

Image Based Relighting of Cultural Artifacts

THÈSE N° 6990 (2016)

PRÉSENTÉE LE 3 JUIN 2016

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE DE COMMUNICATIONS AUDIOVISUELLES
PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Niranjan THANIKACHALAM

acceptée sur proposition du jury:

Prof. M. Pauly, président du jury
Prof. M. Vetterli, Dr P. Prandoni, directeurs de thèse
Prof. H. Maître, rapporteur
Prof. H. P. A. Lensch, rapporteur
Prof. S. Süssstrunk, rapporteuse



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2016

Acknowledgements

I would like to express my deepest gratitude to my two advisors, Prof. Martin Vetterli and Dr. Paolo Prandoni. Martin, thank you very much for giving me the opportunity to do my PhD at LCAV. Thanks a ton for your patience, continued optimism and support, as I attempted to find my footing and for guiding me through this beautiful adventure, while allowing a high degree of independence in research. Your infectious enthusiasm for research serves as a great source of inspiration for your students. Paolo, I absolutely enjoyed your hands-on, practical approach to problem solving, which I believe has influenced me to a great extent. Thanks for magically transforming any piece of convoluted passage that I produced into beautiful text and for all the interesting, thought provoking discussions. Most importantly thanks for helping me look at the big picture, early on in the project. I also sincerely thank Dr. Loïc Baboulaz for his valuable collaboration all these years. While not my official advisor, you guided me in very much the same way. Thanks a lot for the numerous in-depth discussions that helped shape the building blocks of this thesis.

I am very thankful to my thesis committee members, Prof. Mark Pauly, Prof. Henri Maître, Prof. Hendrik Lensch and Prof. Sabine Süsstrunk for agreeing to review my work and taking the time to read my thesis.

I also thank Sabine for the fruitful collaboration, the co-authorship and the many insightful discussions. I would also like to thank Damien Firmenich of the IVRL and Dr. Stefan Trümpler and Dr. Sophie Wolf of the Vitrocentre, Romont, Switzerland for their collaboration in data acquisition and co-authorship.

I am immensely thankful to my officemate, friend and stoic companion: Zhou Xue. I appreciate all the interesting conversations that happened over pizza, cola or coffee, all the great movies and shows that you introduced me to, the light-hearted one-liners and all the running that we never got around to. Many thanks to Jacqueline for helping out with every little thing at LCAV, always with a smile. I consider myself very fortunate to have worked with a multi-talented group at LCAV. Thanks to Hanjie, Gilles, Mitra, Marta, Runwei, Robin and other present and past colleagues at LCAV for their support and the pleasant interactions.

I owe a great deal to my friends, who made life in Lausanne enjoyable. A million thanks to Aarthi, Niketh and Pavi for absorbing me into the clique, and creating a home away from home. Thanks for the delicious food, the many trips, the little adventures in the mountains, the many precious memories and the much needed spice and drama. Thanks in particular for enduring my incessant whining. Thanks, Sathya for all the geek talk and Mathi for being a great travel buddy. A shout out to fellow coffee break enthusiasts Mano, Stefan and others. I would like to thank Dr. Mathew Magimai Doss for his regular mentoring and help. Thanks also to Evelyne for being a wonderful host during my first year in Switzerland.

I am grateful to Prof. Mala John for her encouragement and for inspiring me to pursue higher studies, without which I would have never ventured on this journey. Thanks also to friends back home - Santhosh, Sakthi, Ashwin, Proxy, Mohan, Imran and Suganth for their support and the

amazing company.

Huge thanks to Naveen for being my shoulder angel since high school, for the unwavering support and fair criticism in all of my ventures. It has indeed been quite an adventure from Krishnagiri to where we are today.

I owe whatever success I achieve to the unconditional love, support and dedication of my family. Words cannot adequately describe my gratitude for them and my appreciation of their sacrifices. Thanks Kannan, for being such a great brother, for always looking out for me; milestones you set have also guided me along the way. Thanks Jayasree, for quickly becoming a good friend. For as long as I can remember, my parents Shanthi and Thanikachalam and my grandma Vijayal, have toiled to ensure that my brother and I have access to the best of resources and that the paths we take are obstacle free; all the while making sure that we grow up drenched in love. We will forever remain indebted to you, for the sacrifices you made for the sake of our happiness.

This thesis is dedicated with love, to my family.

Abstract

By incorporating computational methods into the image acquisition pipeline, computational photography has opened up new avenues in the representation and visualization of real world objects in the digital world. For example, we can sample a scene under a few specialized illuminations and a sparse set of viewpoints. We can later, computationally recover the complete light transport properties of the scene. Once we obtain the light transport characterization of cultural artifacts, we can enable users of virtual museums to interact with the artifacts in the same way as we experience these objects in the physical world. In particular, in this thesis, we develop algorithms and tools that facilitate the acquisition of relightable photographs of cultural artifacts, by acquiring their light transport matrix (LTM). A recurrent theme in this thesis is to exploit the low dimensionality of the LTM to develop efficient acquisition strategies for image based rendering.

First, we propose a new acquisition and modeling framework for inverse rendering of stained glass windows. Stained glass windows are a dynamic art form that change their appearance constantly, due to the ever-changing outdoor illumination. They are therefore, an exceptional candidate for virtual relighting. However, as they are anchored and very large in size, it is often impossible to sample their entire light transport with controlled illumination. We build a material specific dictionary by studying the scattering properties of glass samples and exploiting the structure of their LTMs in a laboratory setup. We then pose the estimation of the LTM of stained glass from a small set of photographic observations, as a linear inverse problem that is constrained by sparsity in the custom dictionary. We show by experiments that our proposed solution preserves volume impurities under both controlled and uncontrolled, natural illuminations and that the retrieved LTM is heterogeneous, as in the case of real world objects.

Next, equipped just with a dictionary to describe light transport in stained glass, we focus on the problem of designing a meaningful LTM, for the synthetic rendering of stained glass. Since this is an extremely ill-posed problem, we begin by exploring the physical properties of glass that can be used as constraints in light transport design. We then propose an iterative matrix completion algorithm that generates the LTM of a heterogeneous glass slab, given the dictionary and the physical constraints. We use this synthesis algorithm, in combination with an input texture to simulate stained glass windows in scenarios where inverse rendering is impossible or as an artist’s preview tool. We also introduce a framework for the digital restoration of broken slabs of glass by first acquiring the LTM with inverse rendering and then using the proposed matrix completion framework to repair the fractures.

Finally, we present an easy-to-use, handheld acquisition framework to sample the LTM of more general, reflective scenes. We first non-uniformly sample the scene reflectance by moving the LED attached to a smartphone along an arbitrary trajectory, while simultaneously tracking the position of the LED. The acquired reflectance is resampled to obtain a sparse set of samples on a uniform lattice. Using a compressive sensing framework, we recover an approximation to the uniformly sampled LTM, that is then used in scene relighting. As we used LEDs that

approximate point light sources to sample the reflectance, we show that when the object has a convex surface geometry, the obtained light transport is the Helmholtz dual of the outgoing light field. We thus propose an image based relighting framework for novel light positions, that uses the two plane parametrization common to light field setups.

Key words: Inverse rendering, scene relighting, relightable photographs, stained glass windows, light transport matrix, compressive sensing, dictionary learning, matrix completion, Helmholtz reciprocity, digital restoration, sampling along trajectories, image based relighting.

Résumé

En ajoutant des méthodes de calculs informatiques dans le canal d’acquisition d’image, la photographie dite computationnelle a ouvert la voie à de nouvelles représentations et visualisations d’objets du monde réel dans le monde numérique. Par exemple, nous pouvons échantillonner une scène à partir de différentes illuminations et d’un ensemble clairsemé de points de vue. Nous pouvons ensuite retrouver par calcul la totalité des propriétés de transport de lumière de la scène. Dès lors que la caractérisation du transport de lumière d’un objet culturel est obtenue, il est possible à des utilisateurs de musées virtuels d’interagir avec cet objet de la même façon que dans le monde réel. Dans cette thèse, nous développons des algorithmes et des outils pour faciliter la capture d’images ré-illuminables d’objets culturels en acquérant leur matrice de transport de lumière (MTL). Un thème récurrent de cette thèse est la mise à profit de la faible dimensionnalité de la MTL pour développer des stratégies d’acquisition efficaces.

Nous proposons en premier lieu une nouvelle approche d’acquisition et de modélisation pour le rendu inverse de vitraux. Les vitraux sont une forme d’art dynamique dont les apparences changent à tout instant à cause du changement permanent de l’illumination extérieure. Ils constituent donc un candidat idéal pour la ré-illumination virtuelle. Cependant, comme ils sont emmurés et de grandes tailles, il est souvent impossible d’échantillonner leur transport de lumière avec une illumination contrôlée. Nous construisons un dictionnaire spécifique pour chaque matériau afin d’étudier les propriétés de dispersion de morceaux de verre et d’exploiter la structure de leur MTL en laboratoire. Nous posons ensuite le problème d’estimation de la MTL d’un vitrail à partir d’un ensemble limité de photographies comme un problème linéaire inverse contraint par l’éparpillement des données dans le dictionnaire créé. Nous montrons à travers des expériences que notre solution conserve les impuretés de volume sous des illuminations à la fois contrôlées et naturelles et que la MTL est hétérogène au même titre que les objets du monde réel.

Par la suite, grâce à notre dictionnaire décrivant le transport de lumière d’un vitrail, nous nous focalisons sur le problème de la conception d’une MTL significative pour le rendu d’image synthétique de vitraux. Comme le problème est très mal posé, nous commençons par explorer les propriétés physiques du verre qui peuvent être exploitées pour contraindre la conception du transport de lumière. Nous proposons alors un algorithme itératif de remplissage de matrice qui génère la MTL d’un morceau de verre hétérogène, selon son dictionnaire et les contraintes physiques. Nous utilisons cet algorithme de synthèse avec une texture d’entrée pour simuler un vitrail dans des cas où le rendu inverse est impossible ou comme un outil de prévisualisation pour un artiste. Nous introduisons aussi une solution pour la restauration numérique de morceaux de verre brisés en capturant d’abord la MTL par rendu inverse et en utilisant ensuite l’algorithme proposé de remplissage de matrice pour réparer les brisures.

Pour finir, nous proposons une solution facile d’utilisation et portable pour l’acquisition de la MTL dans des cas de scènes réfléchives. Nous commençons par échantillonner de manière non-uniforme la réflectance de la scène en déplaçant une DEL attachée à un smartphone selon

une trajectoire arbitraire tout en localisant la position de la DEL. La réflectance capturée est ré-échantillonnée pour obtenir un ensemble clairsemé d'échantillons sur un treillage uniforme. En utilisant les techniques de détection clairsemée, nous obtenons une approximation de la MTL échantillonnée uniformément, qui peut être ensuite utilisée pour la ré-illumination de la scène. Comme nous utilisons des DEL qui approximent une source lumineuse ponctuelle pour l'échantillonnage de la réflectance, nous démontrons que lorsqu'un objet a une surface à géométrie convexe, le transport de lumière obtenu est le réciproque de Helmholtz du champ de lumière réfléchi. Nous proposons donc une solution de ré-illumination à partir d'image pour de nouvelles positions de lumière qui utilise la paramétrisation à deux plans courante des systèmes de champ lumineux.

Mots-clés: rendu inverse, ré-illumination de scène, photographie ré-illuminable, vitraux, matrice de transport de lumière, détection clairsemée, apprentissage de dictionnaire, remplissage de matrice, réciprocity de Helmholtz, restauration numérique, échantillonnage sur trajectoires, ré-illumination à partir d'images.

Table of Notations

Inverse and Synthetic Rendering of Stained Glass

\mathbf{y}	Observed Image vector
ℓ	Illumination vector
\mathbf{T}	Light Transport Matrix
\mathbf{Y}	Observation matrix; Ensemble of observed image vectors
\mathbf{L}	Measurement matrix; Ensemble of incident illumination vectors
\mathbf{L}_n	Effective measurement matrix for image pixel n
N	Size of the image plane
M	Size of the illumination plane
P	Number of observations
n	Pixel index of the image plane
m	Pixel index of the illumination plane
\mathbf{t}_n	n^{th} light transport vector
γ_n	Indicator function for the n^{th} image pixel
v	Bandwidth of a perfectly planar but heterogeneous glass slab
δ	Maximum translation caused by refraction at surface irregularities
ν	Bandwidth of heterogeneous glass slab with surface irregularities
τ_n	n^{th} reduced light transport vector
\mathbf{D}	Material specific sparsifying overcomplete dictionary
α_n	Sparse co-efficients of the n^{th} light transport vector in dictionary \mathbf{D}
λ	Wavelength of light
$\ell(\lambda)$	Illumination vector at wavelength λ
$\mathbf{y}(\lambda)$	Image vector at wavelength λ
$\mathbf{c}(\lambda)$	Albedo or texture painted on the surface of glass slab
σ	Material texture, caused by spatially varying light transport
γ_m	Indicator function for the m^{th} image pixel in the dual configuration
τ_m	m^{th} reduced light transport vector of the dual configuration
α_m	Sparse co-efficients of the m^{th} light transport vector in the dual configuration using dictionary \mathbf{D}

Handheld Acquisition of Light Transport

$[x, y]$	Coordinates of the image plane
$[r, s]$	Coordinates of the illumination plane
$\mathcal{T}(r, s, x, y)$	Continuous space light transport field
$\tau_{ij}(r, s)$	Continuous space light transport function at pixel (i, j) of camera
h	Distance between image plane and illumination plane
X	Finite rectangle that is discretized by the image sensor
R	Finite rectangle describing the region of illumination plane that will be discretized
$\Omega_{R,h}$	Solid angle sampled by an illumination plane of size R , placed at a distance h from the image plane
Δ_{rs}	Sampling period of the illumination plane
J	Size of the image plane
M	Size of the illumination plane
(i, j)	Pixel index of the image plane
(m, n)	Pixel index of the illumination plane
\mathbf{T}	Uniformly sampled Light Transport Matrix
$\mathbf{t}_{ij}(m, n), \mathbf{t}_{ij}$	Uniformly sampled light transport function at pixel (i, j)
γ	Trajectory taken by mobile phone
γ_r	Discrete trajectory observed after SLAM
$\mathbf{\Gamma}$	Digitized trajectory after resampling, indicating indices in the lattice where samples have been observed
K	Number of observations, number of non-zero entries in $\mathbf{\Gamma}$
$\tilde{\mathbf{t}}$	Sparsely sampled light transport function
$\tilde{\mathbf{T}}$	Sparsely sampled light transport matrix
\mathbf{H}	Fourier Transform matrix
\mathbf{f}	Fourier coefficients of uniformly sampled light transport function, \mathbf{t}
$\hat{\mathbf{f}}$	Reconstructed Fourier coefficients of \mathbf{t}
$\hat{\mathbf{t}}$	Reconstructed light transport function
$\hat{\mathbf{T}}$	Reconstructed light transport matrix
\mathbf{y}	Image vector
ℓ	Illumination vector

Contents

Acknowledgements	iii
Abstract	v
Résumé	vii
List of Notations	ix
1 Introduction	1
1.1 Motivation	1
1.2 Components of Light Transport	4
1.2.1 Direct Component	4
1.2.2 Scattered Component	5
1.2.3 Secondary Components	5
1.3 Modeling Scene Appearance	7
1.3.1 Classic Forward Rendering	7
1.3.2 Inverse Rendering	8
1.4 Light Transport Matrix	10
1.4.1 Light Transport with Directional Light Sources	11
1.4.2 Light Transport with Structured Light Sources	13
1.5 Thesis outline and Contributions	14
1.5.1 Inverse Rendering of Stained Glass	14
1.5.2 Synthetic Rendering of Stained Glass	15
1.5.3 Handheld Acquisition of Light Transport	15
2 Inverse Rendering of Stained Glass	17
2.1 Introduction	17
2.2 Related Work	18
2.2.1 Inverse Rendering Methods	19
2.2.2 Empirical models	19
2.2.3 Our approach	19
2.3 Digital Modeling of Stained Glass	20
2.3.1 Image Formation Model	21
2.3.2 Structure of the Light Transport Matrix	21
2.4 Sparse Representation of Light Transport	24
2.4.1 Low Dimensionality of Light Transport	24

2.4.2	Learning sparsity inducing dictionaries	25
2.5	Acquisition of Stained Glass	26
2.5.1	Controlled Acquisition	28
2.5.2	Uncontrolled Acquisition	29
2.6	Algorithmic issues	30
2.6.1	Handling surface irregularities	30
2.6.2	Uncontrolled acquisition with sparse regularization	31
2.7	Experiments	32
2.7.1	Acquisition Performance	32
2.7.2	Controlled Acquisition in Romont	36
2.7.3	In-Situ Acquisition of Rose Window, Lausanne	37
2.8	Summary	45
3	Synthetic Rendering of Stained Glass	47
3.1	Introduction	47
3.2	Related Work	48
3.3	Light Transport Design	48
3.3.1	Physical properties of a stained glass slab	50
3.3.2	Synthesis of a Light Transport Matrix	53
3.4	Applications to interactive rendering of stained glass	56
3.4.1	Synthesis of Stained Glass Windows	57
3.4.2	Digital Restoration of Stained Glass	57
3.5	Summary	58
4	Handheld Acquisition of Light Transport	65
4.1	Introduction	65
4.2	Related Work	67
4.3	Light Transport with Point Light Sources	67
4.3.1	Continuous Space Light Transport	68
4.3.2	Uniform Sampling Model for Light Transport	70
4.4	Light Transport Acquisition System	71
4.4.1	Handheld Acquisition	71
4.4.2	Resampling Step	72
4.4.3	LTM Reconstruction	73
4.5	Experiments and Analysis	74
4.5.1	Evaluation of the Tracking Algorithm	74
4.5.2	Experimental Validation against Ground Truth	75
4.5.3	Effect of Sampling Period Δ_r and Path Density	75
4.5.4	Comparison with Other Methods	77
4.5.5	Error Analysis	80
4.5.6	Scene Relighting	82
4.5.7	Mobile Based Acquisition	83
4.5.8	6-dimensional light transport Matrices	83
4.5.9	Discussion	85
4.6	Summary	86
5	Conclusion	89

Contents	xiii
Bibliography	93
Curriculum Vitæ	101

Chapter 1

Introduction

The bird that would soar above the plain of tradition must have strong wings.

Douglas Adams, The Hitchhiker's Guide to the Galaxy

1.1 Motivation

From the pictographs of ancient times to the painted portraits and landscapes of the pre-modern era and the photographs of contemporary times, our pursuit for realism in the graphical representation of visual information has remained alive. Arguably, the rise of consumer level digital photography has made this task effortless. Capturing, processing, sharing and storing high resolution photographs are all increasingly cheaper and accessible to the common man, thanks to the affordability of consumer electronics and the multitude of viewing platforms. But does a photograph truly represent a real-world scene?

Real-world objects exist in 3D space. Their appearance depends upon the direction from which they are viewed as well as the nature of light incident upon them. However, a photograph only describes the appearance of a given scene, under a particular viewing environment, as determined by the direction from which the photograph was taken and the incident light. If these additional dimensions of real-world scene appearance are also digitized and represented in a virtual platform, we can interact with and experience these objects in the virtual world, just as we would in the real world. However, physically sampling scene appearance under both varying viewing angles and incident light is prohibitively complex.

Computational imaging, an emerging paradigm in photography fueled by affordability of computing power, circumvents hardware limitations using algorithmic means. High dynamic range photographs, panoramic images and photospheres are now standard features in most imaging

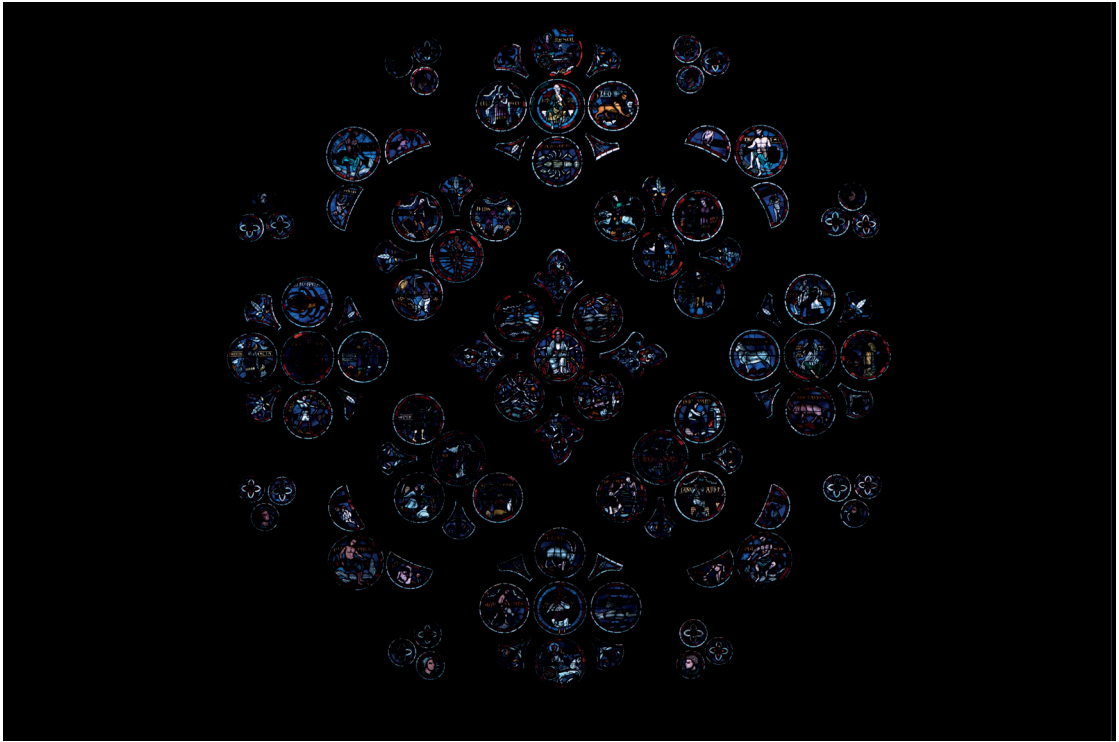
equipment thanks to advances in computational imaging. The Lytro and Raytrix are computational cameras that enable computational refocussing of photographs and acquisition of depth images, by acquiring the light field and intelligently recombining acquired light rays while rendering. New cameras and camera architectures such as the Google Jump have been announced, which enable 360 degree stereo videos as well. These technological advances bring us closer to experiencing 3D photographs of the physical world. One can interact with these photographs, viewing them from different positions in a virtual environment. However, interacting with such 3D photographs in virtual environment is life-like, only when the object blends into the scene by reflecting, refracting or scattering the light present in the virtual environment. Imagine a glass of water whose 3D photograph was obtained by acquiring its light-field. For the light field to ‘blend in’ with a virtual environment of say, a dinner table, the digitized glass will have to refract and reflect light from objects on the table and its background. Without this ability, an acquired light field or a 3D photograph do not interact with the light around it, thus depriving the user of a life-like experience.

Modeling the path of light through an object, requires sampling its appearance under varying illumination; such a venture might seem excessive when digitizing everyday objects. However, digitizing these properties of important historical and cultural artifacts, can help a greater audience to explore and experience these objects using a virtual platform. Truthful digitization of these artifacts can also serve as a means of digitally archiving them. Moreover, a platform that aids in truthful visualization of artifacts under different illuminations can serve as an analysis tool for art historians, curators and general patrons of art to study and infer the physical properties of these artifacts. The focus of this dissertation is on practical methods for relighting of cultural artifacts such as stained glass windows and oil paintings, to enable photographs of the object to interact with virtual illumination.

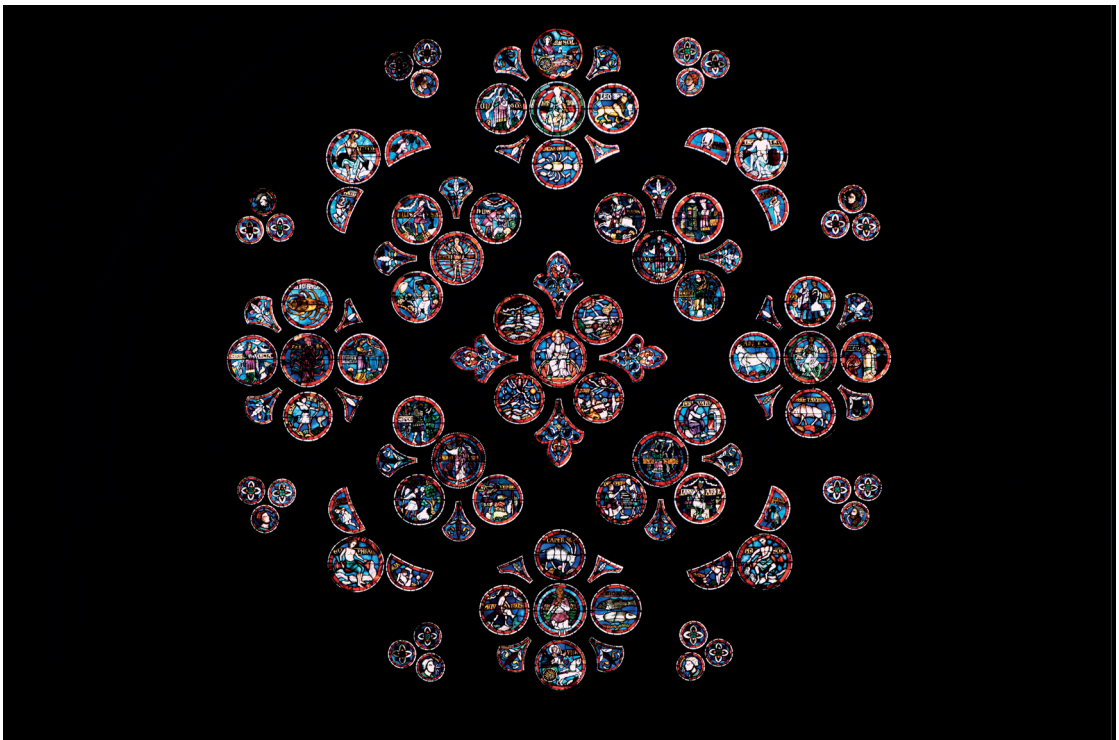
Stained glass windows change in their appearance through the day, as shown in Figure 1.1, due to their translucent nature and location at the interface of constantly varying outdoor illumination and low-lit interiors. The possibility of creating relightable photographs of stained glass windows thus enables us to capture the dynamic nature of this artform in the virtual world. Furthermore, it lets us visualize the appearance of stained glass windows under otherwise impossible viewing environments. However, their large size, stationary nature and the highly heterogeneous nature of their light transport provide unique challenges in acquiring their light transport characteristics. This dissertation presents an approach to digitize the heterogeneous light transport characteristics of stained glass windows.

While oil paintings are largely flat, the presence of different types of ink, often in various blends and layers coupled with the intricate micro-geometry created by paint strokes, result in complex surface reflectance properties. Thus, when a painting is moved relative to a light source or when the illumination is changed, its surface geometry is revealed, which in turn adds to the richness of our visual experience. To facilitate easy, affordable acquisition of these reflectance properties, we present a hand-held strategy to acquire light transport of oil paintings as well as artifacts with more general macro-geometry.

Digitally simulating the appearance of the real world is at the heart of computer graphics, with applications in games, movies, architecture design and virtual reality applications. Thus, the transport of light through scenes of known geometry and material composition, is well studied in computer graphics. We will continue by introducing various components of light transport.



(a)



(b)

Figure 1.1: Appearance of the Rose Window at the Cathedral of Notre Dame, Lausanne. (a) At dawn. (b) At Noon.

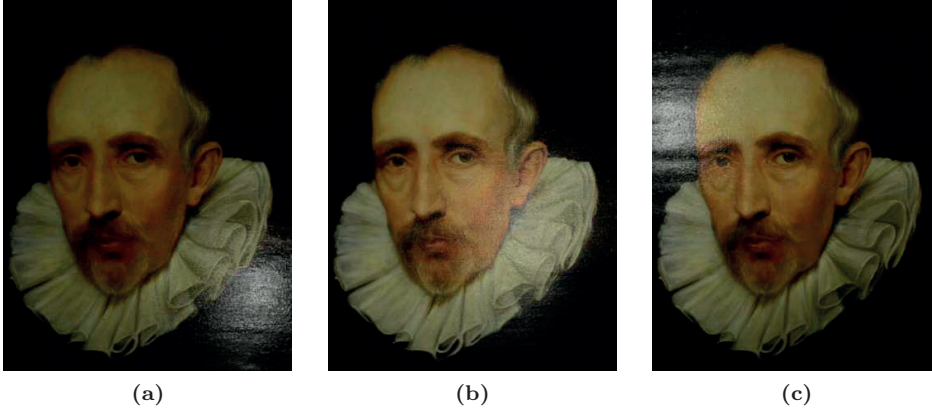


Figure 1.2: Appearance of the potrait of Cornelis van der Geest under different illumination configurations

1.2 Components of Light Transport

Image formation is complex. Scene appearance depends upon the interaction of light with the scene. The path that is taken by light that enters a scene is determined by two factors - the geometry of the scene and the material composition of the scene. A detailed treatment of light transport inside a scene, as dictated by its geometry and its physical properties can be found in [77]. Here we present an overview of the various components. Let x_i and ω_i be the point of entry and direction of incident light while x_o and ω_o are the point of exit and direction of outgoing light. In the following overview, x_i and x_o are defined on the surface of the object.

1.2.1 Direct Component

For reflective objects, the direct component corresponds to the light that is reflected off of the surface of the object as determined by the material composition and the surface normal alone. In diffuse surfaces, irrespective of the direction ω_i of incident light, a fraction of light is scattered uniformly into all the outgoing directions ω_o and is referred to as the *Lambertian component*. When the surface of the object is devoid of any surface roughness, light is reflected in a mirror-like manner (ideal reflection) and contributes to the *specular component*. Often, objects have a rough surface and result in light being reflected along a lobe around the direction of the specular component, resulting in the *glossy reflection* component. The amount of surface roughness determines the width of this lobe; a very rough surface has a wide lobe and a smooth surface results in a very narrow lobe that degenerates to the specular component for mirror-like surfaces. Thus, for each surface point $x_i = x_o$, the direct component is a function of ω_i and ω_o . For transmissive objects, the direct component corresponds to light that is refracted by the object as determined by Snell's Law. Similar to the case of reflective objects, when the object is bounded by rough surfaces, refracted light also occurs as a lobe around the direction of ideal refraction.

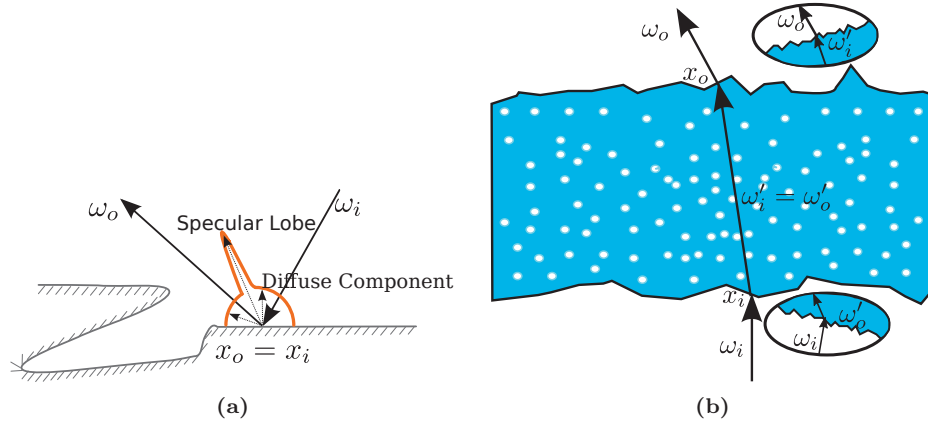


Figure 1.3: Illustration of the direct component of light transport. (a) Reflective Material; Showing the diffuse and specular lobes. (b) Transmissive Material; Showing the refraction events at both interfaces. Micro-structures shown in inset, at each interface can give rise to refraction along more than one direction.

1.2.2 Scattered Component

Numerous real-world objects like marble, wax, milk, stained glass, skin etc are translucent to various degrees. Thus in addition to the direct component, light often enters the surface of these materials and is scattered multiple times before a fraction of it re-emerges outside the material. In case of light that enters and exits on the same side of the object, this is referred to as Subsurface Scattering. When light exits the object through a different side, this is referred to as Volumetric Scattering. In both cases, light entering the object at a given point x_i and a given direction exits at other points x_o , in the neighborhood of the direct component, with the size of the neighborhood determined by how ‘optically dense’ the material is. Thus the scattered component at a given surface point x_o is a function of the point of entry x_i , and the directions ω_i and ω_o .

1.2.3 Secondary Components

When the scene has a convex surface geometry, the direct component and the scattered component are the only contributors to light transport, since light that once exits the scene does not re-enter the scene again. However, in the more general case of objects with a non-convex surface geometry, light that exits from the surface of the object can re-enter the scene acting as an additional source of light, and cause a second set of direct and scattered components, each of which can re-enter the scene again, thus resulting in a recursive process. In layered transmissive materials, light that has entered the object can get reflected at the boundary of each layer back to the same face. Since light that exits the scene at one point re-enters the scene at a different point, and due to the directional and spatial nature of the direct and scattered components, the secondary components of light transport are functions of both the direction and position of the incident light ray. In case of reflective objects, this results in inter-reflections and soft shadows, whereas in case of refractive objects, among other effects, this results in caustics.

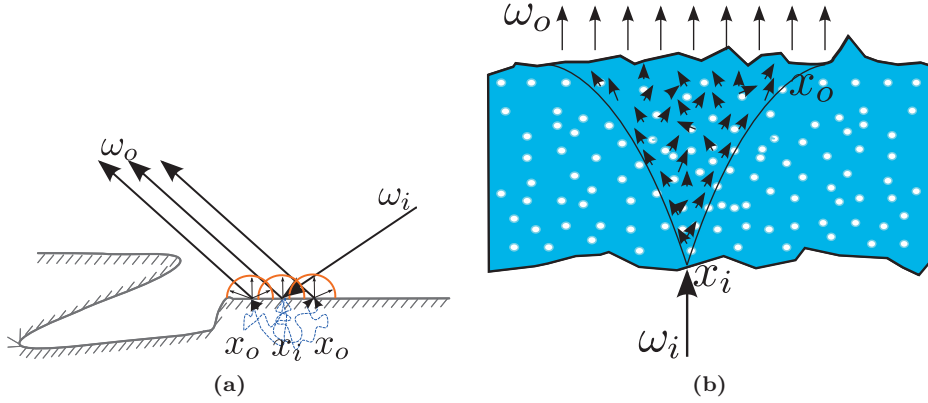


Figure 1.4: Illustration of the scattered component of light transport. (a) Reflective material exhibiting subsurface scattering. (b) Transmissive material exhibiting volume scattering.

Near Field Components

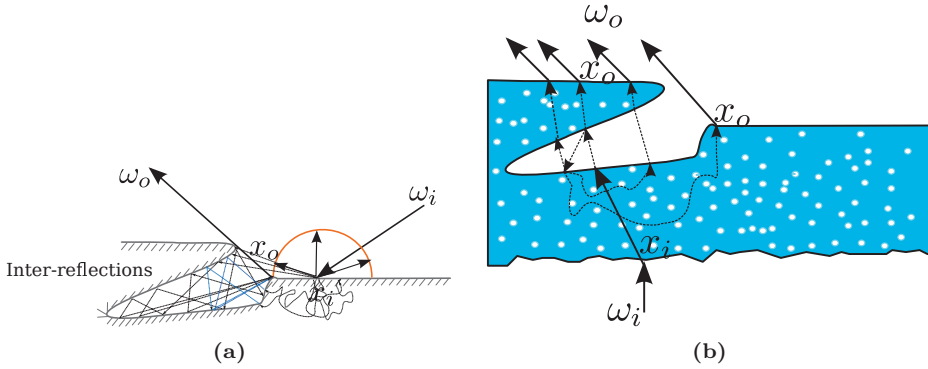


Figure 1.5: Illustration of the Near-field secondary components of light transport. (a) Reflective material exhibiting inter-reflections in a non-convex micro-structure. (b) Transmissive material exhibiting secondary refraction events due to a local non-convex region.

When the scene is largely convex, but contains local non-convex regions, the secondary components for light entering at a given point only occur in a small neighborhood around the point of entry of light, where the size of the neighborhood is determined by the degree of non-convexity. Oil paintings are largely flat, but contain numerous micro-structures owing to paint strokes and are a common example of largely convex objects. Here, light that exits the painting is unlikely to re-enter the scene at a point far-off from the first reflection, however micro-structures in the vicinity of the point of exit of light can reflect light back into the scene. Thus, the near-field secondary components at a given surface point x_o are functions of the direction of incident light ω_i and a small span of x_i in the neighborhood of x_o .

Far Field Components

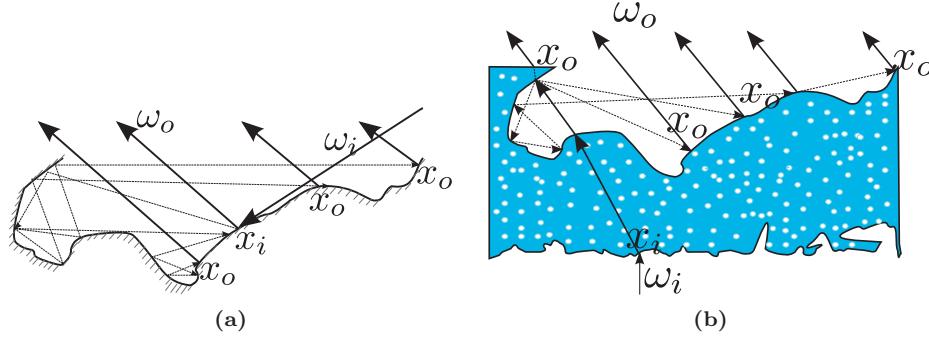


Figure 1.6: Illustration of the Far-field secondary components of light transport. (a) Reflective material exhibiting inter-reflections in a non-convex scene, affecting global light transport. (b) Transmissive material exhibiting secondary refraction events in a non-convex scene, affecting global light transport.

When the scene is non-convex, such as the inside of a room or a pot, light that enters the scene at one point x_i at a given incident direction can exit the scene at a point x_o far-off from x_i after multiple secondary interaction events. Thus this component, at a given surface point x_o is a function of both ω_i and the entire span of x_i .

1.3 Modeling Scene Appearance

Over the years, an exhaustive body of work that aim at providing high quality photo-realistic renderings of both synthetic and real-world objects have emerged. In this section, we briefly review related literature.

1.3.1 Classic Forward Rendering

The goal of classic forward rendering is to simulate the appearance of a scene, given its geometry, material composition and associated light transport models and the illumination sources. For each wavelength of incident light, light transport in the scene is defined completely by the 8-D *Bidirectional Subsurface Scattering Reflectance Distribution Function*, $S(\omega_i, x_i, \omega_o, x_o)$ and

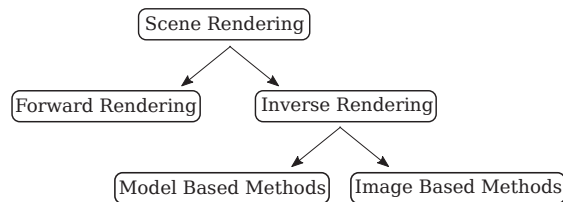


Figure 1.7: A general overview of scene modeling.

the geometry of the scene. Scene appearance, is then defined by integrating the incoming light at each surface point over incoming directions and area A , as given by

$$L_o(\omega_o, x_o) = \int_A \int_{2\pi} S(\omega_i, x_i, \omega_o, x_o) L_i(\omega_i, x_i) (n_{x_i} \cdot \omega_i) d\omega_i dA.$$

When the objects are ‘optically dense’, due to the multiple scattering events, light exiting the object have been shown to be approximately isotropic. This multiple scattering can then be approximated by a diffusion process [37]. In the absence of such analytic [17, 30, 36, 37] or empirical [19] models, light transport through volumes are rendered using the classic volume rendering equation [9].

Often, analytic models for light transport ignore the subsurface scattering term (or subsurface scattering is absent in the material) and assume that light that enters a point exits from the same surface point *i.e* $x_i = x_o = x$. This results in the 6D *spatially varying Bidirectional Reflectance Distribution Function* (sv-BRDF), $\rho(\omega_i, x, \omega_o)$. Alternatively, at each surface point x , light transport is defined by the 4-D BRDF $\rho(\omega_i, \omega_o)$, that describes the diffuse component and the specular lobe for the specific material. The compact representation of light transport associated with each class of materials using parametric models, has received significant focus in computer graphics literature. Numerous analytic, empirical or physics based models [3, 5, 10, 32, 38, 57, 61, 69, 75] have been proposed for various classes of materials. Physically based BRDF models, model the specular lobe at a given surface point as the result of mirror-like reflections at microfacets distributed around its surface normal. Similarly, for transmissive materials, Bidirectional Transmittance Distribution Functions (BTDF) have been proposed [11, 64, 72].

While the scattered and direct components are caused only by light sources defined in the scene, secondary components like inter-reflections and caustics are caused by light re-entering the scene after exiting it. These effects, particularly the far-off components are computationally expensive to be rendered, and numerous approximation methods such as photon mapping, radiosity, metropolis light transport which trade-off accuracy to computational complexity to various degrees have also been proposed.

1.3.2 Inverse Rendering

Inverse rendering refers to the problem of inferring either the light transport properties of a scene or the incident illumination or both, given a set of photographic observations of the scene. In this thesis, we are primarily concerned with the problem of acquiring a scene’s light transport properties, given a set of photographic observations, under a known or measurable set of incident illuminations. The goal of this class of problems can range from recovering the reflectance parameters of a material to digitizing a complex real-world scene by obtaining their light transport characterization. Thus, while classic forward rendering aims at describing synthetic scenes in a photorealistic manner, inverse rendering aims at modeling the appearance of real-world scenes. Inverse rendering can be broadly classified into two categories; methods that ‘model’ scene parameters using analytic models developed for forward rendering and methods that ‘sample’ and represent scene appearance with images.

Model Based Inverse Methods

Model based inverse methods digitize a scene, given a set of photographic observations under known illuminations, by fitting a parametric BRDF model to the acquired data, while simul-

taneously recovering surface normals of the scene. Weyrich et al [77] present an overview of techniques used for acquiring the SV-BRDF of objects.

Common acquisition systems involve a light stage constructed using light sources distributed on a sphere or hemisphere around the object. The object is then photographed under a series of known basis illuminations, which are then used to fit an appropriate parametric BRDF function. Initial setups worked by exhaustively sampling using both varying illuminations and viewing angles [12]. Later approaches [14, 31, 33] only vary the incident illumination, and recover the surface normals followed by fitting a BRDF model. Polynomial texture mapping [44] proposes the use of a biquadratic polynomial to represent the acquired BRDF. These initial methods do not allow independent measurements of specular and diffuse surface reflectances. Recent approaches [24, 26, 27, 43, 70] utilize the low dimensionality of BRDFs in the spherical harmonics basis to achieve faster acquisitions and simultaneously separate diffuse reflections from specular reflections. Other methods [16, 41, 78] sparsely sample both viewpoints and lighting directions and assume a known geometry to extrapolate BRDFs over the entire object. A second class of acquisition systems [23, 34, 40, 60] use a spatially varying light source to sample the reflectance functions.

Similarly, [35] provides an overview of approaches that model light transport in transparent and translucent objects.

While model based methods provide a compact representation for light transport characterization, the approach is affected by a number of factors. Since BRDF models assume that the point of entry and point of exit of light are the same, the models generated by these approaches are erroneous when the object exhibits secondary reflections or caustics, or when the object exhibits heterogeneous volume/subsurface scattering. Moreover, BRDF models are defined for very specific classes of materials and thus the model chosen must be appropriate for the current scene. Also, the BRDF models are often non-linear in their parameters, thus convergence to global optima is also sensitive to initialization.

Image Based Methods

Image based rendering frameworks digitize a scene by treating a scene as a bounding volume, where the scene geometry and material composition remain unknown, as shown in Figure 1.8. The scene’s appearance is then entirely characterized by the 4-dimensional light field $L(x_i, \omega_i)$ that enters the bounding volume, and the 4-dimensional light field $Y(x_o, \omega_o)$ that exits the bounding volume. The 8-dimensional function $S(\omega_i, x_i, \omega_o, x_o)$ then completely describes how light from the incident light field interacts with the scene to generate the outgoing light field. The outgoing light field can be directly sampled by using a light field camera or a camera array [42]. In this thesis, we fix the viewing direction and thus replace the light field camera with a single camera, and focus only on scene relighting. Thus, having fixed the viewing angle, we can sample the light transport $S(\omega_i, x_i, x_o)_{\omega_o}$ from the incident light field to the camera, by varying the incident light field.

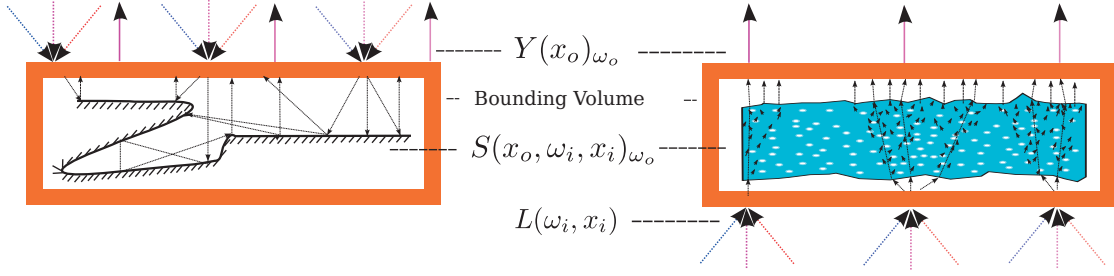


Figure 1.8: An illustration of image based rendering.

A light source that allows us to arbitrarily vary the incident light field is in essence a true 3-D display/screen. Such light sources with good resolving powers along both the directional and spatial dimensions of the light field that they emit, still remain impractical. We are thus limited to using commonly available light sources such as directional lights, point light sources and projectors to describe the incident light field. A sampling setup with an array of directional/point light sources or the projector can only sample a 4-D subspace of the 6-D light transport $S(\omega_i, x_i, x_o)_{\omega_o}$. This digitization thus results in a 4-D Light Transport Matrix (LTM), with each pixel in the observed image described by a discretized Light Transport Function (LTF). The light transport function refers to the slice of $S(\omega_i, x_i)_{\omega_o, x_o}$, that was sampled by the illumination configuration. While image based methods are versatile in their representational power, they are data intensive and do not directly lead to traditional graphics rendering pipelines.

In the following section, we will describe the sampling pattern of common illumination configurations to describe what components of light transport are resolvable in each of the sampling setups.

1.4 Light Transport Matrix

The Light Transport Matrix is an array of arrays, that can either be described as an array of pixel-wise discretized light transport functions or as an array of images, each illuminated by one element of the illumination lattice. The LTM was first introduced as environment matting [79] to represent refraction events in real-world objects. Since then, due to its versatility it has been extended to more general scenarios.

The incident illumination in Figure 1.8 is represented by the array L and the image observed by the fixed camera is represented by the array Y . For convenience in notations, we serialize both these arrays in a column-major format resulting in vectors $\ell \in \mathbb{R}^M$ and $\mathbf{y} \in \mathbb{R}^N$ denoting the incident illumination and observed image respectively. Image formation is then described by the equation,

$$\mathbf{y} = \mathbf{T}\ell.$$

Here, $\mathbf{T} \in \mathbb{R}^{N \times M}$ is the LTM. Under this representation, each column of the LTM represents the serialized image observed when only the corresponding element in the illumination configuration is turned on. Alternately, each row of the LTM represents the serialized discrete LTF, \mathbf{t} . The structure and representational power of the LTM depends upon the illumination configuration that was used to discretize the light transport, as described in the following subsections.

1.4.1 Light Transport with Directional Light Sources

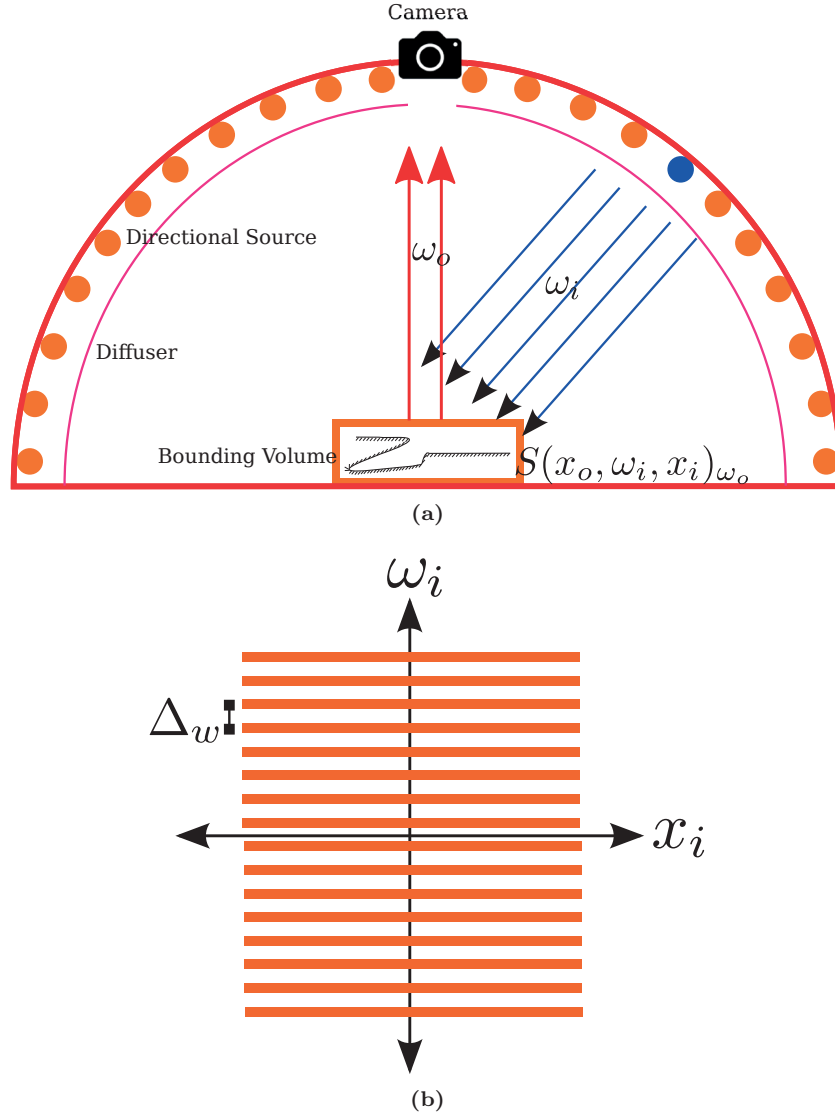


Figure 1.9: (a) Illustration of a typical light stage. (b) The sampling pattern of the light transport $S(\omega_i, x_i)_{\omega_o, x_o}$ at surface point x_o , for light emitted in the direction ω_o . The point of entry of light x_i is unresolvable since the incident light $L(\omega_i)$ is purely directional.

The acquisition setup consists of a lightstage built with directional light sources distributed uniformly on a sphere or a hemisphere, as described in Figure 1.9(a). A diffuser placed in front of the light sources acts as the anti-aliasing filter along the direction of incidence of light. The object is placed in the center of the light stage and a camera is used to record the observed image under each light source. In this setup, when a light source is turned on, the entire object

is illuminated with directional light $L(\omega_i)$.

The sampling periods of the lightstage and the camera are Δ_w and Δ_n while the sampling kernel induced by the diffuser is ψ_w and that of the camera is ψ_n . The sampling operation can then be described as,

$$\mathbf{T}(n, w) = \left\langle \int_{x_i} \langle S(x_o, \omega_i, x_i)_{\omega_o}, \psi_w \left(\frac{\omega_i}{\Delta_w} - w \right) \rangle dx_i, \psi_n \left(\frac{x_o}{\Delta_n} - n \right) \right\rangle, \quad (1.1)$$

$$n = [1 \cdots N] \text{ and } w = [1 \cdots W],$$

where N is the camera sensor size and W is the number of light sources in the lightstage, $S(x_o, \omega_i, x_i)_{\omega_o}$ is the continuous space light transport for fixed outgoing direction ω_o and $\mathbf{T}(n, w)$ is the resulting discrete light transport matrix. The inner product \langle, \rangle is the standard inner product on the vector space of real valued functions, and is defined as,

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx. \quad (1.2)$$

It can be seen that in equation (1.1), at each pixel n the corresponding 4D light transport $S(\omega_i, x_i)_{\omega_o, x_o}$ is projected into a 2D space, as defined by

$$\mathbf{t}_n(w) = \int_{x_i} \langle S(\omega_i, x_i)_{\omega_o, x_o}, \psi_w \left(\frac{\omega_i}{\Delta_w} - w \right) \rangle dx_i, \text{ where } w = [1 \cdots W]. \quad (1.3)$$

We illustrate the sampling operation described in Equation (1.3) in Figure 1.9(b). While the incident light is sampled along the incoming direction ω_i , the integral over x_i , results in all the spatially varying components of light transport for a given point being projected into its corresponding directional components. Thus, if a material only exhibits surface reflections (no subsurface or volume transport of light) and if the scene exhibits a convex surface geometry (no secondary components), image based relighting with the light stage can completely digitize an object's light transport at the chosen resolution. However, if the scene exhibits subsurface or volume light transport or contains a non-convex geometry, then the LTM \mathbf{T} can only be used for relighting with directional lights or with environment maps. Thus relighting with 4D incident light remains unfeasible.

All versions of the light stage [14, 15, 31, 70] can be used to sample the LTM for image based rendering. However, in these cases the acquired data was used to fit parametric models for model based rendering. A detailed account of acquisition, compression and image based rendering of objects digitized using the LTM with lightstages can be found in [45, 48].

In practice, directional light sources are often replaced with light sources that approximate point light sources. Thus, in order to emulate directional light with these isotropic sources, either the size of the scene with respect to the light-stage should be small, or the light stage should have a large radius. In Chapter 4, we propose a new light transport acquisition system that uses point light sources located on a plane.

1.4.2 Light Transport with Structured Light Sources

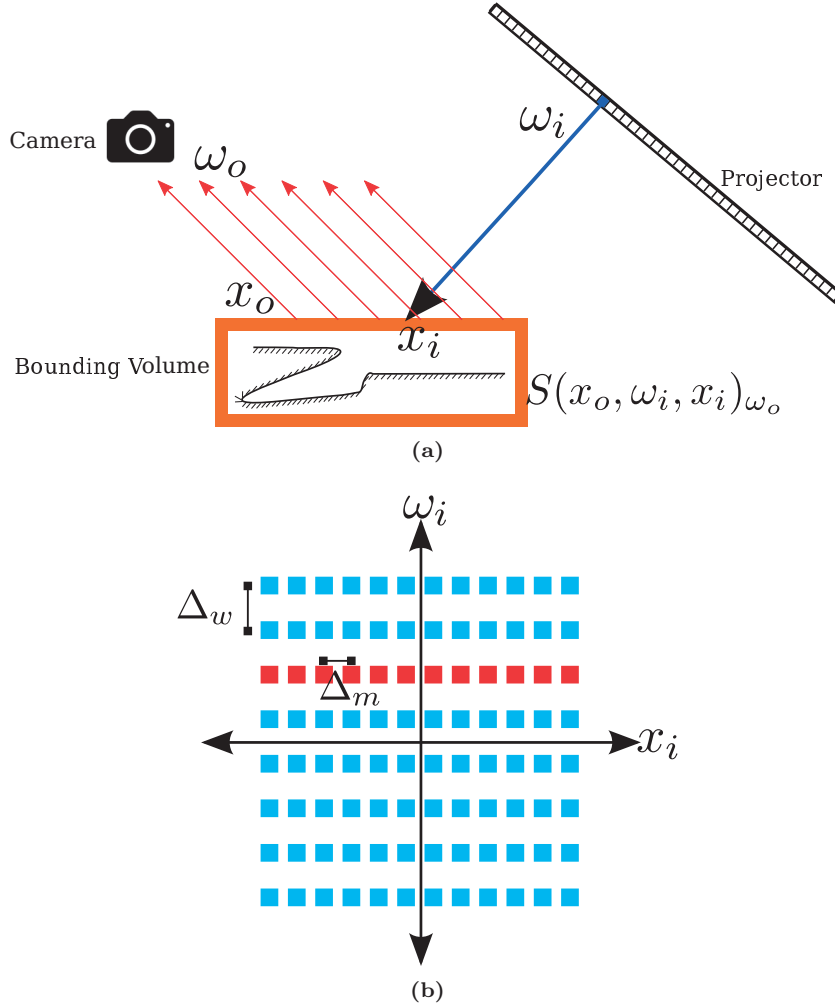


Figure 1.10: Light transport matrix acquisition: (a) Illustration of a typical projector-camera setup for acquiring the LTM. (b) In red, the sampling pattern of the light transport $S(\omega_i, x_i)_{\omega_o, x_o}$ at surface point x_o , for light emitted in the direction ω_o . Note that by repeating the acquisition by placing the projector along different directions, ω_i , the space of all possible incident light fields $L(x_i, \omega_i)$ can be uniformly sampled, as indicated by the pale blue samples.

The acquisition setup consists of a spatially varying light source (a projector or a CRT monitor) and a camera, as described in Figure 1.10(a). The position of the camera and that of the light source with respect to the scene are fixed and describe the viewing direction ω_o and the direction of incident light ω_i respectively. The camera samples the spatially varying radiant light along x_o while the display is used to sample the spatial components of incident light along x_i .

The sampling periods of the projector and the camera are Δ_m and Δ_n , while the corre-

sponding sampling kernels are ψ_m and ψ_n , respectively. The sampling kernel on the incident illumination can be implemented by pre-filtering the displayed illumination pattern. The sampling operation can be described by,

$$\mathbf{T}(n, m) = \langle \langle S(x_o, x_i)_{\omega_o, \omega_i}, \psi_m \left(\frac{x_i}{\Delta_m} - m \right), \psi_n \left(\frac{x_o}{\Delta_n} - n \right) \rangle \rangle, \\ n = [1 \cdots N] \text{ and } m = [1 \cdots N],$$

where N and M are the camera and projector resolutions, respectively. The inner product is defined in Equation (1.2). Here, at each camera pixel n , the light transport $S(x_i)_{\omega_o, \omega_i, x_o}$ is sampled along its spatial dimensions x_i , for the given incident and outgoing light directions ω_i, ω_o . We illustrate this per-pixel sampling operation in Figure 1.10(b). Thus, for the given incident and outgoing directions, this setup captures all the direct, subsurface or volume transport components and secondary components. The LTM acquired with this illumination configuration is thus versatile in its representational power and can handle spatially varying illumination. The setup can be extended to sample the entire 4D incident illumination space defined by ω_i and x_i , by acquiring several LTMs, by moving the projector along a sphere or a hemisphere.

The LTM with structured light was first presented as environment matting [79]. Since then, numerous methods [49, 55, 76] were proposed to enable faster acquisition. More recently, this has been extended to more general scenes. These approaches exploit low dimensionality under sparse transforms [56] or use low-rank approximations [52, 73] or use a combination of spatial and frequency domain sampling [58] to reduce the acquisition time. An acquisition method that moves the projector along various incident illumination directions to sample the entire 4D incident illumination has also been proposed [47]. An external diffuser is used in [52, 73] that results in a sampling kernel that approximates the projector with an array of point light sources. These methods then exploit the low rank nature of the resultant LTM for fast acquisition. The resulting setup, while being relatively compact like the projector camera setups, can capture the effects of varying incident angle like the light-stage setups.

1.5 Thesis outline and Contributions

The thesis focuses on image based methods for acquiring or modeling the light transport characteristics of cultural artifacts, with each chapter focusing on a distinct acquisition scenario. In this section, we present a brief summary of each chapter and its contributions.

1.5.1 Inverse Rendering of Stained Glass

In Chapter 2, we study the digitization of stained glass windows using the light transport matrix model. Stained glass windows are translucent, heterogeneous, very large in size and often cannot be acquired under controlled illuminations. We exploit the structure of the light transport matrix to present a fast acquisition method under controlled illuminations. We then build a material-specific sparsifying dictionary and propose a light transport acquisition system that works both under controlled and uncontrolled illuminations. The main contributions of this chapter are

- A method to acquire the light transport matrix of planar materials under controlled illuminations, without any priors.

-
- A method to obtain a basis with image based priors for translucent material scattering under controlled illuminations.
 - A sparse recovery algorithm for the acquisition of light transport properties of large, mostly planar objects, that preserves volume impurities, to ensure a heterogeneous light transport even under uncontrolled illumination.

1.5.2 Synthetic Rendering of Stained Glass

In Chapter 3, we study the forward rendering problem of stained glass windows, from an image based rendering perspective. Traditionally, volume transport of light is simulated using the classic volume rendering equation. This requires knowing the spatially varying scattering coefficients induced by heterogeneous light transport, as well as models to generate impurities in stained glass. We overcome these difficulties using an image-prior based approach for light transport synthesis. The presented approach has applications in simulating new stained glass windows as well as in the digital restoration of broken stained glass windows. The main contributions of this chapter are

- A matrix completion algorithm for light transport design in heterogeneous glass slabs based on material specific sparsifying dictionaries.
- The first framework to synthesize light transport in complex heterogeneous objects with an image based approach.
- A framework to aid in the digital restoration of broken pieces of stained glass slabs.

1.5.3 Handheld Acquisition of Light Transport

In Chapter 4, we propose an alternate image based rendering scheme for scenes that predominantly exhibit surface reflections. We show that under the special case that the object exhibits convex surface geometry, the acquired light transport is equivalent to that obtained with lightstages. We then propose a strategy for the handheld acquisition of light transport. We also present a reconstruction framework to obtain an approximation to uniformly sampled light transport, given samples along a trajectory. The main contributions of this chapter are,

- Analysis of light transport acquired with point light sources and an image based rendering model.
- A mobile phone based framework for acquiring the reflectance functions of physical objects and its sampling model.
- A reconstruction framework for non-uniformly sampled light transport matrices.
- An extension to acquire 6-D light transport (bidirectional texture function), using a handheld setup.

Chapter 2

Inverse Rendering of Stained Glass

I think the problem, to be quite honest with you, is that you've never actually known what the question is.

Douglas Adams, The Hitchhiker's Guide to the Galaxy

In this chapter, we will discuss our inverse rendering framework for the digitization of stained glass windows.¹

2.1 Introduction

Stained glass is an artistic medium that exploits the scattering properties of colored, translucent glass panes. Stained glass artifacts, which have a millenary tradition, have been (and still are) produced in a surprising variety of forms and techniques but clearly the quintessential examples of stained glass artistry are to be found in medieval buildings such as the Chartres or Canterbury cathedrals or the Sainte-Chapelle. Remarkably, the windows that have survived to our times have done so extremely well and they represent the most chromatically accurate testimony of medieval art available to us. However, much like the architectural structures that they are part of, stained glass windows cannot be moved and their detailed observation is often made difficult by their size and placement; ideally, a high-quality acquisition and rendering toolkit would allow scholars and art lovers to interact with this art form in ways that are simply not possible using direct observation.

Unfortunately, stained glass artwork “resists” many digitization approaches in a number of ways. One well-known difficulty is its high dynamic range, a quality amplified by the contrast

1. The work presented in this chapter was done in collaboration with L. Baboulaz, P. Prandoni, M. Vetterli, S. Trümpner and S. Wolf [67, 68]

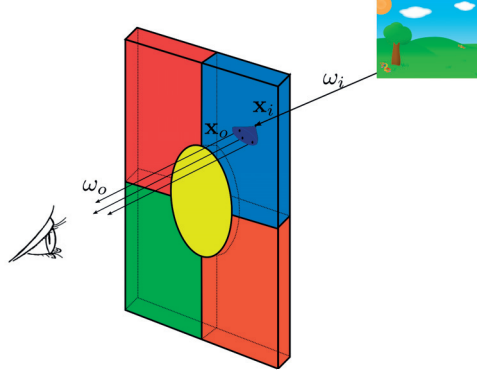


Figure 2.1: Light transport in stained glass is a function of the direction and point of entry of incident light and the direction and point of exit of exitant light, as determined by the bidirectional scattering distribution function.

between the dimly lit interior of a cathedral and the backlit imagery on the windows. But even before we take dynamic range into account, we need to address the fact that stained glass is an eminently non-static medium, designed to be experienced across the many different lighting conditions that change both seasonally and within a single day. Still photographs are therefore a woefully inadequate rendition, since they capture just a single point in the range of all possible illuminations. We are thus interested in obtaining an interactive model of stained glass windows that allows for relighting under dynamic illumination.

In principle, to render any artifact under arbitrary illuminations and viewpoints a sampled version of its 8-dimensional light transport $S(\omega_o, x_o, \omega_i, x_i)$ is required, clearly an impractical proposition. Fortunately, in the case of largely planar stained glass artifacts, we can fix the entry and exit directions as perpendicular to the glass plane; by doing so we simplify the problem to the acquisition of the light transport properties of the object from the back plane of glass to its front.

In this chapter we propose a novel approach called VITRAIL (Volume Impurities and Translucency for Rendering Artifacts with Interactive Lighting) for acquiring and modelling stained glass artifacts that enables interactive rendering capabilities. During the acquisition stage, we obtain image pairs of incident illumination and stained glass appearance. Stained glass is then modelled using light transport matrices obtained by solving a linear inverse problem. Furthermore at this stage, we exploit the known approximate geometry of stained glass to obtain a compact representation for light transport. This in turn facilitates the learning of a sparsity inducing basis for light transport, both of which are utilized as strong regularizers thus allowing for acquisition even under natural uncontrolled illumination. Finally, with our light transport matrix representation, rendering under novel illuminations is obtained with a simple matrix vector product.

2.2 Related Work

The inverse rendering problem under controlled illuminations is a well studied topic in the graphics and vision communities. We briefly review the main methods and outline our approach.

2.2.1 Inverse Rendering Methods

Image based acquisition methods that take the light stage approach [14, 15, 24, 31] and polynomial texture mapping (PTM) [44] capture the spatially varying reflectance of a scene sampling with directional light. These methods are however ill-suited for translucent objects as volumetric and subsurface scattering are functions of spatially varying light. Goesele et al [29] describe a method to digitize translucent objects by scanning its response for various incoming and outgoing angles and interpolating the reflectance function over the object’s surface geometry by assuming a smooth global transport. As described in Chapter 1, LTMs acquired with a spatially varying light source, can directly acquire volume and subsurface scattering components, in addition to direct and secondary components of light transport. It was originally developed to capture refractions in transparent objects [50, 79]. Several approaches have been proposed to speed up the acquisition, including compressive sensing [22, 56], low rank matrix approximation [52, 73] and spectral decomposition [58] have been proposed for solving the inverse problem. Peers et al [53] also designed a spatially programmable curved light stage, thus allowing a hybrid between both approaches, to capture the 6D light transport. These methods are however designed for generic scenes while we exploit the geometry and scattering properties of stained glass for faster acquisitions.

2.2.2 Empirical models

We refer to [35] for a detailed review on the acquisition of transparent, translucent and specular objects. Jensen et al [37] proposed the dipole approximation of the diffusion equation, to model homogeneous subsurface transport for synthetic objects. Since then, several methods have arisen that measure the scattering parameters of materials in terms of a forward and backward scattering coefficient and an absorption coefficient. The most recent work that takes this approach [28], builds a material dictionary on the three scattering parameters. Wang et al [74] model heterogeneous light transport by taking a layered approach to the diffusion equation. They first solve for an inverse diffusion equation, to characterize the heterogeneity in physical samples, which is then used to model synthetic objects. Peers et al [54] describe a compact representation for heterogeneous subsurface transport and represent heterogeneous objects by layers of homogeneous materials. Donner et al [20] take a similar approach to model skin. While these models can be utilized in building virtual stained glass windows, they do not aid in inverse rendering of existing artifacts.

2.2.3 Our approach

We start by acquiring an ensemble of illumination and scene response image pairs under controlled illumination. By exploiting the known approximate planar nature of glass, we solve for a linear inverse problem to obtain the light transport matrix of planar glass slabs. As a one-off preprocessing step, we then learn a sparsity inducing dictionary for light transport. Finally, when faced with large scale digitization of stained glass under controlled or natural illuminations, we recover light transport matrices as a function of acquired image pairs, the sparsifying dictionary and the known compact representation. Given the light transport matrix, rendering under novel illuminations is obtained using a simple matrix vector product. Our approach to digitizing artifacts is unique in that we present the first method that uses image based priors in the form of a dictionary for light transport acquisition. These priors are then used as strong regularizers thus enabling faster acquisition in controlled illumination while allowing for acquisition even under

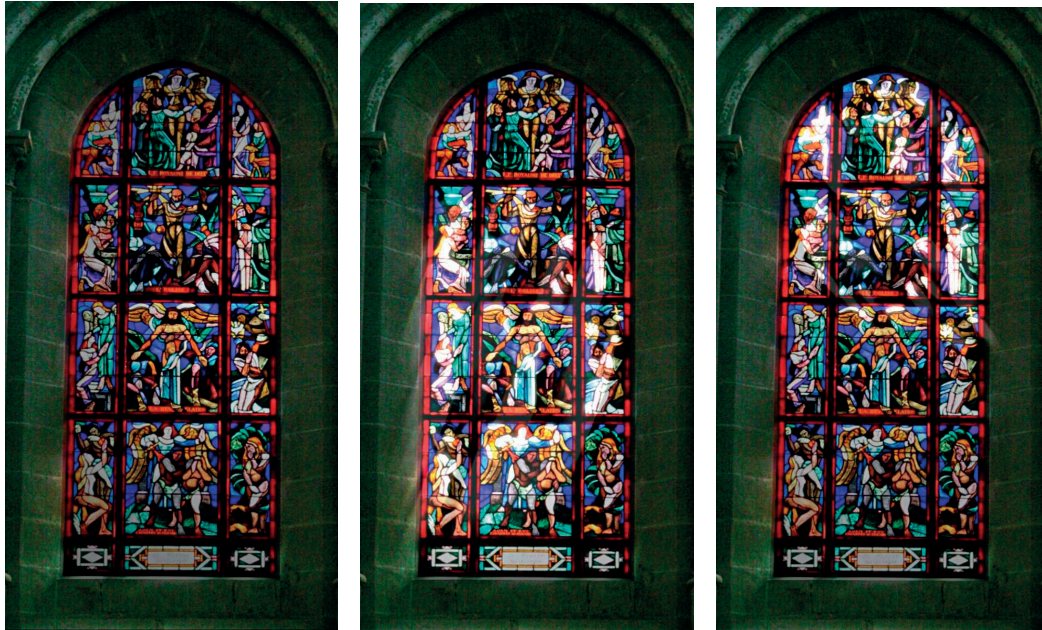


Figure 2.2: A stained glass window from the Cathedral of Lausanne, rendered under a directional light source, using the dual microfacet model [11].

uncontrolled, natural illuminations. This approach preserves heterogeneity including bubbles and corruptions in the digitized artifact.

2.3 Digital Modeling of Stained Glass

As a precursor to our project, we first attempted at rendering stained glass windows using the dual microfacet model [11] for the Bidirectional Transmittance Distribution Function. While this can render stained glass (as shown in Figure 2.2) under directional light sources such as the sun, such a model treats stained glass as a refractive medium.

In contrast to transparent planar glass windows where the transmission of light is dominated by refraction effects, stained glass windows are also translucent and thus light is scattered as it travels through it. The scattering properties of stained glass is not homogeneous either, due to various factors such as, the coloring, the infusion of metallic salts, the presence of air bubbles inside the glass and the irregularity of the surface of stained glass. Furthermore, over centuries, stained glass undergoes transformations of various types, a common one being the corrosion of the surface of the glass exposed to the outside environment (Figure 2.3). This complex blend of factors result in highly heterogeneous light scattering properties. Our goal is to aid in both digitally archiving stained glass windows and in rendering them under dynamic illuminations. Thus, we are interested in preserving the heterogeneity of light transport in stained glass. Since the physics behind the interaction of light with various materials is well understood, it is tempting to pose the inverse problem as one that estimates the exact physical model of the stained glass being digitized. However, this requires the estimation of spatially

varying scattering coefficients, the location, size and shape of air bubbles and other impurities, and the nature of various deteriorations in addition to the structure of surface irregularities. We circumvent these requirements by posing this as an image based rendering problem. In the remainder of this section, we will introduce our image formation model and discuss the structure of light transport.

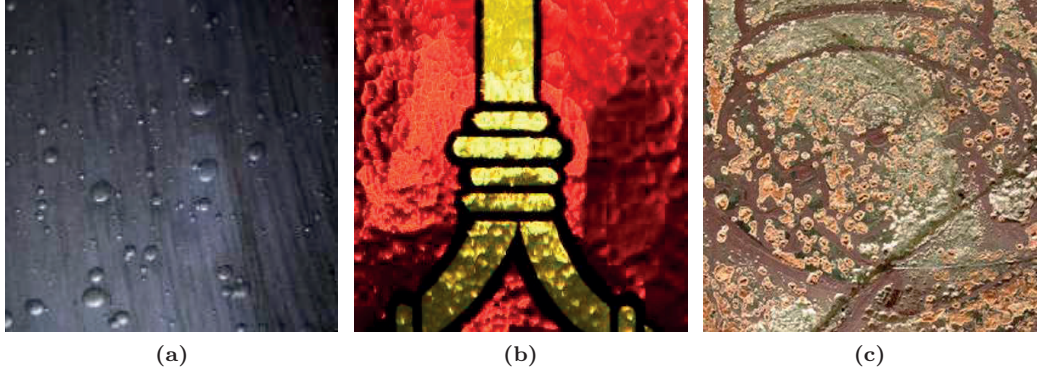


Figure 2.3: Sources of Heterogeneity : (a) A stained glass slab containing air bubbles as volume impurities. (b) The surface structure present on a modern stained glass slab. (c) Corroded outer face of a stained glass slab.

2.3.1 Image Formation Model

We describe image formation in a digitized stained glass using the light transport matrix. We assume a fronto-parallel scenario where the illumination plane, the stained glass window and the camera sensor are parallel to each other. Image formation in each of the color channels (R,G,B) of the camera is then given by

$$\mathbf{y} = \mathbf{T}\boldsymbol{\ell}, \quad (2.1)$$

where $\mathbf{T} \in \mathbb{R}^{N \times M}$ is the light transport matrix, $\mathbf{y} \in \mathbb{R}^{N \times 1}$ is the corresponding color channel of the scene image and $\boldsymbol{\ell} \in \mathbb{R}^{M \times 1}$ is the incident illumination pattern. Both \mathbf{y} and $\boldsymbol{\ell}$ are vectorized versions of the original 2D camera and illumination plane, which have resolution N and M respectively. In case of indoor experiments, the illumination plane is represented by the plane of focus of a projector. The light transport matrix thus defines the transport of light from individual elements on the illumination plane $\boldsymbol{\ell}$ to individual elements (camera pixels) on the image plane \mathbf{y} . Since the projector used to generate the illumination plane is focused on the back plane of glass, while the camera is focused at its front plane, we obtain the complete characterisation of light transport from the back plane of the stained glass to its front plane using the light transport matrix.

2.3.2 Structure of the Light Transport Matrix

The light transport matrix \mathbf{T} is such that its m^{th} column corresponds to the response of the glass to the m^{th} euclidean basis vector \mathbf{e}_m *i.e.*, it represents the discretized point scattering function induced by the glass on the m^{th} illumination element. Thus various kinds of glass have

varied structure in their light transport matrices, as dictated by the point scattering functions. For ease of illustration, let us consider the case where we have a 1D camera and a 1D projector. Then the structure of the resulting 2D light transport matrix is described below.

Transparent, planar glass: In case of a planar, transparent glass, the light transport matrix is an identity matrix since light at normal incidence is transmitted without refraction or scattering events as shown in Figure 2.4(a). The light transport matrix of a colored, transparent glass plane will be a diagonal matrix. A rendered image is shown in Figure 2.5(a).

Transparent, nearly planar glass: A transparent glass with surface irregularities on the other hand induces refraction events as dictated by Snell’s laws. Thus the light transport matrix will still be composed of one non-zero entry per column, but distributed around the leading diagonal, as shown in Figure 2.4(b).

Homogeneous, Translucent, planar glass: A translucent glass with homogeneous point scattering functions will have a light transport matrix that is banded and Toeplitz, as the glass can be modelled as a spatially invariant low pass filter of finite support. A rendered image is shown in Figure 2.5(b).

Heterogeneous, Translucent glass: Consider a completely planar but heterogeneous glass with spatially varying point scattering functions; it will still have a banded light transport matrix but is no longer Toeplitz. The presence of surface irregularities in such a glass will yield a banded light transport matrix with bandwidth slightly larger than for the completely planar case, since refraction at the surface irregularities cause the point scattering functions to be translated from the leading diagonal, as seen in Figure 2.4(d). A rendered image is shown in Figure 2.5(c). Let v be the width of the band (sum of upper and lower bandwidths of the band matrix) induced by a particular type of glass, and δ denote the maximum translation due to surface irregularities, then $\nu = v + \delta$ is the width of the band of the underlying light transport matrix. Note that when we extend this to the real life scenario where we have a 2D camera and a 2D projector, the resulting light transport, when represented as a 4D tensor will still have a banded volume structure along its diagonal. Since we represent the 4D tensor as a 2D matrix by re-organizing the image plane and the illumination plane in a column major format, the resulting light transport matrix has its non zero elements distributed as bands around the leading diagonal.

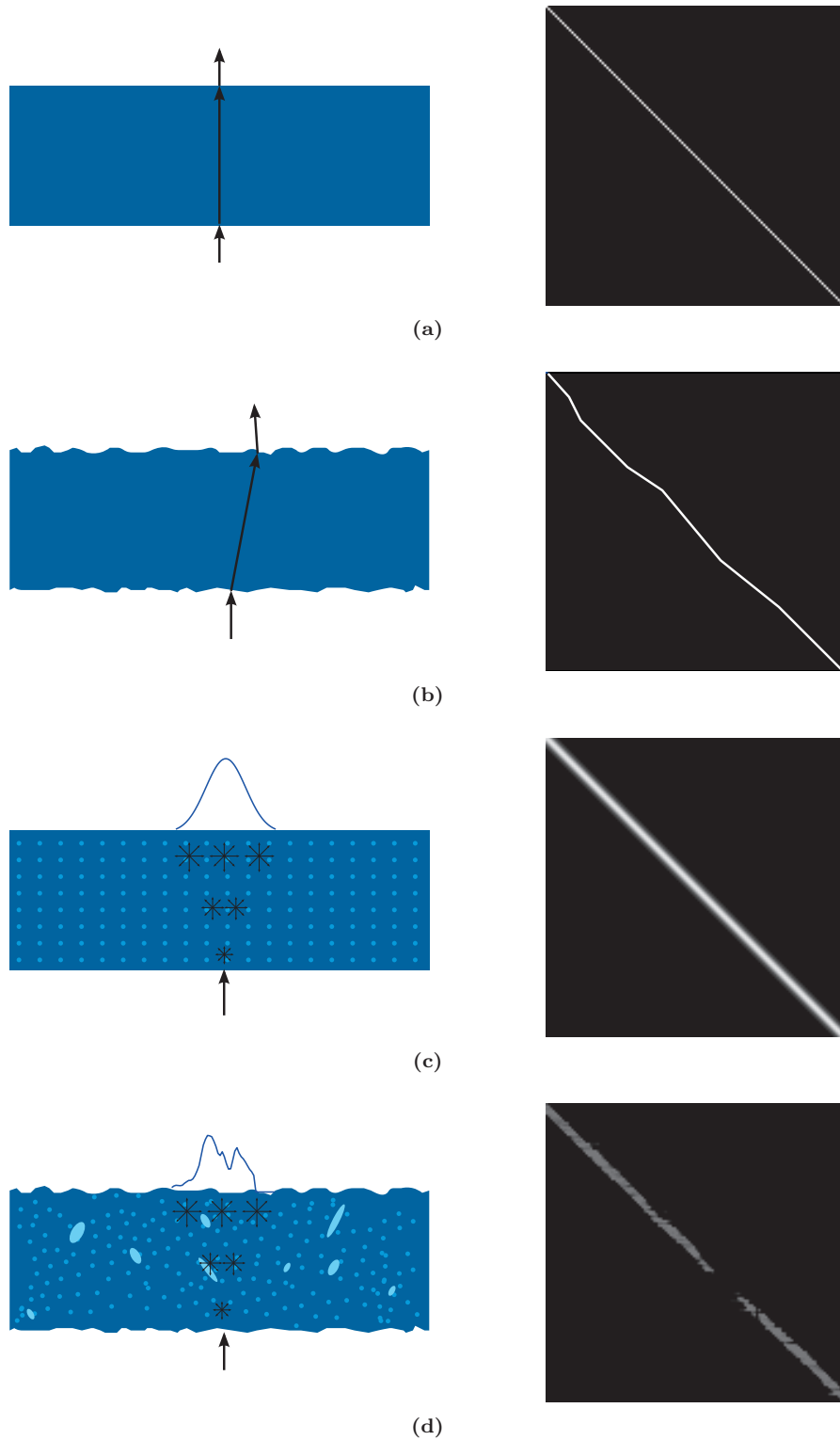


Figure 2.4: Illustration of light transport in various types of glass slabs. 2D-slices of the light transport matrix are shown on the side. (a) A transparent planar glass slab, for light at normal incidence. (b) A transparent glass slab with surface irregularities, for light at normal incidence (refraction is the dominant light transport phenomena). (c) A translucent homogeneous planar glass slab, for light at normal incidence (homogeneous volumetric scattering is the dominant light transport phenomena). A typical stained glass slab, for light at normal incidence (refraction and volumetric scattering occur, heterogeneously).



Figure 2.5: We render a stained glass window assuming different models for light transport. (a) A coloured, transparent model (alpha-matting). (b) A homogeneous translucent model. (c) A heterogeneous translucent model.

The amount of light reaching the n^{th} pixel $\mathbf{y}[n]$ on the camera sensor, is completely defined by the n^{th} row $\mathbf{t}_n = \mathbf{T}[n, :] \in \mathbb{R}^{1 \times M}$ of the light transport matrix², as given by the inner product

$$\mathbf{y}[n] = \mathbf{t}_n \boldsymbol{\ell}.$$

Thus, the n^{th} row of the light transport matrix corresponds to the discretized transmittance function of the surface point at the n^{th} pixel. We refer to \mathbf{t}_n as the n^{th} light transport vector.

2.4 Sparse Representation of Light Transport

Compressive sensing has enabled faster acquisition of signals that have a sparse representation in a known basis, provided that the measurement matrix satisfies the restricted isometry property [21], which ensures that linear measurements of sparse signals are nearly orthogonal. For a detailed review of compressive sensing and its applications, see [6]. It is thus advantageous to find a sparsity inducing basis for light transport vectors.

2.4.1 Low Dimensionality of Light Transport

It has been conjectured [4], and verified by empirical analysis [25] that the bidirectional scattering distribution function in general occupies a low dimensional space. The light transport matrix is a 4D slice of the full 8D bidirectional scattering distribution function obtained by fixing the illumination plane and the image plane. We expect it to lie in a low dimensional space too. While the light transport matrix of stained glass windows is often full rank, from the discussion in the previous subsection, we know that it is a banded matrix. Assume we know the width ν of the banded matrix. We can then define an indicator vector $\boldsymbol{\gamma}_n \in \mathbb{R}^M$ for each light transport vector \mathbf{t}_n , such that

$$\boldsymbol{\tau}_n = \mathbf{T}[n, \boldsymbol{\gamma}_n] \in \mathbb{R}^{1 \times \nu} \quad (2.2)$$

2. Given a matrix \mathbf{A} and scalars n, m , $\mathbf{A}[n, :]$ denotes the n^{th} row of the matrix, while $\mathbf{A}[:, m]$ denotes the m^{th} column of the matrix. If $\boldsymbol{\gamma}$ is a vector that acts as an indicator function, then $\mathbf{A}[n, \boldsymbol{\gamma}]$ denotes the n^{th} row of the matrix, subsampled by $\boldsymbol{\gamma}$.

contains the entire non-zero block of the n^{th} light transport vector. A re-parametrized light transport matrix can be constructed from an ensemble of all the reduced light transport vectors τ_n . As shown in Figure 2.6, the spectral decay of singular values is faster with the re-parametrized light transport matrix, when compared with that of the full light transport matrix. We can thus infer that the light transport matrix does indeed lie in a lower dimensional space and the reduced light transport vectors, τ_n are pre-disposed for building a basis that represents this space.

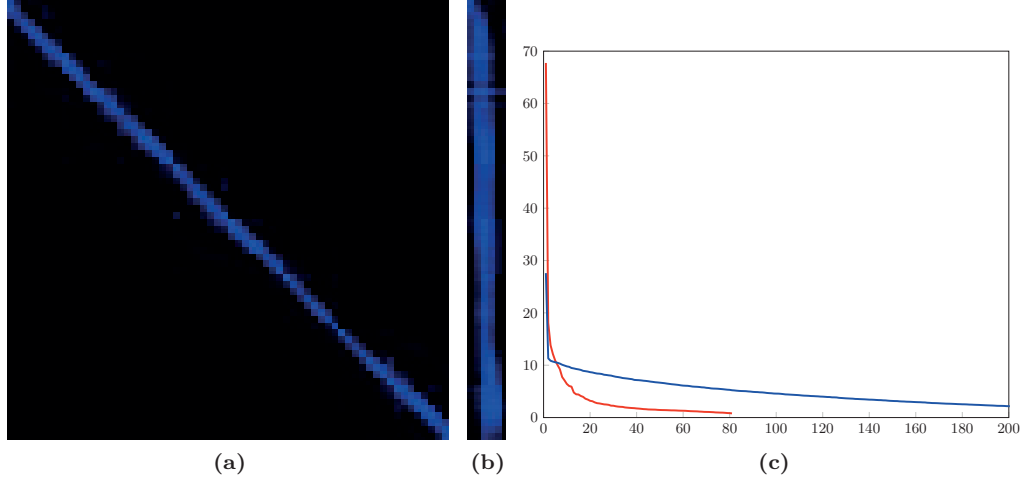


Figure 2.6: Low dimensionality of light transport. (a) A 2D slice of the full light transport matrix of a blue stained glass slab. (b) The re-parametrized light transport matrix, obtained by stacking τ_n . (c) The figure shows the singular value decay of the original light transport matrix and that of the re-parametrized transport matrix. Spectral decay for the 200 largest singular values of full LTM is shown in blue. The spectral decay for the same LTM after re-parametrization is shown in red. Notice the faster decay in the latter case. The prior knowledge about the banded nature of LTM for flat objects can therefore be exploited for a more efficient representation of the LTM.

2.4.2 Learning sparsity inducing dictionaries

Since the light transport properties are material specific, a sparsifying dictionary for each kind of glass is learnt separately. The following conditions are imposed upon the dictionary that is designed:

Condition 1 The basis should span the same low dimensional space that is spanned by the volume scattering component of light transport.

Condition 2 The light transport component should have a sparse representation in the learnt basis.

We start by acquiring the light transport matrices of a few glass planes using the method for controlled acquisition described in the next section. The next step in building the dictionary is a data pooling step. Note that the finite, compact support v of light transport vectors is caused

by volume scattering, whereas the small translations δ are due to the presence of surface irregularities. In order to account for all possible translations, we first recover the volume scattering component by re-organizing each light transport vector \mathbf{t}_n as a 2D image and segmenting the largest connected component in each \mathbf{t}_n . The support v for a given material is the largest spatial support spanned by the observed light transport vectors. To account for surface irregularities, in addition to the extracted τ_n , δ translated versions of each τ_n are pooled into a container matrix $\mathbf{B} \in \mathbb{R}^{v \times C\delta}$ where C is the total number of observed light transport vectors. Here, δ is the largest translation expected due to surface irregularities.

Let $\mathbf{D} \in \mathbb{R}^{K \times v}$ denote the basis being constructed, where K is the number of atoms in the dictionary, then Condition 1 implies that all light transport vectors can be written as

$$\tau_n = \alpha_n \mathbf{D} \quad \forall n \in [1 \cdots C]. \quad (2.3)$$

Furthermore, Condition 2 implies that each $\alpha_n \in \mathbb{R}^{1 \times K}$ is sparse *i.e.*, only a few entries in each α_n are non-zero.

We now learn a dictionary, \mathbf{D} on \mathbf{B} , such that Equation (2.3) is satisfied. This is done traditionally, by alternating between an l_1 -minimization step,

$$\alpha_n = \arg \min_{\alpha_n} \|\alpha_n\|_1 \text{ s.t. } \|\mathbf{b}_n - \alpha_n \mathbf{D}\|_2 \leq \epsilon, \quad \forall n = [1 \cdots C\delta],$$

and an l_2 -minimization step,

$$\mathbf{D} = \arg \min_{\mathbf{D}} \sum_{n=1}^{C\delta} \|\mathbf{b}_n - \alpha_n \mathbf{D}\|_2,$$

until convergence. The resulting dictionary, \mathbf{D} is the required basis in which each $\tau_n \in \mathbf{B}$ has a sparse representation. In the next section, we discuss our framework for the acquisition of light transport matrices.

Algorithm 2.1 Learning priors

```

B = {}
for  $n = 1$  to  $C$  do
     $\tau_n$  = largest connected component  $\{\mathbf{t}_n\}$ 
     $\mathbf{B} = \mathbf{B} \cup \{\delta \text{ translated versions of } \tau_n\}$ 
end for
Init  $\mathbf{D} \in \mathbb{R}^{K \times v}$ 
while Not Converged do
     $\alpha_n = \arg \min_{\alpha_n} \|\alpha_n\|_1 \text{ s.t. } \|\mathbf{b}_n - \alpha_n \mathbf{D}\|_2 \leq \epsilon, \quad \forall n = [1, \cdots C\delta]$ 
     $\mathbf{D} = \arg \min_{\mathbf{D}} \sum_{n=1}^{C\delta} \|\mathbf{b}_n - \alpha_n \mathbf{D}\|_2$ 
end while
return  $\mathbf{D}$ 

```

2.5 Acquisition of Stained Glass

An illustration of our acquisition setup in indoor scenarios is shown in Figure 2.7.

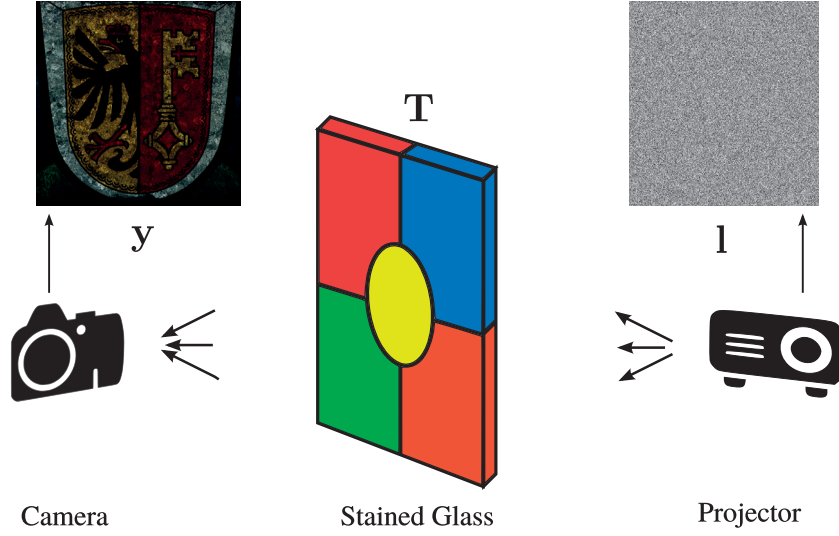


Figure 2.7: Illustration of indoor acquisition setup. The pattern ℓ is projected on the glass, which is observed as the image \mathbf{y} by the camera.

Let $\mathbf{Y} = [\mathbf{y}_1 \ \mathbf{y}_2 \ \cdots \ \mathbf{y}_P] \in \mathbb{R}^{N \times P}$ and $\mathbf{L} = [\ell_1 \ \ell_2 \ \cdots \ \ell_P] \in \mathbb{R}^{M \times P}$ denote the ensemble of observed image and illumination vectors respectively, obtained by stacking these vectors to form matrices. Then from the image formation Equation (2.1), we can write

$$\mathbf{Y} = \mathbf{T}\mathbf{L}. \quad (2.4)$$

If the matrix \mathbf{L} is chosen to be a unitary matrix, the above system of equations can easily be solved by inverting the unitary matrix. However, consider a reasonably modest projector resolution (the discrete illumination plane) of 512×512 . Then $M \simeq 2.6 \times 10^5$ and thus, to solve the linear system (2.4), a staggering 2.6×10^5 observations would be required, which would make such an approach impractical. We therefore look for formulations that exploit our prior knowledge on the light transport matrix of nearly planar glass. Consider the case where the number of observations $P < M$. Equation (2.4) becomes an under-determined linear system with no unique decomposition into \mathbf{T} and \mathbf{L} . In the absence of any priors, a matrix \mathbf{T} such that,

$$\begin{aligned} \mathbf{T} &= \arg \min_{\mathbf{T}} \|\mathbf{Y} - \mathbf{T}\mathbf{L}\|_F \\ &= \arg \min_{\mathbf{T}} \sum_{i=1}^P \|\mathbf{y}_i - \mathbf{T}\ell_i\|_2 \end{aligned} \quad (2.5)$$

is required, where $\|\cdot\|_F$ denotes the Frobenius norm. Solving for equation (2.5) involves optimizing over $N \times M$ variables at the same instance. For a camera with a modest resolution (the discrete image plane) of 1024×1024 , $N \simeq 10^6$ and $N \times M \simeq 2.6 \times 10^{11}$. Clearly, such a problem is intractable. An alternative formulation is

$$\begin{aligned}\mathbf{T} &= \arg \min_{\mathbf{T}} \sum_{n=1}^N \|\mathbf{Y}[n, :] - \mathbf{t}_n \mathbf{L}\|_2 \\ \mathbf{t}_n &= \arg \min_{\mathbf{t}_n} \|\mathbf{Y}[n, :] - \mathbf{t}_n \mathbf{L}\|_2 \quad \forall n = [1, \dots, N],\end{aligned}\tag{2.6}$$

where $\mathbf{Y}[n, :] \in \mathbb{R}^{1 \times P}$ and $\mathbf{t}_n \in \mathbb{R}^{1 \times M}$ are the n^{th} rows of \mathbf{Y} and \mathbf{T} respectively. Thus, the equivalent problem shown above solves for one light transport vector \mathbf{t}_n at a time. Equation (2.6) is referred to as the inverse rendering equation.

In the rest of this section, a practical acquisition strategy for stained glass digitization under controlled and uncontrolled illumination conditions is presented.

2.5.1 Controlled Acquisition

We know from our discussion on its structure that the light transport matrix of stained glass is a banded matrix. Let \mathbf{W}_n be a square window of area ν defined on the 2D illumination plane, around the pixel n of the image plane, obtained by overlaying the image plane on the illumination plane. An indicator function γ_n that determines which light elements from the illumination plane contribute to outgoing light at image pixel n , can then be defined as

$$\gamma_n = \begin{cases} 1 & \text{if } m \in \mathbf{W}_n \\ 0 & \text{otherwise.} \end{cases}$$

Then the ensemble of illumination vectors that contribute to a given image pixel n , $\mathbf{L}_n \in \mathbb{R}^{\nu \times P}$ can be obtained by retaining only the rows of \mathbf{L} whose indices have a non-zero value in γ_n as denoted by,

$$\mathbf{L}_n = \mathbf{L}[\gamma_n, :].$$

The image formation equation for pixel n can now be rewritten as $\mathbf{y}[n] = \tau_n \ell[\gamma_n]$ where $\tau_n \in \mathbb{R}^{1 \times \nu}$ is the reduced light transport vector, as defined in (2.2).

We will now consider two scenarios. A one-off acquisition scenario, where we only acquire one stained glass window of a given type and a large scale scenario, where we are required to acquire several windows of the same type.

One-Off Acquisitions

The inverse rendering equation can be written as,

$$\mathbf{Y}[n, :] = \tau_n \mathbf{L}_n \quad \forall n = [1 \dots N],\tag{2.7}$$

where $\mathbf{Y}[n, :] \in \mathbb{R}^{1 \times P}$ is the vector containing P observed values for pixel n . Since $\nu \ll M$, the above system of equations will have unique solutions when $P > \nu$, while still requiring just $P \ll M$ observations. Thus, by exploiting the banded nature of the light transport matrix, we solve for

$$\tau_n = \arg \min_{\tau_n} \|\mathbf{Y}[n, :] - \tau_n \mathbf{L}_n\|_2 \text{ s.t. } \tau_n \geq 0 \quad \forall n = [1 \dots N],\tag{2.8}$$

where the non-negativity of light transport is also imposed as a constraint.

Compressive Acquisition

If a sparsifying dictionary exists then with a carefully constructed \mathbf{L}_n , we can employ compressive sensing for faster acquisitions. This is particularly attractive in case of large scale digitizations of similar types of stained glass where we can first obtain the light transport matrix of a few stained glass slabs using the one-off approach discussed previously, then learn a sparsifying dictionary from these light transport matrices in order to digitize the rest of the samples via sparse sampling.

Let us now consider perfectly planar, but heterogeneous glass *i.e.*, $\nu = v$. Since we have built the priors such that, $\tau_n = \alpha_n \mathbf{D}$, the above inverse rendering equation (2.7) can be rewritten as

$$\mathbf{Y}[n, :] = \alpha_n \mathbf{D} \mathbf{L}_n \quad \forall n = [1 \cdots N].$$

When the illumination ensemble \mathbf{L}_n satisfies the RIP, the above problem can be solved with $P < S$ measurements by minimizing the l_2 -norm of the residual, such that the coefficients of τ_n in \mathbf{D} form a sparse vector.

$$\begin{aligned} \tau_n &= \alpha_n \mathbf{D} \quad \text{where,} \\ \alpha_n &= \arg \min_{\alpha_n} \|\mathbf{Y}[n, :] - \alpha_n \mathbf{D} \mathbf{L}_n\|_2 \quad \text{s.t.} \quad \|\alpha_n\|_1 < \mu, \\ \alpha_n &\geq 0. \end{aligned} \tag{2.9}$$

In practice for compressive acquisition, we choose the illumination ensemble \mathbf{L} to be a Gaussian random matrix, with entries chosen from the normal distribution. In contrast to previous methods that use compressive sensing for light transport matrix acquisition, we have τ_n to be exactly sparse in \mathbf{D} . Furthermore, by making use of the banded nature of the light transport matrix, we have also reduced the problem size by an order of magnitude. Thus, we can gain in both acquisition time and computational complexity with our approach, in the presence of a controlled light source. We present our method to handle surface irregularities in the next section.

Under controlled illumination when spectrally uniform light is used, the light transport for each color channel (in RGB space) can be obtained by solving for the light transport matrix according to equations (2.9) for each channel independently. Thus for RGB cameras, we solve for the red, blue and green channel light transport matrices independently.

2.5.2 Uncontrolled Acquisition

Let us first see why natural illumination makes inverse relighting a severely ill conditioned problem. We will later show how we can use our framework to regularize the inverse problem.

We know that the incident and the outgoing light-fields in a scene are related by the scene's BSDF. Natural lighting implies that the sun is the only active light source throughout the day, while light reflected from the scene surrounding the object acts as the ambient light. Consider a stained glass window located in an outdoor setting where the BSDF of the scene surrounding the object of interest is $S_E(\omega_i, x_i; \omega_o, x_o)$. Assume that we sample the outgoing luminance of the object at time instant j throughout the day. This would imply that we are sampling the BSDF of the object, $S_O(\omega_i, x_i; \omega_o, x_o)$ at J incident illumination configurations, $L^j(\omega_i, x_i)$. If $L_s^j(\omega_i)$ denotes the sunlight (direct rays of sun are directional), then the incident illumination on the object is given by

$$L^j(\omega_i, x_i) = \langle (1 + S_E(\omega_i, x_i; \omega_o, x_o)), \mathbf{L}_s^j(\omega_i) \rangle.$$

On a given day, the sun spans only a single arc, in the hemisphere that describes the set of all possible illuminations. Thus, the ensemble $\{L^j(\omega_i, x_i)\}$ is well conditioned only when $\{\langle S_E(\omega_i, x_i; \omega_o, x_o), L_s^j(\omega_i) \rangle\}$ is well conditioned. However, several empirical studies have indicated that the BSDF of a natural scene, at the macroscopic level is a very low dimensional function. The ensemble of incident illuminations on the object is thus the sum of two low-dimensional spaces and hence is in turn, low-dimensional. Thus the recovery of $S_O(\omega_i, x_i; \omega_o, x_o)$, by solving for the set of equations:

$$\{Y^j(\omega_o, x_o)\}_J = \{\langle S_O(\omega_i, x_i; \omega_o, x_o), L^j(\omega_i, x_i) \rangle\}_J,$$

is severely ill-conditioned, irrespective of whether we try to recover the light transport matrix or the spatially varying-BRDF/BTDF. For natural illumination, the solution that minimizes the l_2 -norm of the linear system $\mathbf{Y} = \mathbf{T}\mathbf{L}$ is given by the pseudo-inverse of \mathbf{L} .

Now consider the solution obtained by solving the system of equations in (2.9). After imposing the constraint that each τ_n is in the row space of \mathbf{D} , we search for an α_n that minimizes the l_2 -norm of the error, while requiring that the solution is sparse in \mathbf{D} . This new formulation regularizes \mathbf{T} in two ways,

- It ensures that the recovered solution is in the same subspace spanned by the light transport vectors of the material being investigated.
- The recovered light transport matrix is always banded, thus ensuring that it corresponds to an almost planar object.

We can recover the full spectrum light transport, by treating the RGB channels of the incident illumination and outgoing illumination independently and solving equations (2.9) to obtain three light transport matrices, one for each channel. Note that this method solves for differently conditioned light transport at each channel. We present an alternative formulation to deal with varying condition numbers in the next section.

2.6 Algorithmic issues

In this section, we discuss how to handle surface irregularities in glass and the varying condition numbers of different channels of incident natural illumination.

2.6.1 Handling surface irregularities

The convex problem in Equation (2.9) provides the light transport vector when the object has an exactly planar surface geometry. Thus, γ_n for $n = [1, 2, 3..N]$ are exactly known *i.e.*, τ_n occur exactly around the diagonal of \mathbf{T} . In practice, stained glass windows are only near planar, with a varying amount of surface roughness. Since light is refracted at the outer interface of glass before being scattered inside the glass, τ_n occur with a slight offset about the leading diagonal. Here, γ_n is only approximately known. We account for this by defining a series of γ_n^r , $r \in [1, 2, ..\delta]$ about n each defined by translated versions of \mathbf{W}_n . The resultant light transport vector is then defined by the average of the individual τ_n^r . We now redefine $\mathbf{L}_n^r = \mathbf{L}[\gamma_n^r, :]$. The light transport matrix at pixel n is then calculated as

$$\begin{aligned}
\mathbf{t}_n &= \frac{1}{\delta} \sum_r \mathbf{t}_n^r \\
\text{where, } \mathbf{t}_n^r[\gamma_n^r] &= \tau_n^r, \quad \forall r \in [1, 2, \dots, \delta] \\
\tau_n^r &= \alpha_n^r \mathbf{D} \quad \text{where} \\
\alpha_n^r &= \arg \min_{\alpha} \|\mathbf{Y}[n, :] - \alpha_n^r \mathbf{D} \mathbf{L}_n^r\|_2 \quad \text{s.t. } \|\alpha_n^r\|_1 < \mu, \\
\alpha_n^r &\geq 0.
\end{aligned} \tag{2.10}$$

As we include translated versions of sample τ_n while learning \mathbf{D} , each τ_n^r will in-fact still reside in the row space of \mathbf{D} . This step is similar to the translation for shift invariance normally used in implicit dictionary based recovery algorithms. Now, by choosing \mathbf{L} to satisfy RIP, we are in a compressive sensing framework.

2.6.2 Uncontrolled acquisition with sparse regularization

We note that with a controlled illumination, spatially varying white light is used, which ensures that the incident illumination is equally well conditioned over the entire visible spectrum, and the resulting 3-channel transport matrices are in turn the correct transport matrices for the wavelengths corresponding to each of the three channels. With natural illumination on the other hand, depending on incident illumination, which depends on the weather of the day, the illumination ensemble at different spectral wavelengths can have different condition numbers. For example, on a particularly clear day, all wavelengths around blue (and consequently, the blue channel) will have very high condition numbers.

Let us assume that the incident light is spectrally uniform, in other words, the recovered light transport matrix transports the luminance of incident illumination (grayscale) to the RGB channels of the image plane. In this case, the naive approach would be to replace \mathbf{L} by $(\mathbf{L}_r + \mathbf{L}_g + \mathbf{L}_b)/3$, and compute $\mathbf{T}_r, \mathbf{T}_g, \mathbf{T}_b$ by solving for $\mathbf{Y}_r, \mathbf{Y}_g$ and \mathbf{Y}_b . But note that

$$\mathcal{R}(\mathbf{L}_r + \mathbf{L}_g + \mathbf{L}_b) \subseteq \mathcal{R}(\mathbf{L}_r) + \mathcal{R}(\mathbf{L}_g) + \mathcal{R}(\mathbf{L}_b).$$

Let $\mathbf{Y}[n, :] \in \mathbb{R}^P$ be the vector of observed pixel intensities at pixel n in the c^{th} channel. Solving for $\mathbf{Y}[n, :] = \mathbf{t}_n[\mathbf{L}_r \ \mathbf{L}_g \ \mathbf{L}_b]$, where now $\mathbf{t}_n \in \mathbb{R}^{1 \times 3M}$, is better conditioned than solving for $\mathbf{Y}[n, :] = \mathbf{t}_n(\mathbf{L}_r + \mathbf{L}_g + \mathbf{L}_b)$. The part of the illumination matrix that contributes to pixel n is

$$\begin{aligned}
\mathbf{L}_n^{rgb} &= \begin{bmatrix} \mathbf{L}_n^r \\ \mathbf{L}_n^g \\ \mathbf{L}_n^b \end{bmatrix} \in \mathbb{R}^{3S \times P}. \quad \text{Then, for each channel 'c',} \\
\tau_n &= [\alpha_{n1} \mathbf{D} \ \alpha_{n2} \mathbf{D} \ \alpha_{n3} \mathbf{D}] \quad \tau_n \in \mathbb{R}^{1 \times 3S}, \\
\{\alpha\}_{n1, n2, n3} &= \arg \min_{\alpha_{n1}, \alpha_{n2}, \alpha_{n3}} \|\mathbf{Y}[n, :] - [\alpha_{n1} \mathbf{D} \ \alpha_{n2} \mathbf{D} \ \alpha_{n3} \mathbf{D}] \mathbf{L}_n^{rgb}\|_2, \\
&\text{subject to, } \|\alpha_{ni}\|_1 \leq \mu, \ i \in [1, 2, 3], \\
&\alpha_{ni} \geq 0.
\end{aligned} \tag{2.11}$$

Image formation is now described by

$$\mathbf{y}[n] = [\alpha_{n1} \mathbf{D} \ \alpha_{n2} \mathbf{D} \ \alpha_{n3} \mathbf{D}] \ell_n^{rgb}.$$

Light transport recovery under natural illumination is better conditioned with this formulation. While the light transport recovered by this method reveals interesting micro-structures,

visual richness is lost in the rendered images, due to the assumption that the incident illumination is spectrally uniform.

In summary, we can obtain two types of light transport under uncontrolled environment. When solving for equation (2.11), the light transport recovery is better conditioned, by trading off for color in the incident light. If we solve for equation (2.10) instead, both incident and outgoing light are colored, however each channel of light transport can have different accuracies depending on the particular illumination spectrum (e.g. different weather conditions).

2.7 Experiments

The first part of this section presents experimental validation of the acquisition setups described in this work. The next two sections present the application of these methods for the digital acquisition of stained glass windows under controlled and uncontrolled illumination scenarios, respectively.

2.7.1 Acquisition Performance

As discussed in Section 2.5, acquiring the entire light transport matrix by brute force is impractical. We therefore acquire a ground truth LTM for evaluation by fixing the illumination plane resolution (pixel size of the projector) to be the same as our regular setup while restricting the illumination plane to a 32×32 grid. For ground truth acquisition, we project illumination patterns from the Euclidean basis so that the observed images correspond directly to the respective point scattering functions and thus constitute the columns of the light transport matrix. We observed the magnitude of the discrete Fourier transform of the light transport vectors to decay and thus inferred that we sample the illumination plane at above the Nyquist rate.

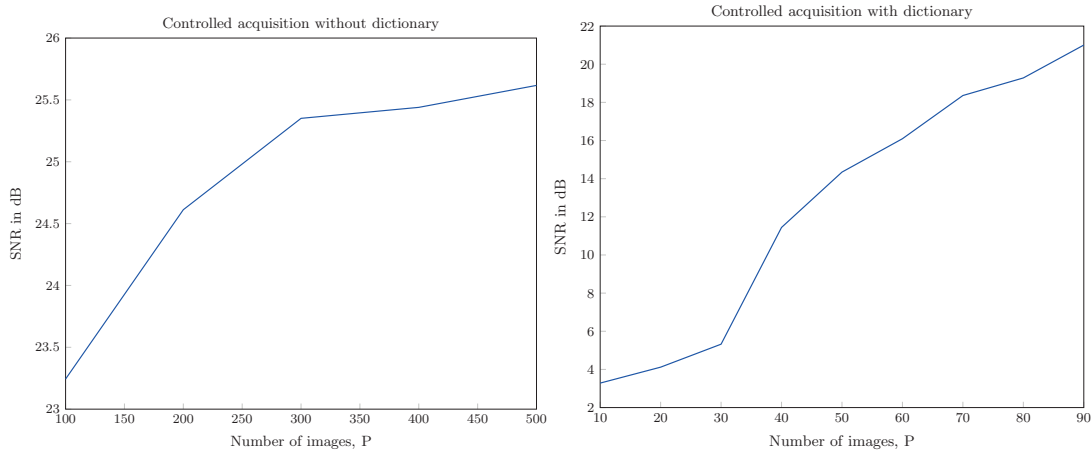


Figure 2.8: Improvement of SNR for increasing number of input images. (a) Controlled Acquisition without a dictionary. (b) Faster Acquisition with dictionary

We now simulate the proposed acquisition of LTM in a controlled environment by multiplying the ground truth with a gaussian random ensemble. We reconstruct the LTM using the controlled

acquisition method described by equation (2.8) and repeat the experiment for various values of $P > \nu$. Figure 2.8(a) shows the corresponding values of SNR defined by $20 \log_{10} \frac{\|\mathbf{T}\|_F}{\|\mathbf{T} - \hat{\mathbf{T}}\|_F}$, where \mathbf{T} is the ground truth LTM and $\hat{\mathbf{T}}$ is the reconstructed LTM. We also acquire a dictionary as described in Algorithm 2.1 and simulate a compressive acquisition using equation (2.9). The SNR for various values of $P < \nu$ is shown in Figure 2.8(b).

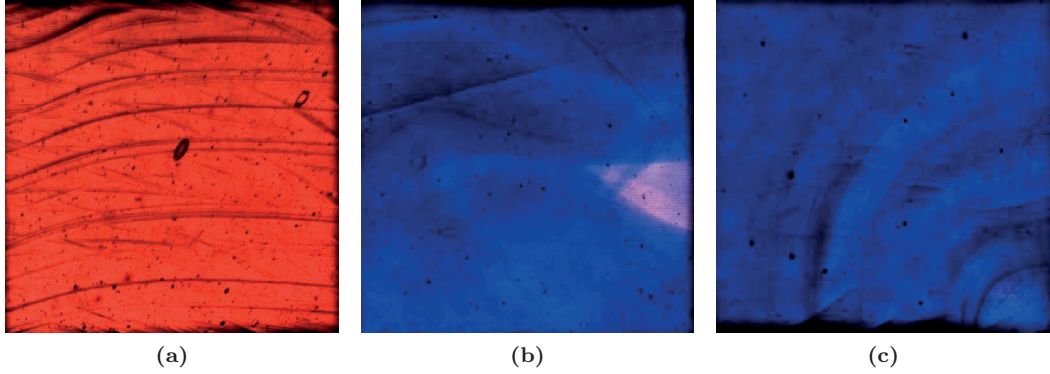


Figure 2.9: Three stained glass slabs, that whose light transport were acquired under controlled illumination, using Equation (2.8).

We acquired three full light transport matrices, each of resolution 262144×16384 ($N \times M$) of three $15\text{cm} \times 15\text{cm}$ stained glass slabs, shown in Figure 2.9 using equation (2.8) by projecting illumination vectors drawn from a Gaussian random ensemble \mathbf{L} . A dictionary $D \in \mathbb{R}^{1024 \times 81}$ shown in Figure 2.10, was then learnt on the light transport matrices of the two slabs shown in Figure 2.9(a) and (b), using Algorithm 2.1. Here, the size of an individual atom in the dictionary is 81 and is determined by the size of the largest connected component in the dataset, while using Algorithm 2.1.

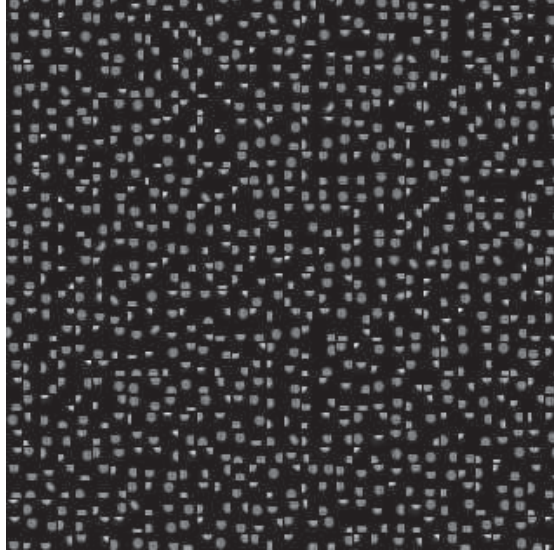


Figure 2.10: An overcomplete basis $\mathbf{D} \in \mathbb{R}^{81 \times 1024}$ with 1024 atoms was learnt from the translated versions of a total of 262144 LTFs. Each basis vector has been re-organized as a 9×9 discretized 2D light transport function.

We then reconstruct the light transport matrix $\hat{\mathbf{T}}$ of the third slab (Figure 2.9(c)) using equation (2.10) for various values of $P < \nu$. We now compare the images rendered with the LTM obtained without the dictionary using Equation (2.8) with those obtained with the dictionary using Equation (2.9). Figure 2.11 shows how the SNR improves for light transport recovery under controlled illumination when the number of illumination vectors used is increased. Here the SNR is defined as $20 \log \frac{|\mathbf{T}\ell|_2}{|\mathbf{T}\ell - \hat{\mathbf{T}}\ell|_2}$, where \mathbf{T} is the reference light transport matrix and $\hat{\mathbf{T}}$ is the recovered LTM.

Figure 2.12 shows that the location of a volume impurity is preserved even when the LTM is recovered from just 5 random measurement vectors. Thus, in the presence of a dictionary, we can choose an arbitrarily small number of measurements, by trading off for SNR. Such an approach will be extremely useful when a large number of objects made of a similar material are digitized.



Figure 2.11: Improvement of SNR for recovered LTMs of a piece of stained glass, as a function of the number of controlled (Gaussian) illumination vectors used.

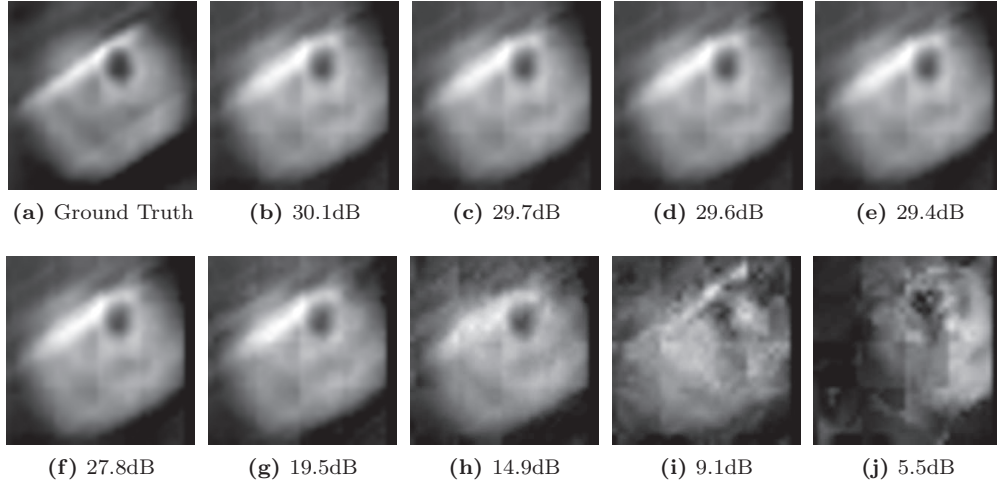


Figure 2.12: PSF under controlled illumination: We show the point scattering function (blue channel) for a point in the blue stained glass exhibiting a ‘bubble’. (a) Ground Truth. (b)-(j) The LTM recovered from 85, 75, 65, 55, 45, 35, 25, 15 and 5 controlled (Gaussian) illumination vectors respectively. The SNR for each recovered PSF is shown in its caption.

For LTM recovery under uncontrolled illumination, since it is impossible to obtain a reference light transport matrix, we use a simulated acquisition for comparison with the reference, \mathbf{T} acquired for the previous experiment. We first obtain an illumination ensemble \mathbf{L} of 105 images of the sky over a period of 36 hours. We synthetically generated the observed image matrix \mathbf{Y} ,

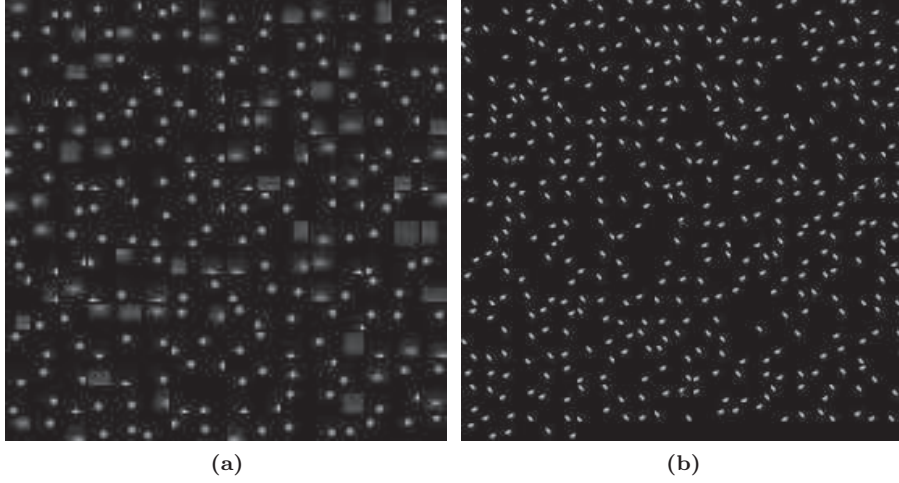


Figure 2.13: Material dictionaries learnt from light transport matrices of different types of glass slabs where individual basis vectors have been re-organized as 2D light transport functions. (a) $\mathbf{D} \in \mathbb{R}^{81 \times 256}$. (b) $\mathbf{D} \in \mathbb{R}^{81 \times 512}$.

by multiplying the illumination ensemble with the reference light transport. We then recover the LTM from this dataset to show that volume impurities are preserved even under natural illumination. Figure 2.14 shows some point scattering functions extracted from LTMs recovered from controlled and uncontrolled illuminations. Note that the volume impurities are preserved even when recovered from uncontrolled illuminations.

In Figure 2.15 we show the volume scattering by a piece of stained glass illuminated by a light source that projects a line onto it. The response for volume scattering recovered under controlled lighting is compared with that obtained from uncontrolled lighting.

We now compare the relighting results of a stained glass artifact, whose light transport was recovered under controlled and uncontrolled illuminations. In both cases, a total of 105 images were used : images drawn from a gaussian random ensemble for controlled illumination and that of the city sky for uncontrolled illumination. Figure 2.16 shows relit images of the stained glass artifact generated with LTM recovered from controlled illumination and natural illumination. It can be seen that even when the LTM is acquired with uncontrolled illumination, the relit images have a reasonably good SNR.

2.7.2 Controlled Acquisition in Romont

We used our acquisition method described by equation (2.8) to obtain the light transport matrices of two exhibits at the Vitromusée, Romont, from a total of 1024 images with each light transport matrix having a resolution of $\simeq 10^6 \times 2.6 * 10^5$ ($N \times M$). One was that of a medieval glass window that was once part of the Cathedral of Notre Dame in Lausanne. Shown in Figure 2.18 are some relighting results obtained with the acquired LTM. The second window belongs to the Renaissance era and is shown in Figure 2.19.

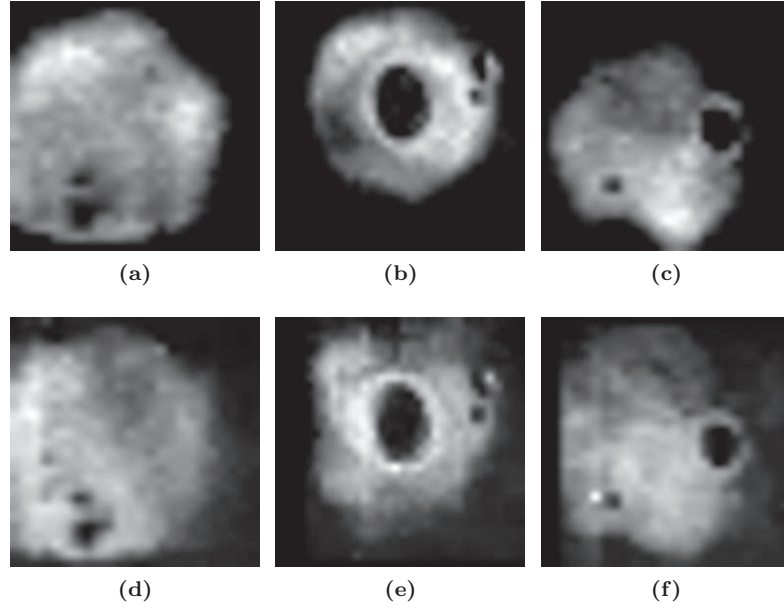


Figure 2.14: Some point scattering functions (blue channel) of a piece of blue stained glass. Subfigures (a),(b) and (c) are extracted from an LTM obtained under controlled illuminations. Subfigures (d),(e) and (f) are the same point scattering functions, when the LTM was recovered under uncontrolled illumination.

2.7.3 In-Situ Acquisition of Rose Window, Lausanne

In this section, we describe an experiment carried out at a 12th century Cathedral of Notre Dame in Lausanne to digitize its rose window. The rose window, assembled between 1231 and 1235 A.D by Pierre d’Arras, is amongst the most important roses in Europe. The window is circular with a diameter of 8 meters. Figure 2.20 shows our acquisition setup at the cathedral. A NIKON D800 was placed inside the cathedral about 42 metres from the rose, focused such that the entire rose is in its field of view. It was programmed to capture the window at 5 exposure steps. Three time-lapse cameras were placed outside the cathedral, just above the Rose. These were programmed to capture the incident light in a synchronized manner with the D800. Both camera setups captured an image once every 12 minutes over a 36 hour time period, thus giving a total of 180 HDR images of the window and the corresponding incident illuminations. In this dataset, 42 image pairs were captured either at night or when the sun was directly behind the window and had to be discarded, thus 138 image pairs were available for analysis.

A dictionary was learnt from the light transport matrices of three separate glass slabs shown in Figure 2.9. We assumed that the Rose window is made of a similar type of glass. We recover the LTM at full sensor resolution, *i.e.*, 7360×4912 by 1920×986 ($\sim 36.1 * 10^6 \times 1.89 * 10^6 : N \times M$). Solving for such a large problem size was made tractable by our modified problem formulation that exploits the banded structure of the LTM. We recover and store the LTM as 16 blocks in order to overcome limitations on the memory.

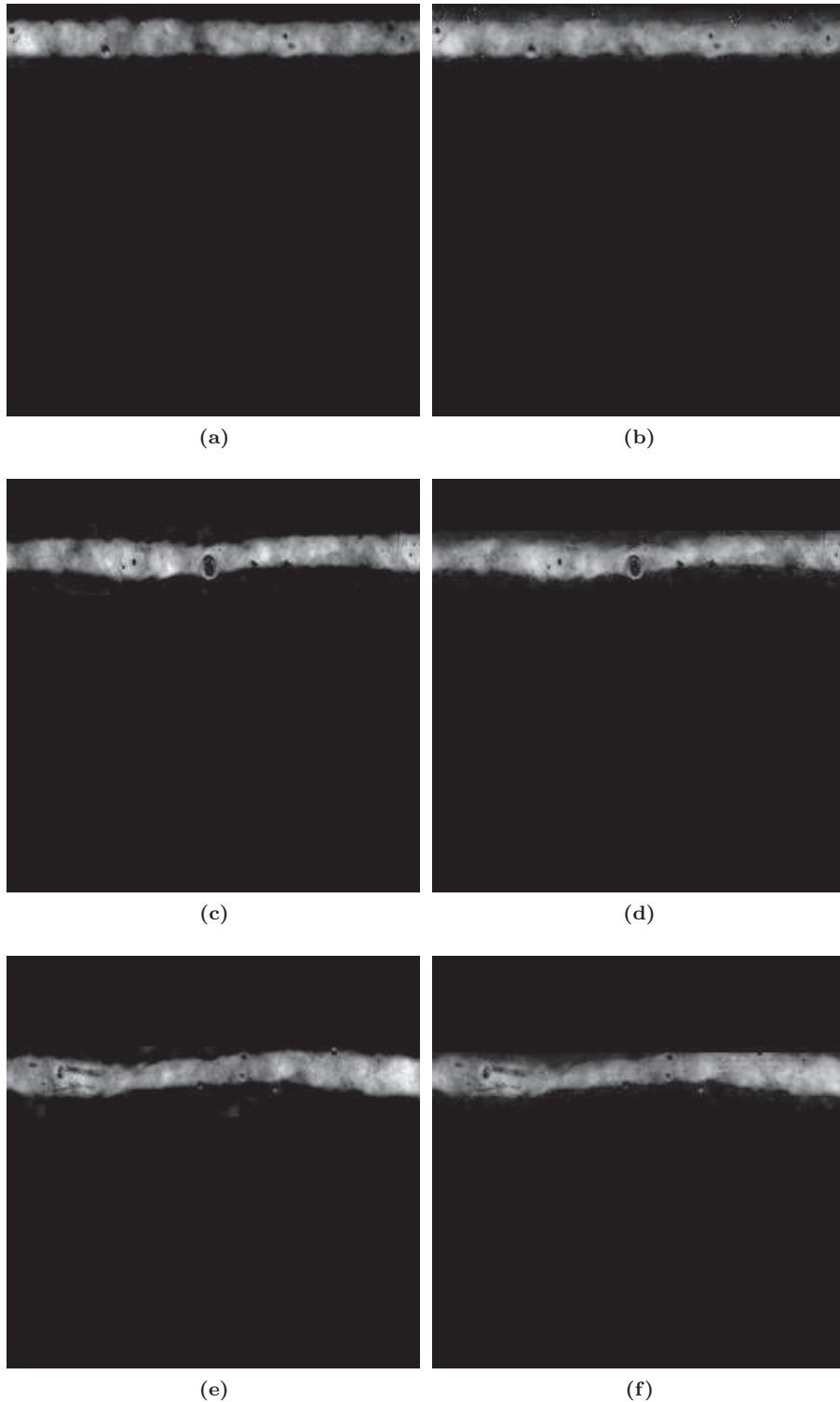


Figure 2.15: Rendered images of a piece of stained glass observed under a light source projecting a line on the glass to demonstrate volume scattering. Subfigures (a),(c) and (e) were obtained by solving the linear system (2.10) under controlled illumination patterns (Compressive sensing). Subfigures (b),(d) and (f) were obtained by solving the same linear system with just 105 natural illumination patterns (Sparse regularization). Notice that the volume impurities are preserved even under uncontrolled illumination. Corresponding values of PSNR are (b) 28.95dB, (d) 32.28dB and (f) 32.95dB.

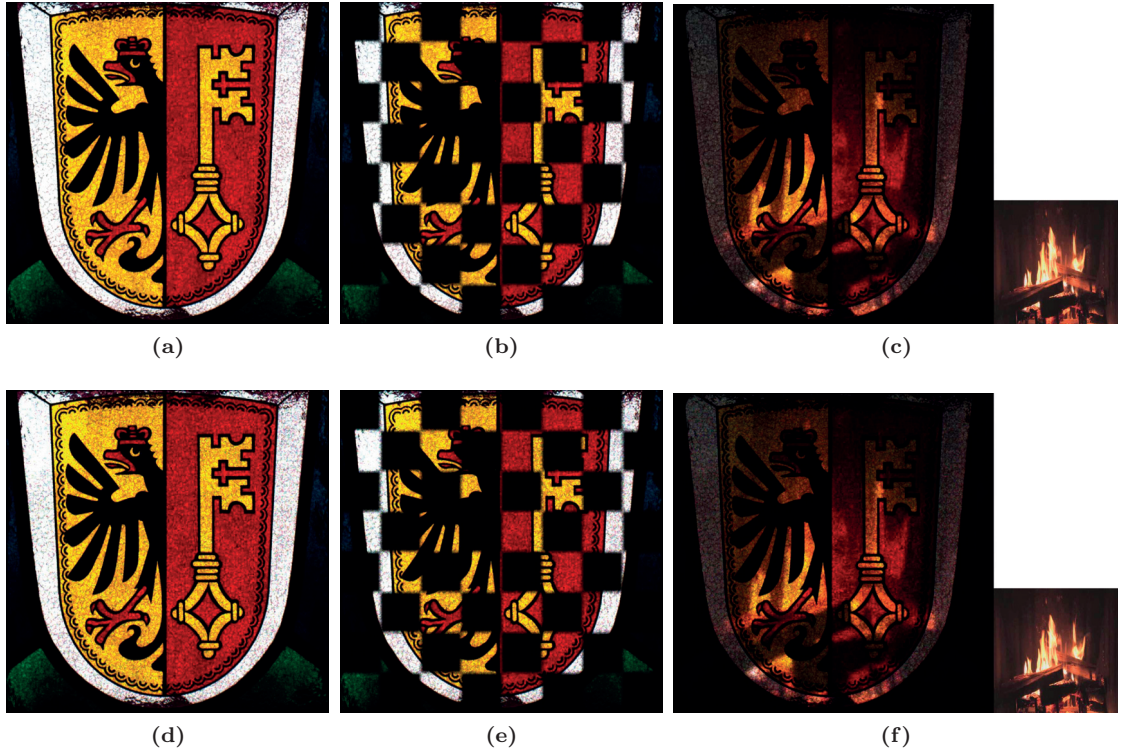


Figure 2.16: Relighting results obtained with the Geneva Flag experiment. The top row contains images rendered from an LTM obtained under controlled illuminations. The bottom row contains images rendered from an LTM obtained under uncontrolled illuminations using Method-1. Subfigures (a) and (d) contain floodlit scenes; $\text{PSNR} = 21.97dB$. (b) and (e) contain images generated when a checkerboard pattern is projected upon the glass artifact; $\text{PSNR} = 24.26dB$. Subfigures (c) and (f) are generated by the illumination pattern shown in bottom-right; $\text{PSNR} = 26.32dB$.

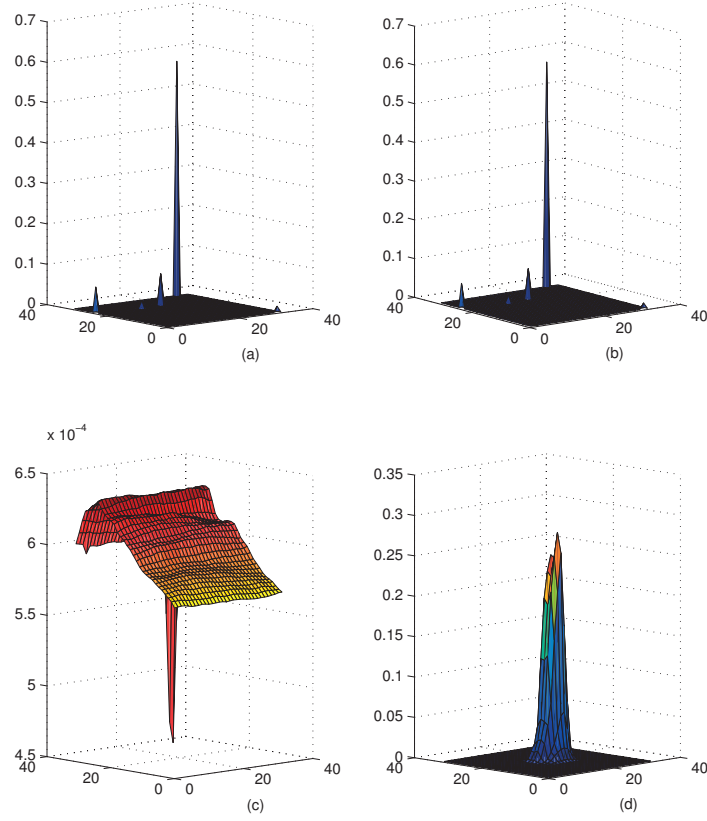


Figure 2.17: The same light transport vector, \mathbf{t}_n recovered from data obtained in an outdoor illumination, using (a) ℓ_2 minimization with non-negativity (b) $\ell_1 - \ell_2$ minimization, with non-negativity (c) block sparse solution under non-negativity and (d) our solution obtained under equation (2.10). Note that in the first three cases, the recovery fails completely as we end up with a light transport vector with non-compact support in each case. None of the heterogeneities were preserved in any of these three cases.



(a)



(b)

Figure 2.18: Relighting results obtained with a medieval window that was once located at the Cathedral of Lausanne, currently exhibited at Romont Vitrocentre. Incident illumination pattern is shown in inset.



Figure 2.19: Relighting results obtained with a Renaissance era exhibit at Romont Vitrocentre. Incident illumination pattern is shown in inset.

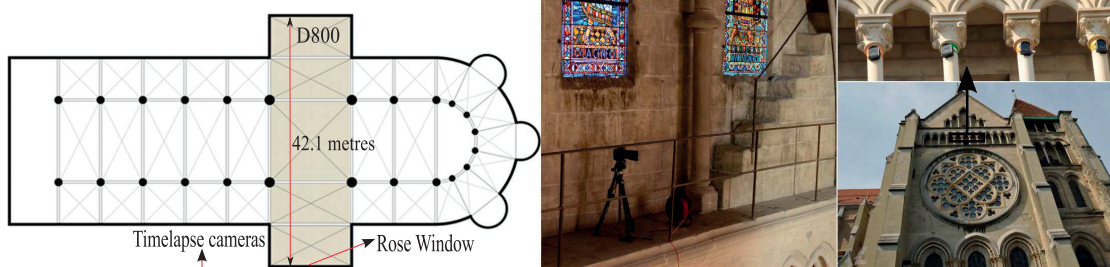


Figure 2.20: The Rose Window acquisition setup. On the left is a schematic of the cathedral. On right, the Nikon D800 focused at the rose window and the time lapse cameras on the exterior are shown.

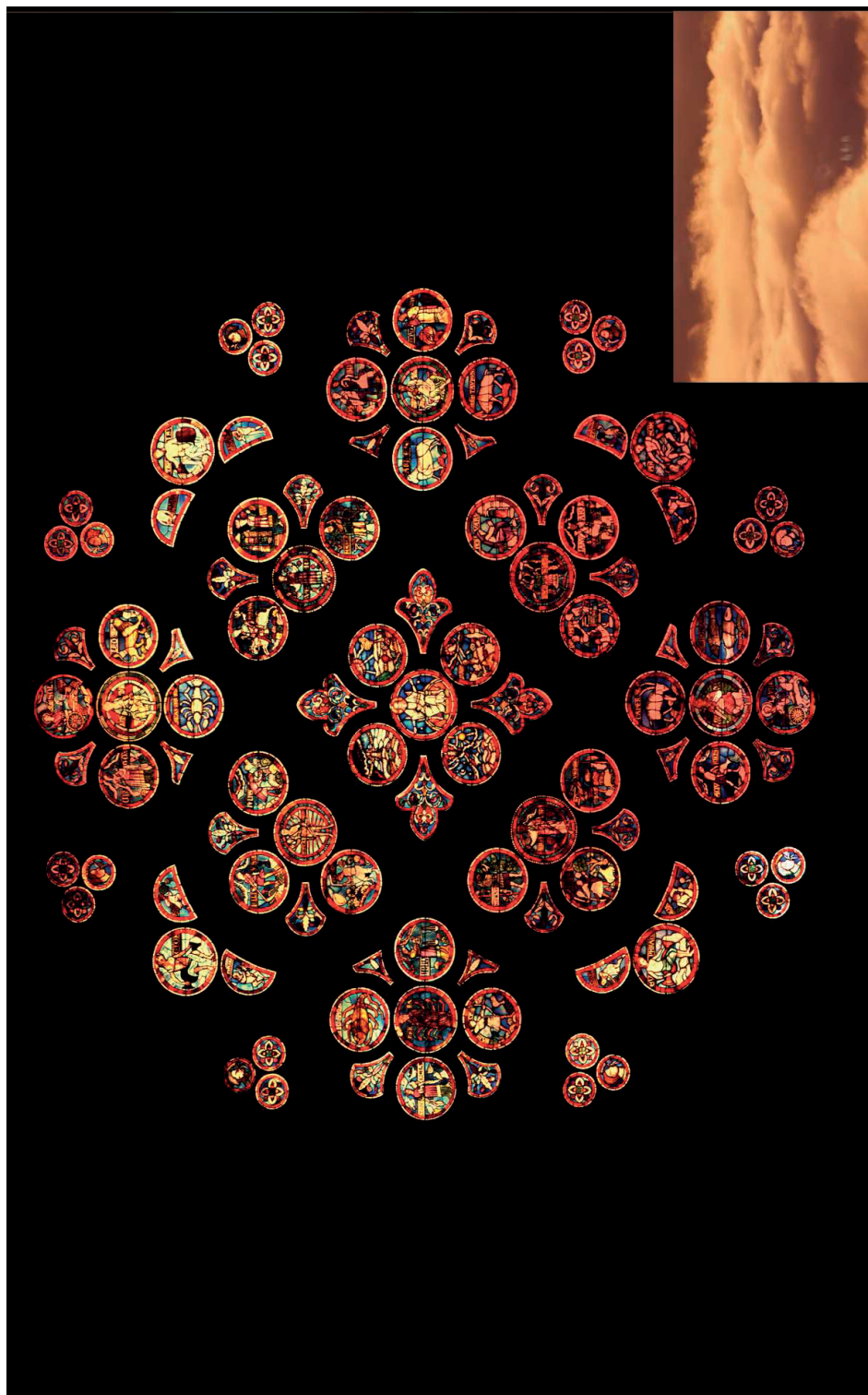


Figure 2.21: Relighting results obtained with the Rose window experiment, obtained by using LTM recovered with equations (2.11). The illumination pattern used for relighting is shown in inset.

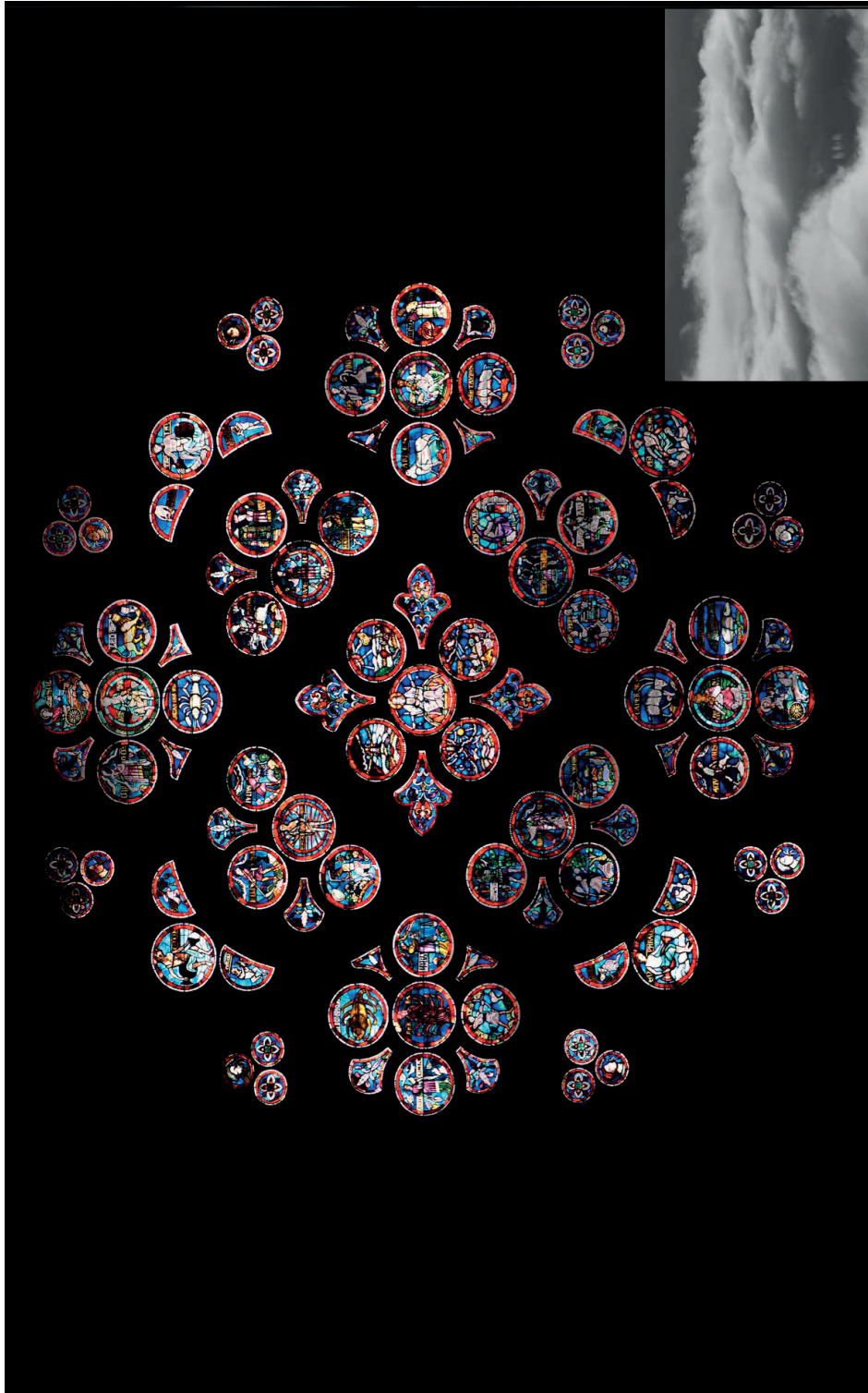


Figure 2.22: Relighting results obtained with the Rose window experiment obtained by using LTM recovered with equations (2.10). The illumination pattern used for relighting is shown in inset.

We recovered the full LTM by solving for equations (2.10). When the experiment was in progress, the sky was clear, and so the blue channel illumination matrix was very poorly conditioned. As seen in Figure 2.23, the rendered images have a washed out appearance at some subfigures of the stained glass. When the LTM was recovered on the same dataset by solving for the system of equations (2.11), we observe a remarkable improvement in the preserved details, as shown in Figure 2.23. Lost textures and details that were not visible in the previous method are restored in this alternative formulation, at the expense of having a monochromatic incoming light.

2.8 Summary

We have presented and validated by experiments a practical workflow for digitizing stained glass, that can be readily utilized in large scale acquisitions. To summarize, when controlled acquisition is possible, we present a method that exploits the banded nature and non-negativity of light transport for fast acquisition. We present a method to learn priors on light transport, which can be used to further reduce acquisition speed under controlled illumination, while serving as a regularizer for uncontrolled acquisitions. Based on methods presented here, we will be able to create a virtual tour of architectural monuments like cathedrals and museums. Visitors of such virtual museums will have the ability to choose the time of the day, the season and the meteorological conditions under which they wish to view the digitized artwork. Given appropriate material dictionaries, the rendered images will be a faithful representation of the visual experience that an actual visitor would encounter.



Figure 2.23: Results from Figures 2.21 and 2.22, shown in full resolution to compare both methods described in this chapter for LTM recovery under uncontrolled illuminations. The left column includes relighting results obtained by using the LTM recovered by solving for equations (2.10). The right column includes relighting results obtained by using the LTM recovered by solving for equations (2.11). Compare the flushed texture of the images on the left with the sharp texture on the right.

Chapter 3

Synthetic Rendering of Stained Glass

I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail.

Abraham Maslow

3.1 Introduction

In Chapter 2, we described a method for inverse relighting of stained glass windows, where the goal was to acquire the light transport characteristics of real stained glass windows from a few photographic observations. While such a framework enables digital archiving and virtual rendering of age-old stained glass windows, it is also extremely interesting to synthesize virtual stained glass windows, particularly given that priors on the light transport are known in the form of dictionaries. Such a scheme, that can synthesize or simulate the light transport properties of virtual stained glass can prove useful as a visualization tool for artists before they assemble the actual stained glass window. Alternatively, it can also be used as a means of synthesizing relightable photographs of an artifact, when only a single diffusely lit photograph is available. Means of synthesizing light transport properties when priors can be learnt, can also serve as an indispensable tool for digitally restoring damaged historical stained glass artifacts, thus enabling us to visualize their intended appearance. Thus, the focus of this chapter is forward rendering of stained glass windows, by ‘designing’ a meaningful light transport matrix, given appropriate dictionaries.

While other approaches that infer scattering parameters of translucent materials from images exist, these are unable to model heterogeneity caused by volume impurities as is present in

stained glass. These methods use the diffusion approximation or the volume rendering equation to render the volume scattering of light in translucent media. Unlike inverse problems, forward problems are typically well defined since they are described by a known physical phenomenon and can be evaluated, given the required parameters. However, since our goal is to utilize the discretized image based priors to simulate light transport, we propose an approach where we directly synthesize the discrete light transport matrix. We first obtain constraints on light transport at each pixel in the image plane based on physical properties of stained glass. We then pose light transport synthesis as a constrained matrix completion problem.

3.2 Related Work

Traditionally, transport of light through volumes can be accurately rendered using the volume rendering equation [9]. To overcome the computational complexity of such accurate simulations, Jensen et al [37] proposed the dipole diffusion approximation to model light transport in homogeneous semi-infinite planes, where light transport was modeled by the diffusion equation. This model was adapted to render subsurface scattering in human faces, by measuring scattering profiles to generate a homogeneous subsurface scattering model, which is then modulated by a spatially varying albedo thus simulating heterogeneity under uniform illuminations. However, this approach is unsuitable for local illumination or for spatially varying light sources. A more accurate multipole diffusion approximation was introduced [17] to model light transport in thin slabs. This enabled the modeling of heterogeneity in skin arising from homogeneous translucent layers [18]. However, spatial variation in scattering parameters was still modeled by modulating with an albedo map. More recently, an approach that acquires the scattering properties of materials by inverting the diffusion equation was presented [74]; these scattering parameters are then used to render the material using the diffusion equation. A layered model for skin has also been proposed [20], where the scattering parameters vary smoothly spatially but are locally homogeneous. While such a model is well suited for skin, stained glass contains volume impurities, air bubbles and corrosion which cannot be modeled by smoothly varying, locally homogeneous scattering models.

In DISCO [29], an approach was proposed to acquire the heterogeneous BSSRDF model from real-world objects, which can be used in inverse rendering. A compact representation for acquired subsurface scattering components has also been proposed [54]. The acquired impulse responses are first approximated by an average homogeneous response, which is then used to obtain a compact factored representation for heterogeneous light transport. Unlike these approaches, our approach can represent the scattering component at each surface point by a linear combination of several individual light transport functions, since it is based on a material dictionary.

In the following section, we first present our approach to design the light transport matrix of glass slabs. We then show its applications in synthesis and restoration.

3.3 Light Transport Design

As seen in Chapter 2, in the absence of significant surface structures, the volume transport of light in stained glass artifacts can be expressed completely using a dictionary of light transport functions, that were learnt on the material in question. Therefore, we started investigating the possibility of rendering synthetic heterogeneous stained glass slabs when priors on the light

transport properties in the form of the material specific dictionary are known, in addition to its approximately planar geometry.

The goal of the inverse rendering problem presented in Chapter 2 was to obtain a light transport matrix \mathbf{T} which generated the observations \mathbf{Y} , when excited by the illumination ensemble \mathbf{L} . This required that for a consistent LTM recovery, \mathbf{T} satisfies the image formation equation,

$$\mathbf{Y} = \mathbf{T}\mathbf{L}.$$

Since in typical acquisition scenarios the above system of equations are under-determined, we introduced additional constraints requiring that $\boldsymbol{\tau}$, the compact non-zero support of each LTF be sparse in the learnt dictionary \mathbf{D} . Thus, we were able to obtain a unique reconstruction of the LTM \mathbf{T} by solving for a constrained optimization problem.

However, synthesizing an LTM given just the dictionary \mathbf{D} , only requires that $\boldsymbol{\tau}$, the compact non-zero component of each LTF has a sparse representation in \mathbf{D} . There exist infinitely many matrices which satisfy the above condition, and as such, the problem is ill-defined.

Consider the following analogy. A monolingual person is given a bilingual dictionary containing words of an alien language. Assuming that the two languages share the alphabet, the person might be able to infer the meaning of already existing literature, using the dictionary. This is similar to the inverse problem. However, in order to create new ‘meaningful sentences’, the person will have to learn the ‘grammar’ of the new language, in addition to their vocabulary.

Similarly, since any light transport function \mathbf{t} , of the required material can be represented using the material specific dictionary \mathbf{D} , in some sense, we have the ‘building blocks’ of the LTM that is to be synthesized. Designing a ‘physically correct’ LTM from the dictionary thus requires further constraints based on the physics of light transport. Typically in forward rendering, the dipole [37] or multipole [17] diffusion approximations define the light transport of homogeneous media, as described by the measured scattering coefficients. Since our priors are in the form of an image-based dictionary, we look for an alternate formulation involving constraints on the structure of LTM inspired by physical characteristics of the object.

We formalize this problem as follows;

Problem Description: Design a *physically meaningful* light transport matrix \mathbf{T} of a heterogeneous stained glass, given:

- its approximately planar geometry and known thickness Δ and
- a material dictionary \mathbf{D} that serves as a sparsity inducing basis for the light transport of the material of interest, also learnt on a sample of thickness Δ .

Note that, if we can design physically correct heterogeneous LTMs for planar glass of thickness Δ , we can generate LTMs for heterogeneous glass of other thickness upto a depth resolution of Δ by cascading several planar LTMs and ignoring boundary effects. This is analogous to approximating a thick heterogeneous slab by several thinner heterogeneous slabs stacked upon each other.

Image Formation Model: The LTM defined in Chapter 2 characterizes both the spectral filtering of incident light caused by infused coloring pigments and the spatial filtering of incident light caused by the volume transport of light in the stained glass slab. In real stained glass, the infusion of color pigments into the glass is also a spatially varying function. In this chapter however, we will assume that stained glass slabs are created by painting colors on colorless, planar,

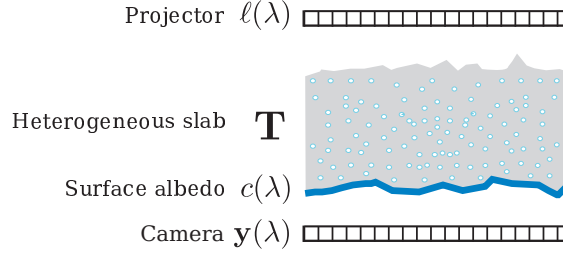


Figure 3.1: Our image formation model with the synthetic stained glass. Here, stained glass is modeled by a heterogeneous glass slab with spatially varying light transport defined by \mathbf{T} whose color is determined by the surface albedo $\mathbf{c}(\lambda)$ given by the texture painted on it.

heterogeneous, translucent glass slabs. Thus in this model the color of stained glass arises from the color painted on the surface of the glass, and is described by a surface albedo $\mathbf{c}(\lambda) \in \mathbb{R}^N$, for each wavelength λ . The LTM of the colorless heterogeneous translucent slab acts as a spatially varying filter that is caused by the volume transport of light. The final appearance of the glass in our model is thus determined by both by the albedo and the colorless glass slab. Thus, image formation with this approximate model is given by

$$\mathbf{y}(\lambda) = \mathbf{c}(\lambda) \odot \mathbf{T}\ell(\lambda), \quad (3.1)$$

where \odot represents element-wise product, λ is the wavelength, and \mathbf{T} is the LTM of a colorless, heterogeneous glass slab. Thus, with this image formation model, we have de-coupled the texture painted on the surface of the object defined by the albedo $\mathbf{c}(\lambda)$, which acts as a filter on the various wavelengths (thus determining the color of the object), from the underlying material slab, which acts as a spatial filter on the incoming light. Therefore in the rest of the section the LTM \mathbf{T} denotes the digitization of the material slab independent of the painted texture. Our goal is thus to synthesize LTMs that approximate colorless, planar heterogeneous glass slabs from the given material specific dictionaries.

3.3.1 Physical properties of a stained glass slab

In this subsection, we will study some physical properties of planar light transport and derive their implications on the light transport matrix. In the next subsection, we will describe a matrix completion approach that is constrained by these physical properties, to design LTMs of planar glass slabs.

Implications of Helmholtz Reciprocity

Helmholtz reciprocity states that a ray of light and its reverse ray encounter the same optical events. Thus, by interchanging the camera and the light sources in a scene, the path taken by light and the amount of energy transferred from one to another remain the same, with only the light direction being reversed. Consider a scene whose light transport from the illumination plane of a projector (ℓ) to the image plane of a camera (\mathbf{y}) is defined by the light transport matrix \mathbf{T} , with image formation described by $\mathbf{y} = \mathbf{T}\ell$. Let this configuration be called the primal configuration. Helmholtz reciprocity implies that when the camera and the projector

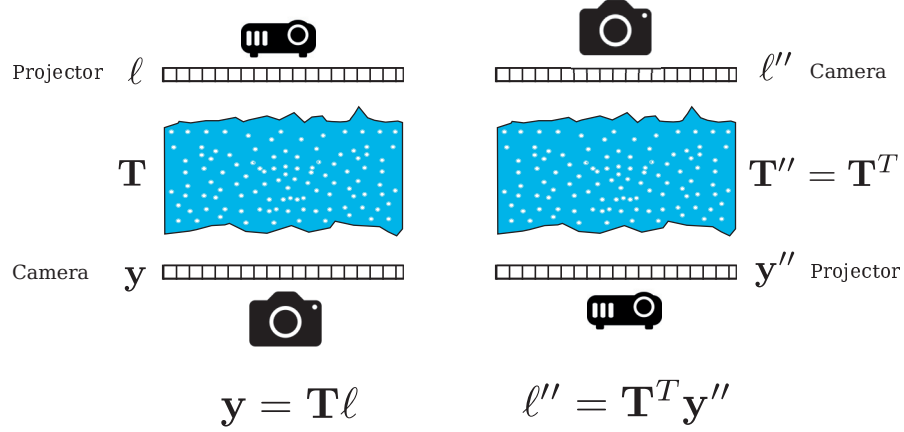


Figure 3.2: Illustration of Helmholtz Reciprocity. If the LTM \mathbf{T} describes light transport in the primal configuration, then its transpose \mathbf{T}^T defines light transport in the dual configuration.

are interchanged, image formation in the new setup (the dual configuration) is described by $\ell'' = \mathbf{T}^T \mathbf{y}''$ [62], as described in Figure 3.2. Here ℓ'' is the image observed by the camera in the new position while \mathbf{y}'' represents the light source configuration described by the projector. Thus if the LTM \mathbf{T} describes light transport in the primal configuration, the LTM of the dual configuration is given by its matrix transpose, \mathbf{T}^T [62]. Thus when the scene is composed of a glass slab, since the material is the same in both the primal and the dual configurations, all physical properties of the LTM \mathbf{T} must also be obeyed by its transpose, \mathbf{T}^T .

Bandedness

As stated in Chapter 2, we know that the light transport functions (LTFs) of almost planar objects have a compact, predictable support on the illumination plane. When these LTFs are vectorized as \mathbf{t} and expressed as columns of the LTM \mathbf{T} , we obtain a banded matrix. As in Chapter 2, we use the indicator function $\mathbf{\Gamma}_i$ to denote this finite support of the LTF \mathbf{t}_i at pixel i . We use the subscript n to denote pixels in the primal configuration. The non-zero component of each LTF \mathbf{t}_n , in the primal configuration can then be defined by,

$$\tau_n = \mathbf{t}_n[\mathbf{\Gamma}_n] = \mathbf{T}[n, \mathbf{\Gamma}_n] \quad \forall n \in [1 \cdots N - 1].$$

Similarly, let us denote the pixels of the dual configuration with the index m . Since \mathbf{T}^T is the LTM of this configuration, each LTF \mathbf{t}_m can then be defined as,

$$\tau_m = \mathbf{t}_m[\mathbf{\Gamma}_m] = \mathbf{T}[\mathbf{\Gamma}_m, m] \quad \forall m \in [1 \cdots M - 1].$$

Sparsity in \mathbf{D}

By construction, the dictionary \mathbf{D} is such that any LTF of the given material has a sparse representation in \mathbf{D} . Thus any τ should have a representation $\tau = \alpha \mathbf{D}$ such that α is sparse. This property, when combined with reciprocity and bandedness yields,

$$\tau_n = \mathbf{T}[n, \mathbf{\Gamma}_n] = \alpha_n \mathbf{D} \quad \text{and} \quad \tau_m = \mathbf{T}[\mathbf{\Gamma}_m, m] = \alpha_m \mathbf{D},$$

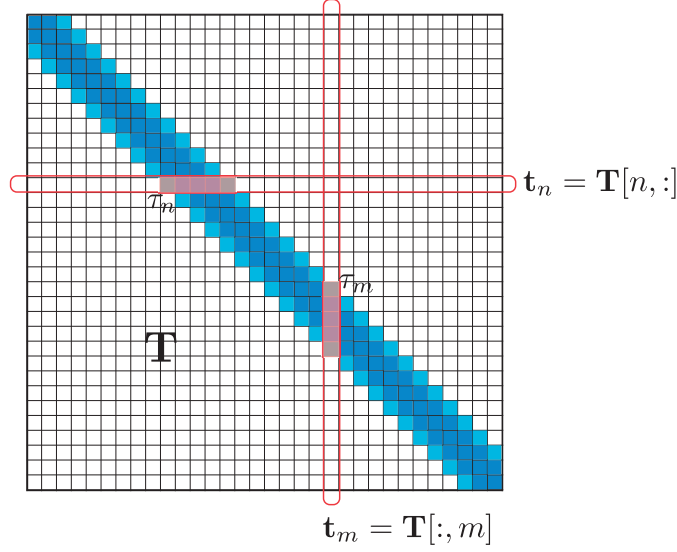


Figure 3.3: Illustration of the banded structure of a 2D LTM.

where α_n and α_m are both sparse.

Non-negativity

Since any light transport function is non-negative, we have

$$\mathbf{t} \geq 0.$$

Constraints on the Material Texture

Let us define the appearance of the colorless heterogeneous glass slab under uniform illumination (floodlit) to be the ‘material texture’, $\sigma \in \mathbb{R}^N$, such that,

$$\sigma = \mathbf{T}\mathbf{1},$$

where $\mathbf{1}$ denotes the uniform illumination. During the design of a synthetic stained glass, while we are always given the painted albedo $\mathbf{c}(\lambda)$ as an input, we might or might not have access to the desired material texture σ . We will consider both cases here.

Let $\sigma_n[n]$ and $\sigma_m[m]$ denote the material texture values of the primal and dual configurations at the n^{th} and m^{th} pixel indices respectively.

When the material texture σ_n for the primal configuration is known, we can impose the following constraint on each LTF \mathbf{t}_n ,

$$\langle \mathbf{t}_n, \mathbf{1} \rangle = \sigma_n[n] \quad \text{where} \quad \mathbf{t}_n = \mathbf{T}[n, :].$$

Since $\langle \mathbf{t}_n, \mathbf{1} \rangle = \|\mathbf{t}_n\|_1 = \|\alpha_n \mathbf{D}\|_1$, the above equation can be recast as,

$$\|\alpha_n \mathbf{D}\|_1 = \sigma_n[n].$$

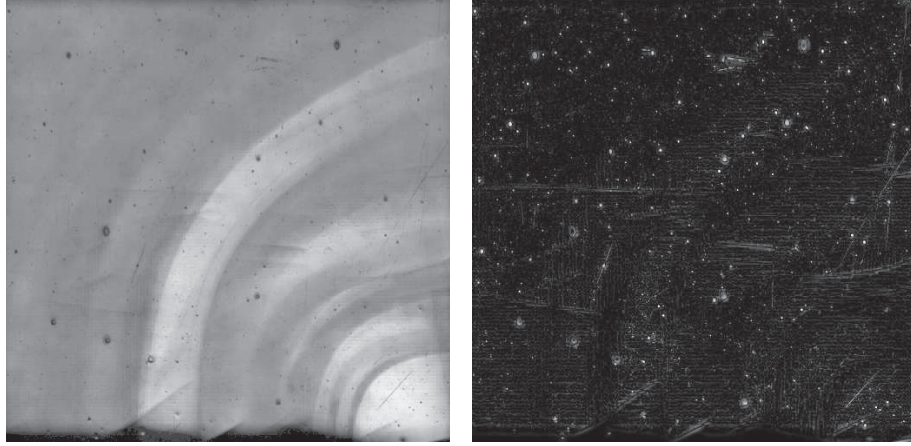


Figure 3.4: Constraints on the material texture. (a) Material texture σ of a real stained glass slab. Subfigure (b) shows the parameter ξ_1 for the same glass slab. Note that the parameter has a low, almost constant value, except in regions with discontinuities in the surface texture.

Similarly, if the material texture σ_m for the dual configuration is known, we can use reciprocity to impose

$$\|\alpha_m \mathbf{D}\|_1 = \sigma_m[m].$$

However, in the absence of such a strong prior on the material texture in one or both the configurations, we can impose other weaker but well known priors. Here, we use the constraint that the material texture must be slowly varying locally. This can be imposed by bounding the difference between the average value of material texture in the neighborhood of pixel n and the value of material texture at pixel n . Thus, for the primal configuration,

$$|\mathcal{M}_w\{\langle \mathbf{t}_q, \mathbf{1} \rangle\} - \langle \mathbf{t}_n, \mathbf{1} \rangle| \leq \xi_1,$$

where $\mathcal{M}_w\{\}$ computes the mean when the indices $q \in W_n$ and W_n is a window defined around pixel n . Let $\mu_n = \mathcal{M}_w\{\langle \mathbf{t}_q, \mathbf{1} \rangle\}$. Then this constraint can be cast as

$$\mu_n - \xi_1 \leq \|\alpha_n \mathbf{D}\|_1 \leq \mu_n + \xi_1.$$

A similar constraint can also be imposed on the dual configuration.

$$\mu_m - \xi_1 \leq \|\alpha_m \mathbf{D}\|_1 \leq \mu_m + \xi_1,$$

where $\mu_m = \mathcal{M}_w\{\langle \mathbf{t}_q, \mathbf{1} \rangle\}$ and $\mathcal{M}_w\{\}$ now computes the mean inside the window W_m defined around pixel m .

In Figure 3.4, we show the material texture σ of a real stained glass window and the parameter ξ_1 derived from it. Note that the parameter is a small constant except in regions with discontinuities in the material texture.

3.3.2 Synthesis of a Light Transport Matrix

In the previous subsection we have introduced some physical constraints on \mathbf{T} and \mathbf{T}^T . In this subsection, we will describe how we model LTM synthesis as a matrix completion algorithm.

Algorithm Overview: Our approach to LTM synthesis is composed of two stages. First, we partially initialize the LTM with LTFs at a random, uniformly distributed subset of pixels in the primal configuration (rows of the LTM). We then use an iterative matrix completion approach, where we update the LTFs of pixels in the neighborhood of recently updated entries, such that constraints presented in the previous subsection are satisfied.

We start by initializing the LTM \mathbf{T} with a small random, uniformly distributed subset of rows \mathbf{t}_n . The non-zero component $\boldsymbol{\tau}_n$ at each row n in the chosen subset is defined as a linear combination of elements from the dictionary \mathbf{D} such that $\tau_n = \mathbf{D}\alpha_n$ where α_n is sparse. If the material texture σ_n is given, we impose the additional constraint on α_n that $\|\alpha_n \mathbf{D}\|_1 = \sigma_n[n]$.

Given this initial partially filled \mathbf{T} we pose LTM design as a matrix completion problem, subject to the constraints stated in the previous section, by alternating between solving for the columns of \mathbf{T} during **loop #1** and rows during **loop #2**. The proposed iterative algorithm for this matrix completion is outlined in Algorithm 1. In the following part, we describe one iteration of the algorithm in detail.

loop #1 At each recently updated row n , we traverse through the columns corresponding to the indicator function Γ_n . At each column $m \in \Gamma_n$ that has not already been filled, we solve Equation (3.3) or (3.2). Let $\tilde{\boldsymbol{\tau}}_m$ denote the entries of $\boldsymbol{\tau}_m$ that have already been computed; *i.e.*, if m_{idx} denotes the entries of $\boldsymbol{\tau}_m$ that have already been set, then $\tilde{\boldsymbol{\tau}}_m = \boldsymbol{\tau}_m[m_{idx}]$ and $\tilde{\mathbf{D}} = \mathbf{D}[:, m_{idx}]$. We can then update the column m as follows.

When the material texture for the dual configuration σ_m is given,

$$\hat{\boldsymbol{\alpha}}_m = \arg \min_{\boldsymbol{\alpha}_m} \|\boldsymbol{\alpha}_m\|_1 \quad (3.2)$$

subject to:

$$\begin{aligned} \|\tilde{\boldsymbol{\tau}}_m - \boldsymbol{\alpha}_m \tilde{\mathbf{D}}\|_2 &\leq \xi \\ \|\boldsymbol{\alpha}_m \mathbf{D}\|_1 &= \sigma_m[m] \\ \boldsymbol{\alpha}_m &\geq \mathbf{0}. \end{aligned}$$

If the material texture for the dual configuration is not known, we instead solve for,

$$\hat{\boldsymbol{\alpha}}_m = \arg \min_{\boldsymbol{\alpha}_m} \|\boldsymbol{\alpha}_m\|_1 \quad (3.3)$$

subject to:

$$\begin{aligned} \|\tilde{\boldsymbol{\tau}}_m - \boldsymbol{\alpha}_m \tilde{\mathbf{D}}\|_2 &\leq \xi_0 \\ \|\boldsymbol{\alpha}_m \mathbf{D}\|_1 &\leq \mu_m + \xi_1 \\ \boldsymbol{\alpha}_m &\geq \mathbf{0}. \end{aligned}$$

Once we solve for Equation (3.3) or (3.2), depending on the availability of σ_m , we can update

$$\boldsymbol{\tau}_m = \hat{\boldsymbol{\alpha}}_m \mathbf{D},$$

$$\mathbf{T}[\Gamma_m, m] = \boldsymbol{\tau}_m.$$

loop #2 During the **loop #2** of each iteration we traverse through columns that were newly filled in the previous loop. At each recently updated column m , we traverse through the rows corresponding to the indicator function Γ_m . At each row $n \in \Gamma_m$ that has not already been

filled, we solve Equation (3.5) or (3.4). Let $\tilde{\tau}_n$ denote the entries of τ_n that have already been computed; *i.e.*, if n_{idx} denotes the entries of τ_n that have already been set, then $\tilde{\tau}_n = \tau_n[n_{idx}]$ and $\tilde{\mathbf{D}} = \mathbf{D}[:, n_{idx}]$. We can then update the row n as follows.

When the material texture for the primal configuration σ_n is given,

$$\hat{\alpha}_n = \arg \min_{\alpha_n} \|\alpha_n\|_1 \quad (3.4)$$

subject to:

$$\begin{aligned} \|\tilde{\tau}_n - \alpha_n \tilde{\mathbf{D}}\|_2 &\leq \xi \\ \|\alpha_n \mathbf{D}\|_1 &= \sigma_n[n] \\ \alpha_n &\geq 0. \end{aligned}$$

If the material texture for the primal configuration is not known, we instead solve for,

$$\hat{\alpha}_n = \arg \min_{\alpha_n} \|\alpha_n\|_1 \quad (3.5)$$

subject to:

$$\begin{aligned} \|\tilde{\tau}_n - \alpha_n \tilde{\mathbf{D}}\|_2 &\leq \xi_0 \\ \|\alpha_n \mathbf{D}\|_1 &\leq \mu_n + \xi_1 \\ \alpha_n &\geq 0. \end{aligned}$$

Once we solve for Equation (3.5) or (3.4), depending on the availability of the material texture, σ_n , we can update

$$\begin{aligned} \tau_n &= \hat{\alpha}_n \mathbf{D}, \\ \mathbf{T}[n, \Gamma_n] &= \tau_n. \end{aligned}$$

Algorithm 3.1 LTM synthesis

```

T = {All zero sparse  $N \times N$  matrix}
init = r random integers in [1,N]
for  $n \in \text{init}$  do
     $\tau_n = \alpha_n \mathbf{D}$     s.t.  $\|\tau_n\|_1 = \sigma_n[n]$                                 loop#0
     $\mathbf{T}[n, \Gamma_n] = \tau_n$ 
end for
while not complete do
    for Each row  $n$  newly updated in loop #0 or #2 do
        for  $m \in \Gamma_n$  do
            Solve Equation (3.3) or (3.2)                                loop#1
             $\mathbf{T}[\Gamma_m, m] = \tau_m$ 
        end for
    end for
    for For each column  $m$  newly updated in loop #1 do
        for  $n \in \Gamma_m$  do
            Solve Equation (3.5) or (3.4)                                loop#2
             $\mathbf{T}[n, \Gamma_n] = \tau_n$ 
        end for
    end for
end while

```

3.4 Applications to interactive rendering of stained glass

We first present some rendered results from the LTMs generated with Algorithm 3.1. We will then present its application to synthetic rendering and digital restoration of stained glass artifacts.

We generated LTMs of colorless glass slabs using the dictionary shown in Figure 2.10 by applying Algorithm 3.1. In Figure 3.5, we render a synthetic glass slab under horizontal line sources, when no material texture was available as an input to the algorithm. Figure 3.6 shows similar renderings of a synthetic glass slab when a material texture for the primal configuration σ_n was available. Thus Equation 3.2 was solved at loop #1 while Equation 3.5 was solved in loop #2. Figure 3.6(a) shows the material texture σ_n used as an input. Figure 3.6(b) and (c) show the synthetic glass slab under flood lights in the primal and dual configurations respectively. Figure 3.6(d)-(i) show renderings of the glass slab under a horizontal line light in the primal and dual configurations.

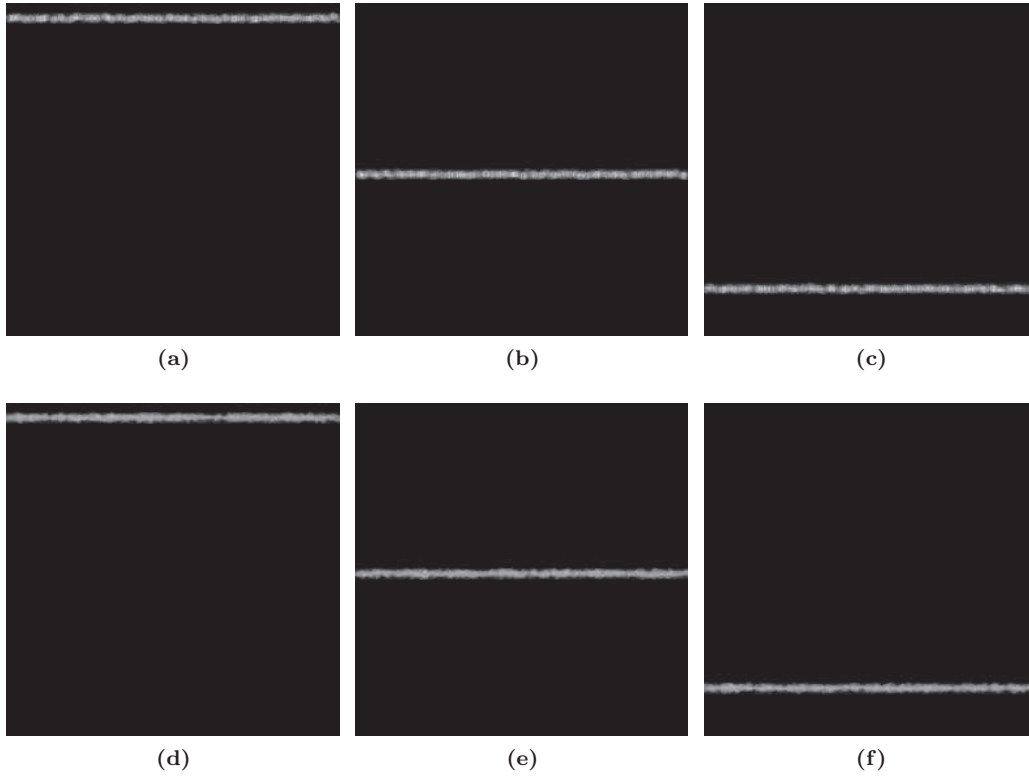


Figure 3.5: A textureless glass slab synthesized using the dictionary shown in Figure 2.10 under Algorithm 3.1. Equations (3.3) and (3.5) were used for synthesis (no material texture for either configuration). Subfigures (a),(b) and (c) show the glass slab as viewed under a moving horizontal line pattern. Subfigures (d),(e) and (f) show the same glass under the dual configuration for the same light sources as (a),(b) and (c) respectively.

3.4.1 Synthesis of Stained Glass Windows

While it is desirable to acquire the light transport properties of real stained glass windows by inverse rendering, visitors to the site often cannot afford to photograph the window against varying illumination over a period of time. However, if using modern High Dynamic Range imaging frameworks they can obtain a single photograph of the window under diffuse illumination, such a photograph can serve as the surface albedo $\mathbf{c}(\lambda)$ painted on the glass, as introduced in our image formation model presented in Equation (3.1). Thus, with such a user generated photograph and an appropriate material dictionary \mathbf{D} , we can generate a synthetic model of the stained glass window using Algorithm 3.1.

We generated a synthetic stained glass artifact from a single floodlit image of a real stained glass artifact containing the flag of Geneva. In Figure 3.7 we compare the real and synthetic stained glass artifacts rendered under similar illuminations. Note that our input to the Algorithm 3.1 only comprises the painted albedo $\mathbf{c}(\lambda)$ and an appropriate dictionary \mathbf{D} .

Alternatively, a stained glass artist might want to visualize the artifact before it has been assembled. In such a scenario, given the artist’s design in the form of the painted texture $\mathbf{c}(\lambda)$ and a dictionary \mathbf{D} learnt apriori on the glass slabs, we can use Equation (3.1) and Algorithm 3.1 to synthesize a relightable photograph of the conceptual stained glass.

In Figures 3.8 and 3.9, we demonstrate simulated rendering of stained glass using material slabs synthesized with our approach. We begin with a painted texture $\mathbf{c}(\lambda)$ shown in Figure 3.8(a). An LTM \mathbf{T} is created by assigning contiguous blocks of LTFs from different material slabs to each glass segment in the painted texture. Figure 3.8(b) shows the simulated glass under flood-light. Figures 3.8(c)-(e) show the same glass under different horizontal line-lights. Notice that each piece scatters light differently. Figure 3.9(a)-(f) compares the rendering of the same glass piece with alpha-matting, with homogeneous transport and with our proposed heterogeneous transport. Figure 3.9 (g)-(i) shows the same heterogeneous glass piece under a dynamic illumination set-up.

3.4.2 Digital Restoration of Stained Glass

Glass is a very stable medium that can retain its structure under stable conditions over a long period of time. Most stained glass windows however, due to their in-situ positioning have been exposed to harsh environmental conditions including: harmful UV radiation from sunlight, extreme weather conditions, biological damage caused by various pests, birds and micro-organisms as well as man-made pollution indoors and outdoors [63]. As a result, stained glass windows get corroded over centuries of existence and sometimes thick layers of dirt eventually become affixed to them. Similar environmental factors can also cause the deterioration of metal comes and frames that hold the various slabs of glass together, which can in turn affect the structural integrity of the stained glass window. While advanced conservation and restoration techniques exist, the general guideline is to be maximally non-invasive [63], retaining original material as much as possible. As a result, when a glass slab is broken or fractured, it is generally preserved as such.

Digitizing the light transport properties of stained glass provides us a platform for restoring such damages by digital means, thus enabling the creation of a restored digital replica of a damaged artifact. In this subsection, we will discuss our approach to restoring a stained glass window with fractures in one of its slabs. An overview of our method is presented in Figure 3.10.

The restoration process is as follows: we begin by acquiring the LTM of the damaged artifact.

We then manually define two indicator functions, one demarcating the fracture and the other describing a region of the stained glass that uses the same type of glass as the broken region. We then learn a dictionary \mathbf{D} on the LTFs of the marked region of glass, using Algorithm 2.1. We then restore the flood-lit image of the damaged artifact by image inpainting to obtain the material texture, σ . For inpainting, we perform a ℓ_1 -regularized frequency domain interpolation on the flood-lit image. Finally, using the restored texture σ and the learnt dictionary \mathbf{D} , we use Algorithm 3.1 to synthesize a new slab for the region containing the fracture. We then embed this synthesized LTM with the original acquired LTM to obtain the LTM of the restored stained glass artifact.

In Figure 3.10(a)), we show a portion of a 19th century stained glass window with fractures in it. Using the two indicator functions (Figure 3.10(c) and Figure 3.10(d)), we obtain the inpainted texture σ_n (Figure 3.10(b)) and the learnt dictionary (Figure 3.10(e)). We compare rendered images with the original LTM and the restored LTM in Figure 3.10(f) and (g). Note that due to interaction of light with the boundaries at the fractures, light transport at pixels surrounding the fracture are also affected in the original LTM. We therefore synthesize the light transport of the entire region around the fractures.

3.5 Summary

In this chapter we introduced a method for the synthesis of light transport matrices of stained glass windows using image based priors that we built for inverse rendering. We generate LTMs using an iterative matrix completion approach, with each row of the LTM being constrained by physical properties at the corresponding pixel. We show that the LTMs thus obtained can be used for simulating light transport, given just a single image of the artifact's texture. We also proposed an approach based on LTM synthesis to digitally restore fractures in stained glass artifacts.

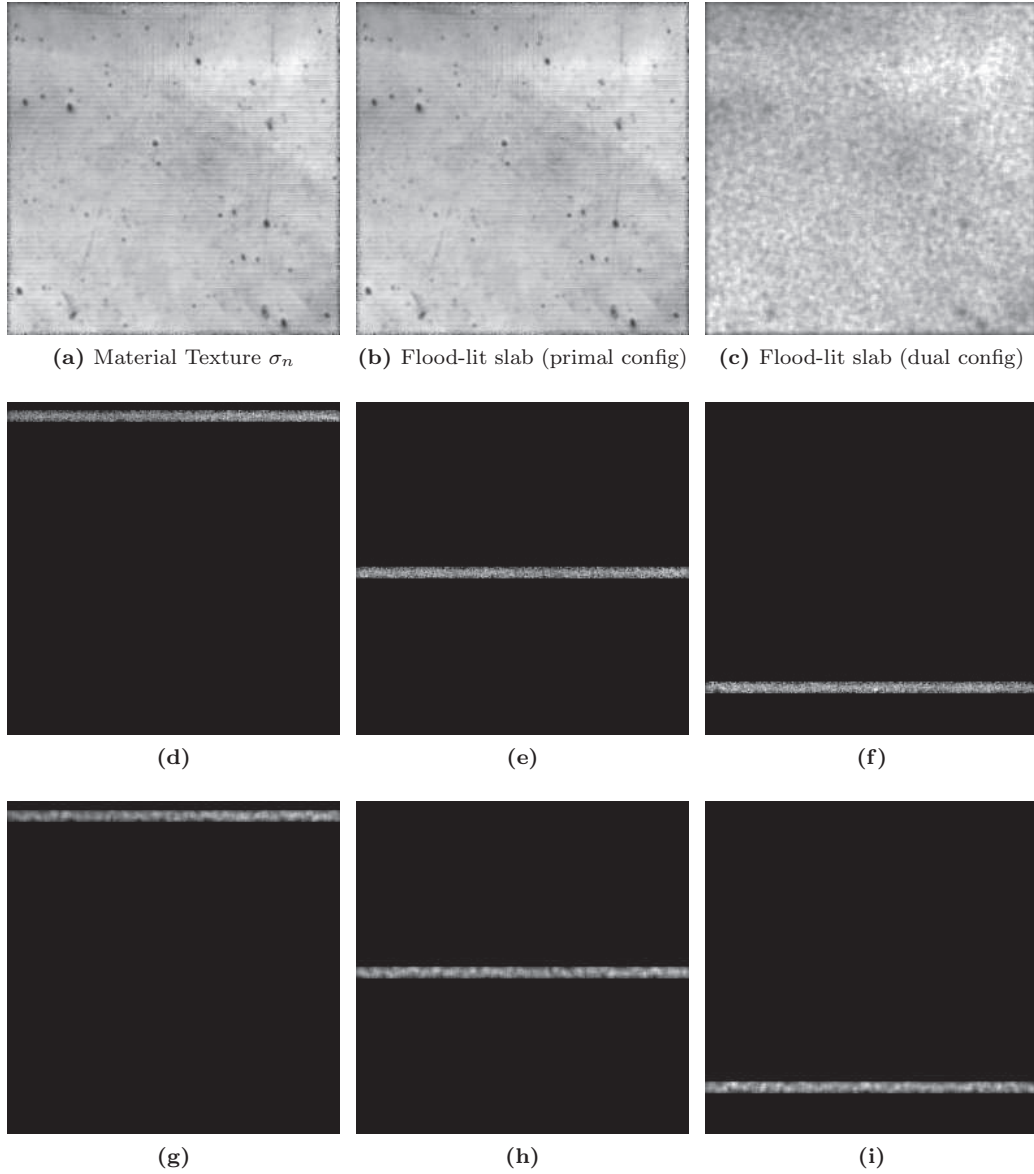


Figure 3.6: A glass slab with no painted texture synthesized using the dictionary shown in Figure 2.10 under Algorithm 3.1. Equations (3.2) and (3.5) were used for synthesis (material texture σ_n was available for primal configuration alone) Subfigures (d),(e) and (f) show the glass slab as viewed under a moving horizontal line pattern. Subfigures (g),(h) and (i) show the same glass under the dual configuration for the same light sources as (d),(e) and (f) respectively.

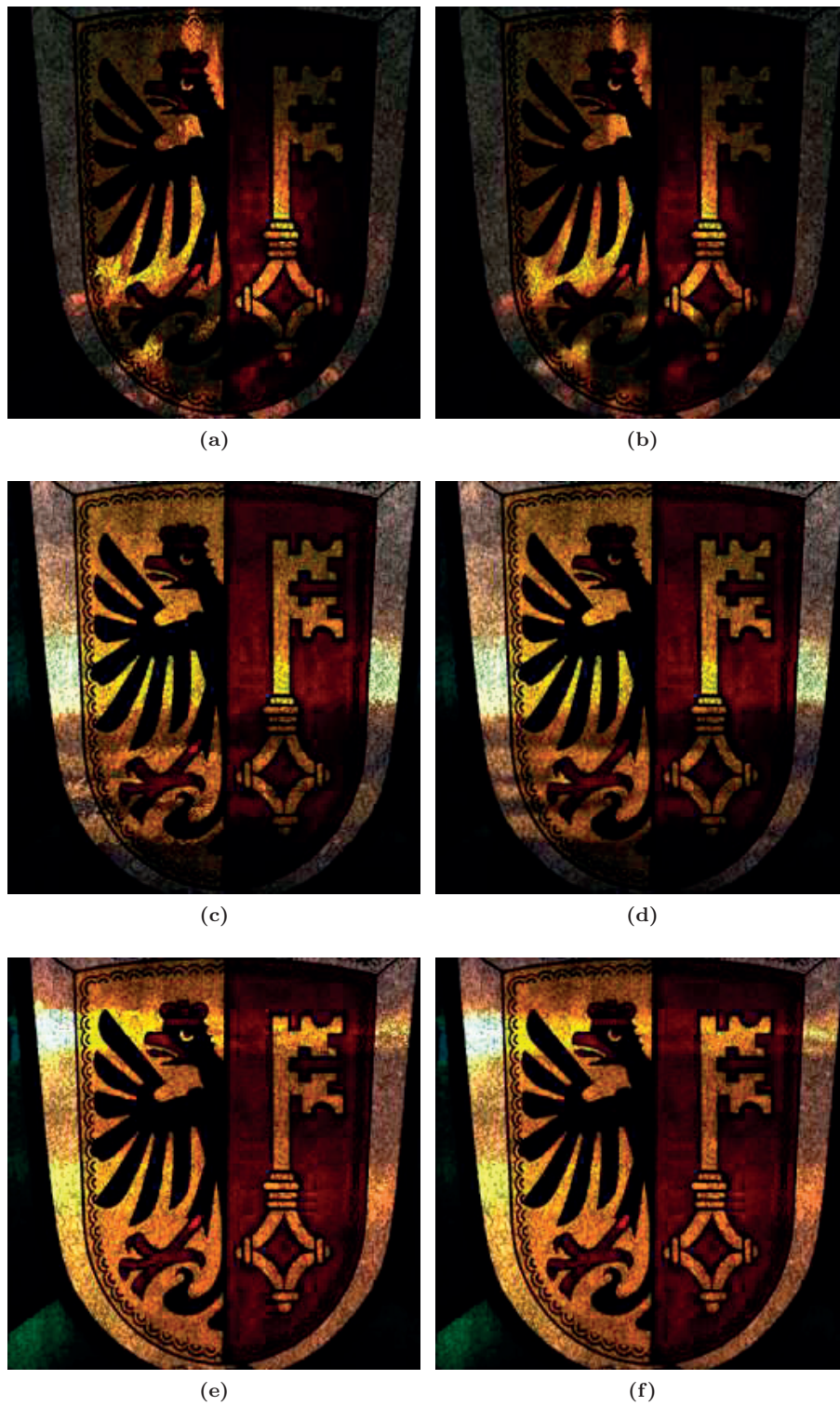


Figure 3.7: A synthetic stained glass artifact rendered from a single flood lit image of a real artifact. Subfigures (a),(c),(e) show the original artifact while subfigures (b),(d),(f) show the synthetic artifact under the same incident illuminations

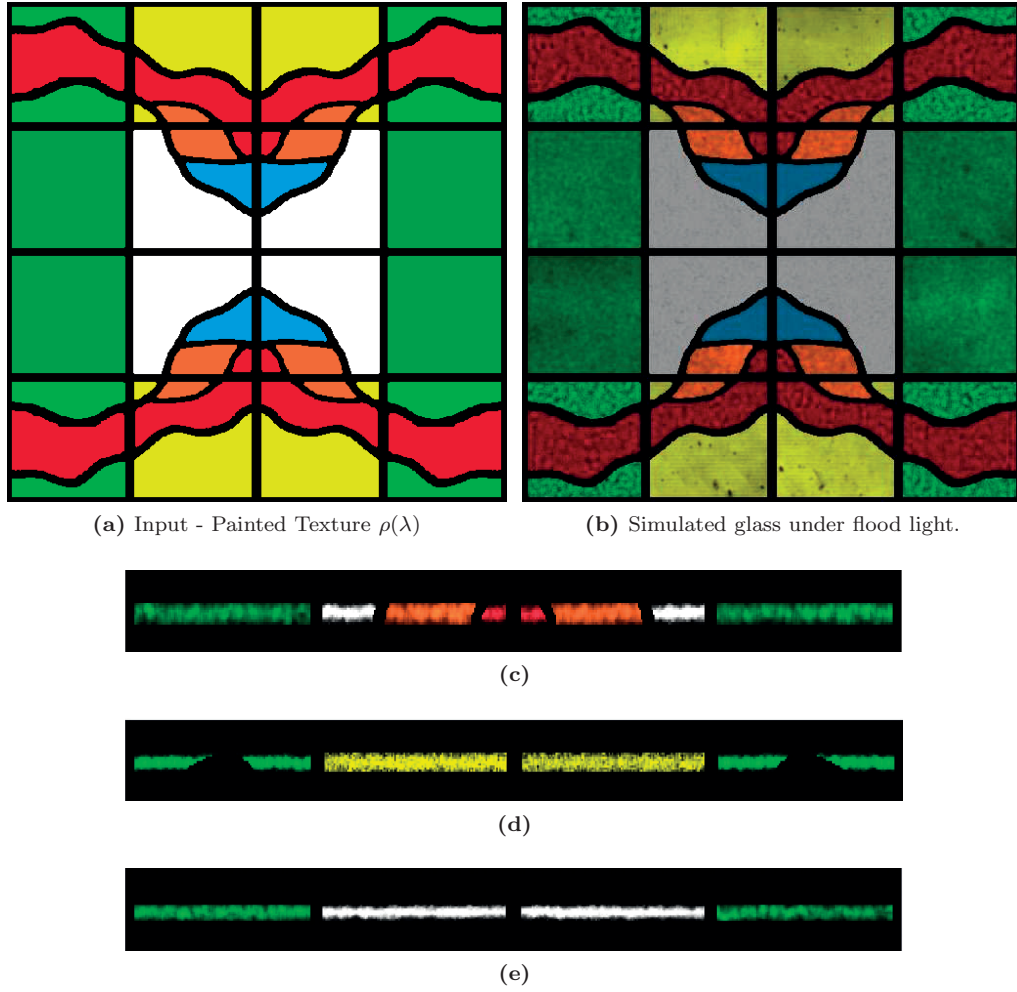


Figure 3.8: A fully simulated stained glass window constructed using seven different glass slabs each of which was synthesized using Algorithm 3.1. Subfigures (c),(d) and (e) show the glass artifact as viewed under a moving horizontal line pattern. Notice different scattering patterns for different pieces of glass.

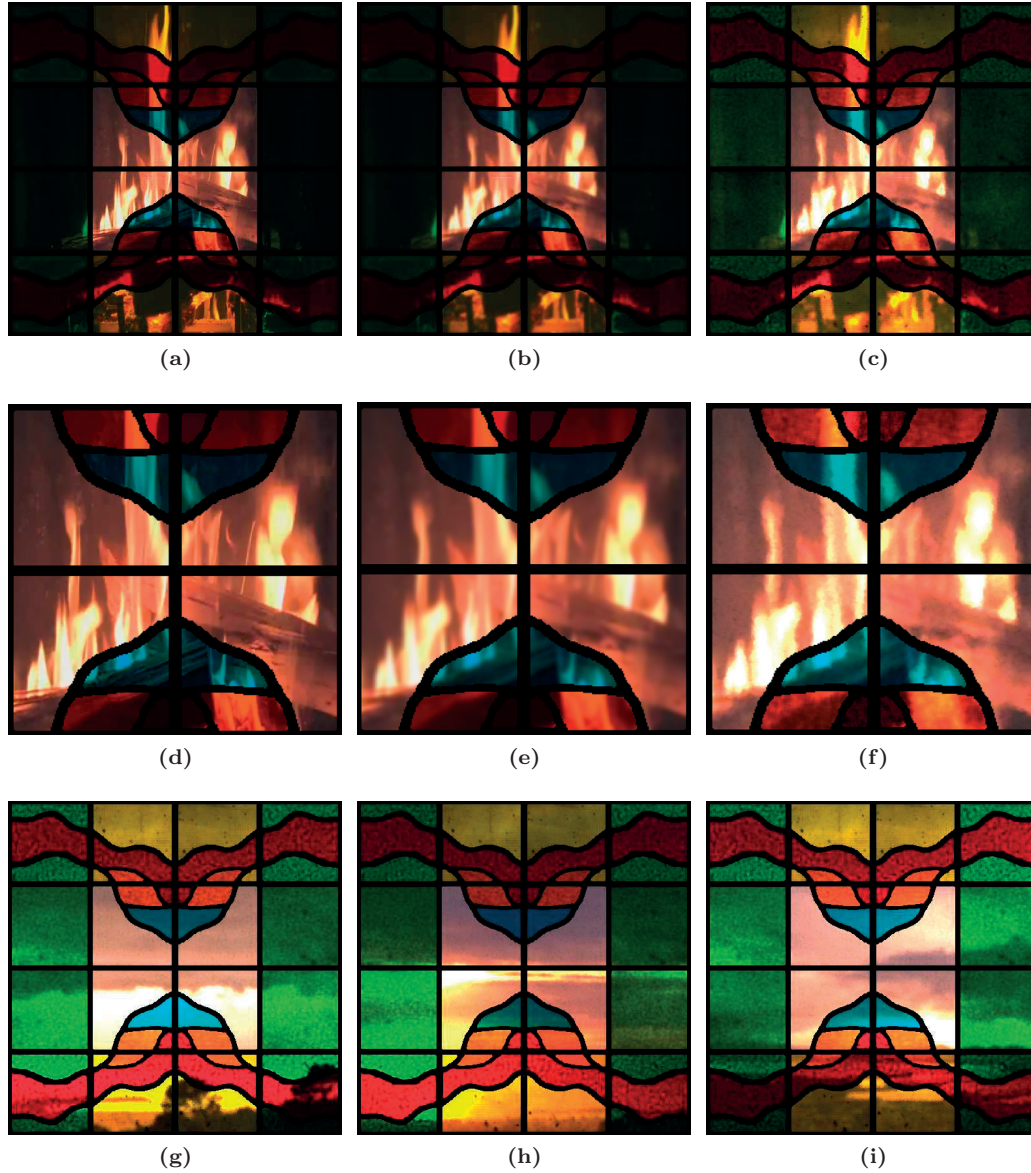


Figure 3.9: The same input texture $\mathbf{c}(\lambda)$ being rendered with (a) alpha-matting (b) homogeneous glass (c) our heterogeneous glass. Subfigures (d),(e),(f) show part of the respective images at a higher resolution. Subfigures (g),(h),(i) show the heterogeneous composite glass rendered under different illuminations.

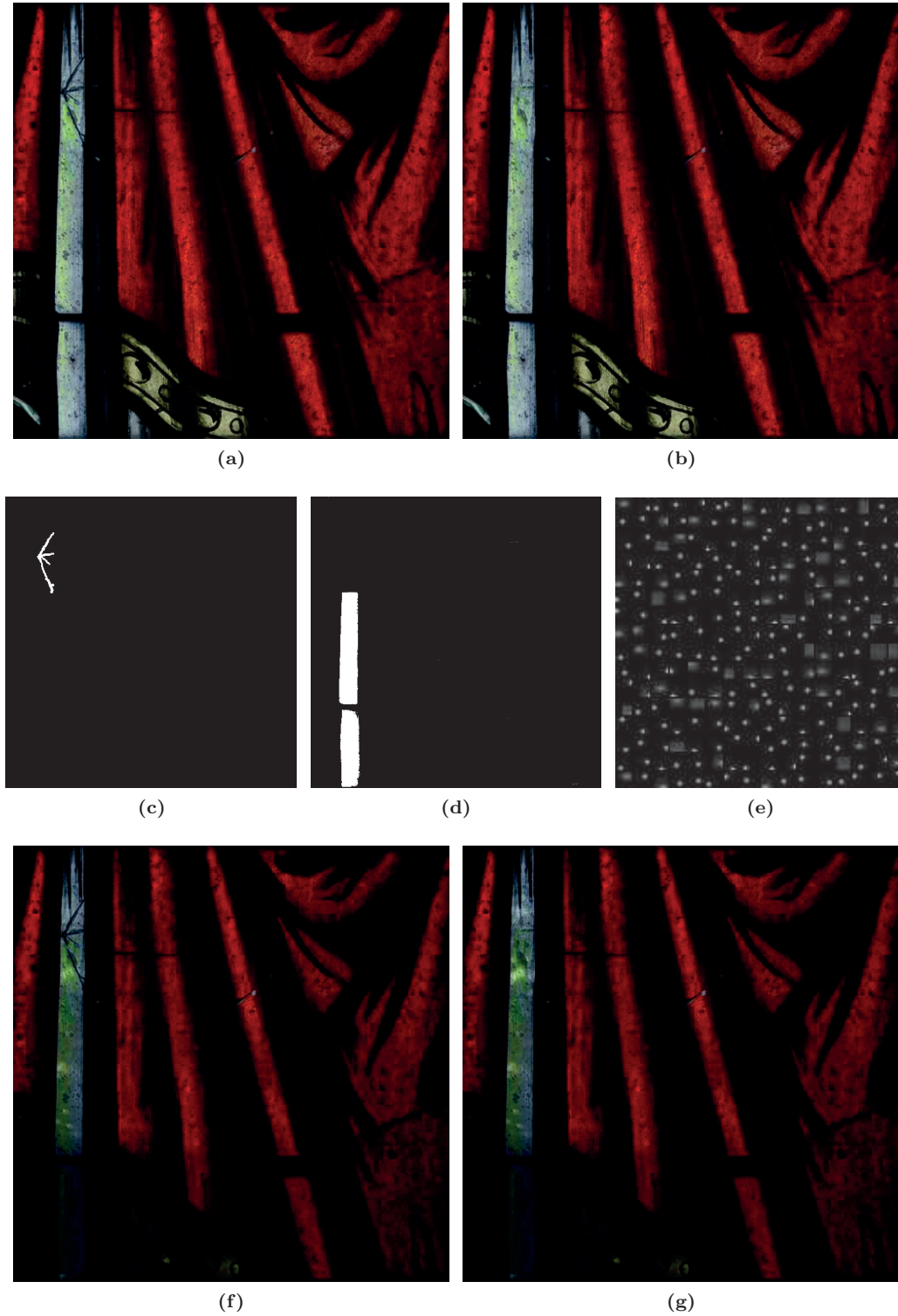


Figure 3.10: Digital restoration of stained glass. (a) The original artifact with fractures in it, under flood lit illumination is shown. (b) The restored texture obtained by inpainting is shown. (c) Subfigure shows an indicator function defining the fracture. (d) Subfigure shows part of the artifact made of the same material as the fractured section. (e) A dictionary learnt on the chosen subregion of the artifact is shown. Subfigures (f) and (g) show the original and the restored artifact rendered under the same illumination.

Chapter 4

Handheld Acquisition of Light Transport

We are stuck with technology when what we really want is just stuff that works.

Douglas Adams, The Salmon of Doubt

In this chapter, we introduce our framework for acquiring and rendering relightable photographs using smartphones.¹

4.1 Introduction

Until now, we focused specifically on the light transport properties of stained glass windows. We will now shift our focus to the more common case, namely reflective scenes. More than a decade ago, the lighting sensitive display [51] was first proposed as a medium for displaying photographs that respond to the viewing environment by dynamically changing their incident illumination. Such relightable photographs are created by acquiring the scene’s reflectance field, so that the scene can be rendered under new, user-defined illumination conditions. As described in Chapter 1, numerous approaches that acquire scene reflectance have been proposed. However, the complexity and the scale of most of the existing acquisition frameworks prevent widescale use beyond specialized studios. In this chapter, we present a novel, affordable, portable and easy-to-use framework for creating relightable photographs, by obtaining the scene’s light transport matrix.

1. The work presented in this chapter was done in collaboration with L. Baboulaz, D. Firmenich, S. Susstrunk and M. Vetterli [66]

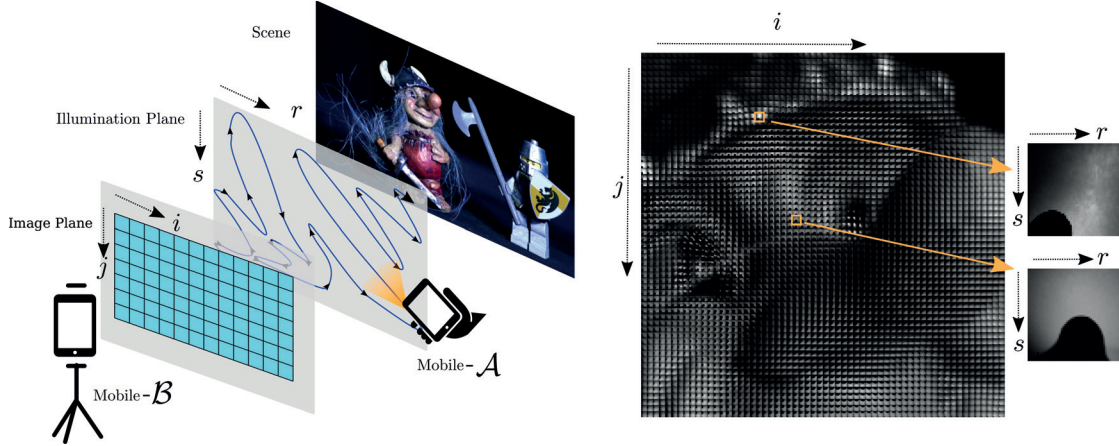


Figure 4.1: Our acquisition system: Mobile \mathcal{B} is placed in a stationary position facing the object. Its camera is turned on and it captures images continuously. The user then scans a slice of the reflectance function by moving mobile \mathcal{A} , with its LED on, in a random trajectory along a plane. On the right, we show the array of array representation of a small part of the LTM from the Toys Dataset. Here, each sub-array is the light transport function of the corresponding pixel. In inset, we show two of these sub-arrays.

We build our acquisition system entirely around the smartphone as shown in Figure 4.1. In our proposed framework, a user samples the reflectance field of the scene by moving a smartphone (mobile \mathcal{A}) with its LED turned on, facing the scene. Simultaneously, a second smartphone or any video-capable camera (\mathcal{B}) records the corresponding scene response. The user interface on the screen of phone \mathcal{A} guides the user by displaying the path taken by the user (See Figure 4.1 for an illustration). We thus obtain a sparse, non-uniform sampling of the reflectance function from the proposed handheld acquisition system. To obtain a meaningful, uniformly sampled light transport matrix (LTM), that we can use for inverse rendering, we then introduce a compressive-sensing based reconstruction strategy. A major difference in our approach, compared with conventional methods that obtain the LTM, is the use of an LED that approximates a point-light source, instead of the more common structured-light (projector) setups or directional light sources used in lightstages. We therefore propose a two plane parametrization for the presented framework, that enables the use of LTMs acquired with point light sources for image based relighting of scenes under more general incident illuminations.

Our proposed acquisition system is easy to use, portable, off the shelf, and is fast, only requiring a few minutes per acquisition. Such a system could enable the creation of user-generated relightable photographs, that have wide reaching applications in virtual tourism and online marketplaces.

We will introduce the two plane parametrization of light transport with point light sources, and discuss image based relighting with this parametrization in Section 4.3. We then present our mobile-based acquisition strategy, its sampling model and the reconstruction methodology for non-uniformly sampled light transport in the next section. We discuss experimental results and future extensions of the current setup in Section 4.5.

4.2 Related Work

In general, the elaborate, specialized setups presented in [24, 31, 44, 70] or projector-camera setups [49, 52, 55, 56, 58, 73, 76, 79] severely limit their usability outside of specialized studios. A simple, easily available hardware setup with an intuitive, fast, and non-restrictive usage scenario is therefore highly desirable.

Polynomial texture mapping [44] has a simple setup, as it can be used with a simple light probe and a user-controlled light source. However, the acquired per-pixel reflectance data is projected into the space of biquadratic polynomials, which results in a loss of specularities. As a compromise, the authors propose a simulated specular reflectance that uses parametric models along with the acquired normal maps. A more recent example is pocket reflectometry [59]. It uses a handheld linear-light source in combination with a reference bidirectional reflectance distribution function (BRDF) chart to acquire the spatially varying BRDF parameters of the scene. This parametric spatially varying BRDF is then used for rendering and is mainly aimed at scenes with convex surface geometry. The free-form light stage [46] uses a free-moving handheld light source in conjunction with four spheres in order to create basis images. A weighted combination of these basis images is then used for image based relighting of the acquired scene. We compare our reconstruction strategy with other reconstruction strategies in Section 4.5. Moreover, unlike these approaches, our approach does not require any special instrumentation apart from the now common smartphone.

Davis et al [13] have proposed a setup to capture light fields using a moving mobile camera. Although we capture the LTM instead of the light field, the two approaches take a similar smartphone-based scanning approach. However, they use bicubic interpolation for their reconstruction, hence a dense sampling of the light field is required during the acquisition stage. In contrast, as we operate in the frequency-domain, our approach allows user to trade-off the density of the trajectory with the resolution of the reflectance function. As we will show in Section 4.5, our work also extends to the reconstruction of light fields from sparse trajectories, thus enabling both dynamic illumination and moving viewpoints.

A rigorous treatment of reconstruction of 2D spatial fields from samples along a trajectory is presented in [71]. However samples are obtained along well-defined curves, hence the approach is unsuitable for handheld sampling. Our acquisition system does not require the user to follow any specific trajectory in order to scan the object, thus making the acquisition more user-friendly.

4.3 Light Transport with Point Light Sources

Conventionally a structured light source (for example, a projector or CRT monitor in conjunction with a diffuser) or a light stage is used to characterize the light transport matrix. While projector-camera setups aid in the acquisition of complex light transport properties, the light stage setups have remained indispensable in acquiring the surface reflectance of objects. In light stage based LTMs introduced in Chapter 1, due to limitations on the size of the region that can be illuminated by purely directional light sources, they are approximated using point light sources placed at a distance from the scene. In contrast, we propose an image based relighting model directly based on point light sources, and can thus benefit from a more compact acquisition scenario, with light sources located closer to the object. Throughout this chapter, we focus on the problem of reflectance acquisition for a fixed view angle and extend it to multiple views in Section 4.5.8.

4.3.1 Continuous Space Light Transport

As introduced in Chapter 1, the light transport $S(\omega_o, x_o, \omega_i, x_i)$ of a general scene is 8-dimensional. In this chapter, since we are primarily concerned with objects that are predominantly reflective, when the scene has a convex surface geometry, light enters and exits the scene at the same point; thus $x_i = x_o$. Thus we are required to sample the spatially varying reflectance $S(\omega_o, x_o, \omega_i)$. Since we fix the viewing angle for now, scene appearance is entirely characterized by the 4D light transport field, $S(\omega_i, x_o)_{\omega_o}$ (Here the outgoing direction ω_o is fixed). Note that, similar to a 4D light field, the 4D light transport $S(\omega_i, x_o)_{\omega_o}$ is defined by a bundle of rays, with each ray defined by a direction ω_i and a point ($x_i = x_o$) on the surface of the object. Thus, similar to a light field, this 4D light transport can be parametrized into a two-plane parametrization. Without loss of generality, the first plane with Cartesian co-ordinates $(x, y) \in \mathbb{R}^2$ is defined to coincide with the focal plane of the camera. As the camera projects this plane to its image plane (sensor or film), we refer to it as the image plane. Let us define the second plane described by Cartesian co-ordinates $(r, s) \in \mathbb{R}^2$ to be parallel to the image plane and located at a distance h from it, in front of the scene. Thus for a fixed viewing angle, image formation is a function of the incident light field that is solely parametrized by planes (r, s) and (x, y) ; this function, represented by $\mathcal{T}(r, s, x, y)$ is the continuous-space light transport field.

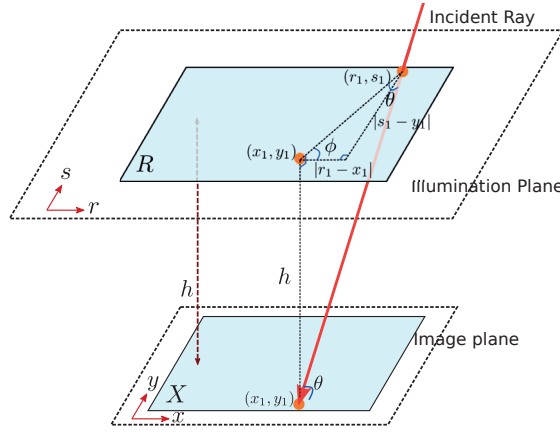


Figure 4.2: Two plane parametrization for continuous space light transport. The figure also describes the relighting procedure for a directional light source (θ, ϕ) .

Given $\mathcal{T}(r, s, x, y)$, the scene can be relit under any incident light field using a linear combination of scene responses to individual light rays in the incident light field. Consider a point light source (isotropic source) located at a point in (r, s) . Under such an illumination, each point in the image plane (x, y) is intersected by a light ray at an uniquely different angle. $\mathcal{T}(r, s, x, y)$ can be obtained by moving a point-light source through the plane defined by (r, s) . As the light source is located on the plane (r, s) we refer to it as the illumination plane. As moving the point-light along infinitely large planes is not possible, we restrict the illumination plane (r, s) to a finite rectangle R and the image plane (x, y) to a finite rectangle X . Note that such a system denotes the Helmholtz dual of the acquisition of a light field using a continuously displaced pinhole camera. Therefore, each point $(x_j, y_j) \in X$ has been observed for the entire range of incident angles subtended by the rectangle R in the illumination plane. Thus for a scene with

convex surface geometry, a point $(x_j, y_j) \in X$ can be rendered under any incident illumination angle θ_i, ϕ_i given by,

$$\theta_i = \arcsin \frac{h}{\sqrt{(x_j - r)^2 + (y_j - s)^2 + h^2}}, \quad \phi_i = \arctan \frac{y_j - s}{x_j - r}, \quad \text{where, } (r, s) \in R. \quad (4.1)$$

The range of available illumination configurations is limited by the dimensions of the rectangle $R : 2a \times 2b$ and its distance h from the scene. For a surface point located directly below the center of the illumination plane, the solid angle swept by the set of available incident light directions is

$$\Omega_{(R,h)} = 4 \arctan \frac{ab}{h\sqrt{a^2 + b^2 + h^2}}. \quad (4.2)$$

Special case When $a = b = h$, $\Omega_{(R,h)} = 2\pi/3$, while the set of all incident directions for a reflective object is 2π as defined by a hemisphere. As can be seen, $\Omega_{(R,h)}$ approaches 2π with increasing a and b or decreasing h .

Given $\mathcal{T}(r, s, x, y)$, we can relight the scene as lit by a *directional light* source along any incident direction that exists inside the solid angle $\Omega_{(R,h)}$, by choosing the incident rays such that equation (4.1) is satisfied (Figure 4.2). Alternately, the scene can be relit with *point light* sources located above or below the illumination plane (r, s) by choosing incident light rays such that they converge to the desired point, as shown in Figure 4.3. Given X and a point source \mathbf{p} located at a distance d from the image plane, the incident rays intersect the illumination plane forming a rectangle defined by the points \mathbf{q}_j given by the equation

$$\frac{\|(x_j, y_j) - \mathbf{p}\|}{d} = \frac{\|\mathbf{q}_i - \mathbf{p}\|}{|d - h|} \quad j \in [1, 2, 3, 4]. \quad (4.3)$$

Image based rendering of the object under point light sources located above or below the

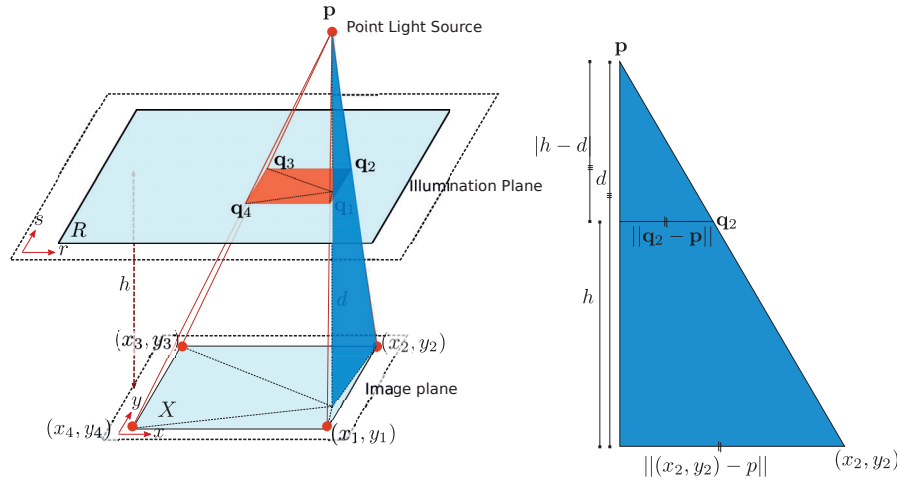


Figure 4.3: Rendering with a point light source at a novel location, given the continuous space light transport.

illumination plane is thus possible by choosing the incident rays such that equation (4.3) is satisfied.

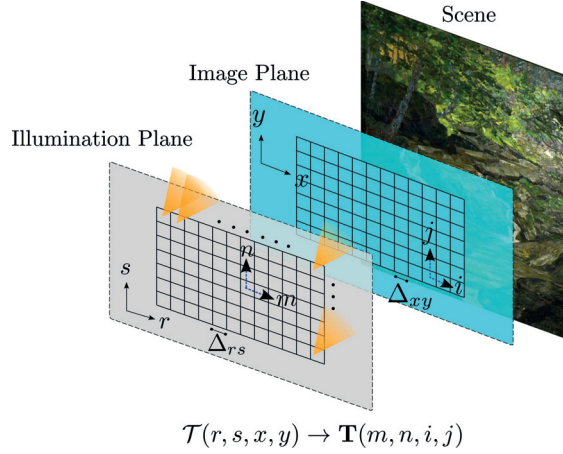


Figure 4.4: Uniform Sampling of the light transport field: The image plane corresponds to the sensor of the camera used. The illumination plane, in our case, refers to an array of point light sources. In conventional projector-camera setups, the illumination plane refers to the focal plane of the projector.

4.3.2 Uniform Sampling Model for Light Transport

Let us now consider a conventional sampling setup for the light transport field, $\mathcal{T}(r, s, x, y)$. Such a sampling setup involves a camera focused on the image plane and an array of point light sources placed on the illumination plane. The camera discretizes the image plane (x, y) , as described by,

$$\tau_{ij}(r, s) \equiv \langle \mathcal{T}(r, s, x, y), \phi_{xy} \left(\frac{x}{\Delta_{xy}} - i, \frac{y}{\Delta_{xy}} - j \right) \rangle, \quad (4.4)$$

where Δ_{xy} and ϕ_{xy} are the sampling rate and spatially invariant sampling kernel of the camera. $\tau_{ij}(r, s)$ denotes the continuous-space light transport function (LTF) at the surface point corresponding to pixel (i, j) on the camera. The inner product is defined in Equation (1.2).

When $(r, s) \in \mathbb{R}^2$ (infinitely large illumination plane) each $\tau_{ij}(r, s)$ corresponds to the complete reflectance function of the surface point corresponding to the pixel (i, j) . However, since $(r, s) \in R$, we only observe a slice of the reflectance function whose angular range is given by $\Omega_{(R,h)}$ in Equation (4.2).

Let us now consider the sampling operation on the illumination plane using an array of point light sources. As each LTF, $\tau_{ij}(r, s)$ at pixel (i, j) undergoes the same sampling operation at the illumination plane, we focus only on the sampling and reconstruction of a general LTF, $\tau(r, s)$. Let $\mathbf{t}(m, n)$ denote the discretized LTF, obtained by uniformly sampling $\tau(r, s)$ with the sampling rate Δ_{rs} and a sampling kernel ϕ_{rs} .

$$\mathbf{t}(m, n) = \langle \tau(r, s), \phi_{rs} \left(\frac{r}{\Delta_{rs}} - m, \frac{s}{\Delta_{rs}} - n \right) \rangle.$$

Thus, given the LTFs $\mathbf{t}_{ij}(m, n)$ for all pixels (i, j) in the camera sensor, the uniformly sampled LTM \mathbf{T} can be obtained, with which we can relight the scene for any incident angles $\omega_i \in \Omega_{(R,h)}$. Note that, much like the camera-array for acquiring light fields, an actual array of point light

sources is complex, non-portable and expensive due to the redundant light sources. In the next section, we will present our proposed sampling strategy that approximates the uniformly sampled LTF.

4.4 Light Transport Acquisition System

In this section, we introduce our easy-to-use, inexpensive, hand-held acquisition system to obtain a LTM that approximates the theoretical model presented in the previous section. Since in our proposed acquisition system we only replace the array of lights with a moving LED, the sampling operation at the image plane is identical to that presented in the uniform sampling case. Thus, we will here only consider the non-uniform sampling and reconstruction of the general light transport function $\tau(r, s)$ (defined in Equation (4.4)).

Our inverse rendering framework is composed of three stages :

- Handheld acquisition
- Resampling step
- LTM reconstruction.

During handheld acquisition, the user samples the scene's light transport along an arbitrary trajectory, resulting in samples located in \mathbb{R} . In the resampling step, these samples are projected onto a lattice, to obtain a non-uniformly sampled light transport, with samples located in \mathbb{Z} . These steps are illustrated in Figure 4.5. Finally, the reconstruction step interpolates the missing samples to obtain an approximation $\hat{\mathbf{T}}$ to the uniformly sampled LTM \mathbf{T} .

4.4.1 Handheld Acquisition

The acquisition begins with the user moving the mobile \mathcal{A} along an arbitrary trajectory γ on a chosen illumination plane, with its LED turned on and facing the object to be scanned. The stationary camera \mathcal{B} then observes $\tau(\gamma)$. While lighting the scene, mobile \mathcal{A} also tracks its own position using its camera. Similar to [13], to perform this tracking we use a mobile implementation of simultaneous localization and mapping (SLAM) technique called parallel tracking and mapping (PTAM)[39]. It works by first initializing a 3D map of the scene using stereo correspondences in an initialization step. Following this, tracking and mapping are done concurrently, with the tracking step computing the current position of the mobile relative to the scene's 3D map. The 3D map is also continuously updated. Thus, PTAM can yield the position of the mobile phone, in real-world co-ordinates, for each observed light position.

To summarize, the acquisition stage involves the following steps.

- Camera \mathcal{B} is placed in a stationary position, with the object of interest in its field of view.
- The user holds mobile \mathcal{A} and marks the object of interest on it. A bounding box corresponding to the illumination plane to be sampled is shown to the user.
- The user must now start moving the mobile \mathcal{A} with its LED on, inside the depicted bounding box. A synchronisation signal sent to mobile \mathcal{B} starts recording a video. The trajectory that the user takes is tracked using the PTAM algorithm and is displayed on the bounding box in real-time. The user continues until the entire bounding box has been spanned.
- Mobile \mathcal{B} outputs a video, each frame corresponding to an observed image. Mobile \mathcal{A} outputs the set of co-ordinates corresponding to the location of the light source on the illumination plane for each frame of the video.

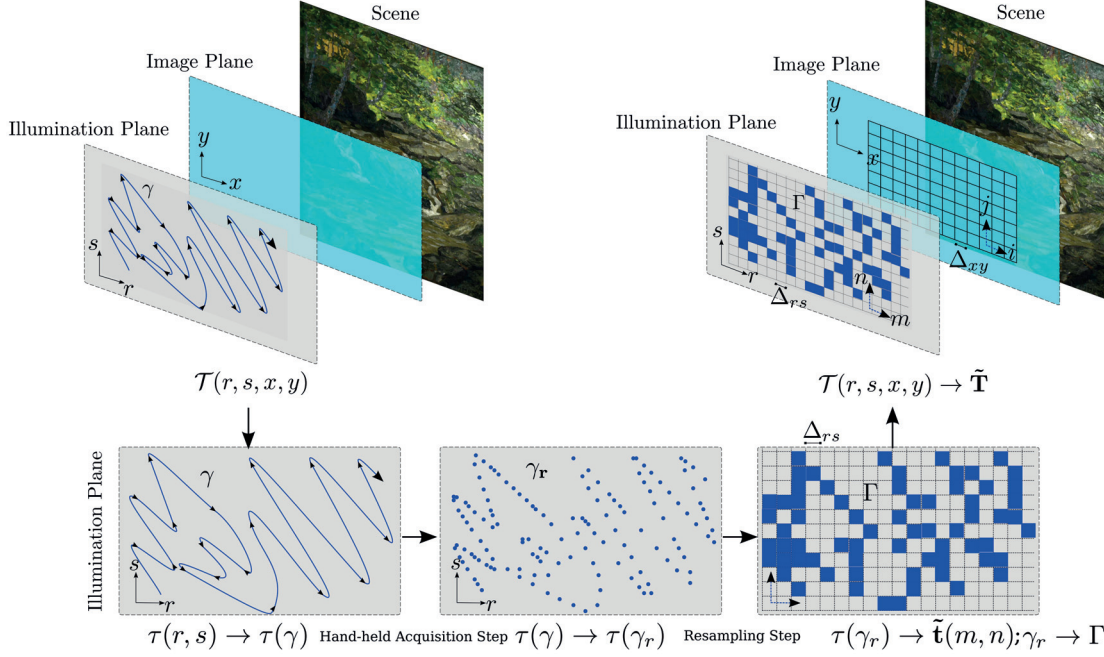


Figure 4.5: Our non-uniform sampling model: The light transport field is sampled by the user by moving mobile \mathcal{A} along a random, uncontrolled trajectory γ . PTAM automatically discretizes this trajectory to yield a light transport field sampled along a discrete trajectory γ_r . This is then resampled into the sampling lattice of the illumination plane to obtain a sparsely sampled LTM

In our proposed acquisition system, the trajectory γ and therefore $\tau(\gamma)$ can only be sampled at non-uniform intervals as a result of,

- the non-uniform rate at which the positions of mobile \mathcal{A} are computed, as determined by its framerate and the PTAM algorithm.
- the uniform sampling at mobile \mathcal{B} , as defined by its video framerate.
- the varying velocity with which the user moves mobile \mathcal{A} .

Thus at the end of the proposed acquisition, only samples $\tau(\gamma_r)$ along a non-uniformly sampled trajectory γ_r defined on (r, s) are obtained.

As discussed later, our reconstruction algorithm requires these samples to be distributed uniformly at random across the desired illumination plane. Hence the user must move on arbitrary trajectories that span the entire illumination plane. To ease this requirement, a user interface displayed on mobile \mathcal{A} indicates the trajectory on the simulated plane in real-time. Tracking mobile \mathcal{A} is the only computation that is required at this stage.

4.4.2 Resampling Step

After acquisition, the light transport has been sampled non-uniformly, with samples not lying on a lattice. We now resample, $\tau(\gamma_r)$ by projecting it onto a lattice defined by a user determined parameter - the sampling period Δ_{rs} . This projection is defined by the binning operation:

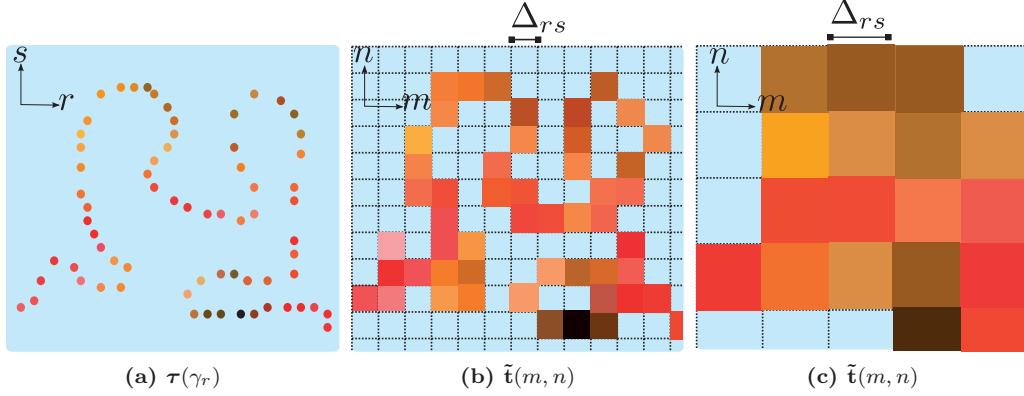


Figure 4.6: An illustration of the resampling step. (a) A non-uniformly sampled LTF, with samples in \mathbb{R} . (b) and (c) The light transport function is resampled with different sampling periods Δ_{rs} by binning. Note that a larger Δ_{rs} has resulted in a more densely sampled lattice, albeit at a loss of resolution, compared with the smaller Δ_{rs} .

$$\tilde{\mathbf{t}}(m, n) = \mathcal{M}(\boldsymbol{\tau}(\gamma_r)W(r - m, s - n)), \quad (4.5)$$

where $W(r, s)$ is the box function defined as

$$W(r, s) = \begin{cases} 1, & \text{if } r \text{ and } s \in [\frac{\Delta_{rs}}{2}, \frac{\Delta_{rs}}{2}) \\ 0, & \text{otherwise,} \end{cases}$$

where, $\mathcal{M}(\cdot)$ is the averaging operator and m, n are indices on the discrete illumination plane. Similarly, the trajectory γ_r can also be projected onto the lattice, to obtain an indicator function $\mathbf{\Gamma}$ that describes the set of elements in the uniform lattice for which samples have been observed. Let K be the number of elements for which samples were observed and M be the total number of elements in the lattice. As can be seen in the Figure 4.6, for a given trajectory, the choice of Δ_{rs} determines the sampling ratio $\frac{K}{M}$, as it affects both K and M . A second factor that determines the sampling ratio $\frac{K}{M}$ is the path density of the trajectory, a metric introduced in [71] as the length of continuous-space trajectory γ per unit area. We study how the choice of Δ_{rs} affects reconstruction and inverse rendering in Section 4.5.3.

Thus, the resampling step yields a sparsely sampled light transport function $\tilde{\mathbf{t}}(m, n)$ at each pixel in the scene with sample locations given by the indicator function $\mathbf{\Gamma}$. The goal of the reconstruction algorithm is then to find $\hat{\mathbf{t}}(m, n)$ an approximation from the sparse set of samples.

For the sake of simplicity, we will express discretized light transport functions $\tilde{\mathbf{t}}(m, n)$ as vectors $\tilde{\mathbf{t}}$, by rearranging them in the column-major format.

4.4.3 LTM Reconstruction

We now obtain an approximation, $\hat{\mathbf{t}} \in \mathbb{R}^M$ from $K < M$ random Euclidean measurements. The reconstruction problem can thus be studied under the sparse sampling framework.

Assuming that the LTFs are S -sparse (or approximately sparse) in the Fourier domain, if \mathbf{f} denotes the discrete Fourier transform of \mathbf{t} such that

$$\mathbf{f} = \mathbf{H}\mathbf{t},$$

where \mathbf{H} is the transform matrix, we obtain an approximation $\hat{\mathbf{f}}$ of \mathbf{f} , by solving for

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \|\mathbf{f}\|_1 \quad \text{subject to} \quad \tilde{\mathbf{t}}_K = \mathbf{H}_K^T \mathbf{f}, \quad (4.6)$$

where \mathbf{H}_K is obtained by sampling the columns of the transform matrix \mathbf{H} using the indicator $\mathbf{\Gamma}$ and $\tilde{\mathbf{t}}_K$ are the K sparse samples from $\tilde{\mathbf{t}}$. It has been shown [8] that when the number of observations K is such that

$$K \geq C \cdot S \cdot \log M, \quad (4.7)$$

then, Equation (4.6) yields the S -sparse approximation of \mathbf{f}_n . Here C is a small constant independent of M . The discrete LTF can be reconstructed by the inverse transform

$$\hat{\mathbf{t}} = \mathbf{H}^T \hat{\mathbf{f}}.$$

We reconstruct $\hat{\mathbf{t}}_{ij}$ at all pixels (i, j) of the camera. The LTM can now be constructed by first vectorizing the image plane and stacking the discretized LTFs of all J pixels of the camera such that,

$$\hat{\mathbf{T}} = \begin{bmatrix} \mathbf{t}_1^T \\ \mathbf{t}_2^T \\ \vdots \\ \mathbf{t}_J^T \end{bmatrix}.$$

Here, each row of $\hat{\mathbf{T}}$ describes how light is transported from the vectorized illumination plane to the corresponding pixel on the vectorized image-plane. Similarly, each column corresponds to the impulse response of the corresponding point-light source. Given $\hat{\mathbf{T}}$, the scene can be relit under novel light positions. We further discuss scene relighting in Section 4.5.6.

4.5 Experiments and Analysis

In this section, we first present the experimental evaluation of the tracking algorithm and then continue with the reconstruction algorithm. We follow by discussing the trade off between the sampling period and reconstruction efficiency and by quantifying the error induced by the reconstruction algorithm. We finally extend the proposed approach to the case of multiple view angles and present demonstrations with mobile-based acquisitions.

4.5.1 Evaluation of the Tracking Algorithm

We implemented the acquisition system presented in Section 4.4.1 on Android smartphones. As part of our reconstruction framework, we require the PTAM algorithm to give accurate estimates of the discrete trajectory $\gamma_{\mathbf{r}}$ of the light source positions in space. To measure the accuracy of PTAM, we compared its estimated coordinates with ground truth data by placing mobile \mathcal{A} on a computer controlled X-Y table that enables precise displacements on a 2D plane.

PTAM consistently predicted the location of the light source accurately with an acceptable root-mean-square error of less than 1mm (0.31mm). In practice, most inaccuracies are caused by the user moving out of the illumination plane while moving the light source. Both the rotation of the light source and its translation along the z-axis can also induce errors. However, PTAM enables us to track these errors and discard unusable data by using an error margin.

4.5.2 Experimental Validation against Ground Truth

We now evaluate our reconstruction strategy by reconstructing the LTM of known objects. For this purpose, first a ground truth LTM \mathbf{T} was obtained for each object by moving a point-light source along a plane by using a computer controlled X-Y table. A simulated indicator function $\mathbf{\Gamma}$ is then described on this lattice to define the sparsely sampled LTM $\tilde{\mathbf{T}}$. We obtain three datasets in this way.

- Dataset, *Notre Dame-a* is a small but specular region of an oil painting, without any downsampling on the image plane (Figure 4.7(a)).
- Dataset *Notre Dame-b* is the complete version of the same oil painting, downsampled to 1024×1024 on the image plane (Figure 4.7(b)).
- Dataset *Toys* is that of two toys with intricate 3D geometry, containing some specular surfaces (Figure 4.7(c)). Note that this scene is highly non-convex, which results in discontinuities the LTFs due to cast shadows. We will further discuss its reconstruction performance in Section 4.5.5.

In our implementation, we chose to move to the YCbCr colorspace and assume that the Cb and Cr channels are illumination invariant. The SNR obtained for each dataset is tabulated in Table 4.1. Here, SNR_Y denotes the signal-to-noise ratio on the luminance channel and is defined as $20 \log_{10} \frac{\|\mathbf{T}\|_2}{\|\mathbf{T} - \tilde{\mathbf{T}}\|_2}$. It can be seen that our recovery algorithm performs consistently well in all three datasets. In this experiment, the sampling ratio $\frac{K}{M}$ in all three datasets was 0.29. Relighting results are shown in Figure 4.15 and 4.16.

Dataset	SNR_Y
<i>Notre Dame-a</i>	29.09 dB
<i>Notre Dame-b</i>	31.88 dB
<i>Toys</i>	20.99 dB

Table 4.1: SNR from various ground-truth experiments. Sampling ratio, $\frac{K}{M} = 0.29$.

4.5.3 Effect of Sampling Period Δ_r and Path Density

The projection of $\boldsymbol{\tau}(\gamma_r)$ into a uniform lattice to obtain the sparsely sampled $\tilde{\mathbf{t}}(m, n)$ as defined in Equation (4.5) is significant as the choice of $\Delta_{r,s}$ determines both the sampling ratio - $\frac{K}{M}$ and the sparsity - S of the signal in the transform domain (due to asymptotic sparsity [2]). In Figure 4.8, we show three discrete trajectories with different path densities projected into uniform lattices of increasing sampling periods. As can be seen from the figure, increasing the



Figure 4.7: A list of objects acquired with our proposed framework.

sampling period Δ_{rs} results in an increase² in the sampling ratio $\frac{K}{M}$.

The success of compressive-sensing based reconstruction requires the following:

- the Restricted Isometry Principle is satisfied which ensures that linear measurements of sparse signals are nearly orthogonal;
- the signal has a sparse representation (or compressible) in the selected transform domain;
- the number of samples K (and therefore, the sampling ratio $\frac{K}{M}$) is such that Equation (4.7) is satisfied.

Hence we expect the reconstruction to succeed when K is larger. In Figure 4.9, the reconstruction accuracy for trajectories with increasing path densities is shown for varying sampling periods. It can be seen that the SNR grows with an increase in both path density and the sampling period, as both tend to increase the sampling ratio.

Although better reconstructions can be obtained by choosing larger Δ_{rs} , a larger sampling interval also results in a smaller resolution in the illumination plane. As can be seen in Figure 4.9 the reconstruction SNR is higher than $20dB$, even with small Δ_{rs} , as long as the path density is not too low. In general, if the user chooses a trajectory with a very low path density, we could increase reconstruction accuracy by trading off signal resolution, whereas a high resolution reconstruction with reasonable accuracy can be acquired when the user chooses a trajectory with

2. Note that, even when the sampling ratio $\frac{K}{M} = 1$, as long as we sample along trajectories, the projection $\tilde{\mathbf{t}}(m, n)$ differs from a uniformly sampled LTF $\mathbf{t}(m, n)$. When $\frac{K}{M} = 1$, $\tilde{\mathbf{t}}(m, n)$ approaches $\mathbf{t}(m, n)$ as the path density approaches 1 (the trajectory is dense).

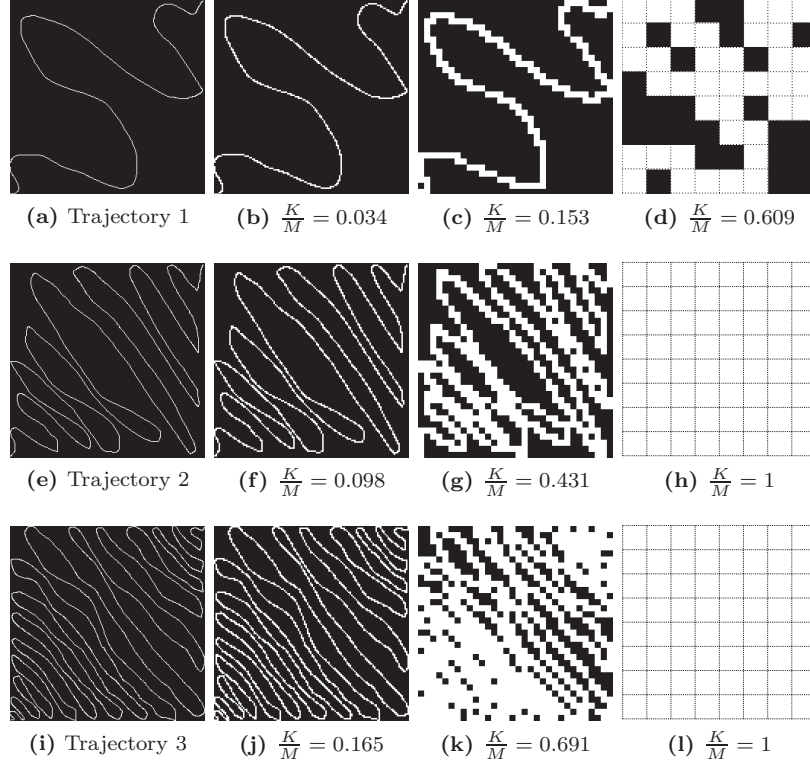


Figure 4.8: Effect of path density of γ_r and Δ_r on the sampling ratio: $\frac{K}{M}$. The first column in this figure shows three different discrete trajectories γ_r in subfigures (a)(e) and (i) with increasing path densities, 0.012, 0.028 and 0.06, respectively. The next three columns represent projections of these trajectories onto uniform lattices of increasing sampling periods: (b), (f), (j) have $2\Delta_r$, (c), (g), (k) have $8\Delta_r$ while (d), (h), (l) have $32\Delta_r$. The corresponding sampling ratios are shown in the respective subcaptions.

better path density. The user interface can be modified to indicate this tradeoff, enabling the user to choose the trajectory accordingly.

4.5.4 Comparison with Other Methods

We now study the improvement in reconstruction with increasing path densities and compare the presented approach with other interpolation strategies.

Bicubic Interpolation

We can use bicubic interpolation to solve for the reconstruction problem as was done for light fields in [13]. Since the samples are irregularly scattered, we can proceed by first doing a Delaunay Triangulation to establish the value of neighbouring samples and then proceed with conventional bicubic interpolation.

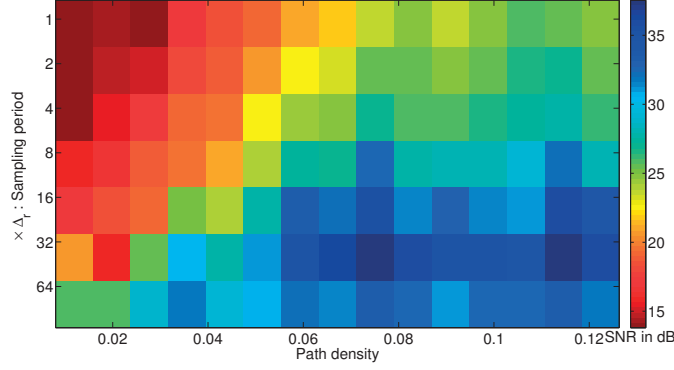


Figure 4.9: Choice of sampling period. We plot the SNR of compressive-sensing based reconstructions for various path densities (along the x-axis) and increasing sampling periods (along the y-axis). As expected, the reconstructions are better for increasing sampling periods and path densities. The results shown are the average SNR of 50 experiments.

Local Polynomial Approximation

In this formulation, we start with the assumption that each vectorized light transport function \mathbf{t} is an analytic function. An analytic function can be approximated locally by a power series. If we now limit the series to P terms, the light transport function can be approximately written as

$$\mathbf{t}(m) \simeq \sum_{j=0}^P a_j (m - m_o)^j$$

for m in the vicinity of m_o . If there are M_k samples that have been observed in the vicinity of m_o , then

$$\mathbf{t}(m_x) \simeq \sum_{p=0}^P a_j (m_x - m_o)^p \quad \forall x \in [1 \cdots M_k].$$

When the above system of equations is solved to obtain $\mathbf{a} = [a_1, a_2 \cdots a_J]$ we can get the P^{th} order polynomial approximation at m_o . If this above system is then solved at every $m_o = [1, \cdots M]$, we can recover the light transport function \mathbf{t} . Thus each light transport function is obtained by solving for M sets of polynomials. A weighted version of the system can be recast as a kernel regression [65] problem.

Low Rank Matrix Completion

It has been previously shown that low-rank approximations of the light transport matrix still preserve key light transport phenomena [73][52]. These methods use specialized setups with two pairs of cameras and projectors in co-axial configuration to directly acquire the low rank approximation of the LTM. However, our hand-held acquisition framework results in a sparsely sampled LTM. We therefore look into matrix completion algorithms instead. Consider a matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$ whose entries $(m, n) \in \Omega$ have been observed. We are interested in finding a matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ with the smallest rank while simultaneously satisfying $\mathbf{C}_{mn} = \mathbf{D}_{mn} \quad \forall (i, j) \in \Omega$.

It has been shown that the following problem is the tightest convex relaxation of the rank minimization problem stated above

$$\tilde{\mathbf{C}} = \arg \min_{\mathbf{C}} \|\mathbf{C}\|_* \text{ subject to } \mathbf{C}_{mn} = \mathbf{D}_{mn},$$

where $\|\cdot\|_*$ is the nuclear norm operator [7], defined as $\|\mathbf{C}\|_* = \sum_p \sigma_p$, with σ_p being the p^{th} singular value of \mathbf{C} . It has been shown that when the number of entries in Ω is large enough, and when the observed entries are distributed uniformly in the matrix, a low rank matrix can be exactly recovered. When the original matrix \mathbf{D} is not low rank, a low rank approximation can be obtained with an upper bound on the error. In our case, though the LTM as a whole is incomplete, with the observed entries randomly distributed along the columns, there is no randomness in the sampling of rows. We therefore look at the individual light transport functions $\mathbf{t}_{ij}(m, n)$. We observe that the entries are now located at random (along Γ). We can thus solve for the following minimization problem

$$\min_{\mathbf{t}_{ij}} \|\hat{\mathbf{t}}_{ij}\|_* \text{ subject to } \hat{\mathbf{t}}_{ij}(m, n) = \tilde{\mathbf{t}}_{ij}(m, n) \quad \forall (m, n) \in \Gamma$$

to obtain a low rank approximation of each light transport function.

Evaluation with Ray-Traced data

To compare the performance of each reconstruction strategy, we first rendered a scene (shown in Figure 4.13) with involved surface geometry by using PovRay, a ray tracing software. We thus obtained the uniformly sampled LTM \mathbf{T} . To simulate sampling as will be in the real-world we chose various trajectories of increasing path density. Shown in Figure 4.10 is the average performance of various approaches discussed for 100 randomly chosen illumination functions from the ray-traced data. Note that when the trajectory taken by the user is sparse, the spatial-interpolation operators fail completely, whereas the compressive-sensing based approach consistently yields a relatively good performance. These performances are for sampling along discrete trajectories, whereas for the purely theoretical case of ‘uniformly random’ samples, all presented algorithms perform significantly better. When the user samples along a dense trajectory, we get good reconstructions, irrespective of the choice of the reconstruction algorithm.

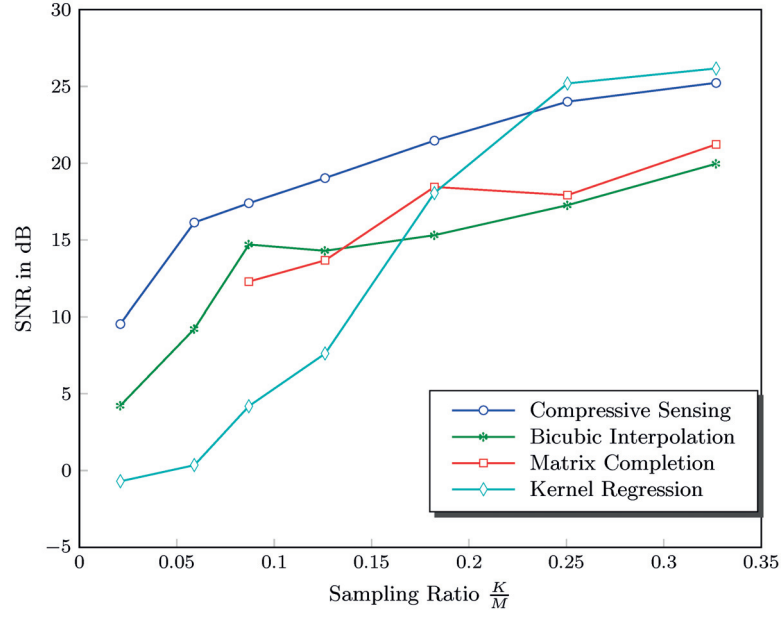


Figure 4.10: Comparison of various reconstruction strategies for increasing path densities.

4.5.5 Error Analysis

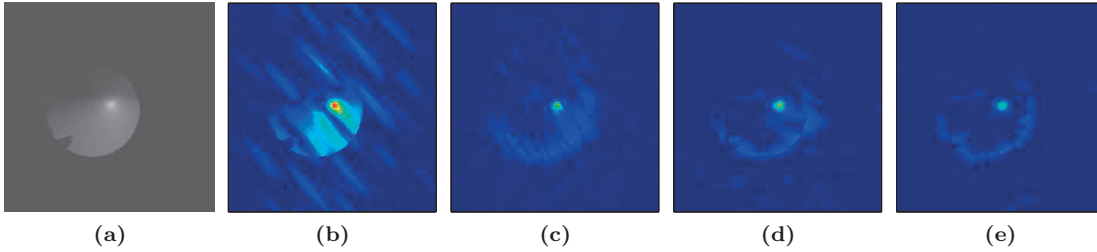


Figure 4.11: The decaying error caused by truncation in the Fourier domain with compressive-sensing. (a) Ground truth LTF exhibiting both a specularity and hard shadows. (b)-(e) Absolute error, for increasing sampling ratios 0.04, 0.08, 0.1, 0.12. SNR increases as 14.08dB, 22.07dB, 22.96dB and 25.33dB, respectively. Note that as the path density is increased, resulting in a higher sampling ratio, the reconstruction error around the discontinuities decreases.

In practice, LTFs seldom have a sparse representation in the Fourier domain. Mirror-like specularities and hard shadows result in discontinuities in LTFs, thus making them non-bandlimited. The compressive-sensing reconstruction therefore recovers the largest Fourier domain coefficients by truncating their Fourier transform. This results in ringing artifacts around these discontinuities due to the Gibbs phenomenon. The variance of these ringing artifacts can however be reduced by increasing the number of recovered components, as is known from classic Fourier

theory. In Figure 4.11, we demonstrate this phenomenon for a single LTF (Figure 4.11(a)) that has discontinuities due to both specularity (at the center) and hard shadows. In Figure 4.11(b), we show the absolute error in reconstruction, for a sampling ratio of 0.04. When this value is increased to 0.08, the reconstruction substantially improves, as seen in Figure 4.11(c). Most of the error energy is now concentrated around the specularity and the border at the shadow. Further increasing the sampling ratio to 0.1 and 0.12 results in higher a SNR and lower error energies around the discontinuities as shown in Figures 4.11(d) and (e).

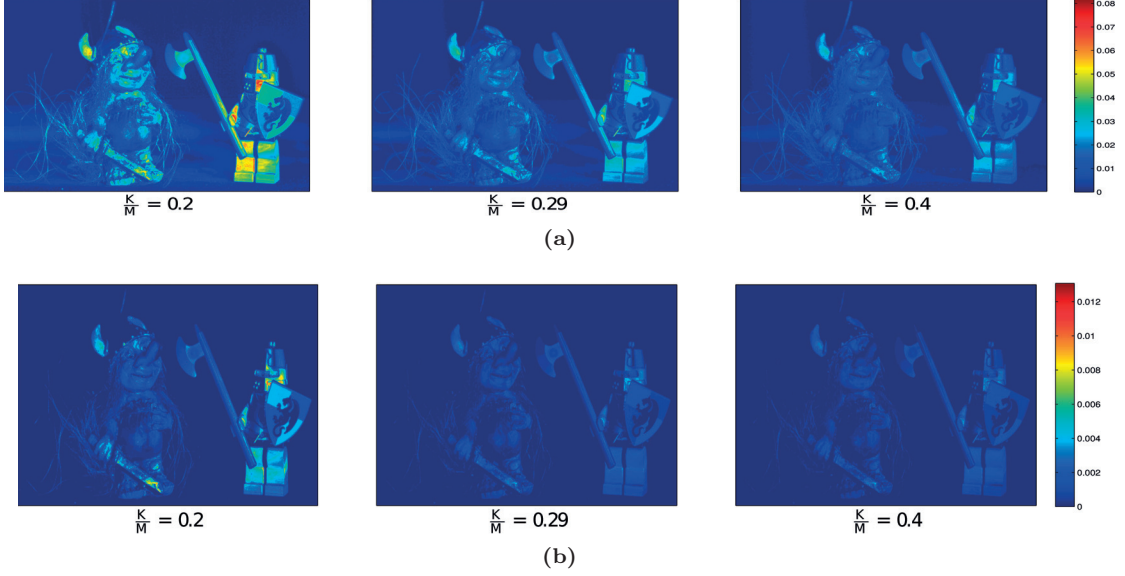


Figure 4.12: Per-pixel mean and variance of absolute error in the reconstruction. (a) The per-pixel mean of the absolute errors from the *Toys* dataset is shown for various trajectories. It can be seen that the mean error falls with increasing sampling ratios. (b) The per-pixel variance of the absolute errors from the *Toys* dataset is shown. It can be seen that the largest error variance occurs around regions exhibiting discontinuities in the form of specularities or hard shadows in their light transport.

The dataset *Toys* consists of large regions with both specularities and hard shadows. In Figure 4.12(a), the per-pixel mean of the absolute error on the luminance channel of the *Toys* dataset, given by $\frac{\langle |\mathbf{T} - \hat{\mathbf{T}}|, 1 \rangle}{M}$ obtained for three different simulated trajectories with sampling ratios $\frac{K}{M}$ being 0.2, 0.29, 0.4, from left to right, is shown. Figure 4.12(b) shows the per-pixel variance of the absolute error $|\mathbf{T} - \hat{\mathbf{T}}|$ for the three simulated trajectories. As can be seen, both the per-pixel mean and variance of the absolute error decrease with increasing path density (and so, increasing $\frac{K}{M}$). The regions in the scene with a large error variance all have a large frequency support in their light transport function; for example, the sword of the troll and the helmet of the soldier are specular. The knee of the soldier has hard shadows ‘cast’ upon it. The horn of the troll and the shield of the soldier have hard shadows ‘attached’ to them. The SNR_y of the entire LTM are 17.899dB, 20.99dB and 21.99dB, respectively, thus indicating a better reconstruction with increasing path densities.

4.5.6 Scene Relighting

Given the light transport matrix \mathbf{T} , the scene can be described under any combination of point light sources in the illumination plane, using the well known image formation equation:

$$\mathbf{y} = \mathbf{T}\boldsymbol{\ell},$$

where $\mathbf{y} \in \mathbb{R}^{J \times 1}$ is the observed (vectorized) image, and $\boldsymbol{\ell} \in \mathbb{R}^{M \times 1}$ is the (vectorized) illumination configuration. Alternatively, the scene can also be relit by recombining rays, as presented in our two plane parametrization in Section 4.3. It can be relit with directional light sources using Equation (4.1). The scene can also be relit under point light sources located outside the illumination plane by choosing incident directions that converge at the desired point in space using Equation (4.3).

We now demonstrate image based relighting with the two plane parametrization. We first described an almost planar scene in PovRay, a ray tracing software and obtained a light transport matrix, by moving a point light source along a rectangle R in a plane defined 10 meters from the scene and parallel to it. Using this light transport matrix, we relight the scene to simulate point light sources located at a second plane 2.5 meters below the original illumination plane by choosing incident light rays as described by Equation (4.3). We compare the images thus rendered via image based rendering to those rendered with raytracing, in Figure 4.13. As expected hard shadows in figure 4.13(b) (ray traced) have been projected as soft shadows in Figure 4.13(a) (our method).

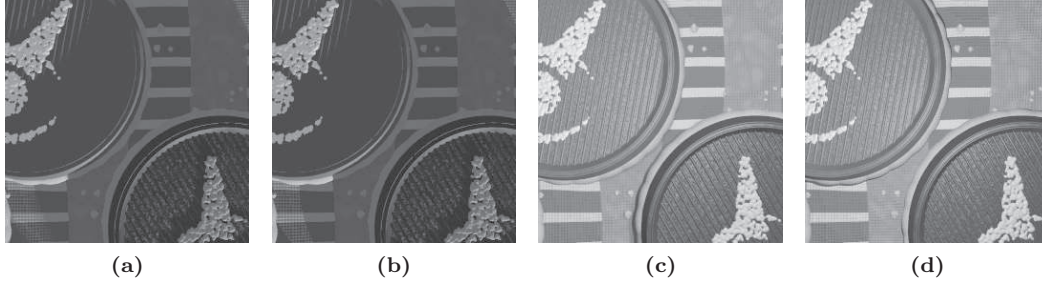


Figure 4.13: Image based relighting with point light sources, compared with images rendered with PovRay. A uniformly sampled light transport matrix was first obtained by fixing the illumination plane at 10 meters from a sample scene in PovRay. (a) and (c) The scene being relit with novel point light source positions by resampling the acquired light transport such that the incident rays converge at points A and B, both located 2.5 meters below the illumination plane. (b) and (d) The same scene rendered with PovRay by placing point light sources at points A and B. Notice that as a result of sampling and interpolation filters, hard shadows in the scene have been rendered as soft shadows in (a).

4.5.7 Mobile Based Acquisition

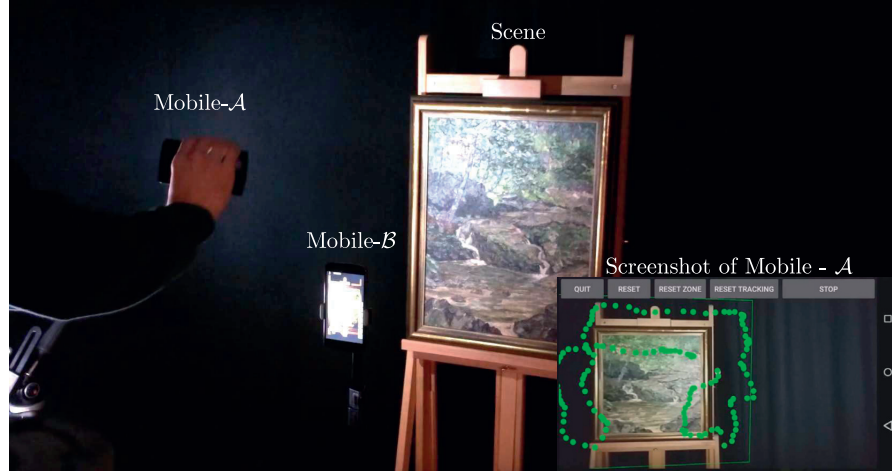


Figure 4.14: Our acquisition system in action: Mobile \mathcal{A} functions as a moving light source while mobile \mathcal{B} records a video of the scene. A screenshot of mobile \mathcal{A} displaying the user's trajectory is shown in inset.

We also tested our reconstruction algorithm using data acquired from our handheld acquisition system implemented on Android smartphones. We obtain the LTMs of three oil paintings and a very specular gold-plated iconograph. The examples presented here contain large amounts of specularities and our algorithm consistently recovers them. Experimental results are shown in Figure 4.17 to 4.20. We also have a simple web-based gallery in [1]

4.5.8 6-dimensional light transport Matrices

We initially assumed a fixed viewing direction, although a captivating interactive rendering requires both changing viewpoints and dynamic illumination. Our reconstruction strategy can be directly extended to incorporate a moving camera and a light source. We simulated an experiment where both mobiles \mathcal{A} and \mathcal{B} are moved, to capture the 6D-light transfer, $\mathcal{T}(r, s, x, y, p, q)$ where (p, q) denotes the plane of movement of mobile \mathcal{B} . Discretizing $\mathcal{T}(r, s, x, y, p, q)$ by extending the framework presented in this chapter, we obtain the 6D LTM, which allows for image-based rendering with both novel viewpoints and illumination configurations. This 6D signal can be interpreted either as an array of light fields, with each light field observed under a different illumination or as an array of LTMs, with each LTM observed under a new viewpoint. With our compressive-sensing based reconstruction, for a sampling ratio of 0.1656, we obtain an SNR of 20.8dB compared to ray traced ground truth data. In Figure 4.21, we show the per-pixel light transfer functions, where the x-axis denotes the light positions and the y-axis denotes the camera positions.

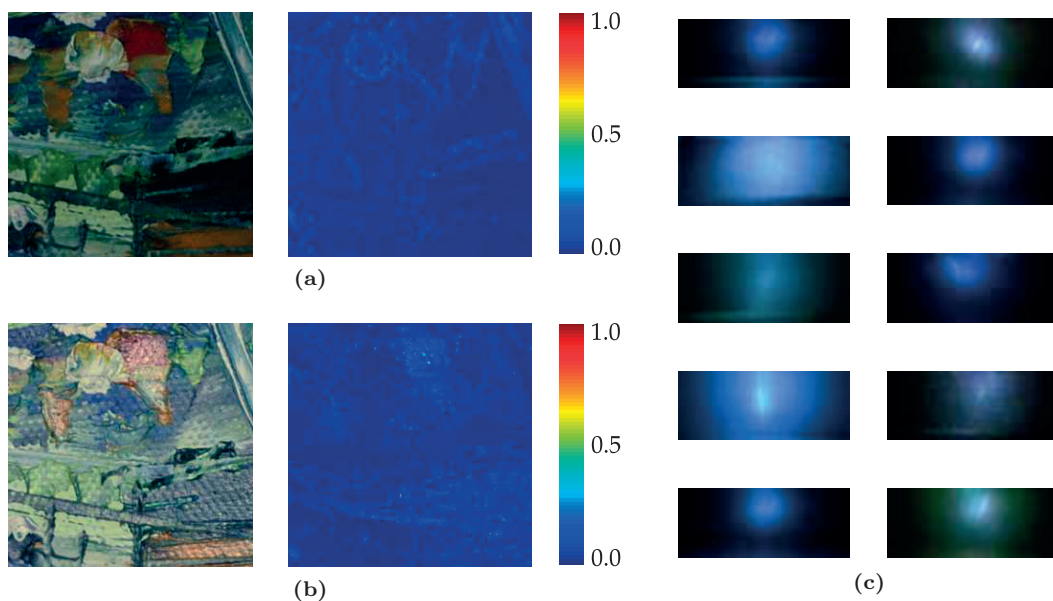


Figure 4.15: Dataset *Notre Dame-a*: (a) and (b) Rendered images under novel light source positions and the absolute error images are shown side by side. (c) Some recovered LTFs exhibiting interesting specularities are shown.

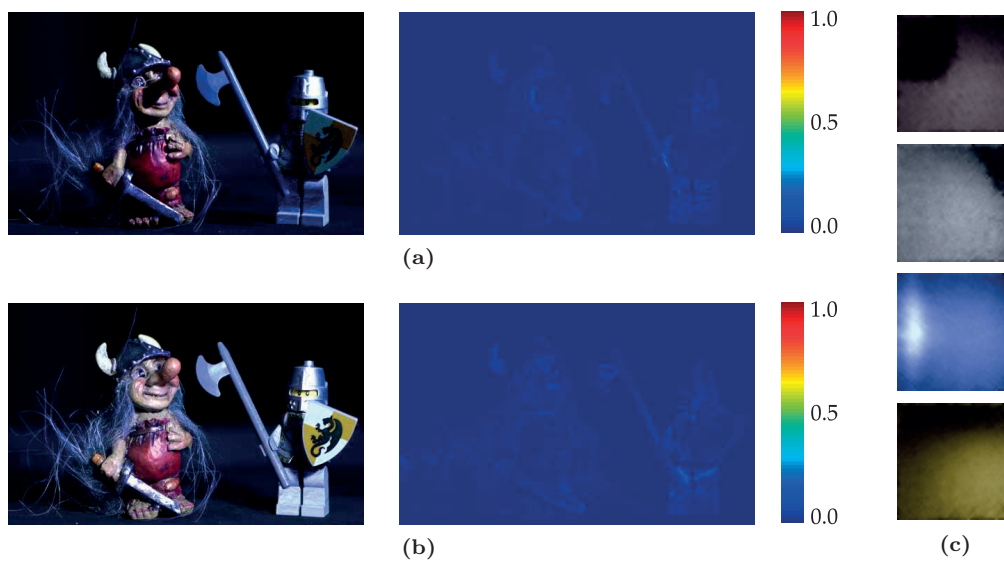


Figure 4.16: Dataset *Toys*: (a) and (b) Rendered images under novel light source positions and the absolute error images are shown side by side. (c) Some recovered LTFs exhibiting specularities and hard shadows are shown.

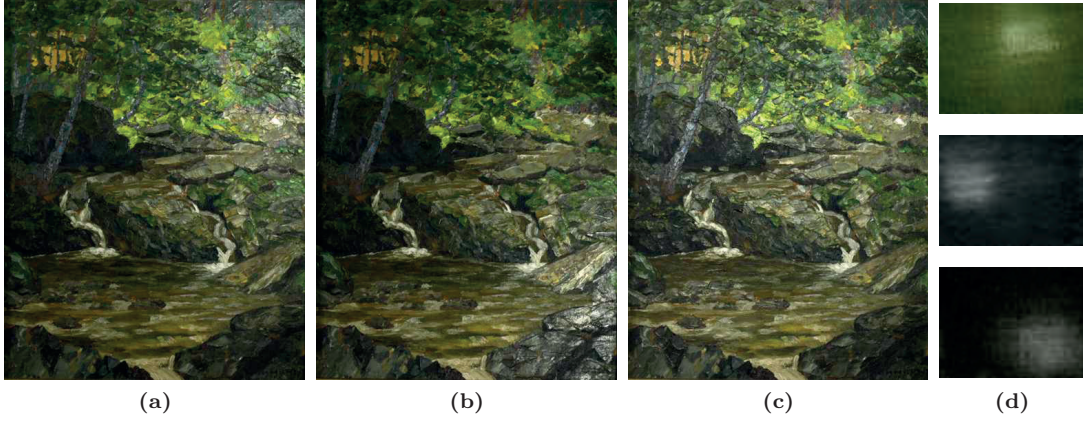


Figure 4.17: (a)-(c) Relighted images of an oil painting by Hohmann under three novel light positions. (d) Some randomly selected light transport functions exhibiting specular structures.

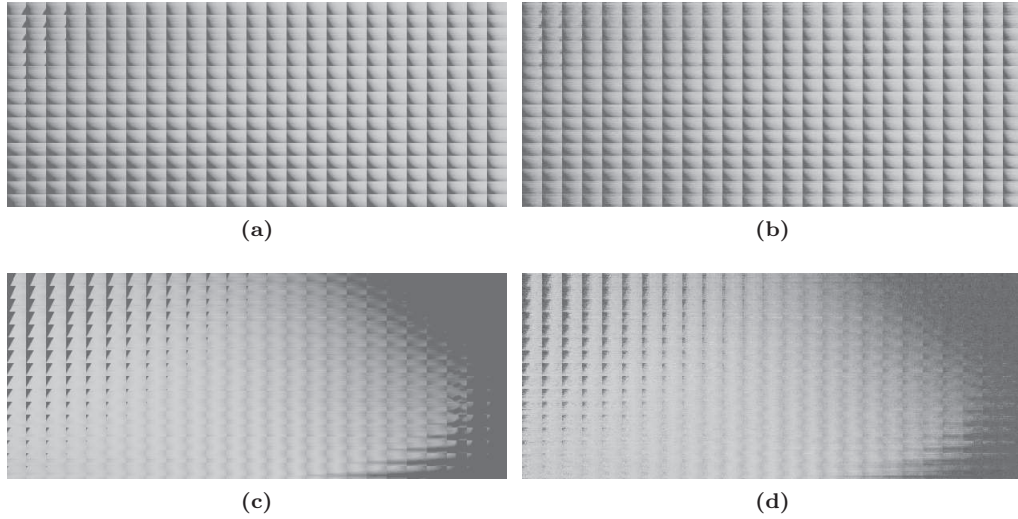


Figure 4.21: Per pixel light transfer where the x-axis denotes light positions and the y-axis denotes camera positions. (a) and (c) Ray Traced signals(Ground Truth). (b) and (d) Reconstruction results obtained with compressive-sensing. $\text{SNR} = 23.1\text{dB}$ and 24.2dB respectively

4.5.9 Discussion

We have presented and evaluated a handheld acquisition and reconstruction strategy for image-based relighting, by reconstructing the LTM of the scene from non-uniformly sampled data.

Our acquisition stage is fast and requires only a few minutes. Since we essentially solve for Equation (4.6) for J pixels, the reconstruction algorithm is a very good candidate for a parallel



Figure 4.18: We acquired a highly specular iconograph that has a gold plating, using our hand held acquisition strategy. Subfigures (a)-(c) show relighted images for three different light positions. Subfigure (d) shows some randomly selected LTFs exhibiting specular structures.

implementation on modern general purpose GPU platforms. This could also be used to design an adaptive algorithm with a varying sampling rate to capture high-frequency components.

During the acquisition of datasets *Notre Dame-a* and *Notre Dame-b*, the mobile \mathcal{A} occluded \mathcal{B} for a few contiguous set of light source positions. Our algorithm successfully fills these occlusions when the simulated path does not contain them.

One practical difficulty we encountered with a completely mobile-based approach is that videos obtained on mobile devices have significant compression artefacts. This in turn can affect our reconstruction algorithm, as it adversely affects the quality of images obtained on mobile \mathcal{B} . However, we see this as a transitional problem and with the evolution of more efficient mobile devices and compression algorithms, this problem will most probably disappear.

4.6 Summary

We have presented a novel portable, intuitive, mobile-based acquisition system for sampling the LTM of real-world objects and a reconstruction framework to interpolate from spatially scattered samples. Unlike earlier methods, our approach does not require any specialized devices in order to sample the reflectance functions and still preserves specularities in the acquired data. Our reconstruction framework is flexible and enables for the resolution of reconstructed LTFs to be traded off with the sampling density with which the trajectories were drawn by the user. We believe this framework provides a feasible approach for everyday digitization and interactive rendering of real-world objects.

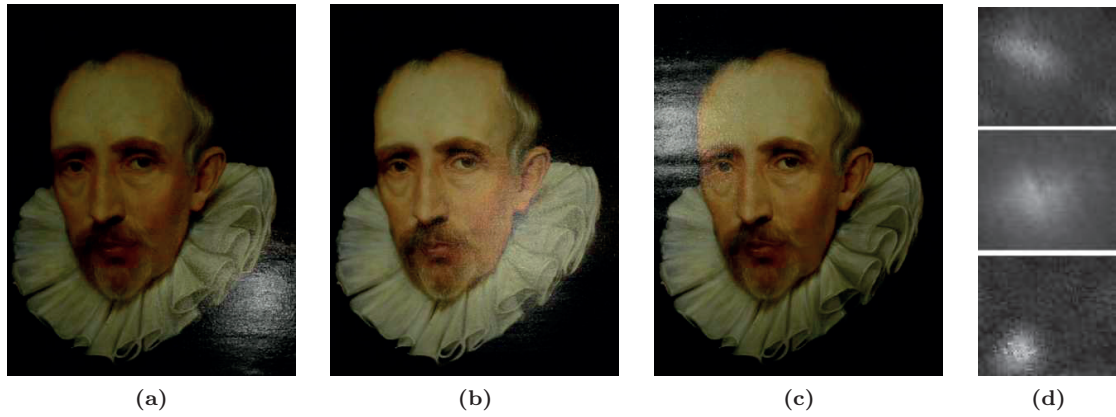


Figure 4.19: We acquired a highly specular oil painting replica of the potrait of Cornelis van der Geest, using our hand held acquisition strategy. Subfigures (a)-(c) show relighted images for three different light positions. Subfigure (d) shows some randomly selected LTFs exhibiting specular structures.

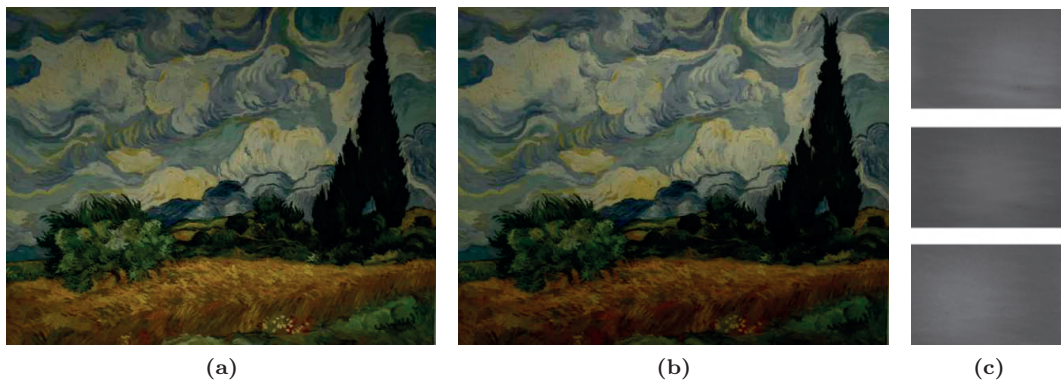


Figure 4.20: We acquired a replica of Van Gogh's 'Wheat Field with Cypresses', using our hand held acquisition strategy. Subfigures (a),(b) show relighted images for two different light positions. Subfigure (c) shows some randomly selected largely diffuse LTFs.

Chapter 5

Conclusion

I may not have gone where I intended to go,
but I think I have ended up where I needed
to be.

*Douglas Adams, The Long Dark Tea-Time
of the Soul*

In this thesis we introduced novel image based approaches for the digitization and visualization of cultural artifacts. Specifically, we exploit the low dimensionality of light transport, to either speed up the acquisition of the light transport matrices of real world objects or to synthesize new objects. Let us now summarize each chapter and discuss future directions in which the presented work can be extended.

Inverse Rendering of Stained glass: We introduced an inverse rendering framework for stained glass artifacts by studying their light transport under spatially varying light. We take advantage of the banded nature of the LTM to find a novel representation for light transport in a sparsity inducing dictionary that can be learnt for individual types of stained glass. We then use this learnt dictionary and the banded nature of light transport to speed up the acquisition process under controlled illuminations. The learnt dictionary also serves as a regularizer for acquisition under uncontrolled illuminations.

The inverse rendering methods presented in this thesis have resulted in a fruitful collaboration with researchers at the Vitrocentre, Romont, Switzerland. The focus has so far been on utilizing the acquired light transport matrix to study stained glass windows as a dynamic artform, whose appearance constantly evolves. This includes investigating the possibility that stained glass artists of the past used the then existing knowledge of varying transmittances in different types of stained glass slabs as a tool in storytelling. The inverse rendering approach presented in the thesis will also be used as a tool for archiving and publishing exhibits in an increasingly digital world.

In future, we should consider the following problem:

- Acquiring the light transport of glass under spatially varying light allows us to characterize the heterogeneous nature of stained glass windows. It is also interesting to characterize the effects of directionally varying incident illumination. Our acquisition setup can be directly extended to sample the 4D incident illumination by moving the projector along a hemisphere. Alternatively, the stained glass sample can be illuminated with a single projector, while being observed by multiple cameras distributed uniformly on a hemisphere, each of which samples the light transport along a particular outgoing direction. Now, by exploiting Helmholtz reciprocity, we can infer the light transport for the dual setup that has one camera and multiple projectors located on a hemisphere. We can then employ the method presented in Chapter 2 to obtain a sparsity inducing dictionary of 4D light transport functions, which can then be used for a complete characterization of light transport in stained glass. Instead of an entirely image based approach, a hybrid approach with a parametric BTDF characterizing the response of directionally varying light and the dictionary characterizing spatially varying light can also be pursued.

Synthetic Rendering of Stained glass: We introduced an image-priors based method for synthetic rendering of stained glass windows that requires only the material dictionary in the form of image priors and the surface albedo of the required stained glass slab. We pose LTM synthesis as an iterative matrix completion problem, with the light transport at each pixel constrained by its physical properties, such as Helmholtz reciprocity, bandedness, sparsity in the known dictionary and non-negativity. Using the presented algorithm, we show examples of synthetic rendering of stained glass from a single image, as well as digital restoration of fractured stained glass windows.

In the future, we should consider the following extensions:

- While we addressed digital restoration of fractured stained glass windows, future extensions should also address restoration of stained glass corroded by environmental factors using priors learnt on well preserved stained glass windows.
- Similar to inverse rendering, a hybrid model that uses a parametric BTDF for directional light and the heterogeneous LTM for volumetric scattering can be considered for synthetic rendering of stained glass windows.

Handheld acquisition of Light Transport: We turned our attention to more general reflective scenes and proposed an acquisition system that is aimed at widening the reach of relightable photographs outside specialized studios. We use two mobile phones; one is stationary and obtains images of the scene while the other is hand-held and moved along an arbitrary trajectory to sample the reflectance function. We re-sample the acquired data to project it onto a uniform lattice. Given this sparsely sampled light transport matrix, we reconstruct an approximation to the ideal, uniformly sampled LTM through compressive sensing. Furthermore we extend the approach to enable interactive rendering with both moving viewpoints and illumination directions, by moving both mobile phones during acquisition. We also present an image based relighting system that uses the two plane parametrization in order to render the scene under arbitrary incident illuminations contained within the illumination plane.

In the future, we should consider the following extensions:

- In the LTM reconstruction problem in Chapter 4, we treat the light transport functions at individual pixels separately. This can be extended by taking advantage of the local

coherence inherent in the light transport of general scenes. For example, we could first segment the image plane into multiple regions, both based on the scene texture and the observed samples. We could then jointly estimate the light transport functions in each region of the scene.

- Although LTFs satisfy our assumption of low dimensionality in the frequency domain, light fields can be better parametrized by straight lines of varying magnitudes and slopes in the epipolar plane. Hence, for the 6D light transport, we could explore a reconstruction strategy that exploits such a parametric form.

With recent advances in Virtual Reality related consumer hardware, users can enjoy immersive visualizations of objects in the virtual world. The inverse rendering framework introduced in this thesis can aid in practical methods for acquisition and modelling real-world objects for interactive rendering and integration into the virtual world.

Bibliography

- [1] “A web-based gallery for relightable photographs,” (Date last accessed 15-Dec-2015). [Online]. Available: <http://go.epfl.ch/relightables>.
- [2] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, “Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing,” *CoRR*, vol. abs/1302.0561, 2013. [Online]. Available: <http://arxiv.org/abs/1302.0561>.
- [3] M. Ashikhmin and P. Shirley, “An anisotropic phong brdf model,” *J. Graph. Tools*, vol. 5, no. 2, pp. 25–32, Feb. 2000. [Online]. Available: <http://dx.doi.org/10.1080/10867651.2000.10487522>.
- [4] P. N. Belhumeur and D. J. Kriegman, “What is the set of images of an object under all possible illumination conditions?” *Int. J. Comput. Vision*, vol. 28, no. 3, pp. 245–260, July 1998. [Online]. Available: <http://dx.doi.org/10.1023/A:1008005721484>.
- [5] J. F. Blinn, “Models of light reflection for computer synthesized pictures,” *SIGGRAPH Comput. Graph.*, vol. 11, no. 2, pp. 192–198, July 1977. [Online]. Available: <http://doi.acm.org/10.1145/965141.563893>.
- [6] A. M. Bruckstein, D. L. Donoho, and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009. [Online]. Available: <http://dx.doi.org/10.1137/060657704>.
- [7] J.-F. Cai, E. J. Candès, and Z. Shen, “A singular value thresholding algorithm for matrix completion,” *SIAM J. on Optimization*, vol. 20, no. 4, pp. 1956–1982, Mar. 2010. [Online]. Available: <http://dx.doi.org/10.1137/080738970>.
- [8] E. J. Candès, “Compressive sampling,” in *Proc. Proceedings of the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures*, pp. 1433–1452, 2006.
- [9] S. Chandrasekar, “Radiative transfer.” Oxford Univ. Press., 1960.
- [10] R. L. Cook and K. E. Torrance, “A reflectance model for computer graphics,” *ACM Trans. Graph.*, vol. 1, no. 1, pp. 7–24, Jan. 1982. [Online]. Available: <http://doi.acm.org/10.1145/357290.357293>.
- [11] Q. Dai, J. Wang, Y. Liu, J. Snyder, E. Wu, and B. Guo, “The Dual-microfacet Model for Capturing Thin Transparent Slabs,” *Computer Graphics Forum*, 2009.
- [12] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, “Reflectance and texture of real-world surfaces,” *ACM TRANS. GRAPHICS*, pp. 730–301, 1999.

-
- [13] A. Davis, M. Levoy, and F. Durand, “Unstructured light fields,” *Comp. Graph. Forum*, vol. 31, no. 2pt1, pp. 305–314, May 2012. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-8659.2012.03009.x>.
 - [14] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, “Acquiring the reflectance field of a human face,” in *Proc. Proceedings of the 27th annual conference on Computer graphics and interactive techniques, SIGGRAPH '00*, 2000.
 - [15] P. Debevec, A. Wenger, C. Tchou, A. Gardner, J. Waese, and T. Hawkins, “A lighting reproduction approach to live-action compositing,” *ACM Trans. Graph.*, vol. 21, 2002.
 - [16] Y. Dong, J. Wang, X. Tong, J. Snyder, Y. Lan, M. Ben-Ezra, and B. Guo, “Manifold Bootstrapping for SVBRDF Capture,” in *Proc. ACM SIGGRAPH 2010 Papers, SIGGRAPH '10*, pp. 98:1–98:10. New York, NY, USA: ACM, 2010. [Online]. Available: <http://doi.acm.org/10.1145/1833349.1778835>.
 - [17] C. Donner and H. W. Jensen, “Light diffusion in multi-layered translucent materials,” *ACM Trans. Graph.*, vol. 24, no. 3, pp. 1032–1039, July 2005. [Online]. Available: <http://doi.acm.org/10.1145/1073204.1073308>.
 - [18] —, “A spectral BSSRDF for shading human skin,” in *Proc. Proceedings of the 17th Eurographics Conference on Rendering Techniques, EGSR '06*, pp. 409–417. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2006. [Online]. Available: <http://dx.doi.org/10.2312/EGWR/EGSR06/409-417>.
 - [19] C. Donner, J. Lawrence, R. Ramamoorthi, T. Hachisuka, H. W. Jensen, and S. Nayar, “An empirical bssrdf model,” *ACM Trans. Graph.*, vol. 28, no. 3, pp. 30:1–30:10, July 2009. [Online]. Available: <http://doi.acm.org/10.1145/1531326.1531336>.
 - [20] C. Donner, T. Weyrich, E. d'Eon, R. Ramamoorthi, and S. Rusinkiewicz, “A layered, heterogeneous reflectance model for acquiring and rendering human skin,” in *Proc. ACM SIGGRAPH Asia 2008 Papers, SIGGRAPH Asia '08*, pp. 140:1–140:12. New York, NY, USA: ACM, 2008. [Online]. Available: <http://doi.acm.org/10.1145/1457515.1409093>.
 - [21] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, 2006.
 - [22] Q. Duan, J. Cai, J. Zheng, and W. Lin, “Fast environment matting extraction using compressive sensing,” in *Proc. Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pp. 1–6, 2011.
 - [23] A. Gardner, C. Tchou, T. Hawkins, and P. Debevec, “Linear light source reflectometry,” in *Proc. In Proceedings of SIGGRAPH 2003, Computer Graphics Proceedings, Annual Conference Series*, 2003.
 - [24] A. Gardner, C. Tchou, A. Wenger, P. Debevec, and T. Hawkins, “Postproduction re-illumination of live action using interleaved lighting,” in *Proc. ACM SIGGRAPH 2004 Posters, SIGGRAPH '04*, 2004.
 - [25] R. Garg, H. Du, S. Seitz, and N. Snavely, “The dimensionality of scene appearance,” in *Proc. Computer Vision, 2009 IEEE 12th International Conference on*, pp. 1917–1924, 2009.

-
- [26] A. Ghosh, T. Chen, P. Peers, C. A. Wilson, and P. Debevec, "Estimating Specular Roughness and Anisotropy from Second Order Spherical Gradient Illumination," in *Proc. Computer Graphics Forum*, vol. 28, p. 4, June 2009.
- [27] A. Ghosh, W. Heidrich, S. Achutha, and M. O'Toole, "A Basis Illumination Approach to BRDF Measurement," *International Journal on Computer Vision*, vol. 90, no. 2, pp. 183–197, 2010.
- [28] I. Gkioulekas, S. Zhao, K. Bala, T. Zickler, and A. Levin, "Inverse volume rendering with material dictionaries," *ACM Transaction on Graphics (Proc. ACM SIGGRAPH Asia)*, 2013.
- [29] M. Goesele, H. P. A. Lensch, J. Lang, C. Fuchs, and H.-P. Seidel, "Disco: Acquisition of translucent objects," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 835–844, Aug. 2004. [Online]. Available: <http://doi.acm.org/10.1145/1015706.1015807>.
- [30] P. Hanrahan and W. Krueger, "Reflection from layered surfaces due to subsurface scattering," in *Proc. Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pp. 165–174, 1993.
- [31] T. Hawkins, J. Cohen, and P. Debevec, "A photometric approach to digitizing cultural artifacts," in *Proc. Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, VAST '01, 2001.
- [32] X. D. He, K. E. Torrance, F. X. Sillion, and D. P. Greenberg, "A comprehensive physical model for light reflection," in *Proc. Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '91, pp. 175–186. New York, NY, USA: ACM, 1991. [Online]. Available: <http://doi.acm.org/10.1145/122718.122738>.
- [33] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler, "A photometric approach for estimating normals and tangents," in *Proc. ACM SIGGRAPH Asia 2008 Papers*, SIGGRAPH Asia '08, pp. 133:1–133:9. New York, NY, USA: ACM, 2008. [Online]. Available: <http://doi.acm.org/10.1145/1457515.1409086>.
- [34] M. Holroyd, J. Lawrence, and T. Zickler, "A coaxial optical scanner for synchronous acquisition of 3d geometry and surface reflectance," in *Proc. ACM SIGGRAPH 2010 Papers*, SIGGRAPH '10, 2010.
- [35] I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich, "Transparent and specular object reconstruction," *Computer Graphics Forum*, vol. 29, no. 8, pp. 2400–2426, 2010. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-8659.2010.01753.x>.
- [36] H. W. Jensen and J. Buhler, "A rapid hierarchical rendering technique for translucent materials," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 576–581, July 2002. [Online]. Available: <http://doi.acm.org/10.1145/566654.566619>.
- [37] H. W. Jensen, J. Stephen, S. R. Marschner, M. Levoy, and P. Hanrahan, "A practical model for subsurface light transport," in *Proc. Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, 2001.
- [38] J. T. Kajiya, "Anisotropic reflection models," *SIGGRAPH Comput. Graph.*, vol. 19, no. 3, pp. 15–21, July 1985. [Online]. Available: <http://doi.acm.org/10.1145/325165.325167>.

-
- [39] G. Klein and D. Murray, "Parallel tracking and mapping for small ar workspaces," in *Proc. Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR '07*, pp. 1–10. Washington, DC, USA: IEEE Computer Society, 2007. [Online]. Available: <http://dx.doi.org/10.1109/ISMAR.2007.4538852>.
 - [40] B. Lamond, P. Peers, A. Ghosh, and P. Debevec, "Image-based separation of diffuse and specular reflections using environmental structured illumination," in *Proc. Computational Photography (ICCP), 2009 IEEE International Conference on*, pp. 1–8, April 2009.
 - [41] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel, "Image-based reconstruction of spatial appearance and geometric detail," *ACM Trans. Graph.*, vol. 22, no. 2, pp. 234–257, Apr. 2003. [Online]. Available: <http://doi.acm.org/10.1145/636886.636891>.
 - [42] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '96*, pp. 31–42. New York, NY, USA: ACM, 1996. [Online]. Available: <http://doi.acm.org/10.1145/237170.237199>.
 - [43] W.-C. Ma, T. Hawkins, P. Peers, C.-F. Chabert, M. Weiss, and P. Debevec, "Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination," in *Proc. Proceedings of the 18th Eurographics Conference on Rendering Techniques, EGSR'07*, pp. 183–194. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2007. [Online]. Available: <http://dx.doi.org/10.2312/EGWR/EGSR07/183-194>.
 - [44] T. Malzbender, D. Gelb, and H. Wolters, "Polynomial texture maps," in *Proc. In Computer Graphics, SIGGRAPH 2001 Proceedings*, pp. 519–528, 2001.
 - [45] V. Masselus, "A practical framework for fixed viewpoint image based relighting," *PhD Thesis*, 2004.
 - [46] V. Masselus, P. Dutré, and F. Anrys, "The free-form light stage," in *Proc. Proceedings of the 13th Eurographics Workshop on Rendering, EGRW '02*, pp. 247–256. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002. [Online]. Available: <http://dl.acm.org/citation.cfm?id=581896.581928>.
 - [47] V. Masselus, P. Peers, P. Dutré, and Y. D. Willems, "Relighting with 4D incident light fields," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 613–620, 2003. [Online]. Available: <http://doi.acm.org/10.1145/882262.882315>.
 - [48] V. Masselus, P. Peers, P. Dutré, and Y. D. Willemsy, "Smooth Reconstruction and Compact Representation of Reflectance Functions for Image-based Relighting," in *Proc. Eurographics Workshop on Rendering*, A. Keller and H. W. Jensen, eds. The Eurographics Association, 2004.
 - [49] W. Matusik, H. Pfister, R. Ziegler, A. Ngan, and L. McMillan, "Acquisition and rendering of transparent and refractive objects," in *Proc. Proceedings of the 13th Eurographics Workshop on Rendering, EGRW '02*, pp. 267–278. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002. [Online]. Available: <http://dl.acm.org/citation.cfm?id=581896.581931>.

-
- [50] W. Matusik, M. H. Pfister, M. R. Ziegler, M. A. Ngan, L. M. Mit, W. Matusik, H. Pfister, R. Ziegler, A. Ngan, and L. Mcmillan, "Acquisition and rendering of transparent and refractive objects," in Proc. *In Proceedings of Eurographics Symposium on Rendering (EGSR*, pp. 267–278, 2002.
- [51] S. K. Nayar, P. N. Belhumeur, and T. E. Boult, "Lighting sensitive display," *ACM Trans. Graph.*, vol. 23, no. 4, pp. 963–979, Oct. 2004. [Online]. Available: <http://doi.acm.org/10.1145/1027411.1027414>.
- [52] M. O'Toole and K. N. Kutulakos, "Optical computing for fast light transport analysis," *ACM Trans. Graph.*, vol. 29, no. 6, pp. 164:1–164:12, Dec. 2010. [Online]. Available: <http://doi.acm.org/10.1145/1882261.1866165>.
- [53] P. Peers, T. Hawkins, and P. Debevec, "A reflective light stage," *Tech. rep., ICT-USC, ICT-TR-04.2006*, 2007.
- [54] P. Peers, K. Berge, W. Matusik, R. Ramamoorthi, J. Lawrence, S. Rusinkiewicz, and P. Dutré, "A compact factored representation of heterogeneous subsurface scattering," *ACM Transactions on Graphics (SIGGRAPH*, vol. 2006, p. 2006, 2006.
- [55] P. Peers and P. Dutré, "Wavelet environment matting," in Proc. *Proceedings of the 14th Eurographics Workshop on Rendering, EGRW '03*, pp. 157–166. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2003. [Online]. Available: <http://dl.acm.org/citation.cfm?id=882404.882427>.
- [56] P. Peers, D. K. Mahajan, B. Lamond, A. Ghosh, W. Matusik, R. Ramamoorthi, and P. Debevec, "Compressive light transport sensing," *ACM Trans. Graph.*, vol. 28, no. 1, Feb. 2009.
- [57] P. Poulin and A. Fournier, "A model for anisotropic reflection," *SIGGRAPH Comput. Graph.*, vol. 24, no. 4, pp. 273–282, Sept. 1990. [Online]. Available: <http://doi.acm.org/10.1145/97880.97909>.
- [58] D. Reddy, R. Ramamoorthi, and B. Curless, "Frequency-space decomposition and acquisition of light transport under spatially varying illumination," in Proc. *ECCV (6)*, pp. 596–610, 2012.
- [59] P. Ren, J. Wang, J. Snyder, X. Tong, and B. Guo, "Pocket reflectometry," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 45:1–45:10, July 2011. [Online]. Available: <http://doi.acm.org/10.1145/2010324.1964940>.
- [60] Y. Schechner, S. Nayar, and P. Belhumeur, "A theory of multiplexed illumination," in Proc. *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pp. 808–815 vol.2, Oct 2003.
- [61] C. Schlick, "A customizable reflectance model for everyday rendering," in Proc. *In Fourth Eurographics Workshop on Rendering*, pp. 73–83, 1993.
- [62] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch, "Dual photography," in Proc. *ACM SIGGRAPH 2005 Papers, SIGGRAPH '05*, pp. 745–755. New York, NY, USA: ACM, 2005. [Online]. Available: <http://doi.acm.org/10.1145/1186822.1073257>.

-
- [63] J. L. Sloan, “Conservation of Stained Glass in America: A Manual for Studios and Caretakers,” in *Proc. Wilmington: Art in Architecture Press.*, 1993.
 - [64] J. Stam, “An illumination model for a skin layer bounded by rough surfaces,” in *Proc. In Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pp. 39–52. Springer-Verlag, 2001.
 - [65] H. Takeda, S. Farsiu, and P. Milanfar, “Kernel Regression for Image Processing and Reconstruction,” *IEEE Transactions on Image Processing*, vol. 16, pp. 349–366, 2007.
 - [66] N. Thanikachalam, L. Baboulaz, D. Firmenich, S. Susstrunk, and M. Vetterli, “Scene re-lighting with smartphones,” *submitted to IEEE Transactions on Computational Imaging*, 2016.
 - [67] N. Thanikachalam, L. Baboulaz, P. Prandoni, and M. Vetterli, “Light transport matrix recovery for nearly planar objects,” vol. 9023, pp. 90 230Q–90 230Q–7, 2014. [Online]. Available: <http://dx.doi.org/10.1117/12.2038909>.
 - [68] —, “VITRAIL : Acquisition, modelling and rendering of stained glass,” *submitted to IEEE Transactions on Image Processing*, 2015.
 - [69] K. E. Torrance and E. M. Sparrow, “Radiometry,” L. B. Wolff, S. A. Shafer, and G. Healey, eds. USA: Jones and Bartlett Publishers, Inc., 1992, ch. Theory for Off-specular Reflection from Roughened Surfaces, pp. 32–41. [Online]. Available: <http://dl.acm.org/citation.cfm?id=136913.136924>.
 - [70] B. Tunwattanapong, G. Fyffe, P. Graham, J. Busch, X. Yu, A. Ghosh, and P. Debevec, “Acquiring reflectance and shape from continuous spherical harmonic illumination,” *ACM Trans. Graph.*, vol. 32, no. 4, pp. 109:1–109:12, July 2013. [Online]. Available: <http://doi.acm.org/10.1145/2461912.2461944>.
 - [71] J. Unnikrishnan and M. Vetterli, “Sampling and Reconstruction of Spatial Fields using Mobile Sensors,” *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2328–2340, 2013. [Online]. Available: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6463464>.
 - [72] B. Walter, S. R. Marschner, H. Li, and K. E. Torrance, “Microfacet models for refraction through rough surfaces,” in *Proc. Proceedings of the 18th Eurographics Conference on Rendering Techniques*, EGSR’07, pp. 195–206. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2007. [Online]. Available: <http://dx.doi.org/10.2312/EGWR/EGSR07/195-206>.
 - [73] J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo, “Kernel Nystrom method for light transport,” in *Proc. ACM SIGGRAPH 2009 papers*, SIGGRAPH ’09, pp. 29:1–29:10. New York, NY, USA: ACM, 2009. [Online]. Available: <http://doi.acm.org/10.1145/1576246.1531335>.
 - [74] J. Wang, S. Zhao, X. Tong, S. Lin, Z. Lin, Y. Dong, B. Guo, and H.-Y. Shum, “Modeling and rendering of heterogeneous translucent materials using the diffusion equation,” *ACM Trans. Graph.*, vol. 27, no. 1, pp. 9:1–9:18, Mar. 2008. [Online]. Available: <http://doi.acm.org/10.1145/1330511.1330520>.

-
- [75] G. J. Ward, "Measuring and modeling anisotropic reflection," *SIGGRAPH Comput. Graph.*, vol. 26, no. 2, pp. 265–272, July 1992. [Online]. Available: <http://doi.acm.org/10.1145/142920.134078>.
- [76] Y. Wexler, A. W. Fitzgibbon, and A. Zisserman, "Image-based environment matting," in Proc. *Proceedings of the 13th Eurographics Workshop on Rendering*, EGRW '02, pp. 279–290. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002. [Online]. Available: <http://dl.acm.org/citation.cfm?id=581896.581932>.
- [77] T. Weyrich, J. Lawrence, H. P. A. Lensch, S. Rusinkiewicz, and T. Zickler, "Principles of appearance acquisition and representation," *Found. Trends. Comput. Graph. Vis.*, vol. 4, no. 2, pp. 75–191, Feb. 2009. [Online]. Available: <http://dx.doi.org/10.1561/06000000022>.
- [78] T. Zickler, R. Ramamoorthi, S. Enrique, and P. Belhumeur, "Reflectance sharing: predicting appearance from a sparse set of images of a known shape," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 8, pp. 1287–1302, Aug 2006.
- [79] D. E. Zongker, D. M. Werner, B. Curless, and D. H. Salesin, "Environment matting and compositing," in Proc. *Proceedings of ACM SIGGRAPH 99*, Computer Graphics Proceedings, Annual Conference Series, pp. 205–214. ACM Press / ACM SIGGRAPH / Addison Wesley Logman, July 1999.

Curriculum Vitæ

Niranjana Thanikachalam

Audiovisual Communications Laboratory (LCAV)
Swiss Federal Institute of Technology (EPFL)
1015 Lausanne, Switzerland

Email: niranjana.thanikachalam@epfl.ch
Website: http://lcav.epfl.ch/people/Niranjana_Thanikachalam

Personal Information

Date of birth: September, 8, 1989
Nationality: Indian

Education

2011–2016 PhD candidate in School of Computer and Communication Sciences,
Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.
2007–2011 B.E in Electronics and Communication Engineering
Madras Institute of Technology,
Anna University, India.

Professional experience

2011–present **Research and teaching assistant,**
Audiovisual Communications Laboratory (LCAV),
Swiss Federal Institute of Technology (EPFL).

Research Interests

- Computational Photography
- Low level computer vision

Publications

Journal papers

- [2]. **N. Thanikachalam**, L. Baboulaz, D. Firmenich, S. Susstrunk and M. Vetterli. *Scene Relighting with Smartphones*, submitted to IEEE Transactions on Computational Imaging 2016.
- [1]. **N. Thanikachalam**, L. Baboulaz, P. Prandoni and M. Vetterli. *VITRAIL : Acquisition, Modelling and Rendering of Stained Glass*, submitted to IEEE Transactions on Image Processing, 2015.

Conference papers

- [2]. S. Trümpler, L. Baboulaz, **N. Thanikachalam**, M. Vetterli and S. Wolf. *Dynamische Bilder. Neue Forschungsansätze zur Wahrnehmung mittelalterlicher Glasmalereien*, accepted to Dritter Schweizerischer Kongress für Kunstgeschichte, Basel, 2016.
- [1]. **N. Thanikachalam**, L. Baboulaz, P. Prandoni and M. Vetterli. *Light Transport Matrix recovery for nearly planar objects*. SPIE Electronic Imaging, San Francisco, California, USA, Proceedings of SPIE, 2014.

Skills

- Programming:** C++ and Matlab.
Working knowledge of Java/Android, Python, HTML/CSS, JavaScript, shell scripting.
Experienced in L^AT_EX.
- Libraries:** OpenCV, OpenGL, OpenCL, Armadillo C++.

