have used written form as a way to communicate their origins, beliefs, and current knowledge. Like oral traditions before it, documentation is another venue where cultural heritage is imparted from one generation to the next. The end of the twentieth century delivered technological developments that heightened awareness of the connection between documents and cultural heritage. Two of the most prevalent developments were widespread access to the Internet and electronic imaging devices.

First, widespread Internet access gives purpose to digitization projects. Once document collections are digitized, the contents may be instantly accessible from anywhere in the world. More recently, broadband Internet access provides a channel for content-rich applications. Photo-realistic visual media can be transmitted without a reduction in quality. Also, new web-based applications have added features and interactivity available to on-line collections. These improvements, in conjunction with search technology, allow exact word searches from millions of physical documents.

Second, low-cost high-quality imaging devices have become more readily available to digitize texts. Flatbed scanners, once considered the best digitization technique for books, are now sold at reasonable prices. Digital still cameras are now available at resolutions much greater than the 300 dots-per-inch (DPI) normally used for document archiving.

The culmination of these technological developments has led to a new era in digital cultural heritage access. Only by accurate document preservation and dissemination can cultural heritage be maintained for future generations.

1.1 The Importance of Document Digitization

The quest for amassing all of the printed works in the world is nothing new. The famous Royal Library of Alexandria, 300BC, began with the goal of acquiring all of the known world's knowledge [71]. With an estimated 400,000 to 700,000 volumes in the collection before the library's destruction, this goal seemed attainable.

Works of cultural heritage provide a guiding line throughout all distinct people groups, each with an enormous amount of knowledge. The concept of gathering all documented knowledge and beliefs still resonates today. Like the Alexandrian library, many organizations are now attempting to acquire all of the world's printed knowledge in digital archives. However, the widespread use of modern technology has transformed both preservation and access to documents of cultural heritage.

1.1.1 Large-Scale Digitization Initiatives

According to the OCLC Online Computer Library Center WorldCat bibliographic database, there are 32 million records describing printed books in American libraries [56]. This translates into at least 32 million unique books that must be digitized in the United States alone. From a global perspective, Lyman et al. have attempted to estimate the number of books printed in the world. Through methodical research and a fair number of extrapolations, the group estimates that the world inventory of original books might be between 74 million and 175 million books [60]. However, this approximation does not include scrolls, manuscripts, and other document forms.

For the millions of books currently available on library shelves, recent projects have begun the daunting task of providing access to these volumes through new digitization and delivery methods. These projects have successfully started a trend to provide the world easy access to all printed materials and thus a direct link to cultural heritage. However, documenting the world's cultural heritage should not be focused solely on the textual information of written or typed volumes. To fully understand what is lacking, one must first recognize the accomplishments that have already been established.

1.1.1.1 Project Gutenberg

One of the pioneer digitization initiatives is Project Gutenberg [86]. Started in 1971, this work focuses on acquiring digital textual versions of public domain books. Driven by volunteers, books are manually converted into ASCII plain text and other eBook formats.

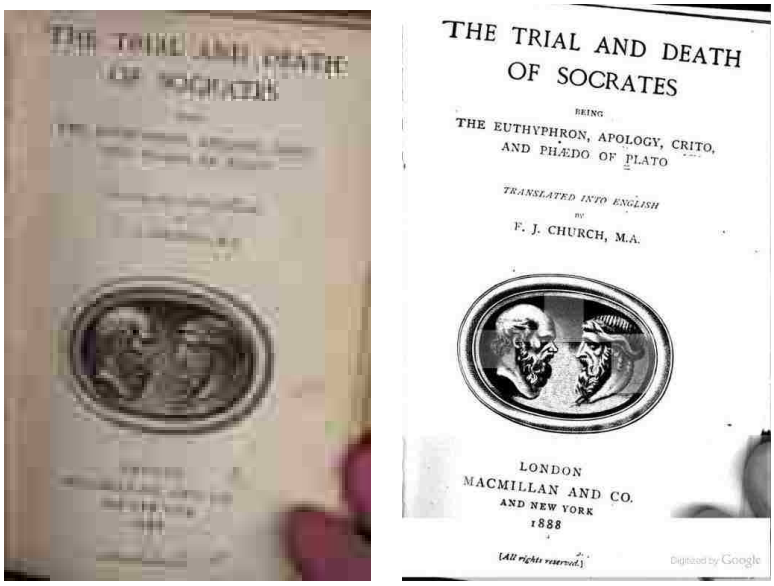
Project Gutenberg started with the goal of providing network access to all books. While much of the content is limited to text, this digital library is accessible from any device with the ability to render text. This lowest common denominator of content excludes no networked user from access.

There are now over 20,000 electronic books available from the project's website. Moreover, other projects are now building off the ground-breaking ideas that brought Project Gutenberg to fruition.

1.1.1.2 Google Book Search

In 2004, Google Inc. began Google Book Search, an effort to digitize all of the world's books. The project claims to be able to digitize one million books per year [34]. However, the company does not provide a total number of books digitized to date.

The goal of Google Book Search is to make textual content in books searchable. This requires a great deal of image processing to accurately segment text from the background or illustrations. The results of post-processing can be seen in Figure 1.1. However, by converting these document images into easily detectable text much of the contextual detail, such as graphics and illustrations, are greatly compromised. Google has made great strides



(a) The original imaged title page. (b) The post-processed version included in the digital library.

Figure 1.1: An example scan from Google Book Search.

toward making this content available. Omitting these contextual details somewhat limit the overall presentation of a volume.

However, it should be noted that Google's digitization technology is not without restrictions. From the Google Book Search Policy [34].

> Google developed innovative technology to scan the content without harming the books. Any book deemed too fragile will not be scanned by Google, but may be treated by expert library staff. Once scanned, all print volumes are returned to the library collections.

Unfortunately, it appears that Google's digital library will not include many fragile works including older volumes that have deteriorated with age.

1.1.1.3 Open Content Alliance

The accessibility of digitized content becomes controversial as more works become digitally available. One the one hand, content access is restricted as opposed to the other mode of thought that makes content completely accessible. For example, Google maintains control and access of the works they have digitized. In contrast, other groups, like Project Gutenberg, maintain freely accessible and reusable public domain works.

Started as a venture between the Internet Archive and Yahoo!, Inc. in early 2005, the Open Content Alliance is also working to make digital collections available. Focusing on improving availability of the content, the OCA has attracted contributors that are willing to donate their already digitized collections. Therefore, much of the collections will be decided upon by the contributors as will the technology employed.

1.2 Digitization is not a One-Size-Fits-All Process

These large-scale document digitization initiatives have exposed complex issues that must be dealt with as the construction of digital libraries continues. Current technology has provided a foundation to start these projects by producing page-turning book and flatbed scanners, both of which are more than satisfactory for most modern bound volumes. However, focusing on only modern texts ignores a significant amount of human knowledge.

The earliest human documents are rock paintings and carvings. Authors of these works made records of heritage and beliefs—the same motivations that compel many modern authors to write. An example petroglyph, an engraved rock surface, used to record a religious deity was created by the Taino people of the Caribbean is shown if Figure 1.2(a).

As documentation moved to typographic recordings, stone continued to be the canvas of choice for many types of legal documentation. This is evident in the vast amount of

(a) Rock art (b) Epigraphy (c) Photography

Figure 1.2: Other types of documents.

epigraphic recordings created in the Greek and Roman empires. Massive pieces of stone were used to record information, see Figure 1.2(b), ranging from historical chronologies to listing donors for building projects.

As new materials such as papyrus and parchment emerged, the creation of documents increased in speed and extent. The composition of these writings came in the form of scrolls and codices. With the volume of books increasing, there were an estimated 30,000 books in Europe before the creation of the moveable-type printing press [41].

Much like early rock art and paintings, original photographs provide another source of historical perspective. Commonly used for many recording purposes, early photographs were often commissioned as a documentation technique. The subjects of these recordings range from family portraits for genealogical purposes to architectural images used to capture economic and locale information. From early photography to current, these visual records provided an irreplaceable glimpse into the past, see Figure 1.2(c).

Expanding our concept of which documents need digitization and preservation also changes the perception of content for standard textual documents. Consequently, by focusing purely on document text, much of the meaning and contextual information is lost. This importance was addressed by Library of Congress Manuscript Digitization Demonstration Project: Final Report [58].

> The committee discussed the degree to which an image need replicate the look and feel of an original manuscript (or other archival) artifact. Although most members agreed that Library of Congress treasures, e.g., drafts of the Gettysburg address, warranted a kind of museum-quality facsimile, there was less consensus that such treatment was warranted for routine documents,

Error	Cause	Effect
Photometric	Non-uniform lighting	inconsistent shading of the content
Geometric	Warping of the substrate	skewing of the content

Table 1.1: The effects of errors on deteriorated document digitization.

> especially the kind of twentieth-century typescripts that form the greatest portion of the Federal Theater Project collection.

This argument resounds for the digitization of all historical documents. Therefore, document digitization cannot be viewed as a standard procedure that excludes many document types and compromises what content is preserved.

1.2.1 Digitization should be a task performed by curators/librarians

As the Google Book Search suggests, some texts are "deemed too fragile" to be handled and scanned by technicians, but "may be treated by expert library staff [34]." This leads to two important questions:

- What digitization technologies are available for fragile volumes?
- Are these technologies available to the expert library staff?

It is the goal of this work to produce tools that are effective at digitizing fragile and deteriorated documents while making them accessible to library staff.

1.3 Effects of Digitization Error

New technology has focused on scanning documents that are not compatible with normal digitization techniques. Due to deterioration and damage, it is impossible to place many older texts in a flatbed or book scanner. Consequently, these degradations not only effect the digitization technology but also greatly reduce the quality of the acquired image.

There are two main sources of error commonly encountered during fragile document digitization: photometric and geometric. Table 1.1 compares the effects of these errors on the digitized document.

In this work, both photometric and geometric error are improved.

1.3.1 Photometric Correction

Photometry measures the perceivable intensities of light by the human eye. The term photometric correction, within the scope of document digitization, applies directly to the

perceived color in a document. Therefore, a photometrically correct document image must refer to an image of a document that contains intensities, or color, across the entire surface that replicates the physical document. However, when a document has become warped from deterioration or damage, creating this photometrically correct document requires uniform illumination across the document—an impractical problem to say the least.

Documents with non-planar shapes require advanced lighting systems to produce acceptable illumination for digitization. Two or more high-powered lights with special diffusing filters must be used to attempt to reduce extreme shading distortions in imaged documents. In contrast to using lighting equipment to illuminate these difficult objects, customized state-of-the-art software developed here may be used to produce the same result without excluding curators from the digitization process.

1.3.1.1 Light Source Calibration

The accurate estimation of scene lighting is required to calculate material properties of the scene. This understanding is used as motivation for the development of a novel light source calibration method as described in Chapter 2.

Using pre-existing document digitization pipelines, a light probe is introduced that requires no *a priori* information from the user. This light probe is made from a simple piece of folded white paper. The composition of the probe is important for two reasons. First, the construction of the probe is simplified to a point that anyone can create one out of, arguably, the most plentiful material in a library or archive. Second, paper itself is viewed as an acceptable object to be placed near priceless works, in contrast to cumbersome specialized rigs.

Using this new and acceptable light probe, the illumination in a scene may be calculated. This type of light source calibration is required for complete archival object digitization. Material information of objects can be determined once the light is calibrated, providing added detail beyond the surface geometry.

Once the illumination is calibrated, there are numerous applications that extend the use of three-dimensional representations. For instance, since the observed appearance of an object has been captured and the scene illumination known, photo-realistic rendering can be accomplished with a dynamic, or moving, light source. Known material parameters provide the intrinsic appearance of a surface when lighting is not included in the rendering. This motivates a direct application for the acquisition of intrinsic object color for items of great importance to this endeavor—deteriorated documents.

1.3.1.2 Photometric Correction of Documents

(a) Illuminated Latin Bible 1407AD by Gerard Brils, Belgium. (Photo Courtesy A. Pingstone 2005)

(b) Illuminated Rossano Gospels 550AD. (Image shows Christ before Pilate).

Figure 1.3: Examples of deteriorated manuscripts.

Current conventions for archival imaging of non-planar documents require expertise that is beyond the reach of many collections. Consequently, many digitization efforts are unable to produce accurate recordings of deteriorated texts, see Figure 1.3. The Latin Bible shown in Figure 1.3(a) contains low-frequency warpage which reduces legibility for some areas of the text. The Rossano Gospels contain high-frequency wrinkles that greatly reduce the quality of the text and illuminations as shown in Figure 1.3(b). However, this work takes warped documents, digitized under imperfect lighting conditions, and restores them with accurate color information when the light source is fully calibrated.

Using light calibration methods in Chapter 2, a novel document photometric correction method is disclosed in Chapter 3. This application provides solutions that bypass the need for professional photographers and expensive lighting equipment.

The result of this work produces a method that corrects lighting and shading in document images without regard for particularities in the document content. Therefore, any type of book or manuscript can be restored with no variation of result if the document contains text, illustrations, or photographs. This process allows the use of simplistic lighting equipment and the novel paper light probe to restore the adverse imaging effects caused by inadequate illumination.

When the document is corrected photometrically, the color distortions perceivable by the observer are diminished. However, for many document applications, the uncorrected geometric distortions still reduce content detection and legibility.

1.3.2 Geometric Correction of Documents

Discussion of the geometry of a document simply refers to its physical shape. Digitizing non-planar documents with standard methods introduces shearing and skewing distortions of the content. For modern texts, damaged volumes can be physically manipulated to reduce these geometric distortions. However, for conserved texts, this type of handling is not an option.

There have been two main approaches to correcting geometric distortions in document images. First, image-based methods rely on a combination of assumptions about the content/context with known constraints on the document. These methods either look at the actual content of the document to determine where distortions occur or at the document composition such as distorted paper edges or bindings. A model is fit to these distortions which can be reversed. This reversion effectively reduces these distortions for many documents.

Second are physically-based simulations of document unwarping. Using the complete three-dimensional geometry of the document, the surface is modeled as malleable material that will deform when interacting with external forces. Many examples exist where the geometry and composition of the document surface is complex enough to prohibit accurate three-dimensional and image-based acquisition. Results that can be acquired with these techniques are discussed and shown in Chapter 4 as well as the unique limitations of each method. A new application of physical-simulation based restoration is introduced in Chapter 5.

The combined application of the photometric and geometric restoration techniques presented in Chapters 3 and 4 correct imaging errors for many documents. However, as technology stands today, there are a large number of documents excluded from these restoration techniques. For example, paper with fine folds and ridges are often difficult to accurately digitize with scanning equipment available to librarians. Also, many types of documents are created by combining various layers of assorted materials. For example, photographic negatives are typically created from at least two layers of plastics.

1.4 A New Paradigm for Document Scanning

Many historical documents contain deteriorations in the surface, such as wrinkles and creases, that have details which are undetectable by imaging devices. In these cases, developing an advanced model for each pixel in the image provides a technique to compensate for lack of resolution in the document sampling.

This pixel-by-pixel image-based acquisition and restoration methodology, discussed in Chapter 5, exploits the transmissiveness of documents to obtain the intrinsic color and distorted shape. The digitization process uses an LCD monitor and a digital video camera both controlled by the same computer to effectively scan the object. The document is placed on top of the display and light patterns are rear-illuminated on the document. As these patterns are displayed, the camera observes the illumination distortions introduced by the document and an overall model is generated to compensate for these discrepancies. The generated model for each pixel allows the photometrically correct content to be extracted directly. In addition, the model provides document shape information that can be used to restore the shape of a geometrically warped substrate with restoration procedures described in Chapter 4.

This image-based technique provides a direct way to generate restored images without knowing any additional information about the document or the scene set up. The transmissive scanning process has another feature. Duplex-sided documents can be scanned in one-pass without the common issue of show-through noise. The applications of this scanner go beyond standard one or two sided documents. Since the scanning technology works completely independent of the document type, more complex documents have been validated for this digitization technique. In fact, even multi-layered documents that remain semi-transmissive can also be digitized and restored with this same approach.

1.5 Moving Beyond Standard Document Archetypes

Thus far, the focus of document restoration research has been on acquiring virtual copies of bound books or manuscript pages. Consequently, restoration techniques for other documents types have been grossly underdeveloped. This lack of research has limited the broader application of virtual document restoration.

An obvious example of these limitations is deteriorated photographic negatives. Many photographic negatives, dating from 1925 to 1955, have deteriorated in various layers making the negatives unprintable. This deterioration has affected many photographic archives

including the Library of Congress, the National Archives, and many state and university collections. The broader implications of these historic collections is unmeasurable.

The standard restoration technique for acetate negatives requires a chemical-based physical separation of layers in a permanent and destructive manner. While this technique provides the restored image, it requires a vast amount of expert time and leaves the photograph forever altered. It should also be noted that the degradation of these negatives can not be stopped, only slowed. Therefore, digitizing these historic photo-negatives to virtual restoration is a technology that must be researched and developed.

Present digitization procedures are not deployable for these multi-layer documents. Due to a great deal of non-uniform separation between layers, the final photographic image contains many channeled photometric distortions (see Figure 1.2(c)). However, the transmissive document scanner introduced in Chapter 5 has been demonstrated to effectively digitized non-standard materials. Since this scanner works in a fully image-based modality, it alleviates many complications that arise with complex materials stacked layer upon layer such as accurate three-dimensional acquisition or feature detection.

In light of these developments, extensions that digitize and restore multi-layer documents may be developed. Consequently, for documents that were completely inaccessible by normal digitization techniques, such as ancient texts with multiple layers of disparate writing, this work's technology will open the door to reveal their contents.

In summary, the following contributions are delivered in this thesis.

- A novel point light source calibration technique that can be performed by library staff.
- A photometric correction technique which uses known illumination and surface properties to remove shading distortions in deteriorated document images.
- An image-based uncalibrated document scanner that utilizes the transmissivity of document substrates to:
  - Extract intrinsic document color information from one or both sides
  - Extract document shape distortion information
- A restoration framework for damaged photographic negatives that corrects photometric and geometric distortions.

Chapter 2

Light Source Estimation

2.1 Introduction

The accurate estimation of scene lighting provides much of the information necessary to calculate material properties of objects in a scene. The lighting models typically estimate light position, or at least illumination direction, and illumination color for one or more light sources.

In the past, techniques use objects themselves to estimate the lighting parameters in a scene. While these methods provide good results, they require the scene properties to be known beforehand. To reduce the constraints on a scene, special light calibration devices, known as light probes, have been developed. Light probes come in a variety of forms and can produce highly accurate light source estimations. As with any specialized equipment, light probes may be easily damaged and unwelcome in areas with conserved priceless works. Therefore current techniques require either a complete understanding of object shapes and reflectance properties in a scene, or specialized, and often complicated, light probes to be inserted in the scene.

In particular, when light calibration is needed for document digitization applications neither of the two previously accepted methods fit well into the scanning paradigm. The document shape and albedo, or intrinsic color, of damaged works do not fit a typical model to aide with the modeling of illumination. Moreover, using a specialized light probe can often be discouraged, since mechanical apparatus could damage the conserved works. Consequently, the light probe introduced in this work falls between the two standard methods for light source calibration.

The explicit intent of the novel light probe presented here is to estimate a point or directional light source color and position. This unique probe is constructed from folded white letter paper with known material properties. The probe's shape is then easily acquired with most three-dimensional scanners and reconstructed using standard three dimensional acquisition techniques. Finally, a high resolution texture image of the surface is obtained.

Melding all of this acquired data, a model is created of the light source position/direction and color using the general irradiance equation.

Light source calibration is vital for accurate object digitization. Understanding the illumination parameters in a scene allows estimation of material properties for objects of interest. This material information provides added dimensions of detail over the base surface geometry. Once the illumination properties of a scene are known, the information can be used in a variety of ways. Photo-realistic rendering can be performed by estimating how the light sources interact with the surface.

More importantly, the material properties of objects provide the intrinsic appearance of a surface without variations due to lighting and other atmospheric effects. The acquisition of object albedo is achieved once excess noise is suppressed. One object type of particular interest to this work are damaged and deteriorated documents, specifically ancient manuscripts and codices. When digitizing these non-planar documents in non-optimal lighting conditions, with the light source fully calibrated, the intrinsic color information may be recovered.

2.1.1 Limitations of current light source calibration techniques

The interaction of light with objects in a scene is an important area of research in Computer Vision. For instance, the estimation of object shape by the observed intensities of its surface was one of the first concepts to produce convincing shape reconstruction results. As a result, this method, Shape-from-Shading, has been an area of major focus in Computer Vision[43, 92] for many decades.

There are four central parameters involved in SfS research: lighting, surface shape, surface reflectance(albedo), and shading. Normally a set of assumptions are made including a Lambertian surface[54] with constant albedo, directional light source or source at the camera optical center, and an orthogonal projection model [24]. In addition Belhumeur et al. [5] have proved that when the light direction and the Lambertian surface reflectance are unknown, the problem is ill-posed with no unique solution. Prados and Faugeras model SfS as a well-posed problem by taking into account the illumination attenuation term [69]. Assuming known light source (at camera optical center), constant Lambertian surface reflectance, and the camera parameters, the shading allows recovery of surface shape from only one image without additional information.

Photometric stereo is another direction taken in providing a Shape-from-Shading solution that lessens the light source assumption by allowing near light sources to be

modeled. Introduced by Woodham [92], this method works with multiple images of a scene under varying illumination condition to add constraints to the reconstruction model. Iwahori et al. [47] performed shape from shading assuming a point light source. To compensate for the added unknowns of a point light source, multiple images are used to obtain the surface shading information. For each of these images the light source position must be manipulated to over-constrain the system. Kim et al. [49] also reconstructed shape using photometric stereo and multiple point light source positions. Koenderink et al. [52] reduce the constraint of knowing the incident light directions for each image acquired. Finally, Frolova et al. [28] extend the previous use of spherical harmonics in photometric stereo to reconstruct depth using point light sources.

Standard SfS assumptions still remain ill-posed without prior information about the construction of a scene. Therefore, many groups have added *light probes* that return the problem to a well-defined state. Hara et al.[40] provide a method to estimate a single point light source position and surface reflectance properties from a single image with known scene geometry. This work assumes detectable surface specularity that can be separated from the diffuse surface reflectance.

Spherical light probes have also provided solutions to light-source calibration. The use of matte spheres that follow Lambertian reflectance properties [98] [96] has also been widely used in light source estimation methods. Hertzmann and Seitz [42] proposed an approach that recovers shape and material by referencing a shading-normal table constructed using a spherical light probe. In their work, distant light and orthographic camera are not assumed, but the approach does remove the Lambertian surface assumption. Powell et al.[68] have introduced a novel calibration object, using 2-3 spheres with known position, to calculate the 3D position of multiple light sources. Also, Takai et al.[84] use the notion of a difference sphere to calculate the position and radiance of multiple light sources. Alldrin et al.[3] use a planar light probe with a multi-layered, transparent medium which differentially absorbs and reflects light. Fourier series coefficients are recovered from the incident lighting parametrized over the plane. Sata et al.[75] have developed robust methods for light estimation using cast shadows of known geometry.

While these methods have produced accurate results for light source calibration, the special light probe requirement adds to the complexity of the system. The solution presented here attempts to achieve the same accuracy without the necessity of complicated light calibration equipment.

2.1.1.1 Appropriateness for Library Collections

While the usefulness of these methods cannot be overstated, their contribution in document imaging is limited to a library archive environment. These limitations led the development of the work presented here. The specific requirements for this environment are a light probe that can be manufactured in-field and providing minimal disruption to current three-dimensional document digitization pipelines.

2.2 Light Source Estimation

Derived from Shape-from-Shading theory, an innovative approach is developed here to estimate lighting condition, including source position and color/intensity, with a folded white-paper light probe. This novel initiative deals with both directional light and point light sources, which is excluded in many existing SfS/photometric stereo algorithms. The method is designed to fit into the standard document acquisition procedures and provide automatic digital restoration. Only a simple piece of white paper, folded at least ten times, is required to estimate the light source including the position for point sources and direction for directional sources, each in the scanner coordinate frame, and its acquired RGB color. The paper is folded such that its surface contains *facets*

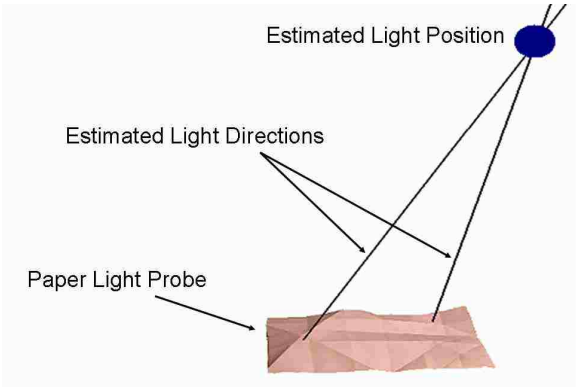


Figure 2.1: A screenshot of our system with estimated parameters labeled.

The folded paper is added to the acquisition scene serving as a *light probe*. The scene and probe are scanned simultaneously, to build a model taking added advantage of the scanning facilities already being used for document acquisition.

2.2.1 Assumptions

As stated previously, a number of assumptions have been used for Shape-from-Shading and light source estimation in current literature. Unfortunately, many of these assumptions do

not hold true in the real-world.

In particular, one common assumption in SfS is the "directional lighting" condition where a light source is at distance from the scene such that all incoming rays of light appear parallel. Consequently, this assumption easily holds for synthetic experiments, but can be hard to reproduce with real-world configurations. The method presented here relaxes this strong constraint by making a reasonable assumption:

For a very small surface region, the illumination from a point light source approximates to directional lighting.

Second, Lambertian surfaces are assumed for both the light calibration target and the document itself. This assumption approximates reflectance properties relatively well for many commonly used paper types, especially historical documents.

In addition, direct illumination is assumed with inter-reflections and shadows neglected. Contingent on this assumption, ambient light must be minimal in the scene, a practical requirement for most document archives. The final assumption of this work is that the attenuation of light during transmission in medium (air) can be ignored.

2.2.2 Estimating Light Source Position

The foundation for this proposed light source calibration method is the general Irradiance Equation:

$$I(u,v) = I_L \rho(x,y,z) \frac{\vec{L}(x,y,z) \cdot \vec{N}(x,y,z)}{|\vec{L}(x,y,z)||\vec{N}(x,y,z)|} \tag{2.1}$$

Where I_L is the light intensity/color, ρ denotes the albedo at the shaded point, L is the lighting direction vector, and N is the surface normal vector at the point.

Under the perspective projection model, let $X = (x,y,z)$ denote a surface point identified by real-world coordinates. It is projected onto image point $i = (u,v)$:

$$i = \mathbb{P}X \tag{2.2}$$

where $\mathbb{P}$ is the camera's projection matrix.

Under the point light source condition in this research, it is assumed that within a "small" region the lighting approximates to directional lighting. This small region on the surface, lit with approximately directional lighting, is referred to here as a *patch*. How small a region should be to satisfy the assumption depends on how far the light source is from the lit object. In the experiments presented here, light sources were placed around 1.5 meters, away from the object, and patches with radiuses about 2.5 centimeters were chosen. How patch size influences the estimation result is discussed in section 2.3.

Suppose N patches are chosen on the calibration target and each patch is illuminated from direction $\vec{L_n} = (lx_n, ly_n, lz_n)$ which is normalized, where $n = 1, 2, ..., N$.

Within patch n, suppose there are M *illumination facets*. An illumination facet is defined as the union of connected polygon faces that share similar normals. Each facet has an intensity I_m and unit normal $\vec{N_m} = (nx_m, ny_m, nz_m)$ where $m = 1, 2, ..., M$. To handle sources of error from real scans, this work has added statistical techniques to compute the normal and intensity for each facet by optimally fitting point samples.

A *Principle Components Analysis Method*[62] is used to compute the normal to the total least-squares fitting plane of a set of j points by finding the smallest eigenvalue of the covariance matrix C_N which is defined as:

$$C_N = \frac{1}{j}\sum_{i=1}^{j}(X_i - \bar{X})(X_i - \bar{X})^T \tag{2.3}$$

where X_i $1 \leq i \leq j$ are 3D points to fit and $\bar{X} = \frac{1}{j}\sum_{i=1}^{j} X_i$.

According to equations 2.1 and 2.2, we can construct an irradiance equation for each patch by assuming the calibration target albedo is white (i.e. $\rho = I_{white}$) to match the color of standard letter paper:

$$I_L I_{white}(lx_n nx_m + ly_n ny_m + lz_n nz_m) = I_m \tag{2.4}$$

Dividing the m^{th} facet by the k^{th} facet yields:

$$\frac{I_m}{I_k} = \frac{lx_n nx_m + ly_n ny_m + lz_n nz_m}{lx_n nx_k + ly_n ny_k + lz_n nz_k} \tag{2.5}$$

which can be written as:

$$\begin{aligned} A_{mk} lx_n + B_{mk} ly_n + C_{mk} lz_n = 0 \\ where: \quad \begin{aligned} A_{mk} &= I_m nx_k - I_k nx_m \\ B_{mk} &= I_m ny_k - I_k ny_m \\ C_{mk} &= I_m nz_k - I_k nz_m \end{aligned} \end{aligned} \tag{2.6}$$

A linear system can be constructed with 3 unknowns: lx_n,ly_n, lz_n for each patch n:

$$\begin{bmatrix} A_{12} & B_{12} & C_{12} \\ A_{23} & B_{23} & C_{23} \\ \cdots & \cdots & \cdots \\ A_{(M-1)M} & B_{(M-1)M} & C_{(M-1)M} \end{bmatrix} \begin{bmatrix} lx_n \\ ly_n \\ lz_n \end{bmatrix} = 0 \tag{2.7}$$

At minimum, four facets are needed to solve the lighting direction for each patch. Usually more facets are used to overcome image noise and other factors such as 3D reconstruction error. Singular Value Decomposition (SVD) is applied to the coefficient matrix to find the null-space of the system and obtain the optimal least-squares solution.

For directional light, one patch is enough to obtain the light direction. Since each ray of light incident on the surface is arriving at the same angle, only one surface normal is necessary to find the light direction. However, for the more difficult case of a point light source, more information is needed. After having obtained the lighting directions for at least two patches, the intersection of the light rays associated with the patches is calculated to obtain the point light source position S, which can be defined through the parametric representation of each light ray:

$$S = C_n + \vec{L_n} t_n \quad n = 1, 2, \ldots, N \tag{2.8}$$

which leads to the linear system:

$$\begin{bmatrix} 1 & -\vec{L_1} & 0 & \ldots & 0 \\ 1 & 0 & -\vec{L_2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 0 & \ldots & -\vec{L_N} \end{bmatrix} \begin{bmatrix} S \\ t_1 \\ t_2 \\ \ldots \\ t_N \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ \ldots \\ C_N \end{bmatrix} \tag{2.9}$$

where C_n is the center of patch n. The solution can be found by multiplying the right side matrix of the equation by the pseudo-inverse of the coefficient matrix.

While initial experiments were performed by hand picking patches, this is burdensome for users. In practice, a RANSAC algorithm is applied to overcome bad patch estimations due to un-modeled reflectance properties and a lack of surface normal variations.

2.2.3 Estimating Light Source Color

With estimated light source position $S = (sx, sy, sz)$, each illumination facet m can vote for a light color estimation by solving I_L from equation 2.4:

$$I_{Lm} = \frac{I_m}{I_{white}(\frac{\overrightarrow{S-C_m}}{|S-C_m|} \cdot \overrightarrow{N_m})} \tag{2.10}$$

where C_m is the center of the facet. The colors voted for are averaged as the final estimation. The operations are applied independently on the R, G, and B color channels. This allows a direct solution for the captured light source color.

2.2.4 System Implementation

The light calibration method described in Section 2.2 is implemented as a graphical interactive application. Upon execution the user is presented with a three-dimensional rendering of the acquired light probe textured with the captured image.

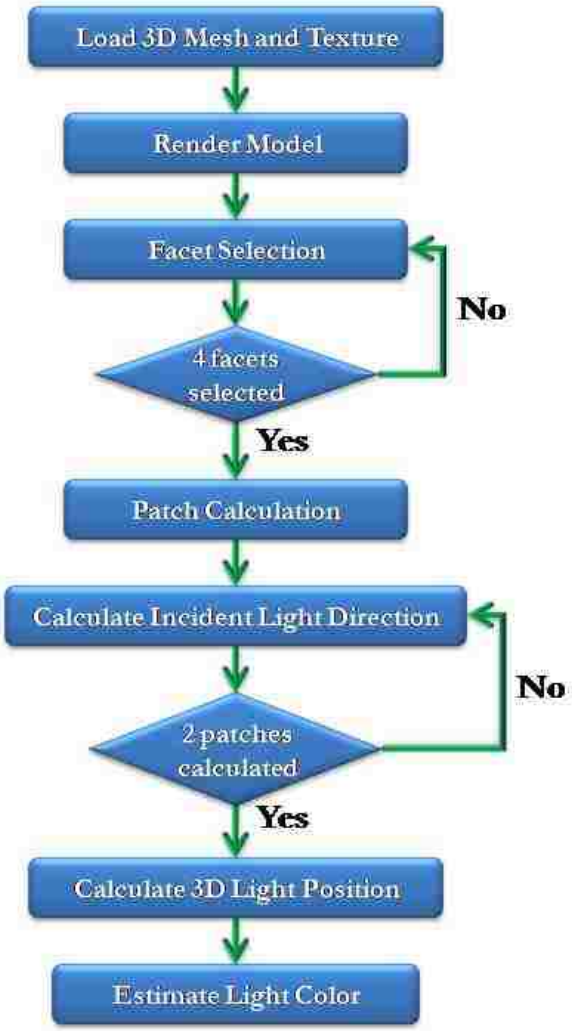


Figure 2.2: Algorithm for light estimation

The user then selects one or more triangle faces with the cursor that will represent a facet. Once the user completes the selection of triangles, the facet is saved for future computation. This step is repeated for at least four facets, the minimum number to estimate a patch. Once the minimum number of facets are reached, the user is polled to continue facet selection or compute the current patch.

When a patch is finalized the incident light ray is rendered onscreen to give the user visual feedback. Patches are selected with the previous process until two patches have been calculated. Then the 3D position of the light source is estimated onscreen. This setup is shown in Figure 2.1. Also at this step in execution, the light source color is calculated depending on the estimated 3D light position.

At this point the user may continue to select facets and compute additional patches while watching the estimated light position update in real-time. Otherwise, the user may save the estimated illumination properties and terminate the application.

2.3 Empirical Analysis

To justify the present method, this initiative presents analysis of various aspects of the light source calibration method with ground-truth data. Using synthetic data, most sources of error are controlled. Errors can be analyzed independently by simulation in a controlled environment. After quantifying the error, results from an actual scan performed with a 3D document scanner are displayed.

2.3.1 Surface Perturbations

All readily available 3D scanning technologies introduce some form of noise in the acquired surface. In particular, high-frequency noise commonly hinders most scanning technologies. These small variations in the surface add increased amounts of error when acquiring three-dimensional data. This error propagates directly to the surface normal for each polygon, which is used directly when estimating the light-source position. To determine how great the influence of high-frequency surface noise is on the final estimation, experimental calculations are performed directly on ground-truth data. Using a random number generator following a uniform distribution, varying levels of random error are added to each dimension of the 3D vertices. Figure 2.3 shows the result of applying these random surface perturbations. As the plot shows, the error follows a linear pattern. Figure 2.4 shows the difference of the ground-truth shaded image with the shaded image of an estimated light position with added error.

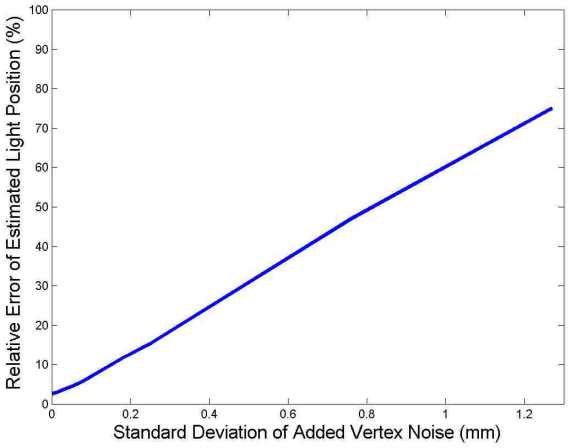
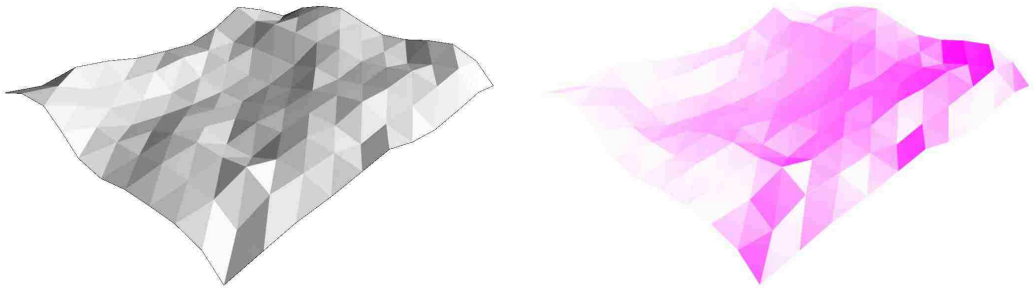


Figure 2.3: Demonstration of the error introduced by random noise in the 3D points of a surface.



(a) The ground-truth shaded surface.

(b) The difference with vertex error added (σ = 1.2).

Figure 2.4: The difference of the ground-truth shaded surface and shaded surface with error added. Note: White (diff= 0) and Red (diff= 255).

2.3.2 Number of Patches Selected

During these experiments, it was determined that the number of patches selected determines the accuracy of the final light position result. To test this result, an increasing number of

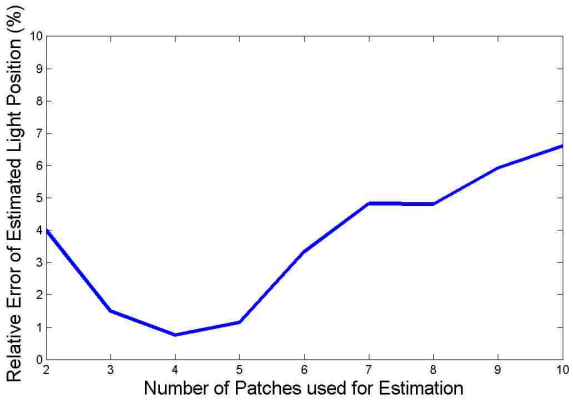


Figure 2.5: Demonstration of the accuracy produced with varying numbers of patches selected.

patches with five facets each were selected. Moreover, the light position was calculated with each increase. The minimum of number of patches must be two, since this method relies on the intersection of at least two rays. Therefore, this test started with two patches and eight more patches were added for a total of ten patches as seen in Figure 2.5. There are two notable features of this plot. The first is the global minimum present when there are between three and five patches. Second, once more than five patches are used, error begins to greatly reduce the overall accuracy of the light estimation. Therefore, the results shown in this presentation all used four patches for the light source estimation.

2.3.3 Number of Facets per Patch

Another parameter in this algorithm is the number of facets selected per patch. The underlying equations require at least 4 facets to solve the set of equations. However, as can be seen in Figure 2.6, once more than 8 facets are chosen per patch, error greatly reduces the estimation. This validates the assumption made here that a small area on the surface can be estimated with directional light, but when this area becomes too large this assumption fails. Furthermore, when only using a small number of facets, their variation in intensity and normals plays a very important role. For example, it should be possible to select four facets with similar intensities and normals. This case is under-constrained and will lead to an invalid solution. However, this may be avoided when a user is selecting

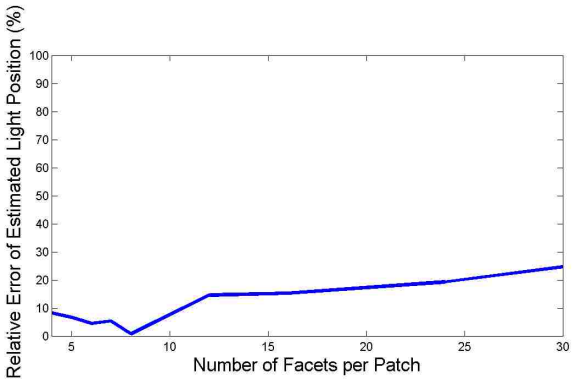


Figure 2.6: Demonstration of the accuracy produced by increasing the total number of facets for every patch.

facets by observation. Also, the RANSAC implementation can avoid these selections after a poor random choice is made.

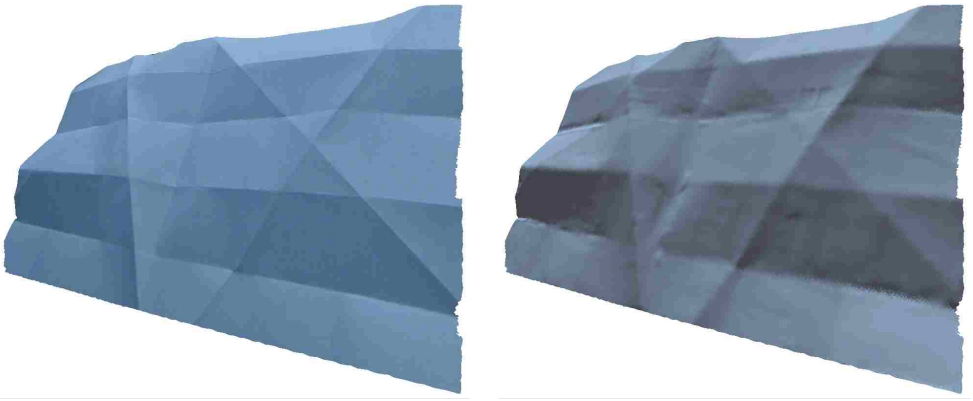
2.3.4 Other Sources of Error

Another source of error can be the location of the patches themselves. Since the algorithm relies on triangulation of directional light rays, a user must keep in mind that patches too close together might produce parallel or near-parallel rays. This issue requires the simple solution of selecting patches with enough relative distance between them to produce accurate intersecting rays.

On another note, there is a different source of error introduced by the texture acquisition. Digital cameras introduce another source of error when imaging a surface. Gaussian noise is an intrinsic property of current image sensor technologies and becomes more apparent when dealing with areas of low illumination in a scene. This can be satisfactorily resolved by averaging the color values across each facet.

2.4 Real-world Analysis

Since performing exact measurements of actual light sources is very difficult, providing numerical analysis of real-world results becomes much harder. However, demonstrating the performance of this light calibration algorithm by visual observation serves as a sufficient example that the results are as expected. In Figure 2.7, the 3D surface of an acquired calibration target is shown. For reference, the acquired texture image, used for calibration, is shown in Figure 2.7(a). The estimated light source position and color is then calculated



(a) The acquired texture image.

(b) The shaded image with light position and color estimated.

Figure 2.7: A comparison between acquired texture image and shaded rendered image with estimated light position.

for this dataset, and the rendered estimation is shown in Figure 2.7(b). As can be seen in Figure 2.7, the shaded surface looks consistent with the acquired texture image.

2.5 Conclusion

The novel light calibration technique presented in this chapter creates a new way for users to estimate the effect that illumination plays in an imaged scene. In particular, the light probe itself is simple to develop and recreate on-site if necessary. This is especially notable when working in conditions where foreign objects are not typically welcome (e.g. Library Archives).

The light calibration method of this work develops an accurate estimation of light position and color. Moreover, the method is presented in an interactive application that gives users real-time feedback showing the accuracy of the current illumination estimations. The results presented here give digitization initiatives new means to estimate illumination properties in the scanning environment without reducing usability for expert library staff.

While this work focused on ease of use for the user, it should be noted that using more than one light source should also be possible. Assuming the same Lambertian properties for the paper light probe it is possible to extend the light source estimation to a non-linear multiple light source system. This is built off of the additive nature of energy when observed by an imaging sensor.

$$f = \Sigma_{j=1}^{n} k_j cos\ \theta \tag{2.11}$$

In the next chapter, the estimated values of this system will be used to correct photometric error introduced when digitizing documents in non-optimal lighting. By calibrating a single point light source, digitization staff will effectively be able to produce the same results as an expert lighting system.

Chapter 3

Photometric Restoration of Documents

3.1 Introduction

Creating standard methods for archival imaging of non-planar documents remains a non-trivial task. Currently, the most widely accepted procedures require using a trained lighting professional to calibrate an advanced lighting system prior to photographing a document. However, these requirements for personnel and equipment are often beyond the budgets of many museums and libraries. Additionally, set up and refinement of equipment in the scene are often tedious and time consuming.

Researchers have worked to the need for professional photographers and expensive lighting equipment using software solutions. These previous techniques try to correct lighting inconsistencies with image-based approaches by estimating variations in shading and attempting to reverse them. However, these methods are limited by the actual content in the document since many portions of a text or graphic may be indistinguishable from lighting error.

The work introduced here provides photometric document restoration in a way that remains ignorant of the content. Using simplistic lighting equipment and a paper light probe, the scene illumination is calibrated using the approach derived in Chapter 2. Using this technology, the shading on the document surface can be directly estimated. Therefore, the adverse imaging effects caused by inadequate illumination of the document can be reversed.

Once the color of the document has been restored, most of the distortion perceivable by the observer has been removed. However, for many applications, such as Optical Character Recognition, the remaining geometric distortions may greatly reduce the detectable content.

3.2 Uniform illumination during document digitization is challenging

Creating an environment with uniform illumination is a difficult–often nearly impossible–task. Even for documents that are nearly flat, but contain small shape distortions, shading

variations are unmistakable for the casual viewer. Refer to Figure 1.3 for an example of subtle photometric distortion.

It must be understood that the intensity variations are caused by the shading of a non-planar surface. Professional photographers may minimize these variations by using a complex network of lighting equipment. Multiple lights must be used to illuminate each angle of a wrinkle or fold. These lights must also eliminate any shadows generated by the surface. Diffusers are typically used to equally distribute the illumination from each light across the document.

While professional photographers normally create an environment that will produce satisfactory results, there are two cases where this is not feasible. First, budget constraints of many museums, libraries, and archives are often limited from accessing the equipment and expertise necessary to generate proper scene lighting. Second, some damaged documents are so severely wrinkled that lighting the scene correctly becomes impossible even with the correct equipment. Both of these cases introduce the need for a virtual solution to document illumination.

3.3 Previous Attempts at Correcting Document Color

The most basic form of photometric correction for documents is binary thresholding. Lu et al. [59] have developed a novel technique to convert many grayscale document images to black and white. However, it is inaccurate to consider thresholding techniques as a form of restoration. Thresholding methods assume that some portion of the image data is noise and remove it from the final "restored" image.

A few works propose methods to address both geometric and photometric correction of documents with constrained geometric types. Zhang et al. [97] use the known shape of an open book on a flatbed scanner to estimate the shading distortion near the book binding. Brown et al. [12] assumes that the boundary of a document is an accurate representation of distortion across the surface. Then the boundaries, with known albedo, are used to calculate and reverse the shading distortion. Zhang et al. [94] use Shape-from-Shading with a Lax-Friedrichs equation to perform photometric correction. The initial step estimates a "shading image" for the underlying distorted substrate and then subtracts the intensity effects from the original image. To handle a larger variety of document forms, Sun et al. [82] presented an algorithm that handles more extensive document types by reducing the shading correction problem to solving a Poisson equation with Dirichlet boundary conditions. This is formulated by relaxing the pixel gradients on the illumination edges that correspond to

edges in the surface depth map. It is developed to handle documents folded with "hard" creases instead of warped with "soft" distortion.

3.4 A New Method for Document Albedo Recovery

The assumption that most documents were originally *flat* underlies this restoration framework. Therefore, digital images/models of documents are usually distorted due to a non-flat(wrinkled) surface. Consequently, non-uniform illumination on these documents creates photometric distortion. To create a procedure for correcting photometric distortion, the way in which light responds to the substrate must be considered.

Many historical documents are printed on early forms of wood pulp paper, papyrus, parchment, or vellum. Over time, most of these substrates have matured in such a way that the surface reflectance properties have been modified. Therefore, instead of the documents producing a glossy finish, which is known as specular reflection, the surfaces produce a dull reflection of light. These diffuse surfaces appear to have the same intensity or shading when viewed from any angle.

This diffuse light reflectance is modeled accurately by the Lambertian reflectance model

$$I_D = I_L cos\,(\theta) \tag{3.1}$$

where I_D is the observed intensity from the illuminated document surface, θ is the angle between the incoming light and the surface normal, and I_L represents the intensity and color of the illuminant. Values of I_L are typically represented in the normalized range of $[0, 1]$, where 1 is the brightest light and 0 emits no energy. Also, since this is a diffuse reflection, there is no need to consider viewing direction. This is because the reflectance only varies directly proportional to the difference in surface and incident light angles and the light intensity.

Therefore, when the light intensity is the maximum, the only factor in reducing the modified intensity of a surface is the relative angle between the incident light rays, $\mathbf{L}$, and the surface, $\mathbf{N}$. Using vector analysis, the value of $cos\,(\theta)$ can be calculated directly for the dot product of these unit normals.

$$I_D = I_L \frac{(\mathbf{L} \cdot \mathbf{N})}{|\mathbf{L}||\mathbf{N}|} \tag{3.2}$$

The observed intensity, I_O, from diffuse reflectance of a point on the surface would then be calculated by $I_O = I_A I_D$ when I_L is 1. Inversely, if the albedo is unknown, it can be calculated as $I_A = \frac{I_O}{I_D}$.

In the case of an extremely distant illuminant and planar document surface, the values of $\mathbf{N}$ and $\mathbf{L}$ remain constant across the entire imaged document. This generates a global scale value for I_D that can be easily adjusted with photo-editing software. However, when a point light source is used to illuminate a document, $\mathbf{L}$ varies for each point on the document surface. Moreover, when the document is non-planar, the surface orientation $\mathbf{N}$ may vary for each point on the document surface. If either a point light source or distorted surface are present in document imaging, obtaining an accurate archival image of a document is difficult. However, when both of these conditions occur simultaneously, a non-linear photometric distortion occurs across the document that is, for all practical purposes, impossible to correct manually.

To further complicate matters, most light sources used to illuminate documents do not produce perfectly white color. Not only does the light source introduce color skew, the imaging device itself will often induce a shift in the color acquisition. Such properties as image temperature and white balance induce added error in the final image. However, if the light source color and intensity (I_L) is also modeled, the following equation can be used for the albedo restoration equation:

$$\begin{bmatrix} \frac{1}{I_{L_R}\frac{(\mathbf{L}\cdot\mathbf{N})}{|\mathbf{L}||\mathbf{N}|}} & 0 & 0 \\ 0 & \frac{1}{I_{L_G}\frac{(\mathbf{L}\cdot\mathbf{N})}{|\mathbf{L}||\mathbf{N}|}} & 0 \\ 0 & 0 & \frac{1}{I_{L_B}\frac{(\mathbf{L}\cdot\mathbf{N})}{|\mathbf{L}||\mathbf{N}|}} \end{bmatrix} \begin{bmatrix} I_{observed_R} \\ I_{observed_G} \\ I_{observed_B} \end{bmatrix} = \begin{bmatrix} I_{albedo_R} \\ I_{albedo_G} \\ I_{albedo_B} \end{bmatrix} \tag{3.3}$$

This work presents a method to automatically correct these distortions using an easy to operate light calibration process introduced in Chapter 2. Once the light source is calibrated and the surface geometry acquired, constraints are added to the previous restoration equation. Since the document geometry was scanned as a step in the document restoration process, the normal, ($\mathbf{N}$) and 3D point ($X_{u,v}$) at each pixel in the image are known. In addition, the light source calibration parameters, S, for the light position in three dimensions, and, I_L, the light source intensity and color are calculated. Then Equation 3.3 may be used to restore the color values for each pixel, $I_{u,v}$ of the image, where $\mathbf{N}_{u,v}$ and I_L are known and $\mathbf{L}_{u,v} = S - X_{u,v}$. The resultant image produced by this process is equivalent to the ideal image of a document with flat surface under optimal lighting.

3.5 Results

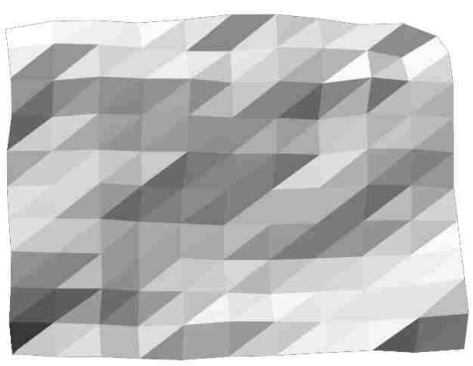
To demonstrate the effectiveness of this photometric restoration approach, restorations were performed on both synthetic and real-world documents. A photometric restoration system

was developed that could input standard 3D meshes and digital texture images from a digital camera. The synthetic document restoration performed with near perfect accuracy compared to the original ground-truth result. The real-world experiments also performed quite satisfactorily. Some loss in quality did occur due to noise in the digitization process.

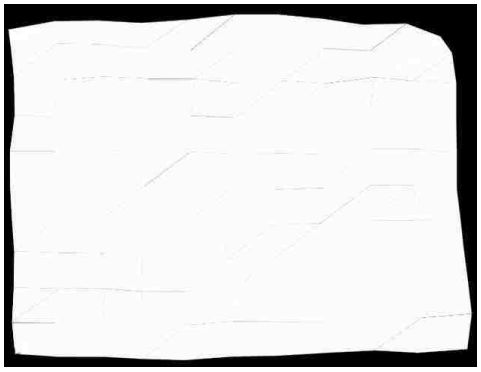
3.5.1 Synthetic Results

Figure 3.1: The synthetic input image of a distorted map document and paper light probe.

To enable proper analysis of the presented photometric correction method, the technique was first performed on a synthetic model with complete ground-truth information known *a priori*. Using Autodesk 3D Studio Max, a virtual scene representing a document digitization environment was generated. The scene included a distorted surface with a document texture map, a distorted white substrate to serve as a light probe, a colored point-light source to illuminate the scene, and a perspective camera. The image of an antique map of New Hampshire was used for the experiment. Surface distortions, mimicking typical

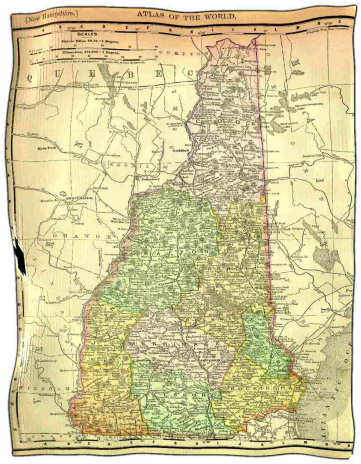


(a) Light Probe



(b) Corrected Light Probe

(c) Original Document



(d) Corrected Document

(e) Ratio Image

Figure 3.2: The results of the synthetic map document scan.

deteriorations of old documents, were added to the planar document using the built-in noise function provided by 3D Studio. The noise was then constrained by the original surface to minimize change in polygon edge lengths during the deformation. This constrained system reduced synthetic stretching or shearing that would not typically occur with most document substrates.

Figure 3.2(a) shows the image of the environment captured by the virtual camera. The combination of the non-planar document surface and point-light source generate numerous shading variations across the document image.

The method described in Chapter 2 is then applied to obtain an estimation for the light position and its color. Using the estimated values, photometric correction was performed that produced a document image free from photometric distortion. Figure 3.2 shows the scanning scene and compares the original distorted document to the corrected result. The document content appears significantly better than the input image: illumination is uniform and albedos are undistorted. In Figure 3.2(d), the corrected result is shown. On the left side of the document, a portion of the surface remains completely black after restoration. This is due to data loss during imaging since the area was under shadow. However, since the surface normal, shape, and light position are known, the system presented here can notify the user of these failures by estimating shadows before restoration is attempted. Users can then take proper action during digitization to eliminate the data loss.

3.5.1.1 Effect of Error on Photometric Correction

The accuracy of the photometric correction system is dependent on the precision of the light source position, light color, and surface normals. To determine the accuracy that is necessary for the light position estimation, a synthetic environment was used with varying position estimations for the light source.

Figure 3.3 shows the pixel-wise error between the ground-truth document image and the photometrically corrected result. The x axis of the plot shows the error of the estimated light position with respect to the overall distance of the light from the document surface. The result demonstrates that when the light source is estimated within 15% of the actual distance relative to the distance of the surface, very low pixel error is generated in the photometric correction.

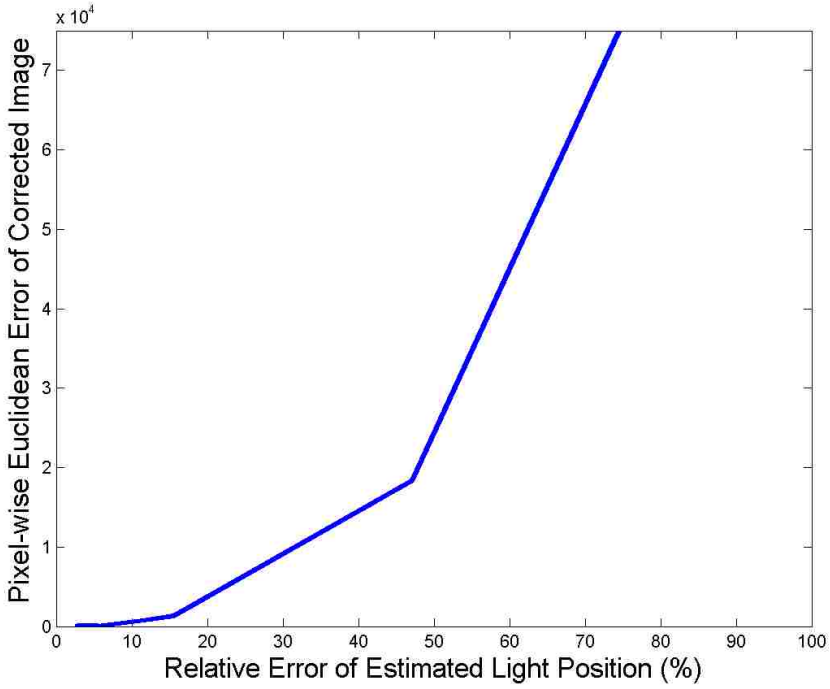


Figure 3.3: Overall photometric correction accuracy dependent on light source estimation.

3.5.2 Real-world Results

In Figure 3.4(a), an image of a wrinkled newspaper is shown in the scanning environment. These non-uniform variations in the surface create a difficult challenge when trying to generate constant surface illumination. The result of this initiative's photometric correction is shown in Figure 3.4(c). To compare the original and corrected result, an image ratio is calculated of the two images using $\frac{I_{observed}}{I_{albedo}}$. Figure 3.4(d) shows the ratio image where lighter intensity values represent areas of greater shading variation.

In Figure 3.6, the image of a heavily worn book cover is shown. This surface is particularly interesting because of water-damage and folding that have occurred to the surface. The result of the photometric correction is shown in Figure 3.6(e). The ratio image in Figure 3.6(f) shows that areas near folds require the greatest correction. However, it should be noted that the 3D scanner used did not have high enough resolution to model all of the surface variations. Therefore, some of the dimples and folds are not completely corrected. However, this is simply a limitation of the scanning device and not a failure of the proposed method.

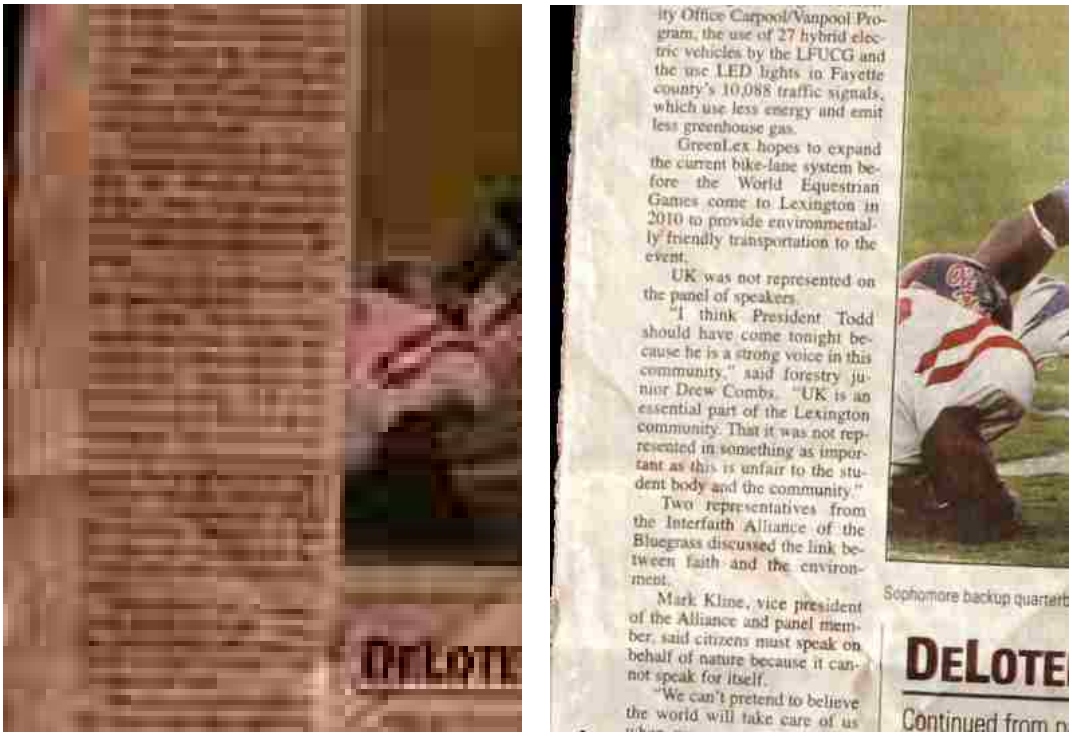
(a) Newspaper in scanning environment

(b) Original Document

(c) Corrected Result

(d) Ratio Image

Figure 3.4: Portion of a deteriorated newspaper clipping.

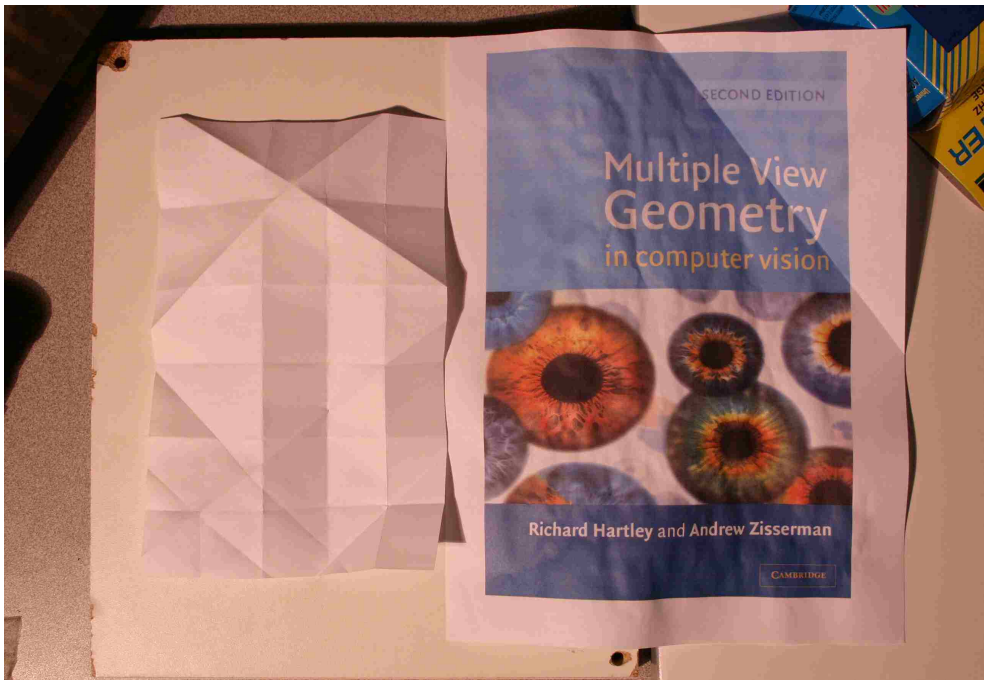


(a) Portion of the original newspaper document (b) Portion of the corrected newspaper document

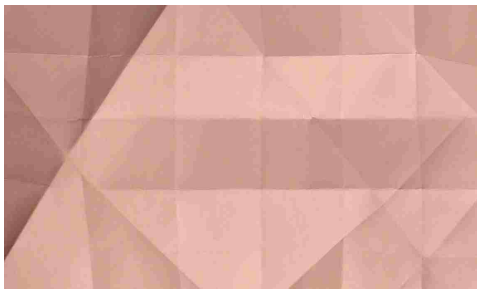
Figure 3.5: A closer view of the newspaper result in Figure 3.4.

3.6 Conclusion

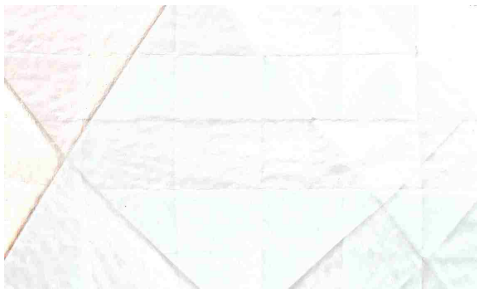
Photometric distortion is one of the most perceivable distortions. The human eye is well suited at discerning small variations in shading, so non-uniform illumination across a document is easily detectable. Once the color of the document has been restored, most of the distortion perceivable by the observer has been removed. However, for many applications, such as Optical Character Recognition, the remaining geometric distortions may greatly reduce the detectable content. Therefore, complete document restoration is performed by augmenting both the shading and shape of a document. In the next chapter, methods for performing virtual geometric correction will be reviewed.



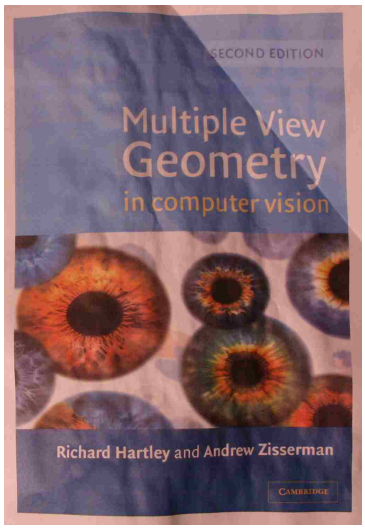
(a) Book cover in the scanning environment



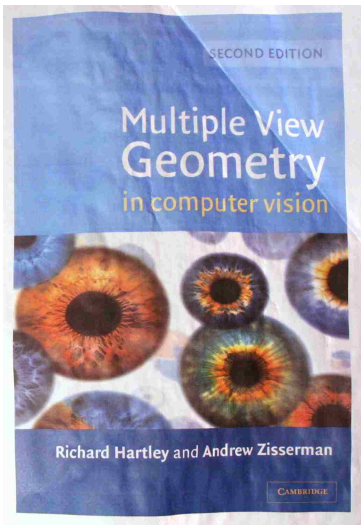
(b) Original Light Probe



(c) Corrected Light Probe



(d) Original Document



(e) Corrected Result

(f) Ratio Image

Figure 3.6: Example of a well-used book cover with water damage and folding.

Chapter 4

Geometric Restoration of Documents

4.1 Introduction

When imaging non-planar documents, distortions in the surface cause shearing and compression of the content in the image. While performing physical flattening on the document works in many cases, this approach is unacceptable for historical documents that may be destroyed if manually manipulated.

Software-based geometric restoration processes have been developed using the observed distortion in the document content. Other processes rely on the known borders of a document to estimate the overall document distortion and reverse the effects. However, both of these methods assume that the distortions will be readily differentiable from the content—an assumption that fails for a large group of documents.

These content-based constraints can be removed by using the complete three-dimensional geometry of the document. Model-based geometric restoration can be performed using physical simulation, non-linear estimation, or iterative methods.

After detailing the benefits of these techniques, the limitations of these methods will be addressed. Particularly, cases when the geometry and composition of the document surface is complex enough to prohibit accurate three dimensional reconstruction.

Geometric correction of distorted documents is typically performed in two steps:

1. Distortion Estimation

2. Distortion Correction

Therefore, most geometric restoration techniques work as a multi-procedure process in line with the two steps previously listed. Since distortion estimation lays the framework for the restoration process, the accuracy produced by the estimation is crucial for the correction to be successful.

4.2 Surface Distortion Estimation

To re-iterate the most important aspect of geometric restoration is developing an accurate estimation of what distortions have occurred in the surface. These estimations range a great deal in form from one project to the next. Document distortion correction techniques have evolved from 2D image processing [16] to 3D model manipulation [90, 97, 9, 12, 82].

Due to the fact that the surface shape distortion induces content distortion and shading irregularities in the captured image, techniques have been developed to tackle geometric rectification. A category of the work assumes parametric models, such as cylinders, for the surface shape [90, 97, 12] while more recently a technique based on physically-based simulation [9] has emerged to flatten arbitrarily warped and crinkled documents.

Geometric document restoration techniques have used various forms of computer vision and image processing techniques to estimate the imaged distortion. These technologies extend in two directions:

- Image-based using a single image of a document
- Metric reconstruction based using advance 3D scanning equipment

Between these two extremes lie numerous other methods that have provided varying results for surface distortion estimation. These techniques will now be reviewed to address the current state-of-art for surface distortion estimation.

4.2.1 Feature-based Estimation

Image-based distortion estimation techniques have been developed to exploit the information provided by a single distorted document image. For modern type-set documents, it can be safely assumed that the lines of text are printed in a uniform manner. Specifically, the vertical distance between textual lines is constant and the lines remain parallel throughout the width of a page.

Cao et al. [14] made use of this assumption to estimate the distortion of bound books. Using feature detection methods, two-dimensional splines are fit to representative lines in the document page. Then, assuming that the vertical distortion is consistent on the page, the conclusion is made that these 2D splines are disparate views of the same 3D spline. Consequently, 3D reconstruction can be performed to estimate the shape of each page. This technique is extremely efficient but limited by the consistency of the document content.

Liang et al. [57] extend feature, or texture, based distortion estimation. Instead of relying solely on the distortion in the lines of text, distortion of the vertical characters is

also exploited. This *texture flow* provides an estimation of distortion in both horizontal and vertical directions. Then a developable surface, that maps directly to a plane, is fit to the textual distortions in an equidistant manner. However, forcing a surface to fit this constraint limits the accuracy of the estimation.

Feature detection is also used heavily in stereo reconstruction methods. However, these will be discussed in section 4.2.4.1.

4.2.2 Boundary-based Estimation

Other distortion estimation techniques attempt to simplify the problem by using only the document boundary as a distortion cue.

Kashimura et al [48] use the top and bottom imaged boundaries to estimate the overall page distortion. Using a calibrated camera, the boundary curves are compared directly to a planar rectangle, the assumed original document shape, to calculate depth variations of a curved page. Since there is assumed to be no shape variation in the vertical direction, the authors successfully demonstrate their technique with images of only the top or bottom portion of a document. While this technique is quite adequate for many bound books, it is unable to handle most other documents, including bound texts with non-uniform distortion from deterioration or damage.

Brown et al. [9, 12] use splines to model the curves of the imaged document boundaries. Tsoi et al. [88] extend the work by using stereo images of the scene. By fitting two-dimensional curves to the document boundary, the imaged distortions may be modeled. The presented result efficiently models the global document distortions such as vertical or horizontal folds as found in a book binding. However, boundary methods are ill-equipped to model localized distortions such as wrinkles or dimples in the surface.

Gumerov et al. [36] use the notion of applicable surfaces to estimate document shape. These reconstructed shapes map to a plane by definition. The synthetic results show promise, however, the system requires extremely accurate boundary detection which may limit real-world applications.

In general, boundary-based estimation limits the amount of distortions that may be modeled. For example, diagonal folds and wrinkling are non-global deteriorations that cannot be estimated by boundary curves.

4.2.3 Shape-from-Shading Reconstruction

One of the foundational initiatives for correcting digitized document images was presented and then extended by Wada et al. [90]. The authors developed a model to handle interreflections between opposing book pages. However, this added further complexity to the reconstruction process. By only scanning bound texts in a flatbed scanner or copying machine, the possible shape constructs were limited to a cylinder.

Tan et al. [85] augment the scanning process of a flatbed scanner to estimate the curvature near the binding of a book. Using the rigid construction of the illuminant and imager in a flatbed scanning device provides a direct way to apply the standard shape-from-shading model:

$$I_o = I_i k \cos\theta \tag{4.1}$$

F. Courteille et al. [19] extends the use of shape-from-shading and the assumption that bound books conform to two parallel cylindrical shapes to generate a distortion estimation. Each of these methods provide reasonable estimations for bound texts when digitized with a flat-bed scanner or copier. However, the basic shape assumptions that are made greatly limit the type of documents that are compatible with these methods.

Prados et al. [69] use a novel SfS formulation to formulate a more general distortion estimation. A point light source is placed near the optical center of the imaging device to create a "well-posed" Shape-from-Shading technique with a single image. The result of this technique is a surface reconstruction of non-planar document based on the document substrate albedo. Like other SfS techniques, an assumption must be made that the document surface contains some area of constant color across the entire surface. Consequently, large amounts of variance on the substrate caused by ink or other foreign materials hamper the results of these SfS estimation met hods.

4.2.4 Photometric Stereo Reconstruction

Extending from Shape-from-Shading research, photometric stereo reconstruction uses at least two scene images, each with a differing illumination, to estimate surface shape.

Cho et al [17] use photometric stereo to estimate the shape of bound text. The document substrate is assumed to have a perfectly white albedo and follow the Lambertian reflectance model. These assumptions require the extraction of a "pure shade image" where the text has been removed through interpolation of the background intensities using histogram analysis. The shade images are estimated for two illumination sources, so that the surface depth can

be directly estimated from the classic photometric stereo model:

$$M = \frac{[L_1 L_2 L_3]^T}{[I_1 I_2 I_3]^T}, \tag{4.2}$$

developed by Horn [43].

4.2.4.1 Stereo Reconstruction

Yamashita et al. [93] and more recently [83] have created a stereo camera system for acquiring depth images of documents. Using two calibrated cameras, three-dimensional points can be reconstructed as a point cloud. Unfortunately, the reconstructions are limited to bound texts like the standard Shape-from-Shading methods previously discussed.

Video mosaicing has also been used to estimate the distortion of documents. Iketani et al. [46] used a portable computer with webcam to obtain multiple low-resolution images of a document. Then, using feature detection, the images are registered and combined to generate a high-resolution document with a shape estimate. This method can reconstruct most document shapes, but it is limited by the accuracy of the feature detection. Moreover, for some documents, features are inconsistent across the substrate, so complete reconstruction will be limited.

4.2.5 Structured-light Reconstruction

Active three-dimensional imaging systems have been used from the beginning of virtual document restoration.

Doncescu et al. [23] utilize a laser stripe to reconstruct the three-dimensional shape of the document surface. Using a fixed incident angle of 45°, the projected laser stripes can be directly used to estimate the document shape. The authors use wavelet analysis to improve the detection of the reconstruction by limiting high-frequency noise. However, this assumption would only work for preserved pages, since high-frequency wrinkles and folds would be truncated.

Pilu [67] use an array of projected laser stripes to triangulate 3D points. These stripes are projected onto the book, orthogonally to the spline, because the shape variation is assumed to be constant along the spline. Then an iterative fitting routine is used to model the reconstructed 3D points as an applicable spline-based surface. However, distortions that occur in a perpendicular direction with the binding will not be accurately modeled, so many distortions will not be digitized.

Brown et al. [9] use a fixed video projector to display patterns of light on the document surface. The system is strongly calibrated as a set of two perspective cameras so that

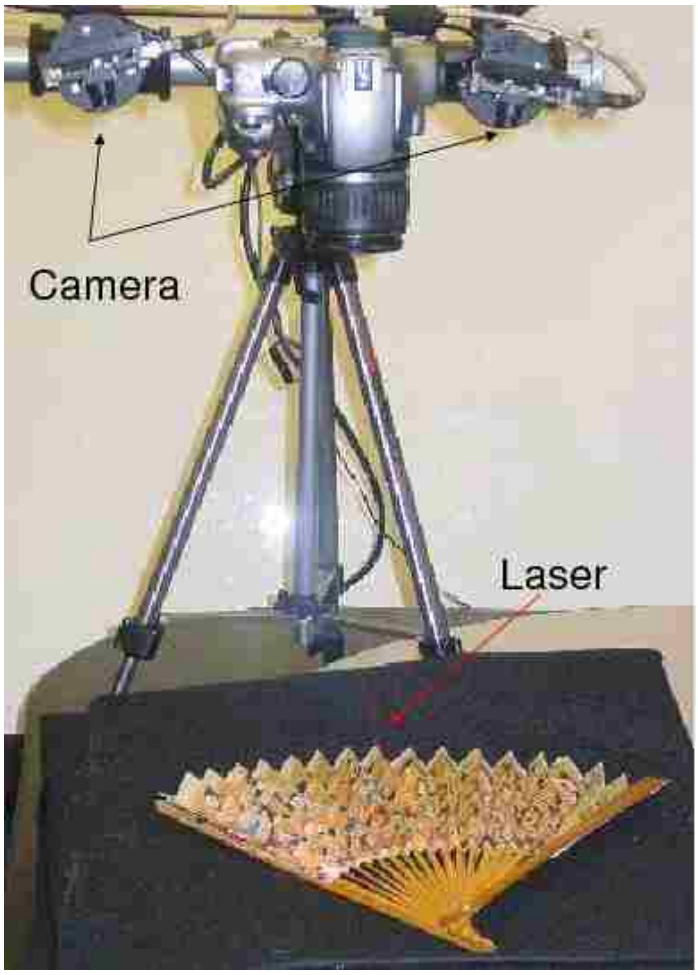


Figure 4.1: A structured-light scanner used for document acquisition.

illuminated projector pixels are detected by the camera during the scanning process. This correlation is then used to directly triangulate the 3D point. The precision of this system is intrinsically related to the lowest resolution of either the projector or camera.

Sun et al. [82] used the scanning technology previously applied in Chapters 2 and 3. A laser stripe is swept across the document surface using a computer controlled pan unit while two strongly calibrated cameras observe the stripe profile and triangulate its position as three-dimensional points. A strongly calibrated texture camera then acquires the color information on the surface. Originally developed to scan rock art [55], this type of structured-light scanner is well suited for document scanning. Figure 4.1 shows a typical document digitization being performed with the scanner. However, as with all types of active scanning systems, variations in surface albedo induce error in the estimated surface.

4.3 Distortion Correction

When the shape of the document substrate has been accurately modeled, surfaces can then be manipulated to conform to the original planar state. The wide range of shape distortion estimation methods available also bring many disparate techniques for restoring document shape. Some of the restoration methods are developed to directly provide only the information necessary for restoration. However, other methods generate complete models of the surface, allowing a variety of restoration methods to be applied.

4.3.1 Cross-Section Document Restoration

For many restoration procedures, the documents are assumed to fit a certain class of texts, for instance bound books. The works that make this assumption [90, 85] reduce the problem from a three-dimensional deformation calculation to a two-dimensional homography mapping.

By fitting a series of equidistant curves orthogonal to the text binding, a direct mapping can be calculated between the curve and a restored plane using similar triangles. Therefore, areas of the document with severe distortion, typically near the binding, are stretched to compensate for the distortion induced by projection.

By calculating a restoration solution using this method, the un-warping of the image reduces to a two-dimensional image warp based on a series of control points. Therefore, the solution is direct and simplistic to calculate. However, the geometric restoration is limited to only a subset of documents, eliminating non-uniformly deteriorated books and single sheet manuscripts.

4.3.2 Boundary-based Restoration

When distortion is modeled as boundary curves, a direct planar mapping may be formulated. Brown et al. [9] use arc-length parametrized splines to model boundary curves. Then a direct mapping is calculated between the spline control points on the curve and control points on a pre-flattened curve.

Tsoi et al. [88] extended the boundary curve-based restoration by using multiple views. Including more than one view of the document improves the overall restoration since super-resolution techniques may be used. These restoration methods are extremely fast, however, only documents and distortion modeled by the boundary are available for this form of restoration.

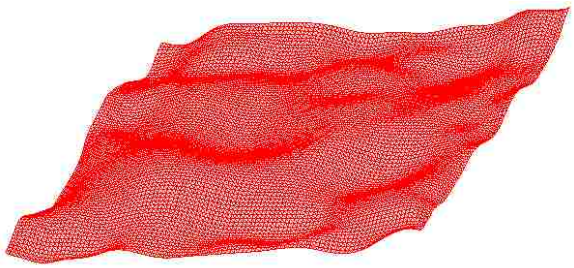


Figure 4.2: A three-dimensional mesh acquired with a structured-light scanner.

4.3.3 Three-dimensional Restoration

Many techniques have opted to acquire as much surface detail as possible by using three-dimensional acquisition. The most full-featured restoration processes to date use complete three-dimensional models of the document surface. Typically generated as polygonal meshes, these models can capture many forms of deterioration that cannot be handled by the previously mentioned techniques. For example, folds, wrinkles, and dimples are all blemishes that may appear non-uniformly on the surface, so parametric representations are unable to model them. A typical polygon mesh reconstructed from a document surface is shown in Figure 4.2.

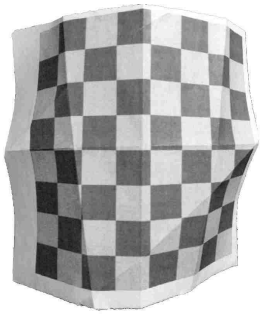
4.3.3.1 Least-squares Conformal Mapping

Brown and Pisula [10] use conformal mapping as a document parametrization method. The 3D shape is mapped to a 2D plane while holding to the conformal constraint. The restorative mapping is calculated by solving a large linear system. By using a linear system, the restoration remains stable unlike many previous relaxation-based techniques. The results of this restoration are effected by the choice of control points. Consequently, a poor choice will cause a poor restoration.

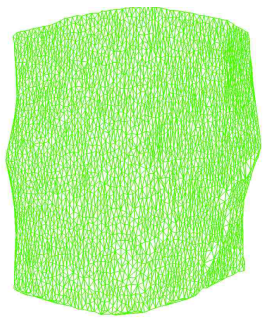
4.3.3.2 Physically-based Simulation

Brown et al. [11] originally introduced the use of particle simulations to restore document shapes. Initially used as a real-time cloth simulation for graphical rendering, this system adequately models other materials well such as paper and papyrus. The work has been extended for more document types by Seales et al. [77]. The basic simulation was also improved by Zhang et al [95].

The simulation system works as an interactive refinement process based on a set of forces applied to a polygonal mesh representing the document surface. Recall the standard

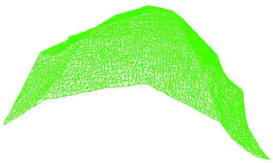


(a) Original document



(b) Original document mesh

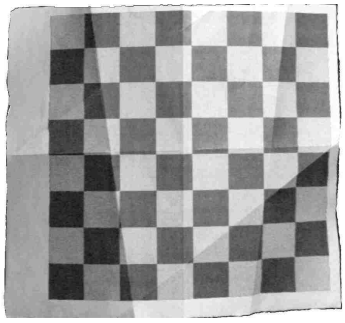
(c) Original document side-view



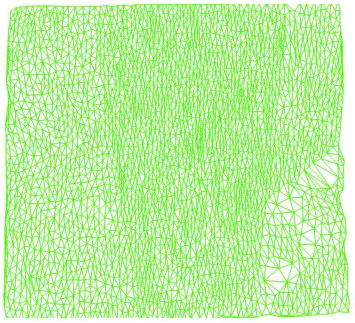
(d) Original document mesh side-view

(e) Flattened document side-view

(f) Flattened document mesh side-view



(g) Flattened document



(h) Flattened document mesh

Figure 4.3: An example virtual flattening procedure.

Newtonian physics model for forces: $F_{int} + F_{ext} = ma$. Here two forces are applied to the mesh: forces from external sources, F_{ext}, and forces from internal sources, F_{int}.

The internal forces are developed as a system of springs. For each vertex in the three-dimensional surface, a set of springs are used to restrict its motion relative to the neighboring vertices. Therefore, F_{int} is modeled as a combination of two forces: f_a and f_b where

$$f_a = \left[k_s(|\vec{x_a} - \vec{x_b}| - r) + k_d \frac{\vec{v_a} - \vec{v_b}\dot{\vec{x}}_a - \vec{x_b}}{|\vec{x_a} - \vec{x_b}|} \right] \frac{\vec{x_a} - \vec{x_b}}{|\vec{x_a} - \vec{x_b}|} \tag{4.3}$$

$$f_b = -f_a \tag{4.4}$$

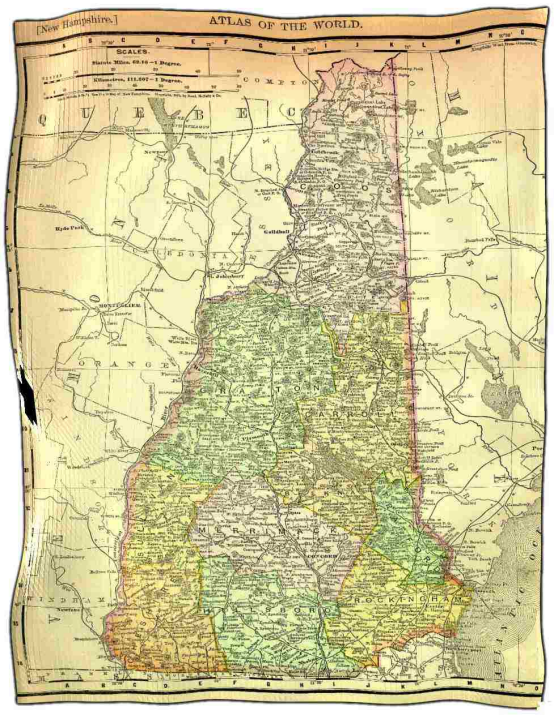
The external forces, modeled asF_{ext} come from two sources. First, a gravitational force, $f = mg$, moves the surface toward the ground plane. This creates a virtual flatten simulation as the document begins to interact with the ground. Consequently, once the surface collides with the ground plane, a drag force, $f = -\mu_d v$, is added to the external force calculation. The drag force models the collision velocity state, v, and separates it into two components: normal velocity $v_n = (N\dot{v})N$ and tangential velocity $v_t = v - v_n$.

These two forces allow a next step to be calculated for each iteration in the flattening simulation. Figure 4.3 shows the geometric restoration process and result. However, it should be noted that the final checkerboard pattern, Figure 4.3(g), still contains photometric errors from non-uniform illumination.

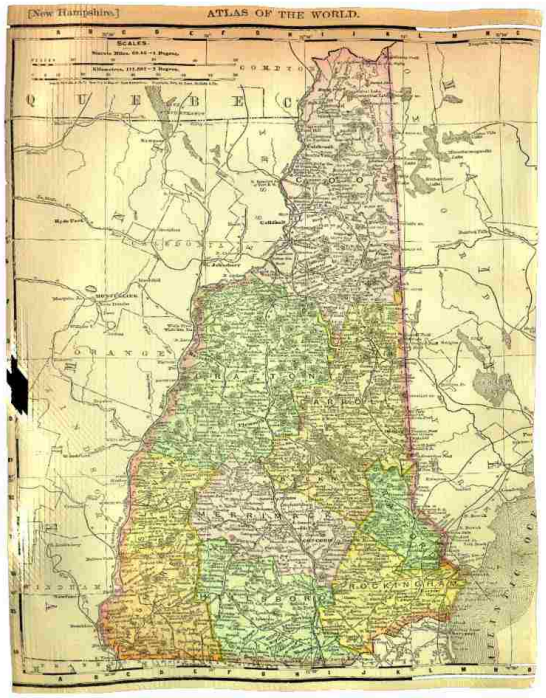
Therefore an implementation of the flattening simulation was used, in conjunction with photometric correction, in this work to complete the restoration process. Consequently, the photometric correction performed in Chapter 3 is then corrected for any geometric distortions. Figure 4.4 show the continuation of the restoration started in Figure 3.2. The result shows that most lines are now straightened effectively removing geometric error. Figure 4.5 shows the correction of the digitized newspaper from Figure 3.4.

4.4 Conclusion

There currently exist a large number of techniques to estimate document distortion and subsequently correct the distortion. Some of these methods make assumptions concerning the composition and layout of a document to achieve faster acquisition and restoration. Other methods perform complete three-dimensional surface reconstructions to attempt to model all forms of surface deviations. Then complex physically-based systems are used to return the documents to their original planar state.



(a) The result of photometric correction presented in Chapter 3.



(b) The final output with photometric and geometric distortion removal.

Figure 4.4: Geometric correction of synthetic result from Chapter 3.

(a) The result of photometric correction presented in Chapter 3.

(b) The final output showing geometric and photometric distortion removal.

Figure 4.5: Geometric correction of synthetic result from Chapter 3.

However, for many forms of deterioration, achieving the accuracy needed to complete restoration is quite difficult. In the next chapter, a new acquisition system is introduced that performs image-based distortion estimation.

Chapter 5

Scanning Documents using Transmissivity

5.1 Introduction

For some document types, full three-dimensional reconstruction is either unnecessary or impractical when attempting digitization and restoration. Many historical documents contain wrinkles, creases, and other high-frequency features either beyond the accuracy of many 3D scanners or require time intensive acquisition procedures. In these cases, a more appropriate approach is to work in a pixel-by-pixel image-based acquisition and restoration methodology. This work develops that methodology by assembling a cost-effective scanning system comprised of a laptop to emulate a *smart* light-table, and a camera to observe illumination changes in the scene.

The technique presented here exploits the transmissivness of documents to obtain a model of the document that allows complete reconstruction of the intrinsic color, content, and distorted shape. Also, to reduce the burden of the system operator, there is no calibration step required before scanning can begin unlike many other document digitization systems mentioned in Chapters 3 and 4.

The transmissive document scanner is designed to accurately digitize and even restore content that is marred by damage and age. The photometrically corrected content is extracted directly during the scanning process while working in a completely image-based realm. Moreover, the obtained shape information can be used to restore the shape of a geometrically warped surface with restoration procedures described in Chapter 4. Consequently, image-based techniques provide a direct way to generate restored images without requiring metric reconstructions that add the overall system complexity.

Exploiting the transmissive nature of most document substrates, this initiative capitalizes on rear-illuminating the document with structured patterns of light while observing the transmitted light with a digital camera. The light patterns are composed of Gaussian-profile stripes in both vertical and horizontal orientations. The distortions of the light stripes are then estimated with a non-linear system and the estimated parameters directly

provide intrinsic content and shape information. These parameters give a basis to begin the photometric and geometric corrections required to restore a document back to its original form.

Another discovered feature of scanning with document transmissivity is the possibility of one-pass duplex-sided scanning without show-through noise. When the opacity of a document is low enough, content from the back surface will often be visible when scanning. However, instead of treating show-through content as noise the data is treated as accurate information from the reverse-side. As a result, a method for acquiring content from both sides of the document in a single scan is presented.

After demonstrating the scanner on both one and two-sided documents, application of this system can be found with other non-traditional document types as well. In fact, even multi-layered documents that remain semi-transmissive can be restored with the developed scanner, specifically non-uniformly warped multi-layer film negatives. The multiple layers of these films are separating, causing both photometric and geometric distortions when re-printing is attempted by normal means. Chapter 6 provides a new mew method for digitizing these negatives that produces unparalleled results.

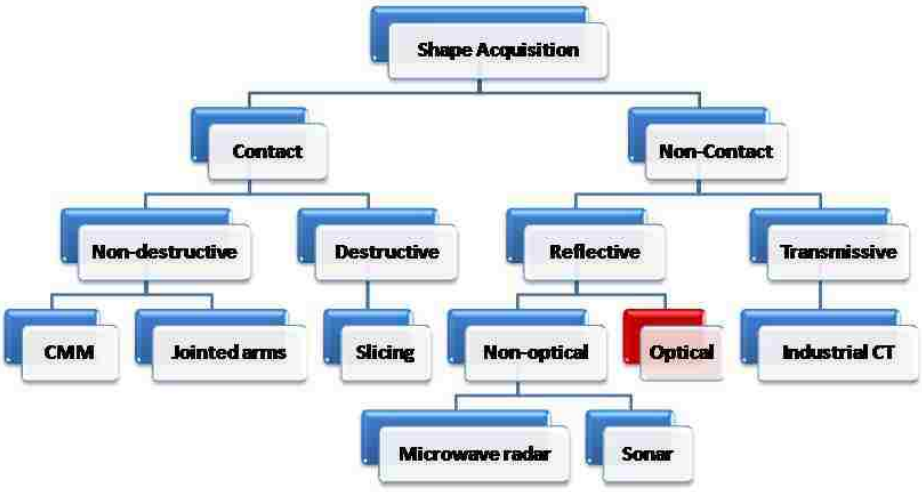
5.2 Previous Contributions

Many forms of shape acquisition have been developed in recent decades. These methods are represented as a taxonomy first organized by Curless et al. [21]. The majority of reconstruction techniques in Computer Vision focus on optical methods based on reflective properties of objects as seen in Figure 5.1(a).

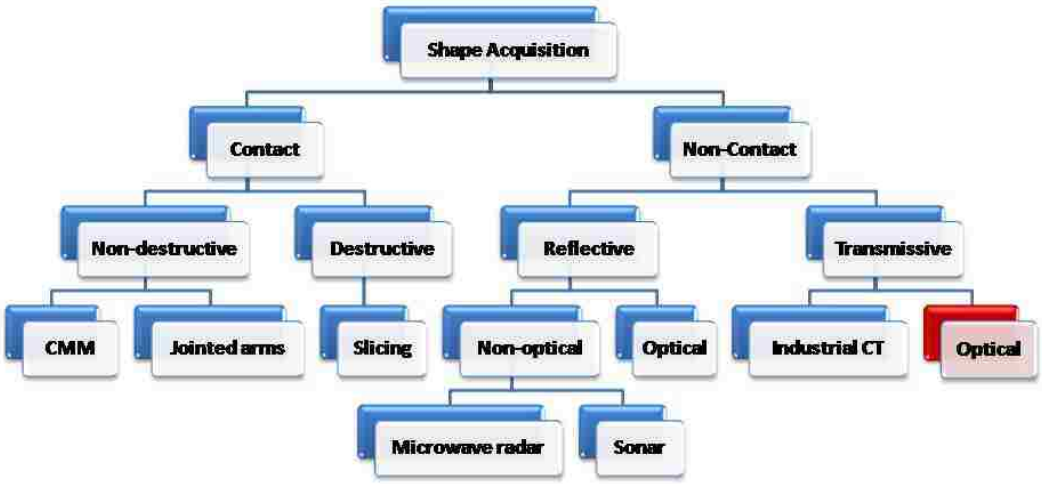
These techniques have been used extensively in the area of document shape reconstruction. As described in Chapter 4, Shape-from-Shading, Photometric Stereo, Stereo Reconstruction, and Structured-light Scanning, all fall into the area of reflective-based reconstruction. Some works have focused mainly on geometric correction of distorted documents [95, 9]. While others have relied on assumed document shapes to provide photometric and geometric corrections for objects, such as bound books [14] and folded documents [12]. Scanning techniques relying on reflective light produce amiable results for many document types. However, there are cases when a great deal of information is lost in the observed image. For instance, when a document is deteriorated and misshapen the document image will be heavily shaded under anything but perfect lighting conditions. When imaging such documents, a solution provided here to correct the photometric distortions is described in chapter 3. In retrospect, it would be more productive to

produce an imaging technique that prevents photometric distortions by design. Also, many unprintable document types are incompatible with standard imagers because of their composition. For instance, documents with more than one layer of substrate greatly reduce the effectiveness of reflective light.

In cases where reflective light acquisition is unsuitable, this work develops a transmissive light scanner. Therefore, the shape acquisition taxonomy must be updated to include new approaches working with transmissive properties as seen in Figure 5.1(b). This work



(a) The original Shape Acquisition Taxonomy introduced by Curless et al. [21].



(b) The updated Shape Acquisition Taxonomy that includes shape acquisition using transmissive properties by optical means.

Figure 5.1: The taxonomy of Shape Acquisition.

provides a novel leap forward by treating many objects as optically transmissive that are typically regarded as opaque.

Raster displays have been used for standard reflective imaging techniques. Schechner et al. [76] and Funk et al. [30] both use computer monitors as controllable light sources. However, their work focused on the reflective light from a surface.

Sen et al. [78] also used raster displays to perform dual photography that is non limited to reflective imaging. The authors defined a process where every pixel in the raster display could be inversely mapped to an imaging device which allows a virtual image to be generated from the display position. While their work is compatible with transparent surfaces, more extensive research exists to acquire the shape and optical properties of general transparent objects. Zonger et al. [99] and Chuang et al. [18] have introduced a novel method for capturing light transport through a scene. By observing the optical changes produced by transparent objects such as drinking glasses and sculptures, new environment mattes may be placed in the scene virtually. Others have focused on estimating the shape of the transparent objects. Ben-Ezra et al. [6] use assumptions about the object shape and the motion of either the object or the background to reconstruct surfaces of transparent objects. The objects that are used are all parametrically modeled to aid in the estimation of the entire object such as a sphere, cube, and cylinder. Morris et al. [64] introduce a technique to acquire the outermost surface of complex transparent objects. Using scatter-trace stereo, a camera observes a scene illuminated by a point-light source. As the light source moves, constraints are added allowing the surface reflectance properties to be calculated for various semi-transparent objects composed of inhomogeneous materials. Less general transparent surfaces have also been calculated. Morris et al. [63] developed a novel method for dynamically calculating the surface normal of rippling water. Their work used two cameras observing a known-calibration target through a known medium such as water.

While previous works have typically focused on general shapes such as glass spheres, wine glasses, or sculptures, other groups have developed techniques for objects in similar forms to documents. For the purpose of rendering plant leaves in real-time, Wang et al [91] develop a system to acquire the reflective and transmissive properties of leaves. Even though the geometric shape was not a focus, the project did produce realistic results for rendering the the various attributes of a leaves including rear-illumination. For documents in particular, Gardner et al. [32] introduced the use of alpha values for documents as a transparency measurement. While the shape of the document was obtained with a reflective light acquisition system, the geometry is augmented with a transmissivity parameter. The

acquisition was a direct calculation of the reduced intensity of an image pixel when a document is placed on a light table. The goal of this work was to improve photo-realistic rendering, and the alpha value allowed realistic rendering with rear-illuminated virtual light sources. However, this project was focused in rendering a document in its current state without mention of restoration.

Other non-optical transmissive scanning techniques have been developed for document scanning. Research by Seales et al. [77] uses x-ray technology to penetrate multiple layers of documents to obtain the content from each individual layer. The scanning technology used by the authors falls into the *Industrial CT* area shown in the shape acquisition taxonomy. However, this advanced technology is limited by its cost and accessibility.

Another issue in standard document scanning is the show-through effect of duplex-printed documents where content from the reverse side is visible. This common issue becomes detrimental to a scan when the content from the reverse side hinders the legibility of the scanned content. Work to develop Image Processing techniques to remove these errors treat the show-though content as noise [79, 87, 66]. Also, document processing corporations, Agfa Corporation [37, 38], Microtek International, Inc. [39, 74], Umax Data Systems Inc. [45], and Xerox Corporation [2], have also focused in improving machinery that will limit the show-through when scanning one side of a document. In contrast, this work demonstrates the ability to scan both sides of a duplex-printed document with a single scanning pass. Moreover, to accomplish the scan of double-sided documents, the show-through data must be treated as accurate data instead of added noise. By using the content of both front and rear sides the show-through effects can be reduced in the resultant scan of both sides.

5.2.1 Physical Model

The solution presented here works on the premise that the composition of most document substrates are composed of numerous non-homogeneous elements. In the case of modern paper, chemical processes are used to render wood pulp to the necessary composition suitable for creating a substrate with the correct balance of rigidity and transmissivity. Consequently, the typical composition of document substrates follow a highly isotropic scattering of transmitted light. As Horn [43] claims, paper can be accurately modeled as a diffuse reflector and consequently a diffuse transmitter of light.

Using the assumption that documents are very non-homogeneous in composition, the scanning method presented here will focus on diffuse transmission. For a single-layer

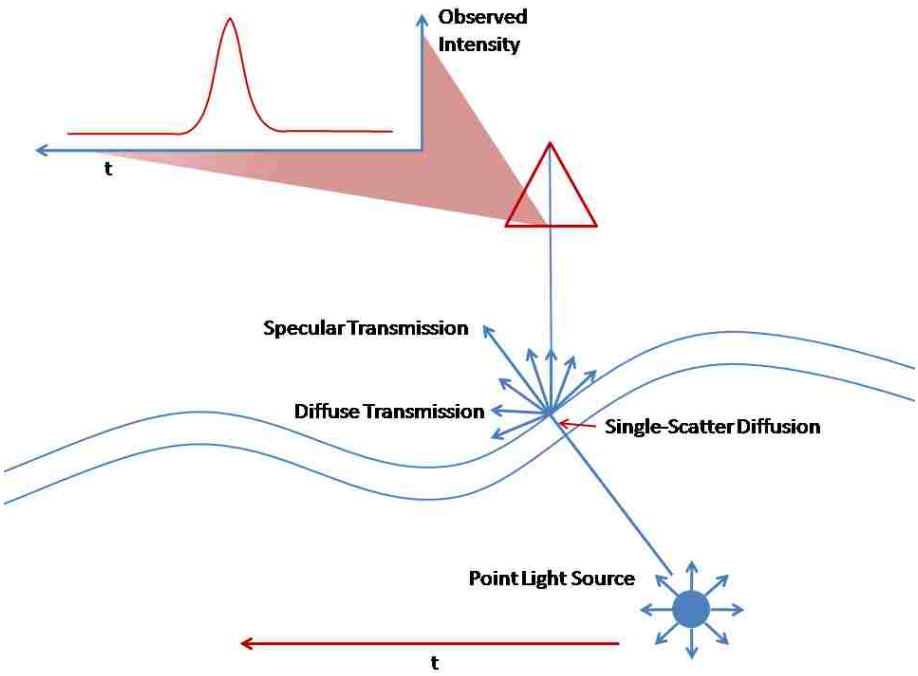


Figure 5.2: Diffuse single-scatter transmission of a back-lit light source.

document, the diffuse transmission of light can be approximated as a single-scatter diffusion. Chandresakar [15] provides an approximation of the single-scattering that occurs in diffuse transmission as:

$$I = I_0 e^{\frac{-\tau}{\mu}} + \frac{1}{4\pi}\omega_0 I_0 \frac{\mu_0}{\mu + \mu_0}\left(e^{\frac{-\tau_1}{\mu}} - e^{\frac{-\tau_1}{\mu_0}}\right) \tag{5.1}$$

where μ_0 is the incoming angle, μ is the outgoing angle, ω_0 is the phase function, I_0 is the incoming intensity, and τ_1 is the material thickness.

The single-scatter transmission has been well studied in the area of computer graphics. Frisvad et al. [27] use Chandrasekhar's work to create an efficient rendering system for thin semi-transparent objects. Moreover, the area of plant/leaf rendering has been thoroughly studied [91, 35, 31, 26] with respect to single-scatter transmission.

Figure 5.2 shows a particular case where a light source is being translated approximately parallel to semi-planar object. For a single-pixel observation, as the light translates, the intensity follows a cosine-like response where the light is incident at an angle parallel to the surface normal. The incoming illumination angle can be calculated as $\mu_0 = l_0 \cdot -n$. Moreover, in the case of diffuse transmission, an assumption can be made that the greatest transmitted intensity will occur when $l_0 \simeq n$. Therefore, for the purpose of this work, μ_0 will be approximated as 1.

For diffuse materials, especially paper, it is safe to model the material as a translucent material with a highly diffuse transmission of light. Therefore, the phase function, can be modeled with isotropic scattering, thus ω_0 becomes a constant 1. The direct transmission can be safely ignored for highly diffuse materials, so single-scattering becomes the only factor in light transmission through the material. Considering the assumptions equation 5.1 can be approximated as:

$$I = \frac{1}{4\pi} I_0 \frac{1}{\mu + 1} \left(e^{\frac{-\tau_1}{\mu}} - e^{-\tau_1} \right) \tag{5.2}$$

where μ is the only varying quantity across the surface while τ_1 and I_0 stay constant.

As will be discussed later in this chapter, the intensity increase generated when the incident light angle, l_0, becomes parallel with the surface normal n will be used to estimate the shape of the document surface. However, for documents that are mostly specular transmissive, directly transmit light, surface normal no longer plays a large role in the intensity of transmitted light. Therefore, the following scanning process only works well for documents that exhibit diffuse transmission.

5.3 Image-Based Document Scanner

By exploiting the transmissive nature of most document materials, the new image-based acquisition technique presented here provides the direct ability to digitize and restore many document types including single-page manuscripts and multi-layer photographic negatives. The design of this scanner hinges on the premise that all necessary information in a document can be obtained through rear-illumination of the substrate with visible light. While there are some documents where this assumption fails, such as with totally opaque materials, the vast majority of documents have some form of translucency.

This scanner enables digitization initiatives of every size to increase the number of documents that may be recorded. Therefore, keeping the technical knowledge of the operator to a minimum is a great advantage, and the scanner itself must be easy to access and engage. Ease of deployment facilitates this accessibility by requiring minimal setup before performing scans.

Additionally, the system requires no calibration in the scanning procedure. Many document digitization systems already mentioned in this work require calibration of both the imaging device(s) and illumination source(s). However, this adds to the complexity of operation and may reduce the number of workers capable of performing a scan. The scanner presented here works in a completely image-based domain, with operations performed on

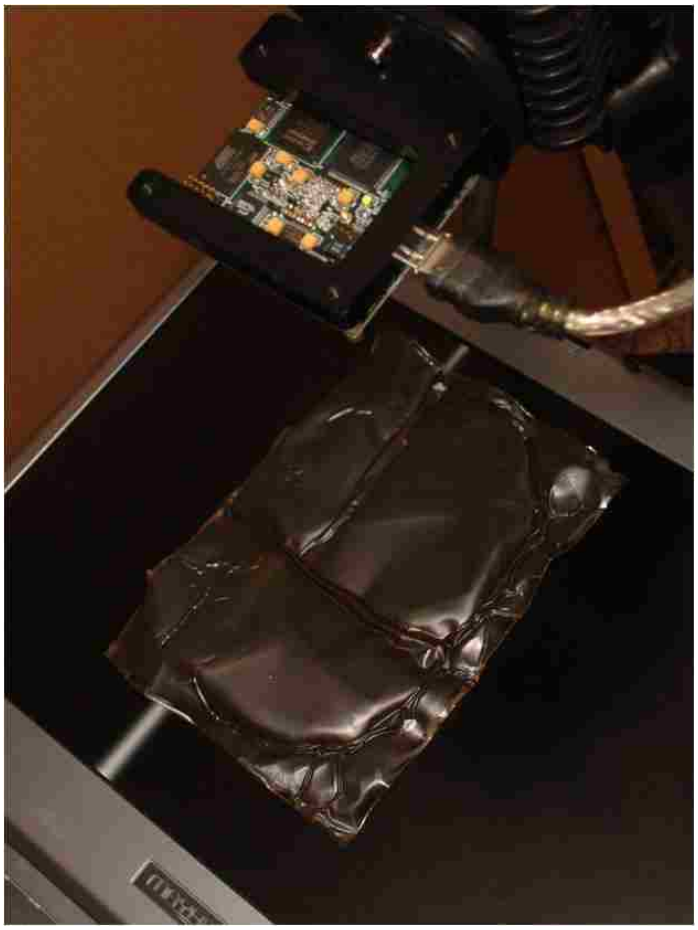


Figure 5.3: An example scanner configuration.

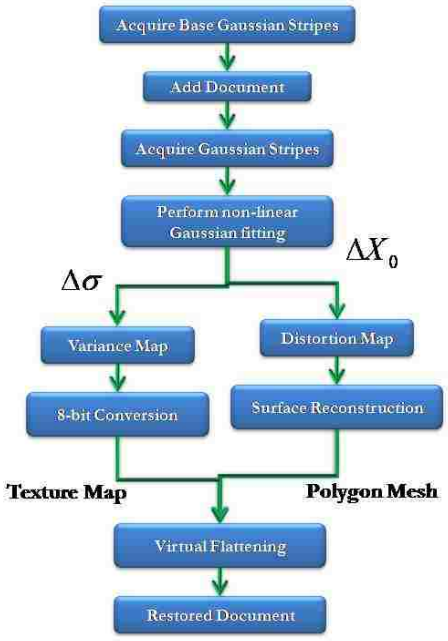


Figure 5.4: The scanning process.

local pixels eliminating the need for global registration or calibration. The scanning is configured by placing a camera above a flat-panel computer monitor as seen in Figure 5.3.

The new image-based scanner obtains the optical properties of the document layers by rear-illuminating the object with time-evolving Gaussian stripes. Two stripes are displayed: a vertical Gaussian stripe given by

$$G(x; x_0, \sigma_x) = de^{\frac{-(x-x_0)^2}{2\sigma_x^2}} \tag{5.3}$$

and a horizontal Gaussian stripe given by

$$G(y; y_0, \sigma_y) = de^{\frac{-(y-y_0)^2}{2\sigma_y^2}} \tag{5.4}$$

where d represents the color depth of the display device.

The acquisition system observes two passes of the horizontal and vertical Gaussian stripes. The initial pass is captured with only the display device in the scene to acquire a base case for the Gaussian parameters σ and X_0. The next pass of the stripes is captured with the document in the scene as shown in Fig. 5.3. After the acquisition of four sweeps, calculations are performed on a pixel-by-pixel basis using the time-evolving Gaussian stripes observed for each one. For each pixel, the intensity values are normalized to one, and the scale factor, α, is saved as the transparency value for that pixel. Then a non-linear Gaussian fit is performed on the normalized intensity values to estimate σ and X_0. This gives two

2D Gaussian functions for each pixel:

$$G(x, y : x_0, y_0, \sigma_x, \sigma_y) = e^{\frac{(x-x_0)^2+(y-y_0)^2}{(\sigma_x+\sigma_y)^2}} \tag{5.5}$$

$$G(x, y : x_0^{'}, y_0^{'}, \sigma_x^{'}, \sigma_y^{'}) = e^{\frac{(x-x_0^{'})^2+(y-y_0^{'})^2}{(\sigma_x^{'}+\sigma_y^{'})^2}} \tag{5.6}$$

The optically distorted Gaussian properties $\sigma_x^{'}$, $\sigma_y^{'}$, $x_0^{'}$, and $y_0^{'}$ are given by Eq. 5.6. The difference between the Gaussian parameters in Eq. 5.5 and Eq. 5.6 gives an estimation of the optical changes due to the object in the scene.

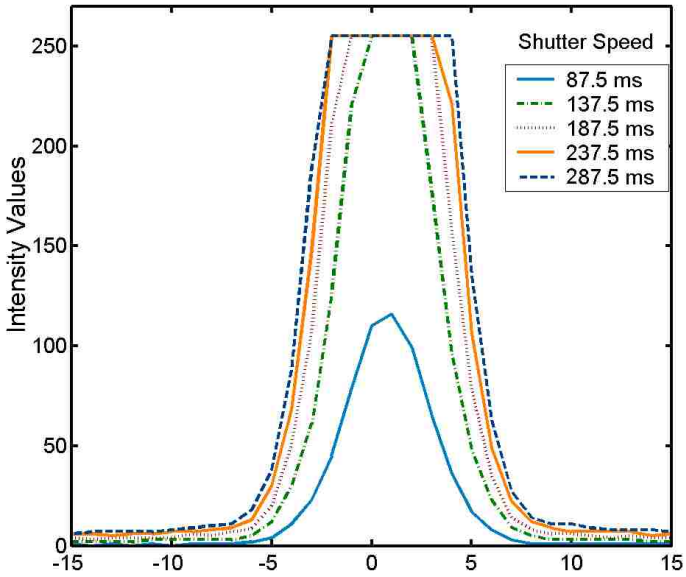
5.3.1 Dynamic Range Considerations

When digital cameras image a scene by taking a digital photograph, an analog to digital conversion takes place. The main technological element in this process is a charge coupled device (CCD). The CCD measures the irradiance, E, for the duration of exposure time (Δt_e) when an image is captured.

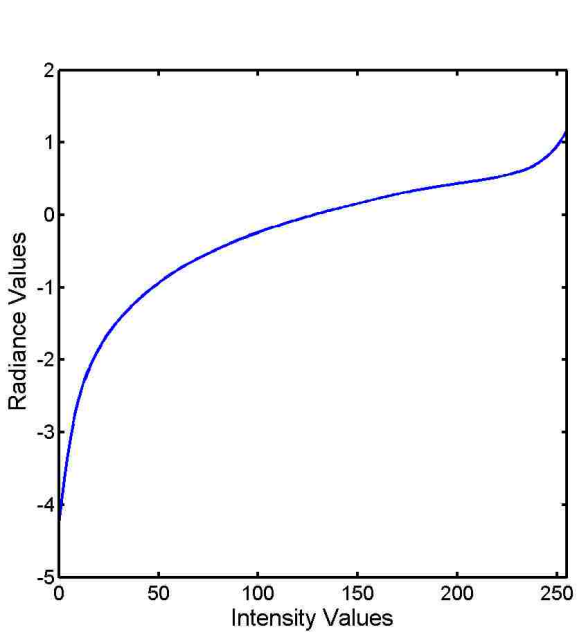
A typical CCD produces 12-bit output, while most image formats use the standard 8-bit values. Therefore a conversion process is required to quantize the information. The mapping from 8-bit to 12-bit is calculated in a non-linear fashion attempting to replicate the response of film. This non-linear mapping, or dynamic range, leads to data loss when multiple radiance values are mapped to the same intensity value. Consequently, in the captured image, the data loss typically appears as saturation, and the limited dynamic range often reduces the quality of a digitized image.

A single image captured from the camera, as seen in Fig. 5.5(a), shows the loss of information due to the dynamic range compression. The radiance values at the peak of the Gaussian stripes are all mapped to the same intensity values by the imaging device which greatly reduces the accuracy of Gaussian fitting algorithms. The intensity profile for one pixel at varying exposure rates is shown in Fig. 5.5(b). In this example, all but the fastest shutter speed suffers from data loss. However, to solely use this exposure rate would also be insufficient since there would be data loss for areas with less transmitted light such as when the document is in place.

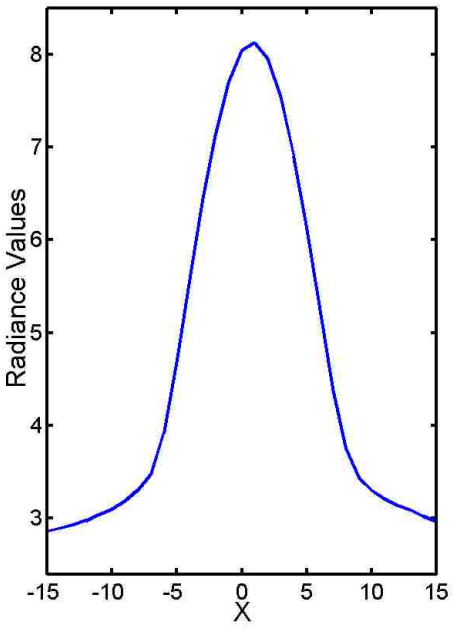
To compensate for this loss of data, High-Dynamic Range Imaging (HDRI) techniques have been developed. In one of the first attempts to correct for dynamic range compression, Madden used the direct output of a CCD to calculate an extended dynamic range image. By using several flux integration times from the CCD, this process can combine multiple images in an added mosaicing process. However, most commercially available cameras do not provide direct access to the CCD, so this technique is not easily applied.



(a) Shows the time-evolving intensity profile of 5 exposures for a single pixel using normal 8-bit images.



(b) The estimated response curve for the pixel in Figure. 5.5(a).



(c) The resultant Gaussian profile in radiance values.

Figure 5.5: The intermediate steps in calculating High Dynamic Range Imaging (HDRI).

Mann et al. [61] developed a technique to combine multiple images taken with a digital camera into a single floating point image related to the CCD pixel response. However, the authors assumed a constant response function that may or may not correctly model the non-linear response curve of a particular camera.

To extend this technique, Debevec et al. [22] also acquire multiple images of the same scene under varying exposure rates. Then the response function for a scene is directly calculated using representative pixels under varying exposures. Once the response function is computed, the set of images can be combined into a floating point radiance map representative of the true radiance in the scene.

In this work, the developments in HDRI are extended to enhance the accuracy of the transmissive light acquisition. Instead of assuming one radiance value per pixel, radiance values are calculated for each position of the Gaussian stripe imaged by the pixel. Starting with the imaging of the same stripe with varying exposure rates, Figure 5.6, a pixel-wise radiance function can be estimated. The intensities of the pixel vary between the 8-bit intensity range of 0-255 as the Gaussian stripe moves across the imaged sensor area. This over-constrains the linear least squares estimation of the response function:

$$O = \sum_{x=1}^{X} \sum_{e=1}^{E} \left(w\left(I_{(x,e)}\right) \left(g\left(I_{(x,e)}\right) - lnE_x - ln\Delta t_e \right) \right)^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} \left(w\left(z\right) g''\left(z\right) \right)^2 \quad (5.7)$$

The weighting function ($w(z)$) was originally described as a simple hat function [22]. However, in this work the compressed Gaussian peaks will play no role in the response curve computation. Therefore, the following weight function w is used:

$$w\left(z\right) = \begin{cases} z+1 & \text{if } z <= \frac{Z_{max}-1}{2} \\ Z_{max} - z + 1 & \text{if } z > \frac{Z_{max}}{2} \end{cases} \quad (5.8)$$

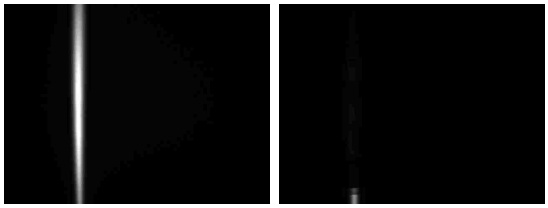
The response function is estimated by choosing a single representative pixel that demonstrates a large dynamic range in the scene. Then the image response curve is defined by Eq. 5.9.

$$g(I_{(x,e)}) = ln\ E_x + ln\ \Delta t_e \quad (5.9)$$

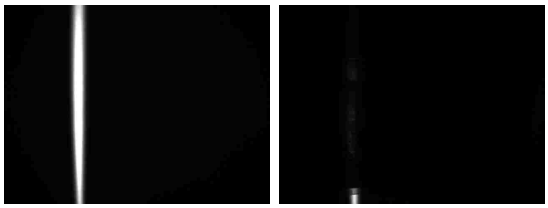
5.3.2 Acquiring Document Content

Traditional document content is printed with ink that is designed to augment how light is reflected from the document surface. Consequently, the composition of ink guarantees the same diffuse reflectance properties for front-incident light and rear-incident light. This fact

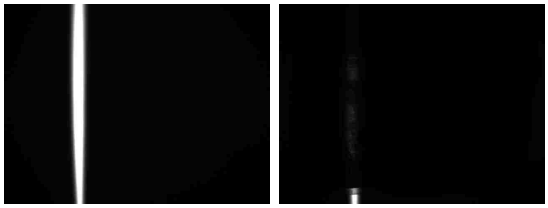
(a) 87.5ms shutter speed



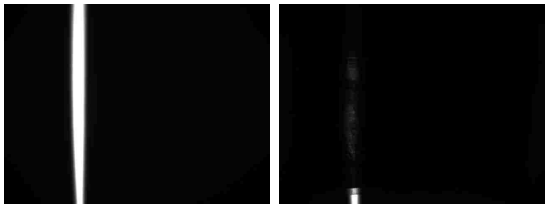
(b) 137.5ms shutter speed



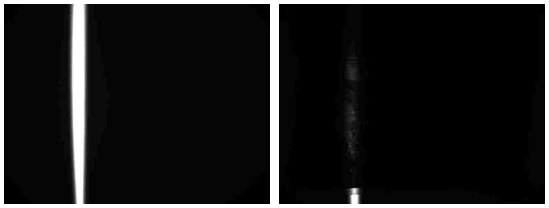
(c) 187.5ms shutter speed



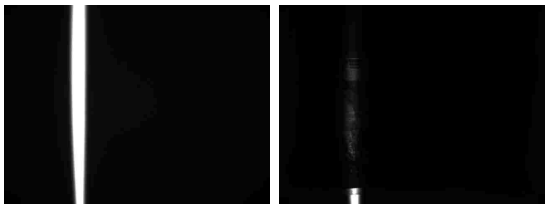
(d) 237.5ms shutter speed



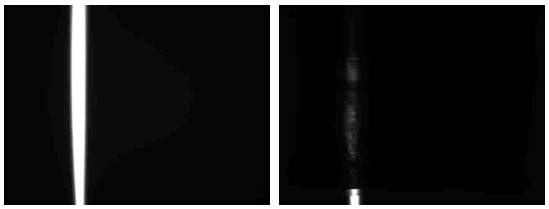
(e) 287.5ms shutter speed



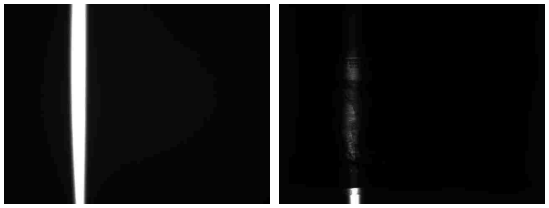
(f) 337.5ms shutter speed



(g) 387.5ms shutter speed



(h) 437.5ms shutter speed



(i) 487.5ms shutter speed

Figure 5.6: An example HDRI capture using nine shutter speeds. The left column shows the base stripe while the right column shows the distorted stripe with document in the scene.

can be observed by holding a single-sided paper document up to a bright light source. The content, formed with ink, will be visible when the document is viewed from the front or back.

The diffuse reflectance property of the ink is estimated in the proposed system by change in Gaussian variance. This variance change ($\Delta\sigma$) of the observed Gaussian stripes for each pixel is defined as the Equation 5.10, and provides extensive information related to the medium imaged by that pixel. When the variance is combined from each pixel into the entire image, the *variance map* is obtained.

$$\Delta\sigma = MAX\left(\left(\sigma_x - \sigma_x^{'}\right), \gamma\left(\sigma_y - \sigma_y^{'}\right)\right) \tag{5.10}$$

This work accommodates the non-square aspect ratio display devices by calculating difference in Dots-Per-Inch (DPI) for the x and y directions. This allows scaling of one axis by γ to remove distortions due to aspect ratio and physical screen dimensions.

Typically, the observed change in variance will be effected by two conditions: opacity and attenuation. When there is a large difference between the base variance, (σ), and modified variance, (σ') this suggests that there is a much higher opacity. Moreover, attenuation will also create a difference in observed variances. However, this effect does not typically generate error in the variance map since attenuation adds only a small low-frequency noise to the result. The variance map is acquired in floating point values, so a conversion step must take place to generate a standard 8-bit greyscale or 32-bit color image. The variance map is converted using the following equation:

$$I(u,v) = \left(V_{(u,v)} + t\right) s \tag{5.11}$$

where t is an intensity translation and s is a scale factor. The values for s and t are determined empirically.

5.3.3 Distortion Shape Estimation

Continuing the discussion from section 5.2.1, the diffuse transmission of light can be used to directly estimate the surface orientation for each pixel observation on the document surface. Moreover, for non-planar documents relative variations in surface orientation provide a direct method to estimate local surface shape variations.

The change between the base position of the Gaussian stripe and the modified position provides a basic light-transport model for one or more layered documents. As the time-evolving Gaussian stripe moves across the display device, the observed transmitted

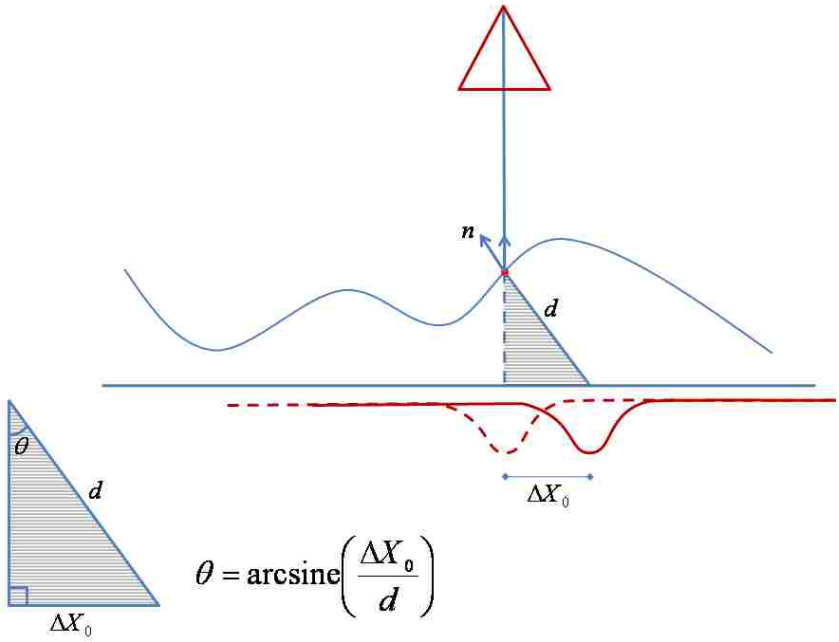


Figure 5.7: Diffuse Transmission of a back-lit light source

intensities will also vary depending on the single-scatter diffusion given in equation 5.1. The shift of the Gaussian peak, given by ΔX_0 and defined by Eq. 5.12, gives a pixel-wise estimate for the change in light transmission through the document medium that may be used to estimate the surface orientation at the outermost layer of a document.

$$\Delta X_0 = \sqrt{(x_0^2 - x_0'^2) + \gamma(y_0^2 - y_0'^2} \tag{5.12}$$

This estimation for each pixel is combined across the entire image to obtain a *distortion map* $D\,(u, v)$.

For a single-layered document, there is only one medium between the light source and the sensor which allows for direct estimation of the surface orientation for each pixel. These pixel-wise orientations, given by $(\Delta x_0, \Delta y_0)$, are used to estimate the orientation of the surface for x and y directions.

$$\Theta_x = arcsine\left(\frac{\Delta x_0}{d_x}\right), \Theta_y = arcsine\left(\frac{\Delta y_0}{d_y}\right) \tag{5.13}$$

However, equations 5.13 have the unknown quantities d_x and d_y since the surface depth remains unknown as shown in Figure 5.7. Therefore, for estimation purposes the mean values of both (Δx_0 and Δy_0) are used for d_x and d_y, so equation 5.13 becomes:

$$\Theta_x \simeq arcsine\left(\frac{\Delta x_0}{(\overline{\Delta x_0})}\right), \Theta_y \simeq arcsine\left(\frac{\Delta y_0}{(\overline{\Delta y_0})}\right) \tag{5.14}$$

To estimate the surface normal, the orientation angles are used in:

$$n \simeq (\Theta_x, \Theta_y, 1)^T \tag{5.15}$$

The normal vector is typically accessed as a unit surface normal where $\sqrt{\left(n_x^2 + n_y^2 + n_z^2\right)} = 1$. Then the surface normal can be defined as:

$$n \simeq \frac{(\Theta_x, \Theta_y, 1)^T}{\sqrt{\left(\Theta_x^2 + \Theta_y^2 + 1\right)}} \tag{5.16}$$

It should be noted that the sign of these normal angles may be globally ambiguous. Similar to the bas-relief ambiguity in shape-from-shading [5], the surface function may be the inverted version of the correct surface.

5.3.3.1 Surface Reconstruction

Once the surface normals are estimated for each pixel, it is straightforward to calculate the surface gradient at these positions. The surface gradient is defined as $\frac{\partial z}{\partial x} \simeq \frac{\Theta_x}{\sqrt{\Theta_x^2+\Theta_y^2+1}}$ in x and $\frac{\partial z}{\partial y} \simeq \frac{\Theta_y}{\sqrt{\Theta_y^2+\Theta_y^2+1}}$ in y. With known surface gradients, an integrable surface reconstruction, introduced by Frankot et al. [25], can be calculated.

Using the Fourier basis function

$$\Phi(u, v, \omega) = e^{(j\omega_u u + j\omega_v v)} \tag{5.17}$$

the surface function $\tilde{z}(u, v)$ can be estimated as

$$\tilde{z}(x, y) = \sum_{\omega \in \Omega} \tilde{C}(\omega) e^{j\omega \cdot (u,v)} \tag{5.18}$$

Therefore if $\tilde{C}(\omega)$ is defined as

$$\tilde{C}(\omega) = \frac{-j\omega_x \tilde{C}_x(\omega) - j\omega_y \tilde{C}_y(\omega)}{\omega_x^2 + \omega_y^2} \tag{5.19}$$

$\tilde{z}$ can be calculated directly in the frequency domain. An example surface is shown in Fig. 5.13(d).

5.3.4 Error

In this work, three major sources of error are encountered. First, since the surface reconstruction is underconstrained, a global average must be used to supplement the initial surface normal estimations. Second, the perspective projection of the imaging device adds low-frequency error in X_0. Third, the dynamic range constraints of the imaging device greatly reduced the accuracy of the Gaussian stripe detection.

5.3.4.1 Surface Normal Inaccuracies

The error introduced by using a globally static triangle edge to calculate the surface normal can be given an error bound. The relative error is calculated as:

$$\frac{\left|\overline{\Delta X_0} - \sqrt{\tilde{z}\,(x,y)^2 + \Delta X_0^2}\right|}{\left|\overline{\Delta X_0}\right|} \leq q\epsilon \tag{5.20}$$

5.3.4.2 Perspective Projection Correction

A global error is introduced into the normal map due to the prospective projection of the imaging system. As the distance from the camera's optical center increases, the angle of incidence on the surface also increases. This creates a systematic shift across the normal map that increases toward the edges of the image. An example of this error when performing a synthetic scan on a plane is shown in Figure. 5.8(d).

To compensate for this error, it is possible to take advantage of the frequency domain where the error occurs. Since the error presents itself as very low-frequency noise, a Gaussian bandpass frequency filter is applied to the Fourier transform of surface normal components in both the $X(u,v)$ and $Y(u,v)$ directions. The bandpass filter as described by Gonzalez et al. [33] uses two Gaussian distributions G_i, the starting frequency, and G_o, the closing frequency.

$$G_i(u,v) = e^{\frac{-(u^2+v^2)}{2\sigma_i^2}} \tag{5.21}$$

$$G_o(u,v) = e^{\frac{-(u^2+v^2)}{2\sigma_o^2}} \tag{5.22}$$

Then the filters are combined to create a single bandpass filter H.

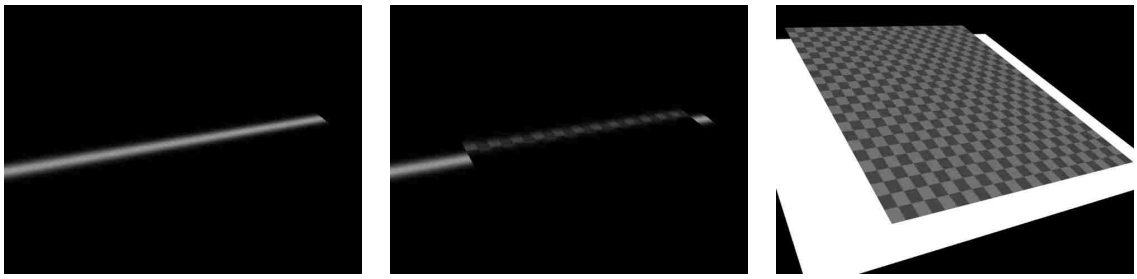
$$H = G_i - G_o \tag{5.23}$$

Once the filter is generated, the surface normals may be filtered using $X'(u,v) = X(u,v)H(u,v)$ and $Y'(u,v) = Y(u,v)H(u,v)$. These processed value are then reduced of the error induced by perspective imaging. Therefore, the surface estimation more accurately portrays the actual document shape configuration.

5.4 Error Analysis

To study the accuracy of the scanning system and investigate sources of error, synthetic scans were performed virtually. Utilizing Autodesk 3D Studio Max, environments, closely matching the real-world scanning compositions, were developed to test various aspects of the system.

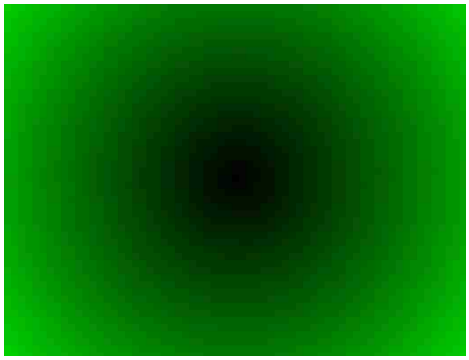
5.4.1 Synthetic Plane



(a) The synthetic light table.

(b) The scanning environment with the plane in place.

(c) The plane.



(d) The distortion map for the plane where lighter intensity shows larger difference in X_0 values.

Figure 5.8: A synthetic scan of a planar object.

The first test of the proposed scanning procedure was built using a plane with textured animation that played the sweeping stripe in both directions and a semi-transparent plane with a checkerboard texture as seen in Figure 5.8(c). This test provided the groundwork for estimating the feasibility of the scanner. The planar test demonstrated the noise introduced by the perspective projection of the imaging device. This can be seen in the distortion map acquired for the plane in Figure 5.8(d).

To correct the low frequency noise, the band-pass frequency filters are applied to the surface normal estimations (ΔX_0). Figure 5.9 shows the resultant surface as the low-frequency band-pass is increased.

5.4.2 Synthetic Sphere

Next a semi-transparent checkered hemisphere was synthetically scanned. This polygonal hemisphere was placed on the rear-illumination source, Figure 5.10(a), while a camera observed each of the light stripe positions as seen in Figure 5.10(b). This scan was

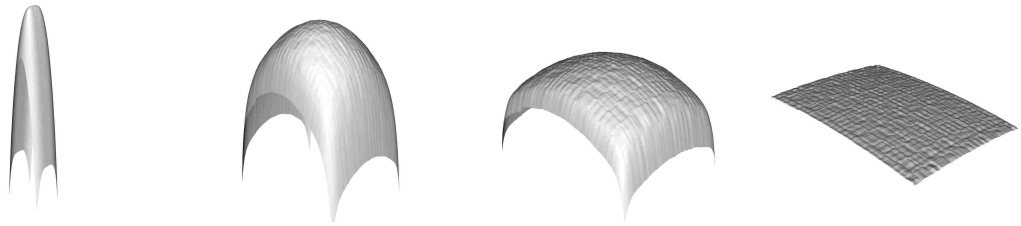
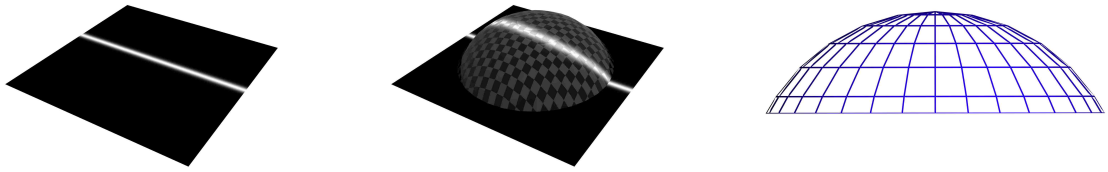


Figure 5.9: Estimated surface shape with decreasing low-frequency band-pass.



(a) Synthetic illumination device (b) Synthetic scan with sphere in place (c) Side view of the spherical object

Figure 5.10: A synthetic scan of a hemi-sphere developed in Autodesk 3D Studio Max.

performed with 600 stripe positions in both the x and y orientations using a virtual camera with 640x480 resolution. Then, once both $\Delta\sigma$ and ΔX_0 are estimated, the surface is reconstructed using the method described in section 5.3.3.1 as shown in Figure 5.11(b).

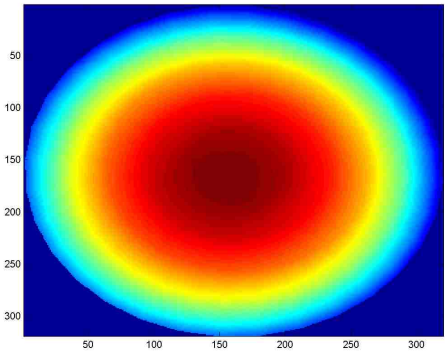
To test the accuracy of the scanning and surface reconstruction, the difference between the actual sphere surface and the estimated surface is shown in Figure 5.11(c). Overall, the results were acceptable for an image-based device. It should be noted that adding two more Gaussian sweeps in diagonal directions would improve the shape estimation.

5.5 Results

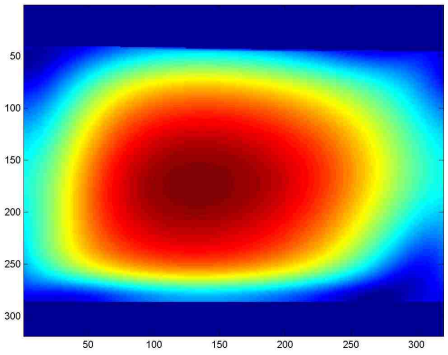
Once the estimation of the synthetic results were satisfactorily obtained, the physical scanner was built using a Windows XP-based 1.6 GHz Pentium M Laptop running at 1024x768 resolution with a 640x480 FireWire gristle camera obtaining the scan images. This setup was used to scan multiple paper documents with successful results.

Since the successfulness of the scanning device depends directly on the transmission of light through the document substrate, it should be mentioned that the paper industry has a standard opacity measure in place. The Technical Association of the Pulp and Paper Industry defines method T425 [1] as:

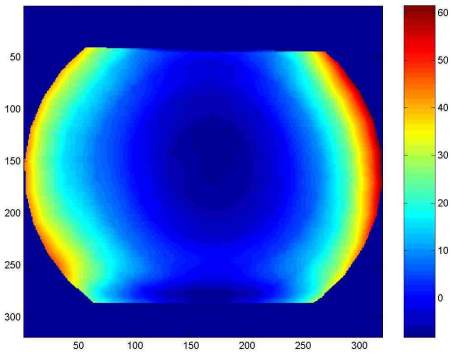
$$Opacity = \frac{R_0}{R_{0.89}} \tag{5.24}$$



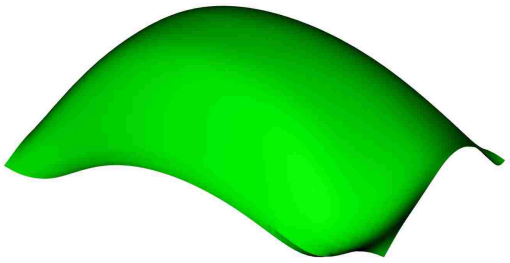
(a) Synthetic sphere depth map



(b) Reconstructed sphere depth map



(c) Absolute difference of depth maps



(d) Estimated sphere shape

Figure 5.11: The analysis of a hemi-sphere developed in Autodesk 3D Studio Max.

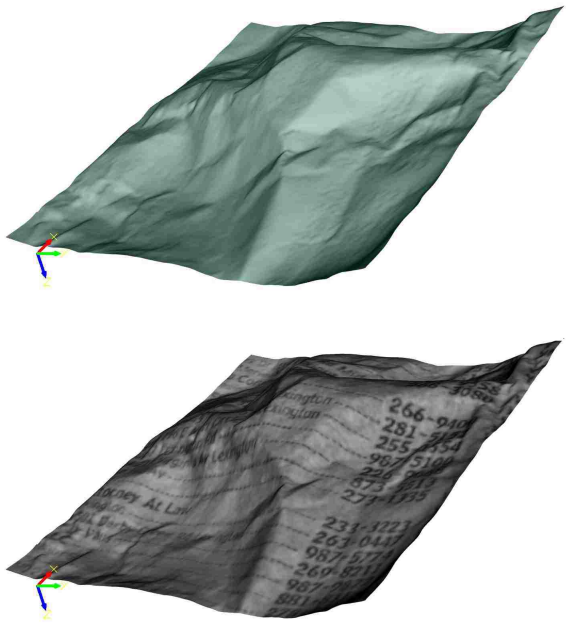


Figure 5.12: The reconstruction document surface.

This is the ratio of reflectance of a document over black backing, R_0, compared to reflectance of a document over white backing, $R_{0.89}$. The opacity is represented as a percentage where 100 is completely opaque and 0 is completely transparent.

5.5.1 Single-layer documents

The first application shown, of this new transmissive scanner, are single-layer documents, including any document on a single piece of substrate. For this application, a snippet of a phone book is used. Directory paper typically has an opacity between 75 and 85 [53], so this type of document works extremely well with the transmissive scanner. The paper has been severely wrinkled, as seen in Figure 5.13(a), to demonstrate distortions common in damaged manuscripts. The scan consisted of displaying 650 horizontal stripes and 400 vertical stripes for base and document cases.

After calculating the variations in the Gaussian Stripe parameters, the document correction techniques can be applied. The intrinsic content of the document is extracted using $\Delta\sigma$ and the conversion given by equation 5.11. Now the photometrically correct content can be used as a texture map on the estimated surface. The distortion map, shown in Figure 5.13(c), calculated from ΔX_0 is used to directly reconstruct the document surface, up to scale using the technique discussed in section 5.3.3.1 as exhibited in Fig. 5.12. This reconstructed surface is used to perform a physical-based simulation of a mass/spring

system, similar to [9]. As discussed in chapter 4, the mass/spring simulation is used to flatten the document and restore the geometric distortions caused by imaging a non-planar surface. The combined result of the total restoration is shown in Fig. 5.13(d).

(a) The original observed image.

(b) The restored content of the document using the variance map ($\Delta\sigma$).

(c) The distortion estimation of the document (ΔX_0).

(d) The final restored document with geometric restoration.

Figure 5.13: The restoration of a crumpled telephone director snippet.

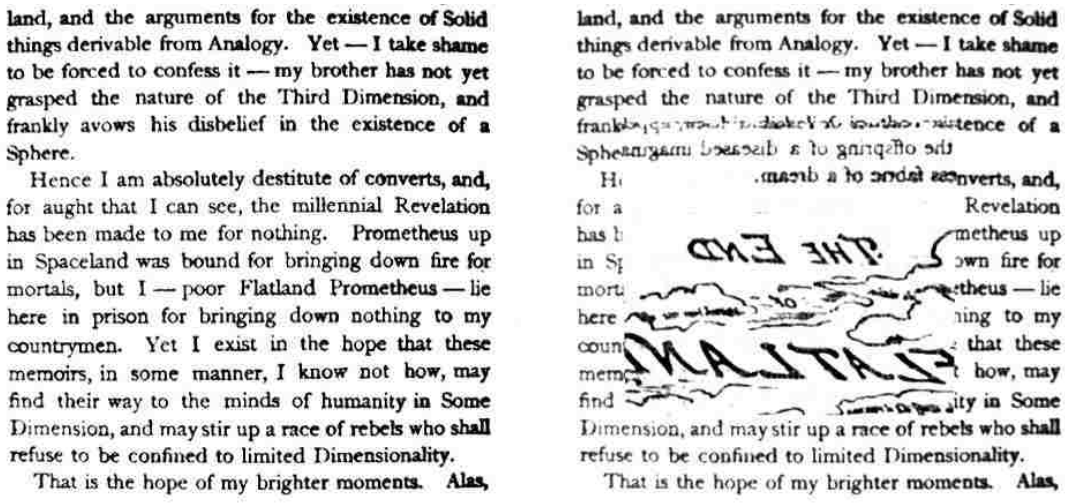
In Figure 5.13(b), obvious noise can be seen in the content estimation from $\Delta\sigma$. This

is due to the composition of the substrate. For lighter weight papers such as telephone directories or newspapers, typically 22.5lb weight or less, the variations in wood pulp play a role in the local transmissivity of a document. The ability to acquire the overall transmissivity of a document opens more applications for the scanning procedure.

5.6 One-Pass Duplex-Sided Scanning

The second application of this transmissive document scanner is one-pass duplex-sided scanning. Single paged documents with content printed on both sides can often be difficult to scan. Typical show-through effects are shown in Figure 5.15(a). Most methods for scanning these documents focus on improving the accuracy for scanning a single side of the substrate. These techniques treat the intensity variations as artifacts that require removal in the final image.

However, instead of treating show-through content as noise, the method presented here relies on the overall transmissivity of a document. As a result, the opacity of each layer can be segmented into a restored image of each side.



land, and the arguments for the existence of Solid things derivable from Analogy. Yet — I take shame to be forced to confess it — my brother has not yet grasped the nature of the Third Dimension, and frankly avows his disbelief in the existence of a Sphere.

Hence I am absolutely destitute of converts, and, for aught that I can see, the millennial Revelation has been made to me for nothing. Prometheus up in Spaceland was bound for bringing down fire for mortals, but I — poor Flatland Prometheus — lie here in prison for bringing down nothing to my countrymen. Yet I exist in the hope that these memoirs, in some manner, I know not how, may find their way to the minds of humanity in Some Dimension, and may stir up a race of rebels who shall refuse to be confined to limited Dimensionality.

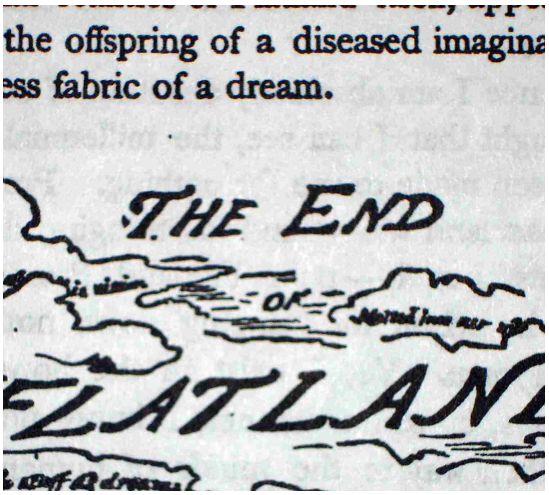
That is the hope of my brighter moments. Alas,

(a) Original image of the front illuminated document.

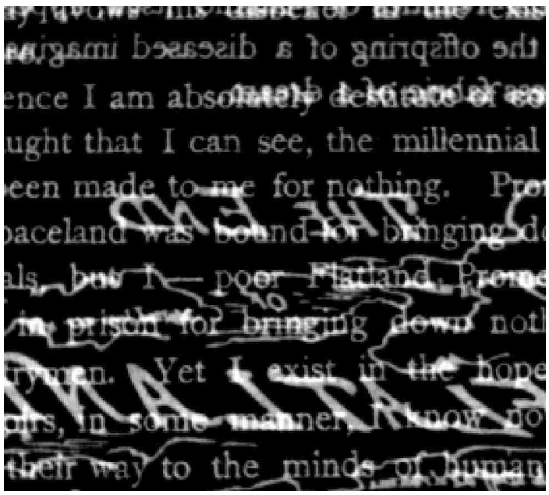
(b) Result showing the reverse content of the document.

Figure 5.14: The results of a duplex-sided one-pass document scan.

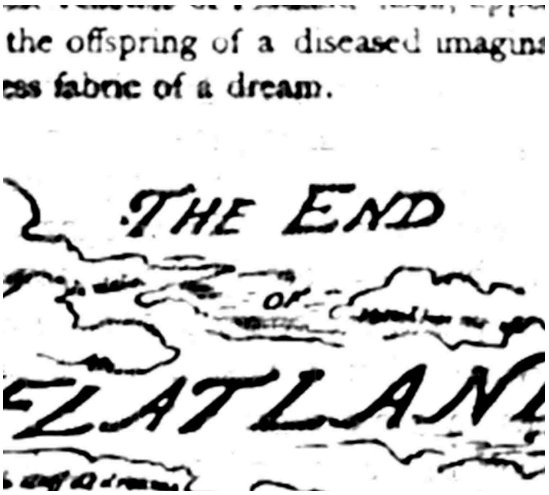
To perform a duplex-sided scan, an additional step must be added to the digitization process. Preceding the standard scanning process the reflected intensities from the front-side of the document must be obtained $intensity_{f_{(u,v)}}$. It is adequate to capture an image of the document with ambient room lighting turned on. Once this image is obtained the scanning process is performed normally. When the scanning is completed the overall document



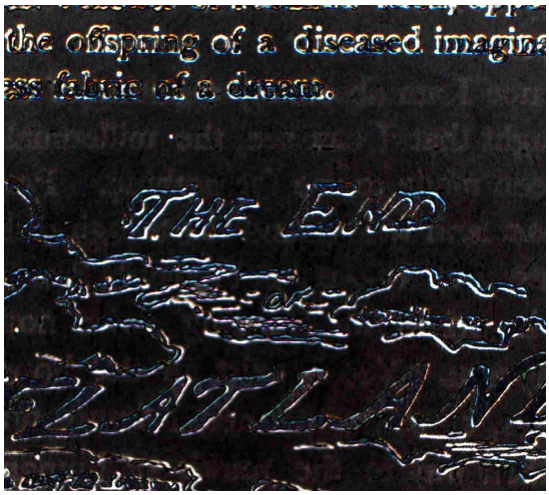
(a) A ground-truth photograph of the reverse side. Notice the *show-through* effects from the opposite side.



(b) The overall document opacity from ΔA.

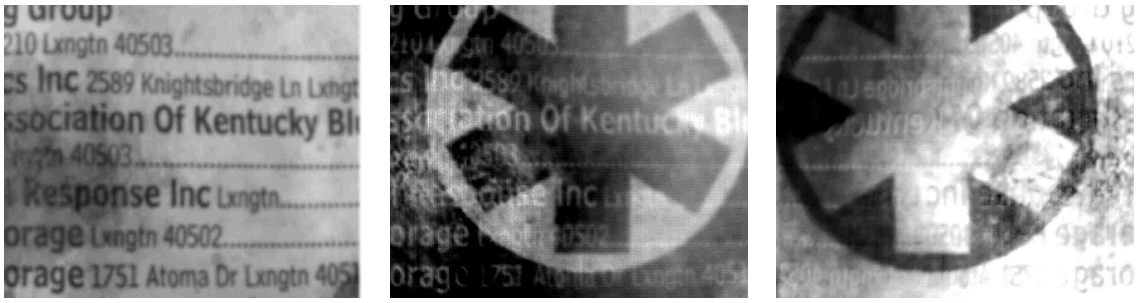


(c) Differentiated content from the opposite side of the document.



(d) The difference of the ground-truth and calculated result.

Figure 5.15: The process of extracting the reverse side content from a duplex-sided one-pass document scan.



(a) Original image of the front illuminated document with severe shape distortion.

(b) Acquired transmissive properties of the document.

(c) Differentiated content from the opposite side of the document.

Figure 5.16: The results of a duplex-sided one-pass document scan.

opacity is given by the variance $\Delta\sigma$. In the case when content appears on both sides of the document, the variance map $\Delta\sigma$ can be segmented into three cases:

- Substrate only
- Substrate and content on either front or reverse side
- Substrate and content on both sides

The content from the back side of the document can then be extracted by subtracting the front surface reflectance from the document opacity.

$$intensity_b = \Delta\sigma - T(intensity_f) \tag{5.25}$$

Using the equation 5.25, where $T(i)$ is an intensity scaling function, an image of only the back side content is generated.

The single scan of a nearly planar book page with duplex-sided print is shown in Figure 5.14. Figure 5.14(b) shows the reconstructed content from the reverse side of the document in the correct orientation and position. To determine the accuracy of this approach, a direct image was taken of the reverse document side, Figure 5.15(a), which gives a ground-truth reference. The estimated content from the proposed method is shown in Figure 5.15(c) and the difference of these two images is shown in Figure 5.15(d).

As another example of the one pass scanning technique is shown in Figure 5.16. This document contains severe wrinkling and water damage. The front side of the document is shown in Figure 5.16(a). Using the variance map estimation, Figure 5.16(b), the reverse content can be extracted. The result shown in Figure 5.16(c) provides a good estimation of the reverse content, however, the noise is greater due to the deterioration and foreign materials (e.g. dried liquids).

5.7 Other Documents

As this work evolved, it became apparent that other document types would now become scannable using the transmissive scanner. Since most document substrates have some intrinsic form of transparency, the scanning technology presented here has widely applicable application. Moreover, documents types that do not fall under the traditional textual forms also need to be considered.

In the next chapter, documents that were developed as intrinsically transmissive objects are explored. These documents, historic photo-negatives, contain equivalent importance to culture and heritage as most paper-based documents. However, like many of the documents discussed here, deterioration has reduced the transmissivity of large areas of the negatives which causes distortions during digitization and printing.

Chapter 6

Photographic Negative Restoration

6.1 Introduction

As stated earlier in this thesis, the majority of document restoration research has focused on digitizing bound books or manuscript pages. For example, the technologies presented in chapters 3 and 4 have used documents that are fully observable with reflective light. However, restoration techniques for other documents, such as deteriorated photographic negatives or laminated works, have not been developed.

This lack of research has limited the broader application of virtual document restoration. For example, many photographic negatives, dating from 1925 to 1955, have severe deterioration in various laminate layers making the negatives unprintable. As these negatives have deteriorated, their lose has affected collections in the Library of Congress, the National Archives, and public and private archives in the United States. It should also be noted that collections in the United States are not the only archives effected. Many international collections have also suffered loss. According to a European study by Klijn et al. [50], two-thirds of the surveyed collections in Europe contain deteriorating negatives. Needless to say, the importance of these historic collections is immeasurable.

Currently, the standard restoration procedure for negatives requires physical separation of layers in a permanent and destructive operation. While these techniques extract the original image, the restoration process requires a great deal of time and leaves the piece permanently altered. Therefore, digitally preserving and restoring deteriorating historic photo-negatives is an urgent challenge that has remained unsolved.

Standard scanning techniques are not suitable for these multi-layer documents, because the non-uniform separation between individual layers generates photometric distortions in the final image. However, as will be shown in this chapter, the transmissive document scanner described in Chapter 5 can be modified to digitize and restore documents other than standard paper-based materials. The digitization and subsequent restoration of damaged photo-negatives is facilitated by the image-based light modeling of the scanner.

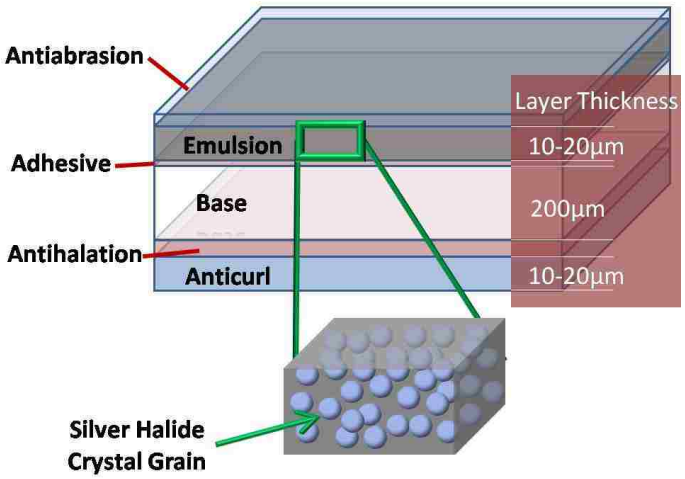


Figure 6.1: The physical composition of film.

Moreover, this work with multi-layer documents opens up the possibility to enhance the technology to scan documents with multiple layers of disparate content.

6.2 Film Review

The basic composition of film is shown in Figure 6.1. The two most important components to be considered in the research presented here are the emulsion layer and base layer. The emulsion layer contains the photographic content of the film, while the base layer provides support and rigidity to the film. Therefore, the base layer itself contains no relevant information but only provides the physical stability necessary to keep the emulsion intact. For negative film, the base layer also possesses transparent properties.

During the standard photographic process, incoming photons hit film for a fixed duration or exposure. As photons reach the film, they cause chemical reactions to occur in the emulsion layer of the film. In the emulsion layer are millions of light sensitive silver halide crystals. As more photons reach a specific area of film, the optical density of the film is increased due the displacement of electrons from silver-halide crystal grains. This process creates a *latent image* in the emulsion layer. To convert the latent image into a permanent usable negative, the film is developed. This development process converts the silver-halide crystal into a more stable metallic silver. In this process, the optical density of the film is increased where more photons originally reached the emulsion layer effectively turning bright areas in the scene very dark due to the metallic silver in the film, hence the term negative.

(a) The emulsion side of a deteriorated negative. (Courtesy: S. Bailey)

(b) The acetate side of the same negative. (Courtesy: S. Bailey)

Figure 6.2: Example of a severely damaged acetate negative.

The evolution of film bases has come in three major forms: cellulose nitrate, cellulose acetate, and polyester. The first and most commonly mass produced film base was cellulose nitrate. As a result, all film manufactured before the mid 1920s is typically a nitrate base. However, as the harmful effects of nitrate became apparent, film manufacturers moved to cellulose acetate as the the film base of choice beginning in 1925. Acetate film, commonly referred to as safety negatives, remained the industry standard for cut-sheet film past the mid 1950s. Because acetate showed many of the stability disadvantages of nitrate film, other base materials were researched. First introduced in 1955 as an improved film base, polyester provided improved physical properties and dimensional stability over acetate films. In the 1960s, polyester replaced acetate as the main film base used for photo-negatives [70].

The acetate film for still photographs discussed in this work was manufactured between 1925 and 1955. Horvath [44] provides extensive documentation to identify these safety negatives through notch markings developed by film manufacturers. However, since the full extent of acetate degradation was not discovered until much later, it is safe to assume that photographs continued to be taken with acetate safety film even after it ceased production. Consequently, over thirty years of photographs exist on cellulose acetate base film and are at risk of total loss.

6.3 Losing Cultural Heritage

The deterioration of vast collections of photographic negatives is widespread. Even under proper conservation, the negatives will continue to decay, greatly reducing the quality of printing. In severe cases, the negatives may not be printable, leaving the content of these images inaccessible.

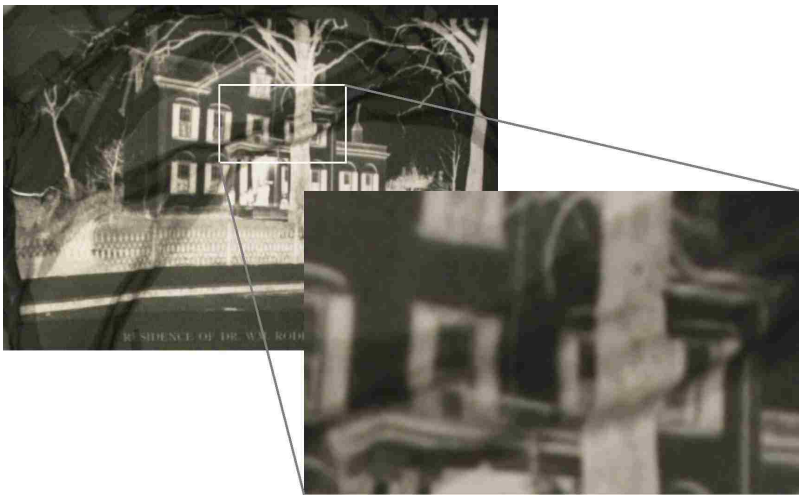


Figure 6.3: An image of a rear-lit deteriorated acetate negative with a magnified view of the photometric and geometric distortion.

According to the Image Permanence Institute, the deterioration of acetate film has been termed *vinegar syndrome*. Reilly expands this term with the following statement.

> Nearly every sizable collection of film has experienced losses due to vinegar syndrome, and many more losses are expected in the future. The main symptoms of this problem are a vinegar-like odor and buckling, shrinking, and embrittlement of the film. All acetate films are susceptible to such deterioration. Whether it happens within a few years or not for centuries depends on the storage conditions in which the film is kept [70].

Horvath [44] performed a study to document and classify the effect acetate decay had on many culturally important collections. According to the author, the degradation of acetate occurs with the introduction of at least one of the following elements: moisture, heat, and/or acid. The deterioration begins with minimal exposure to heat or humidity. Then as the degradation progresses, the acetate itself begins to release acetic acid that speeds the decay process and emits the classic vinegar smell. During the acetate survey, Horvath generated a metric to categorize the level of deterioration that has occurred in each negative.

For negatives classified at deterioration levels 1–3, standard printing procedures will produce a visually accurate positive photograph of the negative. Moreover, level 4 negatives may need added post-processing to produce an adequate archival photograph. However, for negatives categorized as level 5–6, no previous techniques exist to print the negative.

Negatives that have deteriorated to the point where the emulsion has delaminated from areas of the acetate base provide a particularly difficult case for archivists. Separations

Deterioration	Type	Description
Level 1	No deterioration	flat negative.
Level 2	Curling	Slight or moderate symmetric edge curl; smooth surfaces on both emulsion and base sides
Level 3	Smell	the negative smells distinctly of acetic or butyric acid
Level 4	Warpage	wavy non-symmetric portions in edges and surface of the negative
Level 5	Bubbles	minimal separation between the emulsion and base or between the base and the anti-curl backing of the film
Level 6	Delamination	Separation of the emulsion, base, and anticurl layers; from slight localized separation to massive separation. A white crystalline exudate is often seen under the separated portions of the base material

Table 6.1: The states of acetate film deterioration.

create a network of wrinkles in the emulsion typically referred to as *channels* [72]. Channels present themselves, during viewing, as darkened regions where the separation of layers causes increased attenuation and refraction of light between the negative layers causing an overall reduction in imaged intensity. Beyond the intensity distortions created by the deterioration, these negatives also demonstrate severe dimensional distortion. Horvath noted that most negatives measured .008 to .010 inches in thickness at local measurements using calipers. However, for negatives that have reached level 6 degradation, the global height variations of the film can be greater than 0.5 inches.

The number of film collections containing negatives with level 5–6 damage was widespread and increasing when Horvath first surveyed the problem in the United States in 1987. While it was initially thought that the deterioration of acetate film was isolated to a narrow set of negatives from the late 1940s to early 1950s, it is now understood that these deteriorations are possible in all acetate films. In fact, the deterioration has been well documented even under proper conservation [44]. Consequently, the timeframe when acetate film was the industry standard has produced millions of negatives that are now at risk of total loss.

In the most severe cases, safety negatives can no longer be used to print usable images as shown in Figure 6.3. Once negatives reach level 6 degradation, the cost to retrieve a high quality print is overwhelming. Utilizing normal reprint measures to acquire a print of the negative results in highly distorted images.

6.4 Previous Photograph Restoration Techniques

Document restoration procedures typically fall under two categorizations: Physical and Virtual. Both of these categories have typically been applied to most document types with varying degrees of success.

6.4.1 Virtual Restoration

There has been a significant contribution of work focusing on the restoration of deteriorated photographs. Digital Inpainting, first introduced by Bertalmio et al. [8], provides an efficient procedure for restoring areas of loss in digital images. Inpainting has been improved in many ways since Digital Inpainting was first developed. Bertalmio et al. [7] continued their work in image restoration with improvements in inpainting. Roth et al. [73] use a learning framework based on Markov Random Fields to develop a set of image priors that may be used to restore data loss in photographs of natural scene. Criminisi et al. [20] extends inpainting by adding texture synthesis to the region filling. This results in the ability to remove large objects from a scene by replacing them with a synthesized texture using inpainting.

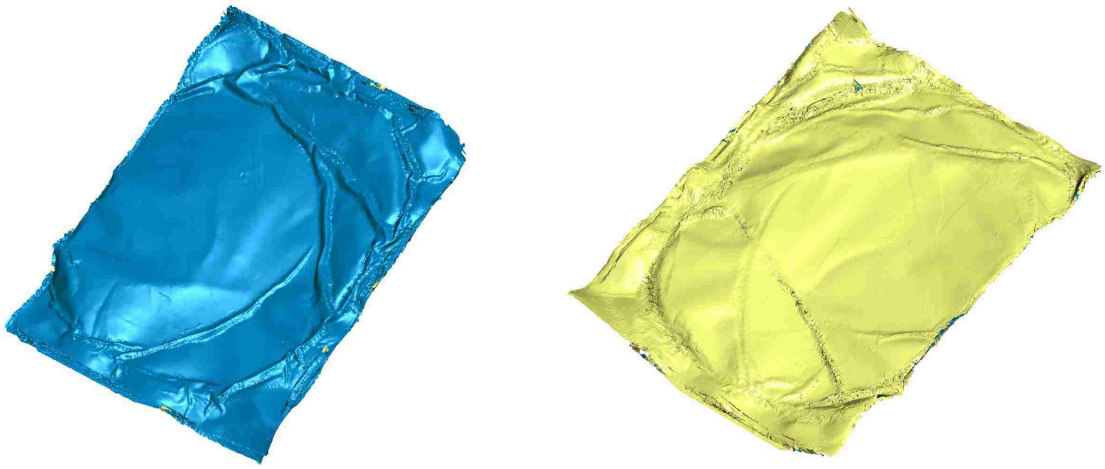
However, it must be noted that these procedures assume total loss for areas in need of restoration. The area of damaged photograph restoration has been addressed specifically with image and signal processing. Aridzzone et al. [4] provided a content based representation to assist in automatic and semi-automatic restorations.

Bruni et al. [13] use reflective light to detect blotches that have not fully covered the intrinsic content. Then they extract the underlying content to remove the deterioration.

Stanco et al. [80] provide an overview of standard photograph damage and deterioration types. Then the authors present some possible image-based restoration techniques using standard image processing. However, these methods generally focus on a scanned image of a photograph and usually only handle standard photographic prints.

Moving beyond standard image processing techniques other groups have used the intrinsic properties of the film base to assist in virtual restoration. Stanco et al. [81] have worked to restore glass plate negatives that have been fragmented due to cracking. Unlike acetate film, glass film typically does not have degradation qualities. However, the base itself is fragile and breakage is quite common. To virtually restore these negatives, they are photographed and re-aligned using rigid transformations.

The restoration of deteriorated acetate film can not be performed with any of these solutions. Shrinkage of the film base distorts a negative print in two distinct forms. Channels



(a) A high-resolution 3D scan of a deteriorated negative.

(b) The 3D scan of the acetate side from the same negative.

Figure 6.4: Example of a high-resolution 3D scan.

formed between the emulsion and base generate non-uniform intensity variations referred to as photometric distortion. Also, as the base layer shrinks, the emulsion layer remains constant in size which causes the emulsion to buckle and wrinkle. These shape changes cause geometric distortion in the imaged negatives.

6.4.2 Physical Restoration

The best case scenario for acetate film collections is to preserve collections in a state that negates the need for any type of restoration procedure. However, as noted earlier, the deterioration of acetate is relentless even under proper conservation. Moreover, for many collections, damage to the film has already removed the possibilities for standard access.

Currently, the standard restoration procedure for negatives requires physical separation of layers in a permanent and destructive operation. A professional conservator must strip the emulsion layer from the degraded film base using physical and chemical manipulation. Then the emulsion is reattached or duplicated onto another sheet of film [44]. While these techniques extract the original photometrically and geometrically restored image, a great deal of time is required to restore the photograph leaving the piece permanently altered. As with any conservation process, these invasive restoration methods may cause damage to the photographic content. Therefore, an accessible solution to digitally preserve and restore deteriorating historic photo-negatives must be developed.

6.5 A Complete Transmission Model for Film Negatives

Developing a virtual restoration solution for the film negatives began as an extension of the methods described in Chapters 3 and 4. The first attempt to solve the restoration problem requires accurate scanning of both the emulsion and base layers. The surface information was acquired using a Faro Laser ScanArm with a claimed precision of less than 0.01mm. Results of this scan are shown in Figure 6.4. While the emulsion gelatin acts as a typical lambertian surface, the acetate is highly specular and semi-transparent in delaminated areas. These factors add a error to the reconstructed acetate surface.

Since the structure of the already described negatives consists of the same composition and ordering of materials, the acetate negative may be physically modeled as a system of two thin films. This assumption allows the generation of a physically-based simulation to estimate the effect the materials play on the light transmitted through the respective layers.

6.5.1 Refractive layer estimation

The acetate layer is modeled as a minimally absorbing material with high transparency. Therefore, the model to calculate the transmissivity ($\mathcal{T}$) of the acetate must also include the reflectivity ($\mathcal{R}$) of the layer. The relationship between $\mathcal{T}$ and $\mathcal{R}$ is defined by the the law of conservation of energy as given in Equation 6.1.

$$\mathcal{T} + \mathcal{R} = 1 \tag{6.1}$$

Specifically, for the acetate layer we define $\mathcal{T}_a$ and $\mathcal{R}_a$ as the transmissivity and reflectivity, respectively. These coefficients are defined as follows:

$$\mathcal{T}_a = \frac{n_3 \cos\theta_3}{n_1 \cos\theta_1} \frac{t_{12}^2 t_{23}^2}{1 + r_{12}^2 r_{23}^2 + 2 r_{12} r_{23} \cos 2\beta} \tag{6.2}$$

$$\mathcal{R}_a = \frac{r_{12}^2 + r_{23}^2 + 2 r_{12} r_{23} \cos 2\beta}{1 + r_{12}^2 r_{23}^2 + 2 r_{12} r_{23} \cos 2\beta} \tag{6.3}$$

where $\beta = \frac{2\pi}{\lambda_0} n_2 h \cos\theta_2$, $t_{12} = \frac{2 n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2}$, $t_{23} = \frac{2 n_2 \cos\theta_2}{n_2 \cos\theta_2 + n_3 \cos\theta_3}$, $r_{12} = \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2}$, and $r_{23} = \frac{n_2 \cos\theta_2 - n_3 \cos\theta_3}{n_2 \cos\theta_2 + n_3 \cos\theta_3}$. Using Equation 6.1, the intensity of light transmitted through the acetate layer may be calculated by $1 - \mathcal{R} = \mathcal{T}$. This allows the final transmission intensity to be calculated by $I_i' = (1 - \mathcal{R}_a)^2 I_i$

Modeling the Emulsion Layer

The model for the emulsion layer must include an absorption coefficient ($\mathcal{A}$) that directly corresponds to the intensities of the original film negative. This consideration may be used

in the law of conservation of energy to modify the previously given formula 6.1 to allow absorption to be modeled.

$$\mathcal{T} + \mathcal{A} + \mathcal{R} = 1 \tag{6.4}$$

Therefore, for the emulsion layer the following values are defined; transmissivity $\mathcal{T}_e$, absorption $\mathcal{A}_e$, and reflectivity $\mathcal{R}_e$. To model the reduction in transmitted energy, these properties are applied in progression. The incoming energy intensity I_i is first reduced by the reflectance induced by the emulsion. This gives $I_i' = (1 - \mathcal{R}_e)^2 I_i$ as the intensity entering the emulsion layer where R_e is defined similar to 6.3. Next, absorption is considered. According to the Beer-Lambert law, the absorption of light modeled as:

$$\frac{I_{outgoing}}{I_{incoming}} = e^{-\alpha\ell} \tag{6.5}$$

where α is the absorption coefficient, which relates directly to the original negative intensities, and ℓ is a global value that includes the thickness of the emulsion layer and its density. This provides an estimation of the transmitted light of the emulsion layer with $I_i' = e^{-\alpha\ell} (1 - \mathcal{R}_e)^2 I_i$

6.5.2 Initial Physical Model

This work models an estimated virtual image of a negative when it is lit from the acetate side by combining the physical effects attributed to the acetate and emulsion layers together as light is transmitted through a film negative. Using the transmitted light from the acetate layer, $(1 - \mathcal{R}_a) I_i$, as the incident energy on the emulsion layer, the final transmitted energy estimate is $I = e^{-\alpha(\mathbf{x})\ell} (1 - \mathcal{R}_e)^2 \left((1 - \mathcal{R}_a)^2 k((d)) \right)$ where $k((d))$ represents the extinction coefficient between the acetate and emulsion layers.

Using this formulation for estimating the final transmission of light through the film negative, the intrinsic intensities of the undeteriorated negative represented by $\alpha(\mathbf{x})$ is estimated. By solving for $\alpha(\mathbf{x})$, these intensities are calculated, using the following equation:

$$\alpha(\mathbf{x}) = -\ell ln \left(\frac{I_{observed}(\mathbf{x})}{(1 - \mathcal{R}_e)^2 (1 - \mathcal{R}_a)^2 k(\mathbf{d})} \right) \tag{6.6}$$

Equation 6.6 is calculated for each pixel in the observed image to estimate the original intensities of the undistorted negative image.

The initial tests suggested that the results were highly dependent on the accuracy of the acquired surface geometry and the camera calibration. Figure 6.6(b) shows the estimated effect the negative has on the transmission of light. From observation, it can be seen that

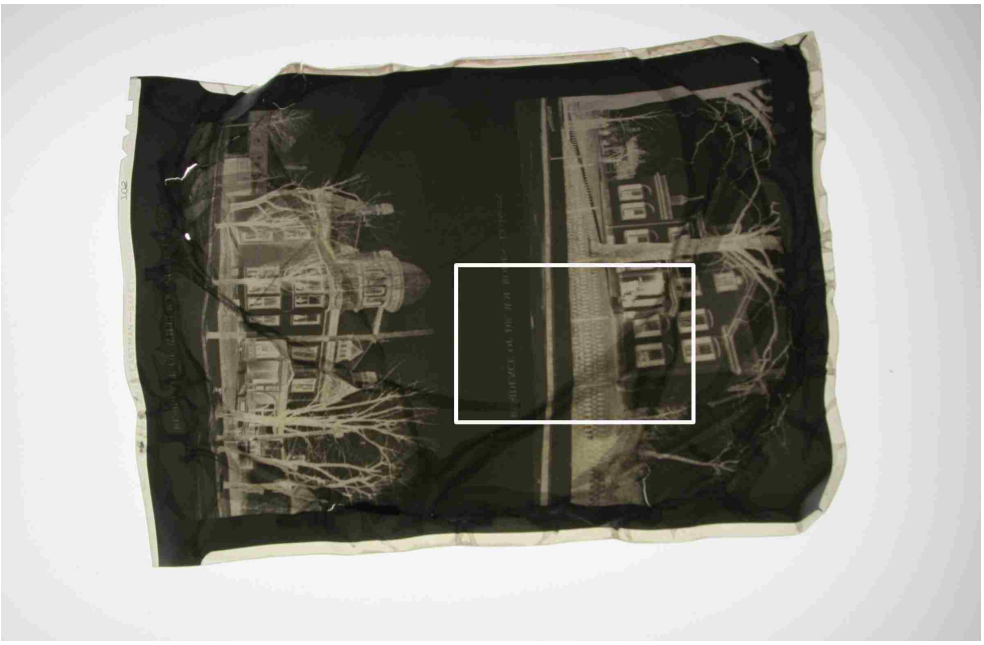


Figure 6.5: Cellulose acetate negative.

(a) Zoom-in of original image

(b) Estimated transmission through acetate and emulsion layers

(c) Initial correction attempt

Figure 6.6: Initial correction attempt.

the areas of reduced intensity in the negative image closely match that of the estimation. However, when directly applying the algorithm to the negative image the result is not quite what was expected. Figure 6.6(c) shows the result of the attempted correction. At first glance, it can be seen that the camera registration is not exact. This creates a misalignment with the negative image that creates error in the corrected areas. Also, much of the high-frequency noise was removed to reduce scanning error in the 3D scan. However, these first tests demonstrate that the accuracy of the scanned surface, especially high-frequency variations, are necessary to correctly model the light transmission.

These deficiencies in the shape acquisition lead to a dilemma in developing a restoration procedure. While scanning could be improved to add accuracy to the reconstruction of the film layers, there will always be features that lie under the precision of a scanner. Therefore, the research shifted to a novel acquisition paradigm using image-based scanning.

6.6 Image-Based Modeling of the Negative Restoration

Image-Based Modeling and Rendering has been widely research in recent years. One of the great benefits of working in the image-based realm is the reduction of requirements for complete three dimensional data that is either unavailable or unattainable.

The area of material model formulation, using image-based methods, and rendering has been a widely research area in computer graphics. Wang et al. [91] produced a real-time renderer of plant leaves that included global illumination effects. This work is of particular interest due to the application of an image-based acquisition technique to reconstruct the transmissivness and reflectivity of the leaves. Devices have been built to acquire the material properties of various types of documents and other materials. Gardner et al. [32] introduce a linear light source gantry that obtains a Bidirectional Reflectance Distribution Function (BRDF) of an object while providing depth and opacity estimates. Also, Mudge et al. [65] use a *light dome* to obtain reflectance properties of various materials. These works, and many others, show the possibilities of photo-realistic rendering of acquired objects. However, the negative restoration problem requires a much greater focus on obtaining the transmissive properties of material modeled as a Bidirectional Transmittance Distribution Function (BTDF). Furthermore, the purpose of this proposed work is not to realistically render a material in a synthetic scene, but to restore a negative to its original form by estimating the material changes caused by deterioration.

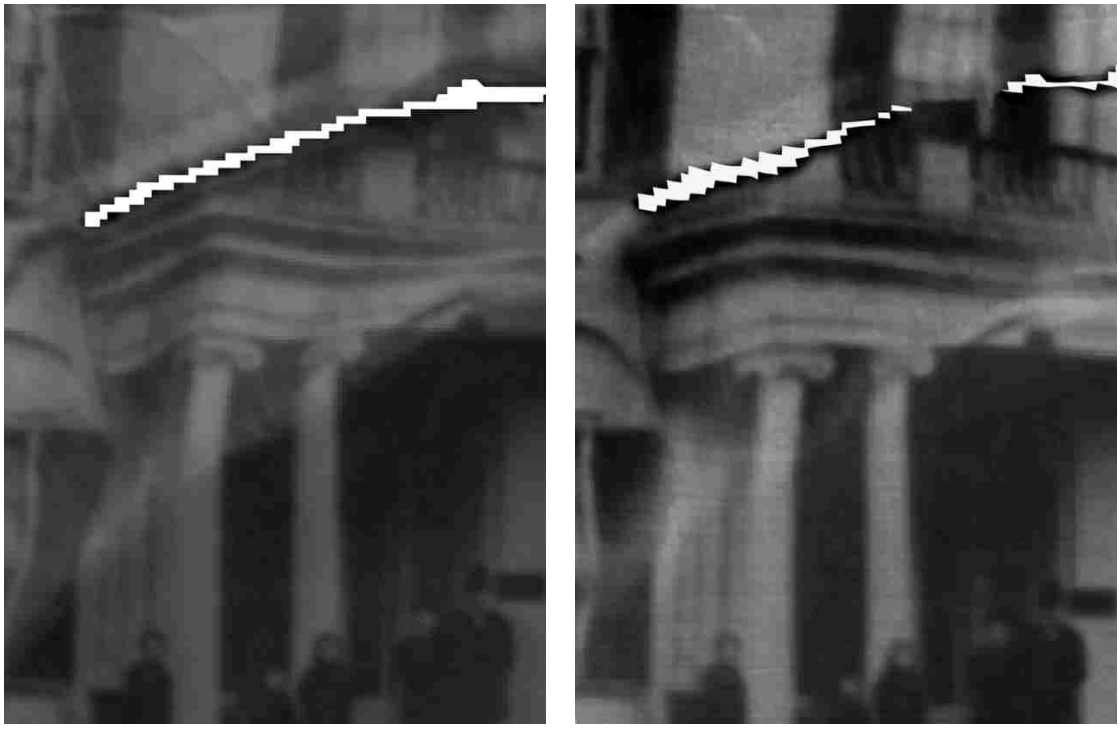
(a) Original negative image

(b) Original positive image

(c) Distortion estimation using (ΔX_0)

(d) Corrected image using variance Map ($\Delta\sigma$)

Figure 6.7: A photographic recording of a home in Lexington, Kentucky USA.



(a) A region with information loss due to cracking (marked with white)

(b) The restored region

Figure 6.8: A close-up of the original negative from Figure 6.7.

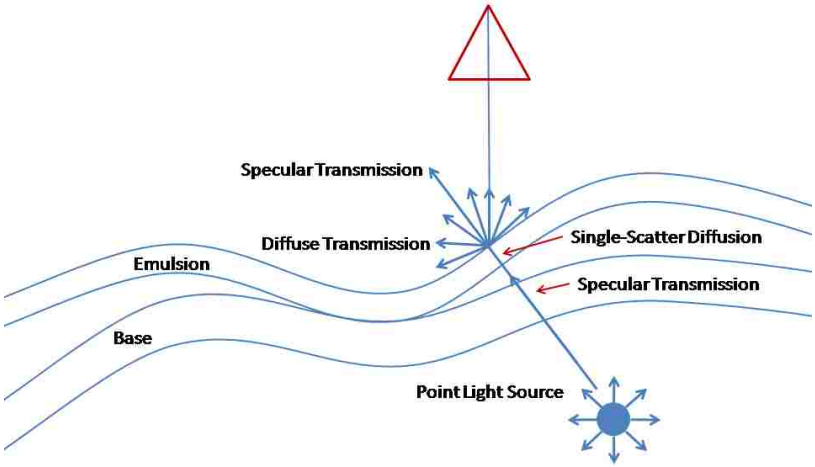


Figure 6.9: The transmission of light through a deteriorated photographic negative.

6.6.1 Physical model for approximated light transport

As explained in section 6.2, the two main components of a photonegative are the acetate and emulsion layers. For this model, to accurately model the transport of light through a photonegative, some assumptions must be made. First the acetate layer is a highly transparent material that remains highly specular. For this work, it is safe to ignore the reflective and refractive properties of the acetate layer and to only focus on direct transmission as shown in Figure 6.9. However, by ignoring these properties, the scanning results do not produce the same results as presented in Chapter 5. The attenuation is affected by the acetate layer, but $\Delta\sigma$ no produces a response highly tuned toward the emulsion layer. Second, the emulsion layer, by design, is highly isotropic which leads to a fair approximation of a diffuse transmitter. While this assumption works well for most of the emulsion layer, there are areas with very little density in the negative that appear nearly transparent. In these areas, the diffuse approximation leads to some error.

6.6.2 Time-evolving Gaussian Stimuli

This work has developed a novel image-based acquisition system to acquire the BTDF for transmissive documents. By using time evolving 1D Gaussian profiles as a structured-light pattern, it is possible to obtain non-linear least-squares estimates of the negative transmission function.

$$G(x : A_x, x_0, \sigma_x) = A_x e^{\frac{-(x-x_0)^2}{2\sigma_x^2}} \tag{6.7}$$

$$G(y : A_y, y_0, \sigma_y) = A_y e^{\frac{-(y-y_0)^2}{2\sigma_y^2}} \tag{6.8}$$

These stripes are observed twice to obtain relative difference when the scene is empty and when the negative is present on the scanner. Consequently, the three relative changes in the Gaussian parameters can be estimated as:

$$\Delta A = \sqrt{(A_x^2 - A_x'^2) + \gamma(A_y^2 - A_y'^2}} \tag{6.9}$$

$$\Delta\sigma = MAX\left(\left(\sigma_x - \sigma_x^{'}\right), \gamma\left(\sigma_y - \sigma_y^{'}\right)\right) \tag{6.10}$$

$$\Delta X_0 = \sqrt{(x_0^2 - x_0'^2) + \gamma(y_0^2 - y_0'^2}} \tag{6.11}$$

The parameters given in equations 6.9, 6.10, and 6.11 are estimated for each pixel in the scanned image of the negative. As a result, the system generates pixel-wise estimations for overall opacity (ΔA), content density ($\Delta\sigma$), and shape distortion (ΔX_0).

6.6.3 High-Dynamic Range Capture

Accuracy in the Gaussian fitting is reduced largely due to the compression of the negative opacity values to an 8-bit grayscale image representation. Therefore, High Dynamic Range Imaging is used to compensate for intensity compression during imaging.

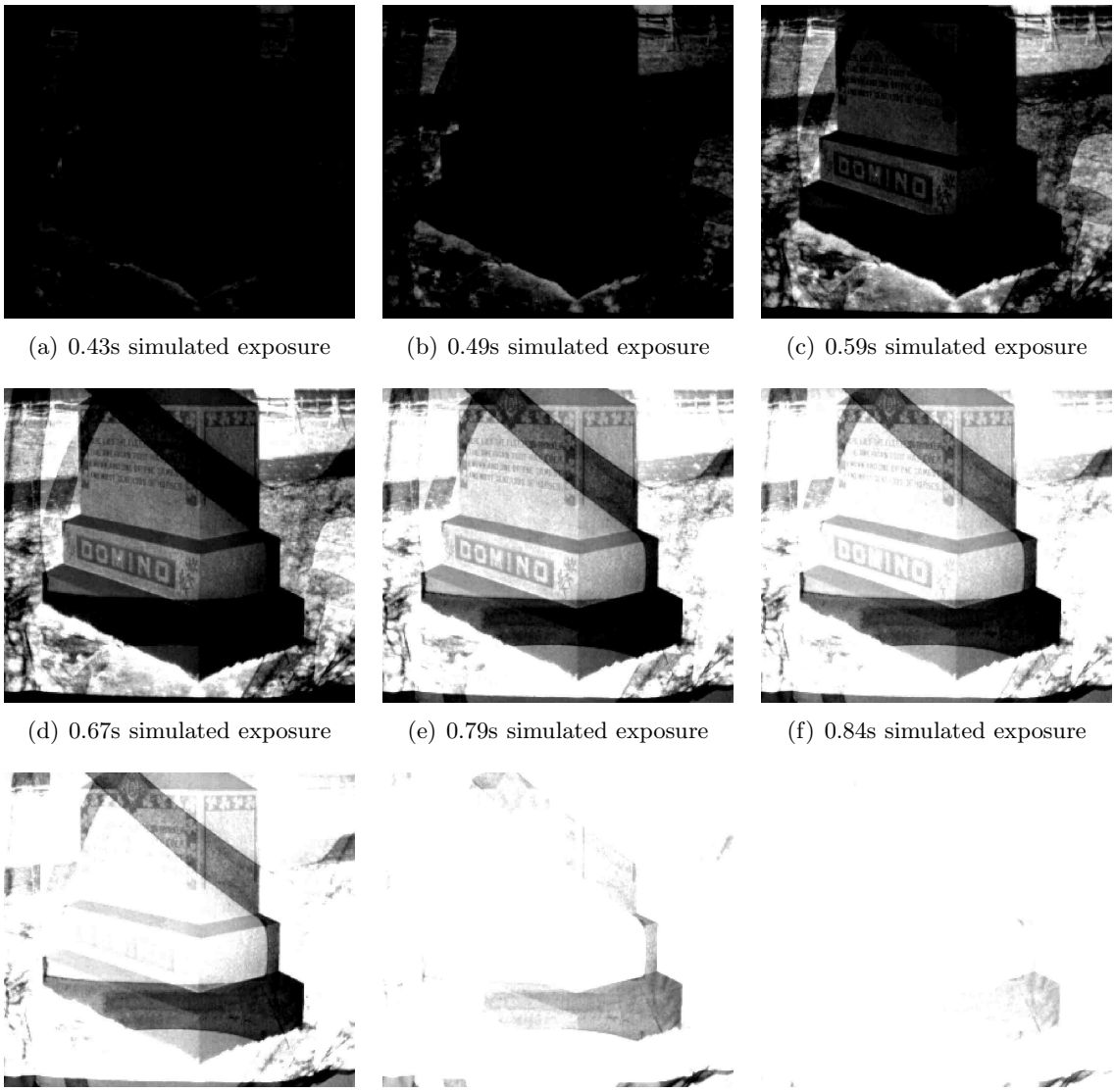
6.6.4 Detecting Deterioration with Virtual Exposures

As an application of the High-Dynamic Range Imaging, the channels formed during the delamination of the acetate film layers can be detected by simulating exposure changes directly from the estimated variance map $\Delta\sigma$. The radiance mapping function is used to estimate the virtual exposures.

$$I = f(E\Delta t) \tag{6.12}$$

However, instead of directly using the radiance value E, the variance value from the Gaussian estimation, $\Delta\sigma$, is used. Therefore, $I = f(\Delta\sigma\Delta t)$ where $f(x) = x + i_{shift}$ and i_{shift} is a global scaling of all the values.

Figure 6.10 shows the estimated images with various exposure times between 0.43–0.94 seconds. The key property to notice in the virtual exposures is the residual intensity reduction incurred by emulsion channels. The intensity effects are still obvious in the longer exposures where most image data has become saturated.



(a) 0.43s simulated exposure (b) 0.49s simulated exposure (c) 0.59s simulated exposure

(d) 0.67s simulated exposure (e) 0.79s simulated exposure (f) 0.84s simulated exposure

(g) 0.88s simulated exposure (h) 0.92s simulated exposure (i) 0.94s simulated exposure

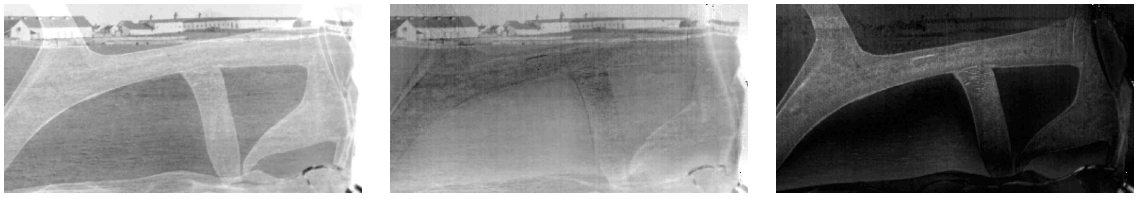
Figure 6.10: Simulated exposure settings using the acquired variance map $\Delta\sigma$.

The lack of light transmission through the channels is a consistent property among all severely deteriorated negatives. Therefore, channels can be detected in an analogous manner to the watershed method.

6.6.5 Virtual Restoration

The virtual restoration introduced here is implemented in two independent steps to correct the distortion caused by the acetate deterioration. First the intensities of the image are restored with photometric correction. Second, the skew of the document is corrected with geometric correction.

6.6.5.1 Photometric Correction



(a) Image obtained from ΔA (b) Image obtained from $\Delta\sigma$ (c) $|\Delta A| - |\Delta\sigma|$

Figure 6.11: Photometric correction from $\Delta\sigma$.

The photometric distortion occurs when there is a separation between the acetate and emulsion layers of the document. As described in Chapter 3, the photometric consistency of a document is imperative for digital document preservation. In the case of photo-negatives, channeling greatly reduces the printability of the content. However, unlike other photograph restoration techniques, this method does not assume that these areas with low intensity contain data loss. The main variation in intensity, due to channels, is caused by refraction, interreflection, and attenuation of light as it passes between the acetate and emulsion layers.

The intrinsic intensities of a document can be directly estimated by $\Delta\sigma$. Figure 6.11 shows the difference between Gaussian variations in amplitude ΔA and variance $\Delta\sigma$. The areas of larger difference have greater intensity in Figure 6.11(c). This allows a restoration of the original negative appearance including the difficult uniform intensity acquisition. Through empirical analysis, this result shows that the change in the σ parameter, the variance, is mostly effected by transmissive changes in the material. Consequently, optical distortions play a relatively small role in the variance of a Gaussian distribution.

6.6.5.2 Geometric Correction

As Chapter 4 discusses, shape distortions often impede the proper document digitization of damaged documents. These distortions are common with all document types when imaging a non-planar surface. Specifically concerning photo-negatives, the acetate shrinkage causes the emulsion layer to curl and wrinkle forming deep channels in the surface.

The distortion map that is generated by ΔX_0 is used to estimate the distorted shape of the negative. Using the integrable surface estimation discussed in Chapter 4, a polygonal mesh can be estimated up to scale. This estimated geometry is then used to a starting point for a mass-spring virtual flattening simulation as described in Chapter 5.

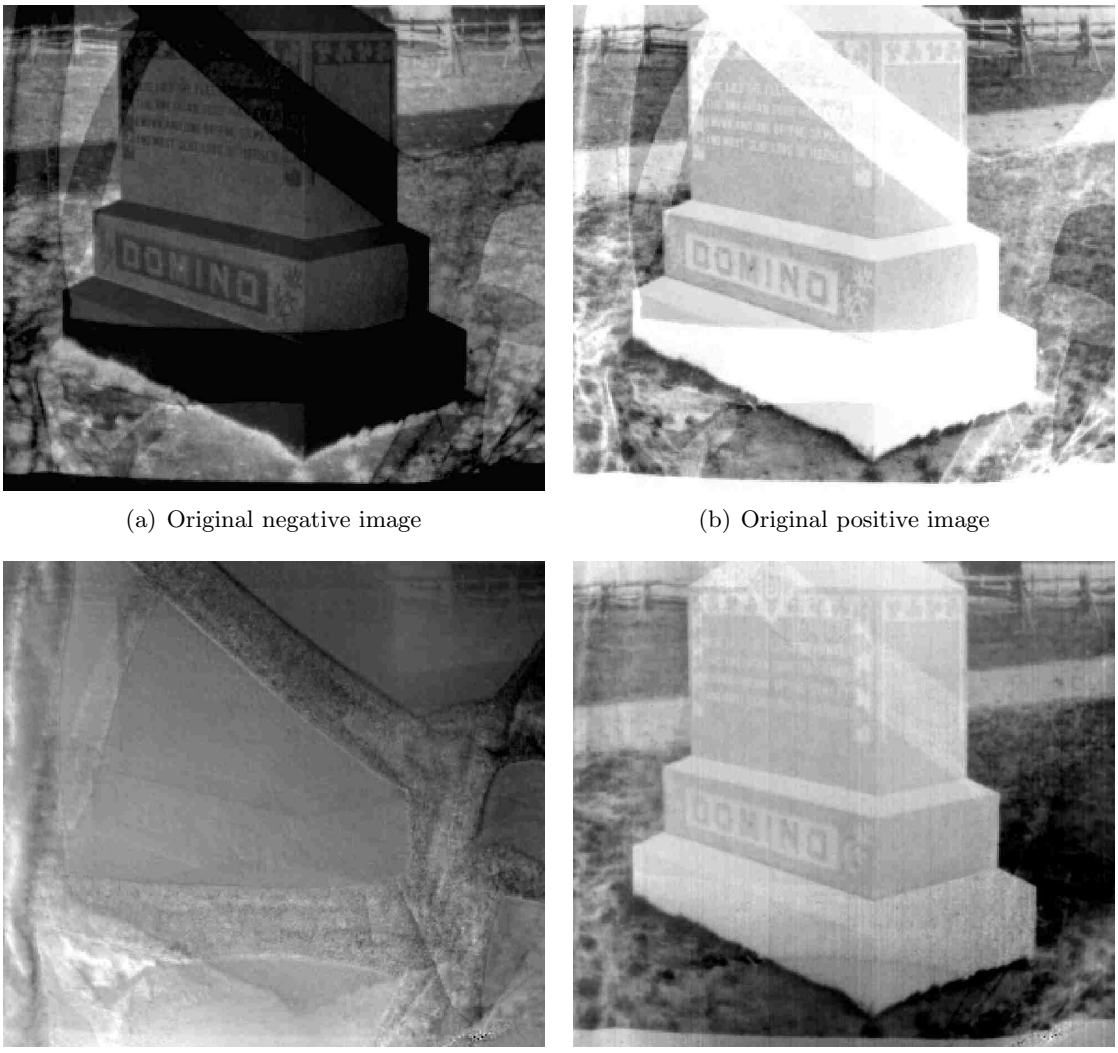
6.7 Results

The complete restoration system uses both photometric and geometric correction to automatically remove distortions introduced by the acetate deterioration. The photometrically corrected result for each of the negatives are shown in Figures 6.12(d), 6.7(d), 6.13(d), and 6.15(d). The photometric correction provides a fully automated way of restoring original negative content. Example distortion maps are shown in Figures 6.12(c), 6.7(c), 6.13(c), and 6.15(c). The estimated shape for the negatives in Figure 6.13 and 6.15 are shown in Figure 6.14 and 6.16 respectively.

A close-up of the resultant geometrically flattened document is shown in Fig. 6.8. In Figure 6.8(a) and Figure 6.8(b), a crack in the emulsion layer, which contains the photographic content, is marked as solid white. This area contains some information loss, where the emulsion gelatin has chipped away but much of the content remains. It can be seen through the geometric flattening process that both sides of the crack are brought back together as an added result of the restoration.

6.8 Conclusion

The novel acquisition and restoration system presented here provides the first known solution to digitize and restore deteriorated photographic negatives. The fully automated photometric correction process delivers restored content without the tedious post-processing step required in most photographic restoration techniques. The geometric correction utilizes a novel distortion estimation to facilitate surface reconstruction and follow-on virtual flattening. The successful restoration results of the multi-layer negatives opens new areas of application for other deteriorated multi-layer documents.



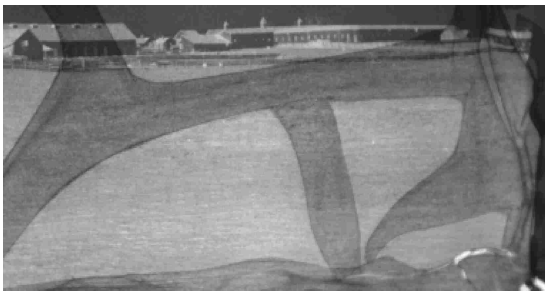
(a) Original negative image

(b) Original positive image

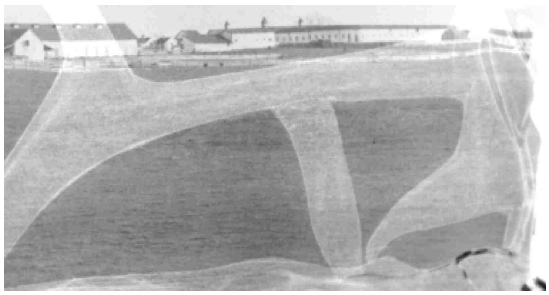
(c) Distortion Map (X_0)

(d) Corrected image using variance Map ($\Delta\sigma$)

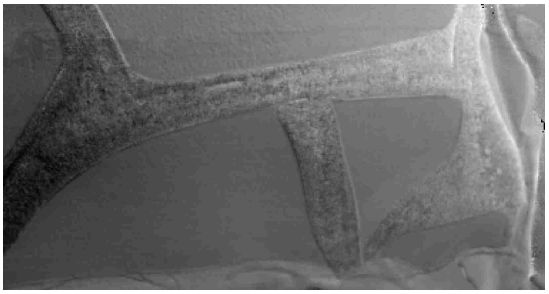
Figure 6.12: A photographic recording of a gravestone. The negative shows severe deterioration.



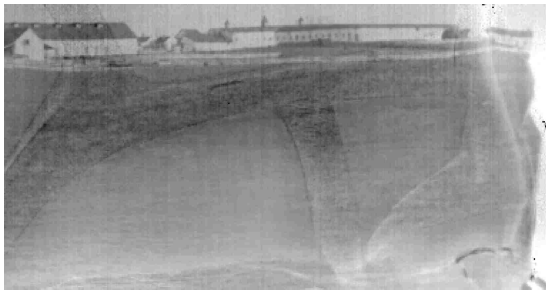
(a) Original negative image



(b) Original positive image

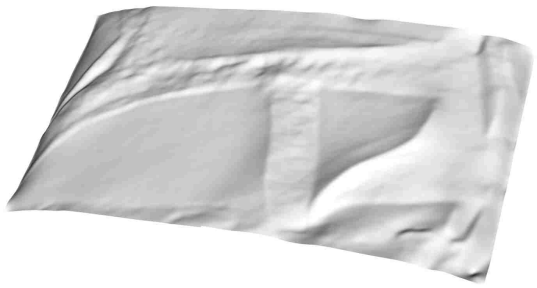


(c) Distortion Map (X_0)

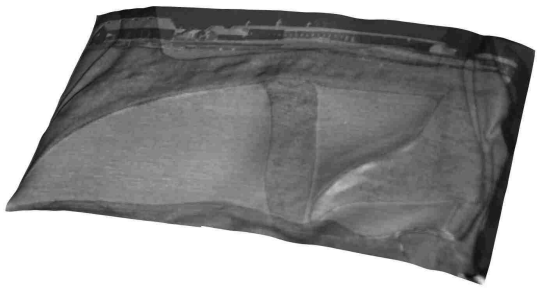


(d) Corrected image using variance Map ($\Delta\sigma$)

Figure 6.13: Example deteriorated negative.

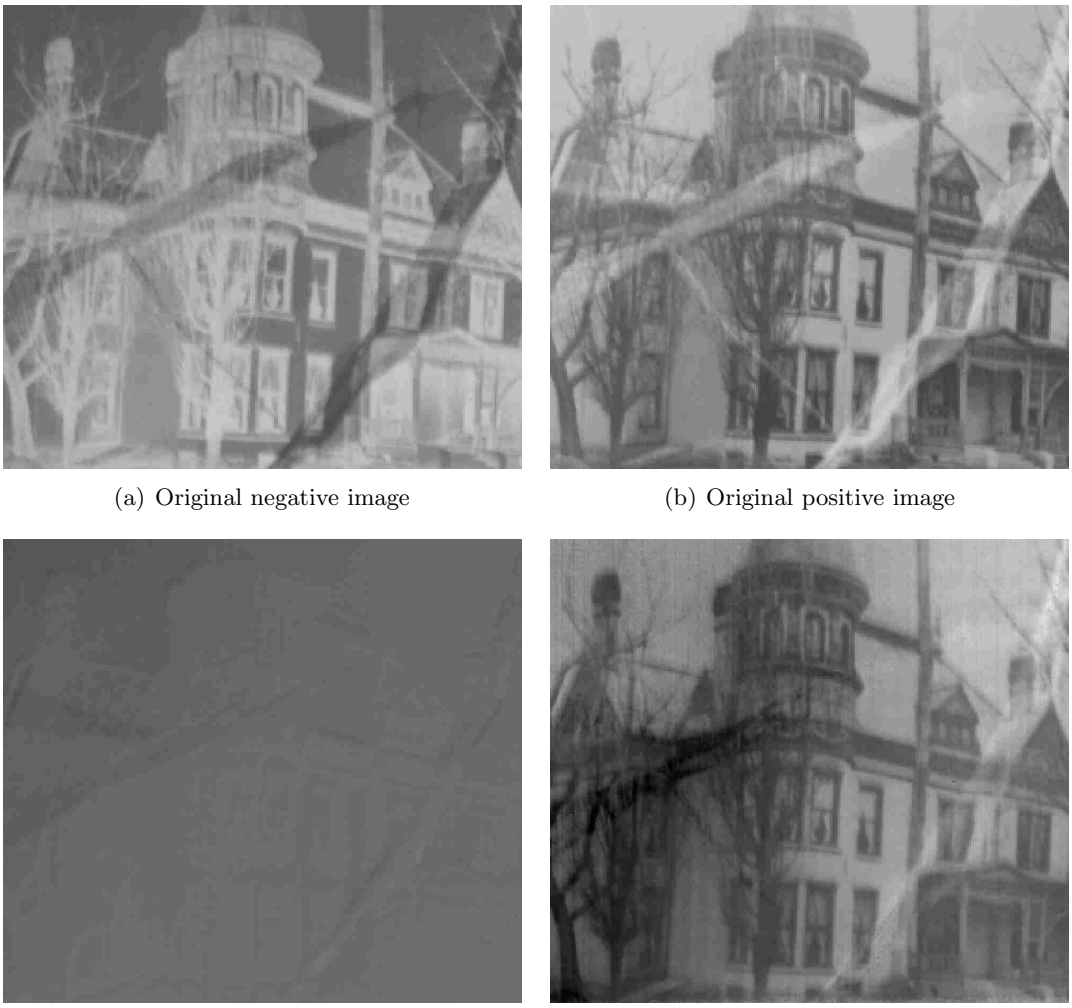


(a) Estimated Surface



(b) Estimated Surface with texture

Figure 6.14: The shape of the negative from Figure 6.13.



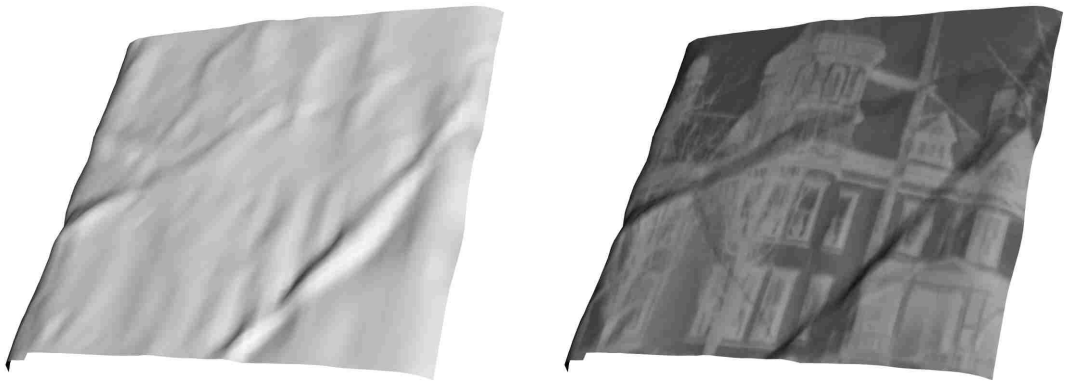
(a) Original negative image

(b) Original positive image

(c) Distortion Map (X_0)

(d) Corrected image using variance Map ($\Delta\sigma$)

Figure 6.15: Another architectural photographic record from Lexington, Kentucky USA.



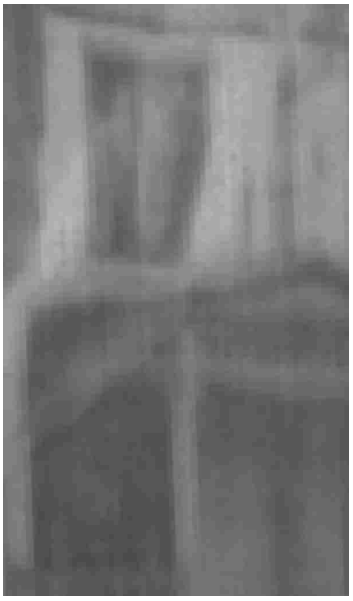
(a) Estimated Surface

(b) Estimated Surface with texture

Figure 6.16: The shape distortion estimation of negative from Figure 6.15.

(a) Photometrically corrected negative with geometric error.

(b) The negative with both photometric and geometric error corrected.



(c) A close-up of a heavily distorted window.

(d) The corrected result. Notice the window now returns to the original rectangular shape.

Figure 6.17: The geometric correction of negative from Figure 6.15.

Chapter 7

Conclusion

In conclusion, the developments presented in this dissertation provide new technologies that may be used directly in current large-scale digitization initiatives.

First, a novel point light source calibration technique was presented that can be performed by library staff. Then, a photometric correction technique which uses known illumination and surface properties to remove shading distortions in deteriorated document images can be automatically applied with geometric restoration.

Also, an image-based document scanner that utilizes the transmissivity of document substrates was introduced. This scanner was shown to accurately extract intrinsic document color information from one or both sides and obtain the document shape distortion information.

Finally, a restoration framework for damaged photographic negatives that corrects photometric and geometric distortions was developed. The novel acquisition and restoration system presented here provides the first known solution to digitize and restore deteriorated photographic negatives.

7.1 Summary

This work presents novel methods for digitizing and restoring cultural heritage documentation. Librarians, curators, and archivists need easy access to tools that enable new forms of virtual restoration without damaging priceless artifacts.

The novel light calibration technique presented in Chapter 2 provides a simple to use system for non-technical users, while providing a new method for users to estimate the effect that illumination plays in an imaged scene. Past limitations incurred by using special calibration equipment are eliminated by using a piece of folded white letter paper as a light probe. This is especially notable when working in conditions where foreign objects are not typically welcome (e.g. Library Archives).

The light calibration method presented here develops an accurate estimation of light position and color. Moreover, the method is developed as an interactive application that gives users real-time feedback of the estimated parameters. As the light source is calibrated, the estimated parameters are introduced into the rendered scene. This feedback gives users a direct understanding of the accuracy in the current illumination estimations. The calibration technique this dissertation provides gives digitization initiatives new means to estimate illumination properties in the scanning environment without reducing usability for expert library staff.

Expanding upon the developed light calibration technique, Chapter 3 introduces a new photometric correction technique for distorted documents. The estimated light source parameters are used to correct photometric error introduced when digitizing warped documents in non-optimal lighting. By calibrating a single point light source, digitization staff will effectively be able to produce the same results as an expert lighting system.

Photometric distortion of an imaged document is one of the most obvious sources of error. The human eye is well suited for discerning small variations in shading, so non-uniform illumination across a document is easily detectable. Consequently, using the presented photometric correction technique, this error will be removed. Therefore, once the color of the document has been restored, most of the distortion perceivable by the observer has been removed. However, the remaining geometric distortions may, in some cases, greatly reduce the legibility of the content.

Complete document restoration is performed by augmenting both the shading and shape of a document. In Chapter 4, methods for performing virtual geometric correction are reviewed. There currently exist a large number of techniques to estimate document distortion and subsequently correct the distortion. Some of these methods make assumptions concerning the composition and layout of a document to achieve faster acquisition and restoration. Other methods perform complete three-dimensional surface reconstructions to attempt to model all forms of surface deviations. For each of these methods, the goal is to obtain an accurate representation of the surface distortions in the document.

Once the distortion is estimated, restoration techniques can be applied to correct the shape distortions. This work focuses on physically-based systems to model the warped surface in order to return the document to the original planar state. However, for many forms of deterioration, achieving the three dimensional accuracy needed to complete restoration is quite difficult. This limitation is addressed by developing a new document scanner paradigm.

In Chapter 5, a new acquisition system is introduced that performs image-based geometric and photometric distortion estimation. By relying on the assumption that most document substrates have some form of opacity, visible light may be transmitted through the document to obtain various properties of the text. These estimated parameters provide invaluable information concerning various document types including shape distortion information and intrinsic color information. The document properties allow for various applications to be performed on the digitized content such as one-pass duplex-sided scanning, virtual flattening of warped documents, and photometric correction of documents with inconsistent shading.

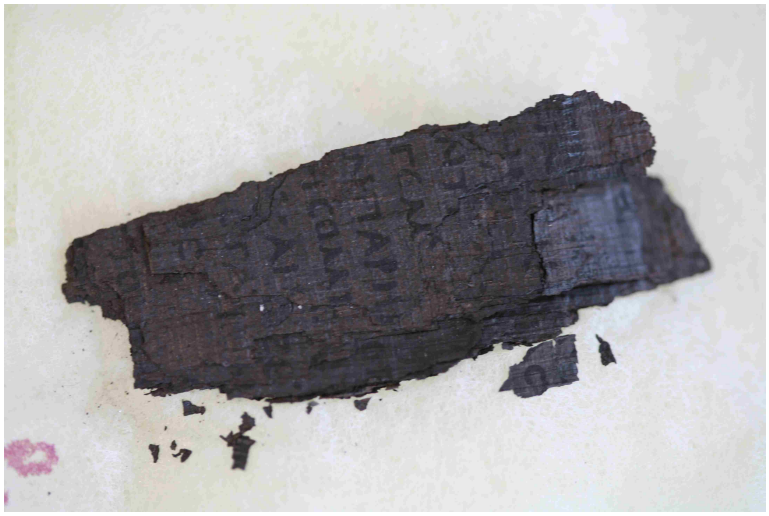
As this work evolved, it became apparent that other document types would now become scannable using the transmissive scanner. Moreover, documents types that do not fall under the traditional textual forms also need to be considered in the digitization and restoration framework. In Chapter 6, documents that were developed as intrinsically transmissive objects are explored. Historic photonegatives contain equivalent importance to culture and heritage as most paper-based documents. However, like many of the documents discussed here, deterioration has reduced the transmissivity of large areas of the negatives which causes distortions during digitization and printing.

7.2 Future Work

The introduction of a transmissive light document scanner in this dissertation has established a new paradigm for document digitization and restoration. The document scanner has been successfully demonstrated to scan and restore single-layer documents with both single and duplex sided content. Documents with multiple layers of substrate have been shown to be compatible with the scanning process as well. Photographic negatives which contain a single layer of content and multiple layers of translucent materials can be digitally restored.

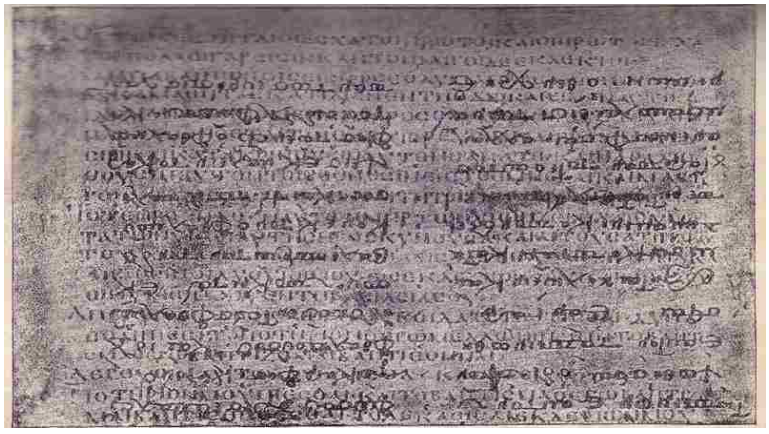
7.2.1 Accessing unique document compositions

The progress made in scanning multi-layer documents has opened new possibilities for previously in-accessible texts. There are many historical documents that, for various reasons, have many layers adhered together. For example, large collections of documents exist that were carbonized in the aftermath of volcanic eruptions. A notable example of this deterioration are Herculaneum scrolls, Figure 7.1(a), made illegible by the 79AD eruption of Mount Vesuvius.



(a) Example of carbonized papyrus from Herculaneum (79AD).

(b) Birchbark letter (12th Century).



(c) Codex Ephraemi showing Matthew 20:16-23 [29]

Figure 7.1: Example inaccessible documents.

A more common example comes from medieval codices. During the time when paper itself was a commodity, documents would be cut and reused to form book backings. Consequently, multiple layers of texts could be glued together to create rigid documents.

These documents may have as little as two or as many as five or more layers irreversibly stuck together. While there exist efforts to physically separate layers to extract the content, it may take years to completely restore one text. Moreover, the content discovery process typically leaves the original document destroyed. Consequently, this form of destructive exploration is not acceptable for the vast majority of conserved ancient texts.

Other documents materials generate very well preserved results, but without suitable digitization methods. Birchbark texts, as seen in Fig. 7.1(b), show an example letter written on tree bark.

Another document type of extreme importance are palimpsests. These parchment-based documents were reused multiple times for often disparate content. The surface of a page was abraded to remove the visible ink from the substrate. Therefore, scribes would then start with almost blank pages for a new volume.

However, it has been discovered that often times the original text still remains in some fashion on the substrate. For instance, the Archimedes palimpsest contains original works of Archimedes, but has been overwritten with a 13th century prayerbook. Research has been performed by Knox et al. [51] to incorporate multi-spectral imaging into a segmentation algorithm to separate the layers of text. All of the work thus far on palimpsests has focused on reflective energy imaging using visible light or X-ray.

Using a technique similar to the scanning method introduced in Chapter 5, may present a novel digitization method for many of these multi-layer documents. Documents that retain some transparency to visible light through the layers of substrate are accessible with the presented method. As an extension of the duplex-sided scanning, the results presented here already demonstrate effectiveness in separating multiple layers of ink.

7.2.2 Increasing Accuracy

As technology develops, new hardware may be directly incorporated in the transmissive scanning framework. Since the scanning systems presented here operate without a constraint on display resolution or imaging resolution, higher resolution technologies may be used without any change to the software scanner. This will directly address error introduced by lack of spatial resolution in scanning.

Bibliography

[1] TAPPI T 425. Opacity of paper (15/d geometry, illuminant a/2 degrees, 89reflectance backing and paper backing), test method t 425 om-06.

[2] Thomas Acquaviva. Reproduction apparatus transmissivity sensor system for duplex documents. United States Patent 5,710,957, Jan 1998.

[3] Neil G. Alldrin and David J. Kriegman. A planar light probe. In *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2324–2330, Washington, DC, USA, 2006. IEEE Computer Society.

[4] E. Ardizzone, H. Dindo, U. Maniscalco, and G. Mazzola. Damages of digitized historical images as objects for content based applications. In *Proceedings of the 14th European Signal Processing Conference, Eusipco 2006*, Florence, Italy, September 2006. EURASIP.

[5] Peter N. Belhumeur, David J. Kriegman, and Alan L. Yuille. The bas-relief ambiguity. *Int. J. Comput. Vision*, 35(1):33–44, 1999.

[6] M. Ben-Ezra and S.K. Nayar. What does motion reveal about transparency? In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1025–1032 vol.2, 13-16 Oct. 2003.

[7] M. Bertalmio, L. Vese, and G. Sapiro S. Osher. Simultaneous structure and texture image inpainting. In *Computer Vision and Pattern Recognition*, volume 2, pages 707–12. IEEE Computer Society, June 2003.

[8] Marcelo Bertalmio, Guillermo Sapiro, Vicent Caselles, and Coloma Ballester. Image inpainting. In Kurt Akeley, editor, *Siggraph 2000, Computer Graphics Proceedings*, pages 417–424. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, 2000.

[9] Michael Brown and W. Brent Seales. Image restoration of arbitrarily warped documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10):1295–1306, October 2004.

[10] Michael S. Brown and Charles J. Pisula. Conformal deskewing of non-planar documents. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1*, pages 998–1004, Washington, DC, USA, 2005. IEEE Computer Society.

[11] Michael S. Brown and W. Brent Seales. Document restoration using 3d shape: A general deskewing algorithm for arbitrarily warped documents. In *ICCV*, pages 367–375, 2001.

[12] M.S. Brown and Y.C. Tsoi. Geometric and shading correction for images of printed materials using boundary. *IEEE Trans. Image Proc.*, 15(6):1544–1554, June 2006.

[13] Vittoria Bruni, Andrew Crawford, Anil C. Kokaram, and Domenico Vitulano. Digital removal of blotches with variable semi-transparency using visibility laws. In *BVAI*, pages 254–263, 2007.

[14] H. Cao, X.Q. Ding, and C. Liu. A cylindrical surface model to rectify the bound document image. In *ICCV '03: Proceedings of the Nineth IEEE International Conference on Computer Vision*, pages 228–233, 2003.

[15] Subrahmanyan Chandrasekhar. *Radiative transfer*. Dover Publications, 1960. Unabridged and slightly revised verson of the work first published in 1950.

[16] A. Chaudhuri and S. Chaudhuri. Robust detection of skew in document images. *IEEE Trans. on Image Proc.*, 6(2):344–349, 1997.

[17] Seong Ik Cho, Hideo Saito, and Shinji Ozawa. A divide-and-conquer strategy in shape from shading problem. In *CVPR '97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, page 413, Washington, DC, USA, 1997. IEEE Computer Society.

[18] Yung-Yu Chuang, Douglas E. Zongker, Joel Hindorff, Brian Curless, David H. Salesin, and Richard Szeliski. Environment matting extensions: towards higher accuracy and real-time capture. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 121–130, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[19] Frdric Courteille, Alain Crouzil, Jean-Denis Durou, and Pierre Gurdjos. Shape from Shading for the Digitization of Curved Documents. *Machine Vision and Applications*, 18(5):301–316, October 2007.

[20] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 13(9):1200–1212, 2004.

[21] Brian Curless and Steven Seitz. 3D photography. *ACM Siggraph '00 Course Notes Course No. 19*, August 2000.

[22] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, August 1997.

[23] A. Doncescu, A. Bouju, and V. Quillet. Former books digital processing: image warping. In *DIA '97: Proceedings of the 1997 Workshop on Document Image Analysis*, page 5, Washington, DC, USA, 1997. IEEE Computer Society.

[24] M. S. Drew. Photometric stereo without multiple images. In B. E. Rogowitz and T. N. Pappas, editors, *Proc. SPIE Vol. 3016, p. 369-380, Human Vision and Electronic Imaging II*, pages 369–380, June 1997.

[25] Robert T. Frankot and Rama Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 10(4):439–451, 1988.

[26] O. Franzke and O. Deussen. Accurate graphical representation of plant leaves. In *International Symposium on Plant Growth Modeling, Simulation, Visualization, and their Applications*, 2003.

[27] Jeppe Revall Frisvad, Niels Jorgen Christensen, and Peter Falster. Efficient light scattering through thin semi-transparent objects. In *GRAPHITE '05: Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 135–138, New York, NY, USA, 2005. ACM.

[28] Darya Frolova, Denis Simakov, and Ronen Basri. Accuracy of spherical harmonics approximation for images of lambertian objects under far and near lighting. In *ECCV 2004*, pages 574–587, 2004.

[29] Henry Frowde. *The S.S. Teacher's Edition: The Holy Bible*. University of Oxford, 1896.

[30] Nathan Funk and Yee-Hong Yang. Using a raster display for photometric stereo. In *CRV '07: Proceedings of the Fourth Canadian Conference on Computer and Robot Vision*, pages 201–207, Washington, DC, USA, 2007. IEEE Computer Society.

[31] B. Ganapol, L. Johnson, P. Hammer, C. Hlavka, and D. Peterson. Leafmod: A new within-leaf radiative transfer model. *Remote Sensing of Environment*, 63:182–193, 1998.

[32] Andrew Gardner, Chris Tchou, Tim Hawkins, and Paul Debevec. Linear light source reflectometry. *ACM Trans. Graph.*, 22(3):749–758, 2003.

[33] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[34] Google, Inc. Cic/google book search project frequently asked questions.

[35] Y. M. Govaerts, S. Jacquemoud, M. M. Verstraete, and S. L. Ustin. Three-dimensional radiation transfer modeling in a dicotyledon leaf. *Applied Optics*, 35:6585–6598, November 1996.

[36] Nail A. Gumerov, Ali Zandifar, Ramani Duraiswami, and Larry S. Davis. 3d structure recovery and unwarping of surfaces applicable to planes. *Int. J. Comput. Vision*, 66(3):261–281, 2006.

[37] Kevin J. Haded, Mark G. Brook, Mark E. Tellam, John F. Omvik, and Joseph A. Wheeler. Scanning system for scanning reflective and transmissive images. United States Patent 5,691,824, Nov 1997.

[38] Kevin J. Haded, Mark G. Brook, Mark E. Tellam, John F. Omvik, and Joseph A. Wheeler. Scanning system for scanning reflective and transmissive images. United States Patent 6,151,139, Nov 2000.

[39] Loi Han. Multi-resolution transmissive and reflective optical scanner. United States Patent 5,907,411, May 1999.

[40] Kenji Hara, Ko Nishino, and Katsushi Ikeuchi. Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4):493–505, 2005.

[41] Harry Ransom Center: The University of Texas at Austin. The digital gutenberg project.

[42] Aaron Hertzmann and Steven M. Seitz. Shape and materials by example: A photometric stereo approach. volume 01, page 533, Los Alamitos, CA, USA, 2003. IEEE Computer Society.

[43] B.K.P. Horn. Shape from shading: a method for obtaining the shape of a smooth opaque object from one view. *MIT Project Mac Technical Report TR-79*, 1970.

[44] David G. Horvath. *The Acetate Negative Survey: Final Report.* The University of Louisville, Ekstrom Library, Photographic Archives, Louisville, KY 40292, February 1987.

[45] Yin-Chun Huang and Chien-Liang Yeh. Image scanning device with switching mechanism for selectively scanning reflective and transparent document. United States Patent 6,445,480, Sep 2002.

[46] Akihiko Iketani, Tomokazu Sato, Sei Ikeda, Masayuki Kanbara, Noboru Nakajima, and Naokazu Yokoya. Video mosaicing for curved documents based on structure from motion. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 391–396, Washington, DC, USA, 2006. IEEE Computer Society.

[47] Y. Iwahori, H. Sugie, and N. Ishii. Reconstructing shape from shading images under point light source illumination. In *IEEE ICPR 90*, volume I, pages 83–87, 1990.

[48] M. Kashimura, T. Nakajima, N. Onda, Hideo Saito, and S. Ozawa. Practical introduction of image processing technology to digital archiving of rare books. In *Internationa Conference on Signal Processing Application Technology (ISPAT98)*, October 1998.

[49] Byungil Kim and Peter Burger. Depth and shape from shading using the photometric stereo method. *CVGIP: Image Underst.*, 54(3):416–427, 1991.

[50] Edwin Klijn and Yola de Lusenet. *In the Picture. Preservation and digitisation of European photographic collections.* European Commission on Preservation and Access, 2000.

[51] Keith Knox, Charles Dickinson, Lichao Wei, Jr. Roger L. Easton, and Robert Johnston. Multispectral imaging of the archimedes palimpsest. pages 206–210, 2001.

[52] Jan J. Koenderink and Andrea J. Van Doorn. The generic bilinear calibration-estimation problem. *Int. J. Comput. Vision*, 23(3):217–234, 1997.

[53] Inc Kruger. Paper specifications, 2007.

[54] J. Lambert. Photometria sive de mensura et gradibus luminus colorum et umbrae. *Eberhard Klett*, 1760.

[55] George V. Landon and W. Brent Seales. Petroglyph digitization: enabling cultural heritage scholarship. *Machine Vision and Applications*, 17(6):361–371, December 2006.

[56] Brian Lavoie, Lynn Silipigni Connaway, and Lorcan Dempsey. Anatomy of aggregate collections: The example of google print for libraries. *D-Lib Magazine*, 11(9), September 2005.

[57] Jian Liang, Daniel DeMenthon, and David Doermann. Flattening curved documents in images. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2*, pages 338–345, Washington, DC, USA, 2005. IEEE Computer Society.

[58] Library of Congress. Library of congress manuscript digitization demonstration a project: Final report. October 1998.

[59] Shijian Lu and Chew Lim Tan. Thresholding of badly illuminated document images through photometric correction. August 2007.

[60] Peter Lyman and Hal R. Varian. How much information. Technical report, 2003.

[61] S. Mann and R. W. Picard. On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. In *The Society for Imaging Science and Technology 48th Annual Conference*, pages 422–428, May 1995.

[62] N. J. Mitra, A. Nguyen, and L. Guibas. Estimating surface normals in noisy point cloud data. *Special issue of Int. J. Computational Geometry and its Applications*, 14(4–5):261–276, 2004.

[63] Nigel J. W. Morris and Kiriakos N. Kutulakos. Dynamic refraction stereo. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision*, pages 1573–1580, Washington, DC, USA, 2005. IEEE Computer Society.

[64] Nigel J. W. Morris and Kiriakos N. Kutulakos. Reconstructing the surface of inhomogeneous transparent scenes by scatter trace photography. In *Proc. 11th Int. Conf. Computer Vision*, October 2007.

[65] Mark Mudge, Jean-Pierre Voutaz, Carla Schroer, and Marlin Lum. Reflection transformation imaging and virtual representations of coins from the hospice of the grand st. bernard. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, pages 29–39, 2005.

[66] B. Ophir and D. Malah. Show-through cancellation in scanned images using blind source separation techniques. In *ICIP07*, pages III: 233–236, 2007.

[67] M. Pilu. Undoing page curl distortion using applicable surfaces. In *IEEE Computer Vision and Pattern Recognition Conference*, December 2001.

[68] Mark W. Powell, Sudeep Sarkar, and Dmitry Goldgof. A simple strategy for calibrating the geometry of light sources. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(9):1022–1027, 2001.

[69] Emmanuel Prados and Olivier Faugeras. Shape from shading: a well-posed problem? In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, California*, volume II, pages 870–877. IEEE, June 2005.

[70] J. M. Reilly. Ipi storage guide for acetate film. 1993.

[71] Charles Rollin. Silas Andrus, 1815.

[72] Mark Roosa. Care, handling, and storage of photographs. 1992.

[73] Stefan Roth and Michael J. Black. Fields of experts: A framework for learning image priors. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2*, pages 860–867, Washington, DC, USA, 2005. IEEE Computer Society.

[74] John S. Rubley and Loi N. Han. Transmissive/reflective optical scanning apparatus. United States Patent 5,574,274, Nov 1996.

[75] Imari Sato, Yoichi Sato, and Katsushi Ikeuchi. Illumination from shadows. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(3):290–300, 2003.

[76] Y.Y. Schechner, S.K. Nayar, and P.N. Belhumeur. A Theory of Multiplexed Illumination. In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 808–815, Oct 2003.

[77] W. Brent Seales and Yun Lin. Digital restoration using volumetric scanning. In *JCDL*, pages 117–124, 2004.

[78] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual photography. *ACM Trans. Graph.*, 24(3):745–755, 2005.

[79] G. Sharma. Show-through cancellation in scans of duplex printed documents. *IEEE Transactions on Image Processing*, 10(5):736–754, 2001.

[80] F. Stanco, G. Ramponi, and A. de Polo. Towards the automated restoration of old photographic prints: a survey. *EUROCON 2003. Computer as a Tool. The IEEE Region 8*, 2:370–374 vol.2, 22-24 Sept. 2003.

[81] F. Stanco, L. Tenze, G. Ramponi, and A. de Polo. Virtual restoration of fragmented glass plate photographs. *Electrotechnical Conference, 2004. MELECON 2004. Proceedings of the 12th IEEE Mediterranean*, 1:243–246 Vol.1, 12-15 May 2004.

[82] Mingxuan Sun, Ruigang Yang, Yun Lin, George V. Landon, W. Brent Seales, and Michael S. Brown. Geometric and photometric restoration of distorted documents. In *Proceedings of the 10th IEEE International Conference on Computer Vision*, volume 2, pages 1117– 1123, October 2005.

[83] Yusuke Suzuki, Atsushi Yamashita, and Toru Kaneko. Correction of geometric and photometric distortion of document images using a stereo camera system. In *Proceedings of IAPR Conference on Machine Vision Applications (MVA2007)*, pages 215–218, May 2007.

[84] Takeshi Takai, Koichiro Niinuma, Atsuto Maki, and Takashi Matsuyama. Difference sphere: An approach to near light source estimation. *cvpr*, 01:98–105, 2004.

[85] Chew Lim Tan, Li Zhang, Zheng Zhang, and Tao Xia. Restoring warped document images through 3d shape modeling. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(2):195–208, 2006.

[86] The Project Gutenberg Organization.

[87] Anna Tonazzini, Luigi Bedini, and Emanuele Salerno. Independent component analysis for document restoration. *International Journal on Document Analysis and Recognition*, 7(1):17–27, March 2004.

[88] Yau-Chat Tsoi and Michael S. Brown. Multi-view document rectification using boundary. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, June 2007.

[89] United Nations Educational Scientific and Cultural Organization. Division of cultural heritage programme 2004-2005. 2005.

[90] Toshikazu Wada, Hiroyuki Ukida, and Takashi Matsuyama. Shape from shading with interreflections under a proximal light source: Distortion-free copying of an unfoldedbook. *Int. J. Comput. Vision*, 24(2):125–135, 1997.

[91] Lifeng Wang, Wenle Wang, Julie Dorsey, Xu Yang, Baining Guo, and Heung-Yeung Shum. Real-time rendering of plant leaves. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses*, page 5, New York, NY, USA, 2006. ACM Press.

[92] Robert J. Woodham. *Photometric method for determining surface orientation from multiple images*. MIT Press, Cambridge, MA, USA, 1989.

[93] Atsushi Yamashita, Atsushi Kawarago, Toru Kaneko, and Kenjiro T. Miura. Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system. In *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 1*, pages 482–485, Washington, DC, USA, 2004. IEEE Computer Society.

[94] Li Zhang, Andy M. Yip, and Chew Lim Tan. A restoration framework for correcting photometric and geometeric distortions in camera-based document images. October 2007.

[95] Li Zhang, Yu Zhang, and Chew Tan. An improved physically-based method for geometric restoration of distorted document images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, November 2007.

[96] Yufei Zhang and Yee-Hong Yang. Multiple illuminant direction detection with application to image synthesis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(8):915–920, 2001.

[97] Zheng Zhang, Chew Lim Tan, and Liying Fan. Estimation of 3d shape of warped document surface for image restoration. In *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 1*, pages 486–489, Washington, DC, USA, 2004. IEEE Computer Society.

[98] Qinfen Zheng and Rama Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):680–702, 1991.

[99] Douglas E. Zongker, Dawn M. Werner, Brian Curless, and David H. Salesin. Environment matting and compositing. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 205–214, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

Vita

George V. Landon, Jr. was born July 11, 1981 to George and Kathleen Landon. He was born and raised in Alexandria, Kentucky within Campbell County. George received a BS in Computer Science from the University of Kentucky, Lexington, Kentucky in December 2002. He is currently working as an Assistant Professor in Computer Science at Eastern Kentucky University, Richmond, KY. Professional publications

- George V. Landon and W. Brent Seales. Petroglyph digitization: enabling cultural heritage scholarship. *Machine Vision and Applications*. 17:6, pages 361-371, Springer Berlin / Heidelberg, December, 2006.

- George V. Landon, Yun Lin, and W. Brent Seales. Towards Automatic Photometric Correction of Casually Illuminated Documents, Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on , pp.1-8, 17-22 June 2007

- Saikat Chakrabarti, **George V. Landon**, Mukesh Singhal. Graphical Passwords: Drawing a Secret with Rotation as a New Degree of Freedom, in Proc. of the Fourth IASTED Asian Conference on Communication Systems and Networks (AsiaCSN 2007), Phuket, Thailand, April 2-4 2007, ACTA Press, pp. 561-173.

- M. Sun and R. Yang and L. Yun and **George Landon** and W. B. Seales and M. S. Brown. Geometric and Photometric Restoration of Distorted Documents. In International Conference on Computer Vision 2005, volume II, pages 1117-1123, 2005.

- George V. Landon and W. Brent Seales. Building and Visualizing 3D Textured Models for Caribbean Petroglyphs. In 21st Congress of The International Association of

Caribbean Archaeology (Trinidad and Tobago), July 2005.

- George V. Landon and W. Brent Seales. New methods for rock art recording and virtual analysis. In Michael A. Cinquino, Michele H. Hayward, and Lesley-Gail Atkinson, editors, *Caribbean Rock Art*. The University of Alabama Press. In Press.

- George V. Landon and W. Brent Seales. New Methods for Rock Art Recording and Virtual Anaylsis. *Symposium on Latin American Rock Art: Conservation and Research Status.* 2006 meeting of the Society of American Archaeology.

- W. Brent Seales and **George V. Landon**. The Museum and the Media Divide: Building and Using Digital Collections at the Instituto de Cultura Puertorriqueña. In *D-Lib Magazine*, volume 11, May 2005.

- Yun Lin, **George V. Landon**, and W. Brent Seales. Developing a Framework for Acquisition, Preservation, and Dissemination of 3D Documents. Technical Report TR-463-06, Department of Computer Science, University of Kentucky, Lexington, KY, December 2006.