VITRAIL: Acquisition, Modeling, and Rendering of Stained Glass

Niranjan Thanikachalam, Loïc Baboulaz, Paolo Prandoni, Stefan Trümpler, Sophie Wolf, and Martin Vetterli, *Fellow, IEEE*

***Abstract*— Stained glass windows are designed to reveal their powerful artistry under diverse and time-varying lighting conditions; virtual relighting of stained glass, therefore, represents an exceptional tool for the appreciation of this age old art form. However, as opposed to most other artifacts, stained glass windows are extremely difficult if not impossible to analyze using controlled illumination because of their size and position. In this paper, we present novel methods built upon image based priors to perform virtual relighting of stained glass artwork by acquiring the actual light transport properties of a given artifact. In a preprocessing step, we build a material-dependent dictionary for light transport by studying the scattering properties of glass samples in a laboratory setup. We can now use the dictionary to recover a light transport matrix in two ways: under controlled illuminations the dictionary constitutes a sparsifying basis for a compressive sensing acquisition, while in the case of uncontrolled illuminations the dictionary is used to perform sparse regularization. The proposed basis preserves volume impurities and we show that the retrieved light transport matrix is heterogeneous, as in the case of real world objects. We present the rendering results of several stained glass artifacts, including the Rose Window of the Cathedral of Lausanne, digitized using the presented methods.**

***Index Terms*— Banded matrices, computational relighting, cultural artifacts, dictionary learning, light transport, sparse recovery, stained glass.**

I. 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 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. This problem of acquiring the reflectance or (in our case) the volumetric scattering properties of real-world objects to enable rendering in virtual environments is termed as inverse rendering. The workflow for inverse rendering typically involves an acquisition stage with a multiple illumination setup followed by a modelling stage where the acquired data is used to obtain reflectance models by exploiting prior knowledge on reflectance properties and in the final rendering stage, the object is rendered in virtual environments. This is in contrast with generic rendering, where a virtual object is 'synthesized' from its 3D geometry and known or assumed reflectance functions.

In principle, to render any artifact under arbitrary illuminations and viewpoints a sampled version of its bidirectional scattering distribution function (BSDF) is required (Figure 1). The BSDF, however, is an 8-dimensional function that relates the incident light field to the outgoing light field for each possible entry and exit point and for each entry and exit angle of the light rays; the full acquisition of the BSDF is 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.

Manuscript received June 28, 2015; revised May 18, 2016; accepted June 9, 2016. Date of publication July 11, 2016; date of current version August 3, 2016. This work was supported in part by Google through the Google Focused Research Award and in part by the Swiss National Science Foundation under Grant 20FP-1_151073. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mujdat Cetin.

N. Thanikachalam, L. Baboulaz, P. Prandoni, and M. Vetterli are with the School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland (e-mail: niranjan.thanikachalam@epfl.ch; loic.baboulaz@epfl.ch; paolo.prandoni@epfl.ch; martin.vetterli@epfl.ch).

S. Trümpler and S. Wolf are with Vitrocentre Romont, Romont 1680, Switzerland (e-mail: stefan.truempler@vitrocentre.ch; sophie.wolf@vitrocentre.ch).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the author. The material contains two versions of the same video, a high-resolution video and a low-resolution video. The total size of the videos is 446 MB. Contact niranjan.thanikachalam@epfl.ch for further questions about this work.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2585041

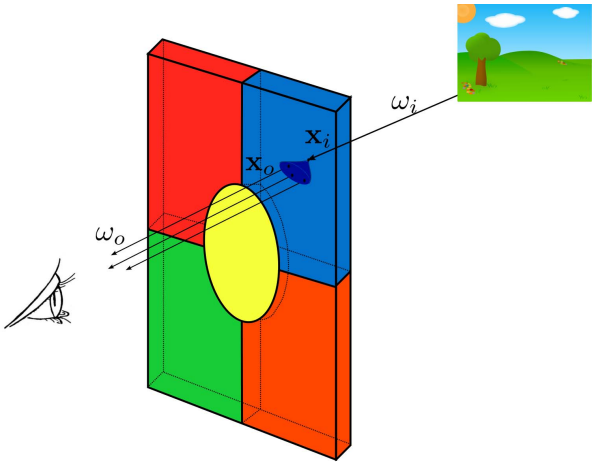


Fig. 1. Light transport in stained glass is a function of the direction and point of entry of incident light and the direction and point of exit of exitant light, as determined by the bidirectional scattering distribution function.

In this paper 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, in the absence of strong directional light. Finally, with our light transport matrix representation, rendering under novel illuminations is obtained with a simple matrix vector product.

The main contributions of this paper are:

- 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.

II. Related Work and Background

The inverse rendering problem under controlled illuminations is a well studied topic in the graphics and vision communities.

A. Inverse Rendering Methods

Image based acquisition methods that take the light stage approach [1]–[4] and polynomial texture mapping (PTM) [5] capture the spatially varying reflectance of a scene sampling with directional light, sometimes exploiting low dimensionality in the spherical harmonics basis. Tunwattanapong *et al.* [6] use an illumination device that projects continuous spherical harmonics in conjunction with multiview stereo to extract the reflectance functions of 3D objects. In these methods, acquired data is then typically used to obtain parametric reflectance functions and can model specularities, diffuse reflections, inter-reflections and soft shadowing to a resolution limited by the angular resolution of the light stage. These methods are however ill-suited for translucent objects as volumetric and subsurface scattering are functions of spatially varying light. Goesele *et al.* [7] 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.

The light transport matrix (LTM) captures the scene response to a spatially varying light source using a projector-camera setup; it can recover both diffuse and specular reflection or refraction phenomena in the scene for the specific incoming and outgoing light directions considered, in addition to subsurface scattering and volume scattering. It was originally developed as environment matting to capture refractions in transparent objects [8], [9]. More recently, to speed up acquisition, several approaches including compressive sensing [10], [11], low rank matrix approximation [12], [13] and spectral decomposition [14] have been proposed for solving the inverse problem. Peers *et al.* [15] 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.

B. Empirical Models

We refer to [16] for a detailed review on the acquisition of transparent, translucent and specular objects. Jensen *et al.* [17] 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 [18], builds a material dictionary on the three scattering parameters. Wang *et al.* [19] 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.* [20] describe a compact representation for heterogeneous subsurface transport and represent heterogeneous objects by layers of homogeneous materials. Thus when the exact layered nature of the composite material is known, as is the case with modelling synthetic scenes. Donner *et al.* [21] 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.

C. Our Approach

We start by acquiring an ensemble of illumination and scene response image pairs under controlled illumination.

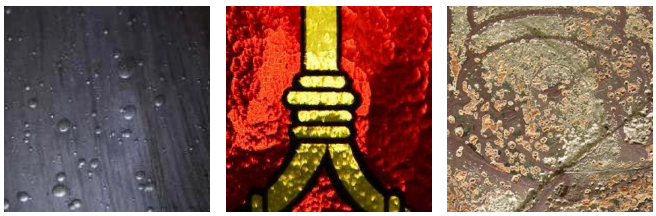
TABLE I
TABLE OF SYMBOLS

$\mathbf{y}$	Observed Image vector
$\boldsymbol{\ell}$	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
$\boldsymbol{\gamma}_n$	Indicator function for the n^{th} image pixel
υ	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
$\boldsymbol{\tau}_n$	n^{th} reduced light transport vector
$\mathbf{D}$	Material specific sparsifying overcomplete dictionary
$\boldsymbol{\alpha}_n$	Sparse co-efficients of the n^{th} light transport vector in dictionary $\mathbf{D}$

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 by 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 for compressive acquisition in controlled illumination. When the incident illumination is ill-conditioned, these priors can be exploited for sparse regularization. This approach preserves heterogeneity including bubbles and corrosions in the digitized artifact.

III. Digital Modelling of Stained Glass

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. Stained glass windows also feature grisaille paint added as an essential artistic tool to the surface of the glass, often even on both sides. 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). This complex blend of factors result in highly heterogeneous light scattering properties. 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. A list of symbols used in this paper is presented in Table I. In the remainder of this section, we will introduce our image formation model and discuss the structure of light transport.



(a) (b) (c)

Fig. 2. 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.

A. 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. The light transport equation is then given by

$$\mathbf{y} = \mathbf{T}\boldsymbol{\ell}, \tag{1}$$

where $\mathbf{T} \in \mathbb{R}^{N\times M}$ is the light transport matrix, $\mathbf{y} \in \mathbb{R}^{N\times 1}$ is the 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.

B. 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.

a) 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 3(a). The light

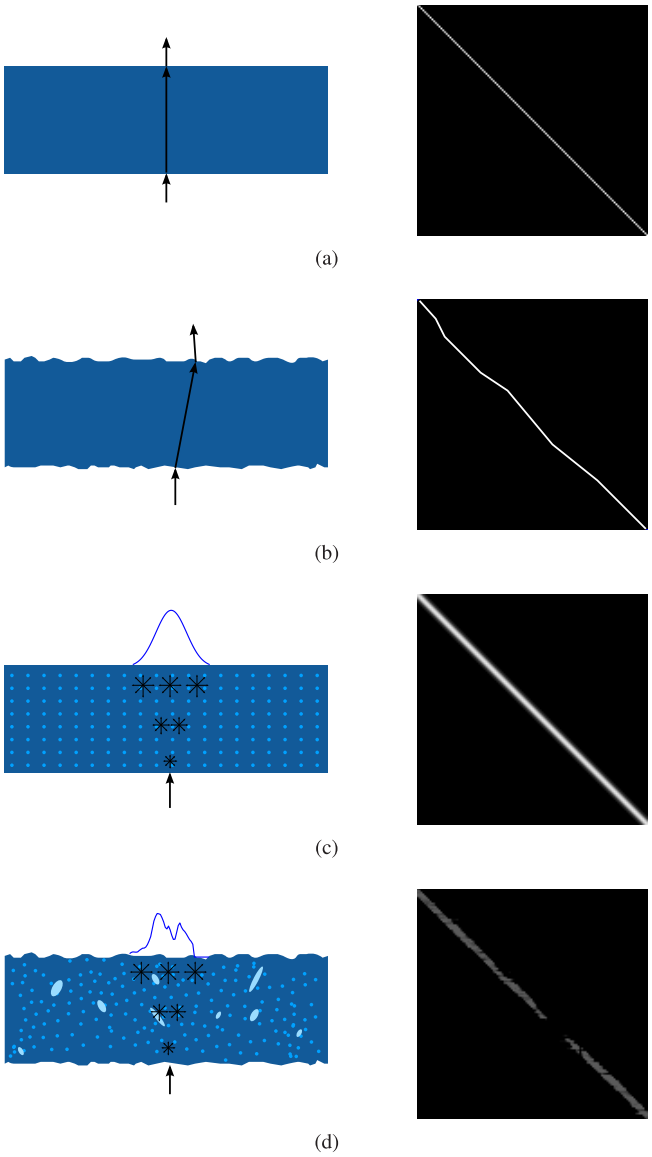


Fig. 3. Illustration of light transport in various types of glass slabs. 2D-slices of the light transport matrix are shown on the side. (a) A transparent planar glass slab, for light at normal incidence. (b) A transparent glass slab with surface irregularities, for light at normal incidence (refraction 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).

transport matrix of a colored, transparent glass plane will be a diagonal matrix. A rendered image is shown in Figure 4(a).

b) 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 3(b).

c) 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 4(b).

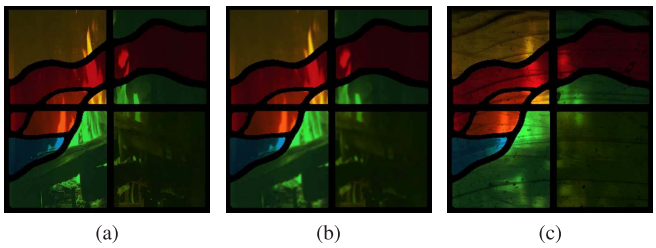


Fig. 4. The same stained glass rendered assuming (a) a coloured, transparent model (alpha-matting) (b) a homogeneous translucent model (c) a heterogeneous translucent model.

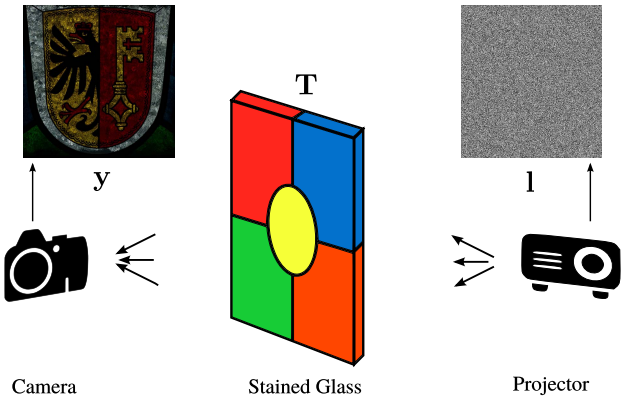


Fig. 5. Illustration of indoor acquisition setup. The pattern $\boldsymbol{\ell}$ is projected on the glass, which is observed as the image $\mathbf{y}$ by the camera.

d) 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 3(d). A rendered image is shown in Figure 4(c). Let υ 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 $\nu = \upsilon + \delta$ is the width of the band of the underlying light transport matrix.

The amount of light reaching the n^{th} pixel $\mathbf{y}[n]$ on the camera sensor, is completely defined by the n^{th} row $\mathbf{T}[n,:] \in \mathbb{R}^{1\times M}$ of the light transport matrix, 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.

In the following section, we will introduce a method for controlled acquisition of the light transport of stained glass by exploiting its structure.

IV. Controlled, One-Off Acquisition of Stained Glass

An illustration of our acquisition setup in indoor scenarios is shown in Figure 5.

Let $\mathbf{Y} = [\,\mathbf{y}_1\,\mathbf{y}_2\cdots\mathbf{y}_P\,] \in \mathbb{R}^{N\times P}$ and $\mathbf{L} = [\,\boldsymbol{\ell}_1\,\boldsymbol{\ell}_2\cdots\boldsymbol{\ell}_P\,] \in \mathbb{R}^{M\times P}$ denote the ensemble of observed image and illumination vectors respectively, obtained by stacking these vectors to form matrices. Then from the image formation equation (1), we can write

$$\mathbf{Y} = \mathbf{TL} \tag{2}$$

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 $M \simeq 2.6\times10^5$ and thus, to solve the linear system (2), 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 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,

$$\begin{aligned}\mathbf{T} &= \arg\min_{\mathbf{T}} ||\mathbf{Y} - \mathbf{TL}||_F \\ &= \arg\min_{\mathbf{T}} \sum_{i=1}^{P} ||\mathbf{y}_i - \mathbf{T}\boldsymbol{\ell}_i||_2\end{aligned} \tag{3}$$

is required, where $||\cdot||_F$ denotes the Frobenius norm.. Solving for equation (3) 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 $N\times M \simeq 2.6\times10^{11}$. Clearly, such a problem is intractable. An alternative formulation is

$$\begin{aligned}\mathbf{T} &= \arg\min_{\mathbf{T}} \sum_{n=1}^{N} ||\mathbf{Y}[n,:] - \mathbf{T}[n,:]\mathbf{L}||_2 \\ \mathbf{T}[n,:] &= \arg\min_{\mathbf{T}[n,:]} ||\mathbf{Y}[n,:] - \mathbf{T}[n,:]\mathbf{L}||_2 \quad \forall n = [1,\cdots N],\end{aligned} \tag{4}$$

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 (4) is referred to as the inverse rendering equation.

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 $\boldsymbol{\gamma}_n$ that determines which light elements from the illumination plane contribute to outgoing light at image pixel n, can then be defined as

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

Then each light transport vector $\mathbf{T}[n,:]$ is such that

$$\boldsymbol{\tau}_n = \mathbf{T}[n, \boldsymbol{\gamma}_n] \in \mathbb{R}^{1\times\nu} \tag{5}$$

contains the entire non-zero block of the n^{th} light transport vector. The ensemble of illumination vectors that contribute to a given image pixel n, $\mathbf{L}_n \in \mathbb{R}^{\nu\times P}$ can then be obtained by retaining only the rows of $\mathbf{L}$ whose indices have a non-zero value in $\boldsymbol{\gamma}_n$ as denoted by,

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

The image formation equation for pixel n can now be rewritten as $\mathbf{y}[n] = \boldsymbol{\tau}_n\boldsymbol{\ell}[\boldsymbol{\gamma}_n]$ where $\boldsymbol{\tau}_n \in \mathbb{R}^{1\times\nu}$ is the reduced light transport vector, as defined in (5).

The inverse rendering equation can be written as,

$$\mathbf{Y}[n,:] = \boldsymbol{\tau}_n\mathbf{L}_n \quad \forall n = [1\cdots N], \tag{6}$$

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

$$\boldsymbol{\tau}_n = \arg\min_{\boldsymbol{\tau}_n} ||\mathbf{Y}[n,:] - \boldsymbol{\tau}_n\mathbf{L}_n||_2 \text{ s.t } \boldsymbol{\tau}_n \geq 0 \quad \forall n = [1\cdots N], \tag{7}$$

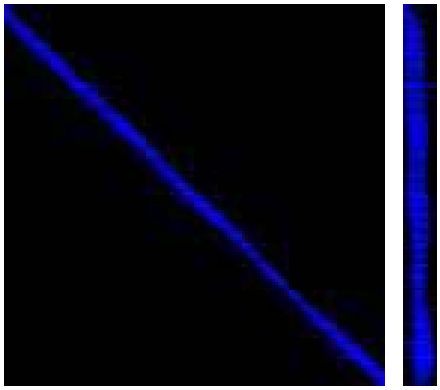
where the non-negativity of light transport is also imposed as a constraint. This framework of solving an inverse problem at each pixel n is similar to the standard light transport acquisition frameworks [10], [11]. However by exploiting the bandedness of the light transport matrix using the indicator function, we are able to solve for an over-determined linear system ($P > \nu$) although the original light transport matrix was under-sampled ($P < M$).

V. Sparse Representation of Light Transport

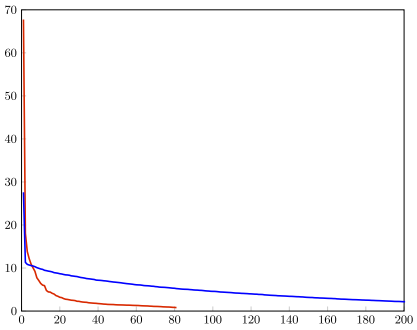
In recent years, compressive sensing has enabled faster acquisition of signals that have a sparse representation in a known basis, provided that the measurement matrix satisfies the restricted isometry property [22], which ensures that linear measurements of sparse signals are nearly orthogonal. For a detailed review of compressive sensing and its applications, see [23]. It is thus advantageous to find a sparsity inducing basis for light transport vectors.

A. Low Dimensionality of Light Transport

It has been conjectured [24], 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 on its structure, we know that it is a banded matrix. A re-parametrized light transport matrix can be constructed from an ensemble of all the reduced light transport vectors $\boldsymbol{\tau}_n$. As shown in Figure 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, $\boldsymbol{\tau}_n$ are pre-disposed for building a basis that represents this space.



(a) (b)



(c)

Fig. 6. (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 $\boldsymbol{\tau}_n$. (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 better representation of the LTM.

B. 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 one-off method described previously. The next step in building the dictionary (illustrated in Figure 7) is a data pooling step. Note that the finite, compact support υ of light transport vectors is caused by volume scattering, whereas the small translations δ are caused due to the presence of surface irregularities. We first recover the volume scattering component by re-organizing each light transport vector $\mathbf{T}[n,:]$ as a 2D image and then segmenting the largest connected component to recover the contiguous non-zero component in each $\mathbf{T}[n,:]$. The support υ 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 $\boldsymbol{\tau}_n$, δ translated versions of each $\boldsymbol{\tau}_n$ are pooled into

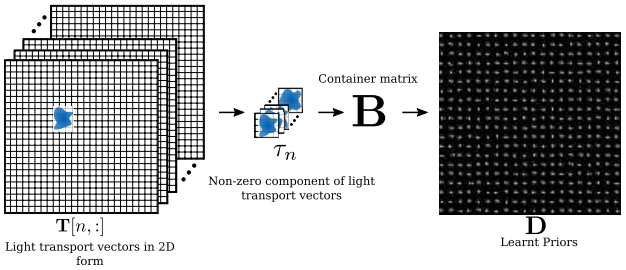


Fig. 7. Schematic representation of learning the sparsity inducing priors. The non-zero component $\boldsymbol{\tau}_n$ of each 2D light transport vector $\mathbf{T}[n,:]$ is first obtained by connected component analysis. Translated versions of these non-zero blocks are then used to populate a container matrix $\mathbf{B}$, which is then used to learn the priors.

a container matrix $\mathbf{B} \in \mathbb{R}^{\upsilon \times C\delta}$ where C is the total number of observed light transport vectors. δ is the largest translation expected due to surface irregularities.

Let $\mathbf{D} \in \mathbb{R}^{K \times \upsilon}$ 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

$$\boldsymbol{\tau}_n = \boldsymbol{\alpha}_n \mathbf{D} \quad \forall n \in [1 \cdots C], \tag{8}$$

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

We now learn a dictionary, $\mathbf{D}$ on $\mathbf{B}$, such that (8) is satisfied. This is done traditionally, by alternating between an l_1-minimization step,

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

and an l_2-minimization step,

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

until convergence. The resulting dictionary, $\mathbf{D}$ is the required basis in which each $\boldsymbol{\tau}_n \in \mathbf{B}$ has a sparse representation. In the next section, we describe the usage of the learnt priors in light transport acquisition.

VI. Acquisition of Stained Glass With Sparsifying Priors

We will now consider two scenarios. A controlled acquisition where we are required to acquire several windows of the same type and an uncontrolled acquisition, under outdoor illumination conditions.

A. Controlled, 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.,* $v = v$. Since we have built the priors such that, $\boldsymbol{\tau}_n = \boldsymbol{\alpha}_n \mathbf{D}$, the above inverse rendering equation (6) can be rewritten as

$$\mathbf{Y}[n, :] = \boldsymbol{\alpha}_n \mathbf{D} \mathbf{L}_n \quad \forall n = [1 \cdots N], \tag{9}$$

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 $\boldsymbol{\tau}_n$ in $\mathbf{D}$ form a sparse vector.

$$\boldsymbol{\tau}_n = \boldsymbol{\alpha}_n \mathbf{D}$$

where,

$$\begin{aligned} \boldsymbol{\alpha}_n &= \arg\min_{\boldsymbol{\alpha}_n} ||\mathbf{Y}[n, :] - \boldsymbol{\alpha}_n \mathbf{D} \mathbf{L}_n||_2 \\ \text{s.t} \quad &||\boldsymbol{\alpha}_n||_1 < \mu, \quad \boldsymbol{\alpha}_n \geq 0. \end{aligned} \tag{10}$$

In contrast to previous methods that use compressive sensing for light transport matrix acquisition, we have $\boldsymbol{\tau}_n$ 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 appendix on Algorithmic issues.

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 (10) for each channel independently. Thus for RGB cameras, we solve for the red, blue and green channel light transport matrices independently.

B. Uncontrolled Acquisition

The image formation model defined in equation (1), describes the transport of spatially varying incident light, whose direction of incidence at each surface point on the object is given by its location with respect to the projector. Unlike this indoor scenario, incident illumination in the outdoor case consists of both strong directional light from the sun and environmental light reflected from scene behind the stained glass window.

Over any period of time, the component of the incident light corresponding to direct sunlight varies directionally. Consequently our light transport model does not capture this component of image formation.

As shown below, we can however use our image formation model defined in equation (1) to approximately describe the transport of environmental light from behind the stained glass window. Note that stained glass windows are planar and not too optically dense unlike other translucent media such as marble or wax. For a given viewing angle (camera position) as illustrated in Figure 8, each surface point on the stained glass (pixel in the observed image) is only influenced by environmental light from a small contiguous region of the scene behind it. Thus, it can be seen that, under outdoor illumination and in the absence of a strong directional light source, the light transport matrix is still banded.

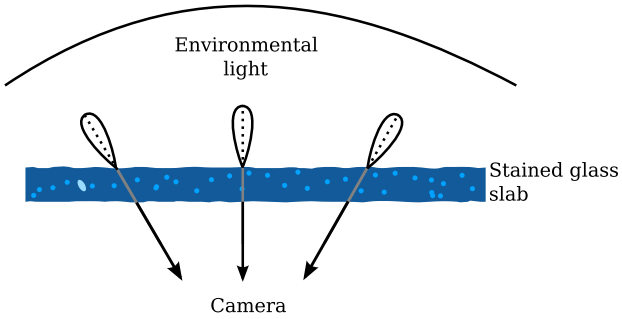


Fig. 8. Illustration of image formation with environmental light. Since stained glass is not too optically dense, in the absence of strong directional light, each pixel in the camera is influenced only by environmental light from a small area behind it.

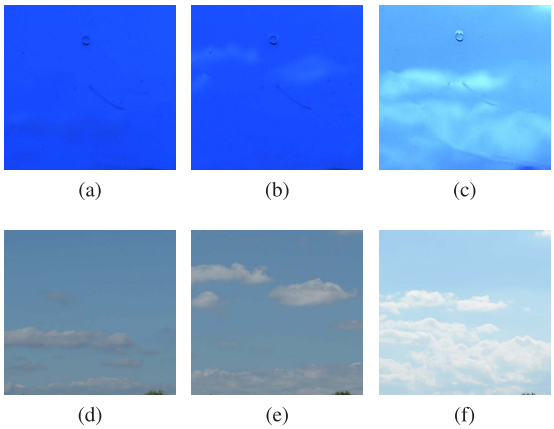


Fig. 9. Stained glass slab under outdoor illuminations. Subfigures (a), (b), (c) show the appearance of stained glass placed outdoors at 10:00AM, 1:00PM and 5:00PM respectively. Subfigures (d), (e), (f) show the incident environmental illumination. Note that in subfigure (c), the sun started to appear on the upper right corner of the region directly behind the stained glass slab and hence the directional components of light transport such as the specularity in the air bubble are visible. The effects of environmental illumination that can be approximated with out light transport model are present in all the image pairs.

In Figure 9, we show observed images of a stained glass slab placed outdoors and the corresponding incident environmental illumination, at different times in a day. It can be seen that the bandedness assumption on the light transport matrix is true. Furthermore, we observe that the strong directional light from the sun contributes significantly for only a small range of incident directions when the sun is directly behind the stained glass.

Let $\mathbf{Y}$ and $\mathbf{L}$ be the ensembles of observed images and the region of environmental illumination that contributes to light transport, over a period of time. The conventional solution that minimizes the l_2-norm of the linear system $\mathbf{Y} = \mathbf{TL}$ is given by the pseudo-inverse of $\mathbf{L}$. However, since the appearance of the scene behind the stained glass window does not change significantly under natural illumination, the matrix $\mathbf{L}$ is ill-conditioned. As a result, the inverse rendering problem is now ill-posed.

Consider the solution obtained by solving the system of equations in (10). After imposing the constraint that each $\boldsymbol{\tau}_n$ is in the row space of $\mathbf{D}$, we search for an $\boldsymbol{\alpha}_n$ 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 (10) 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 appendix on algorithmic issues.

VII. 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.

A. Acquisition Performance

As discussed in section IV, 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. 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 (7) and repeat the experiment for various values of $P > \nu$. This is the over-determined case and Figure 10(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 on a similar dataset as described in Algorithm 1 and simulate both one-off and compressive acquisition using equations (7) and (10), for the under-determined case: $P < \nu$. The SNR from both the proposed methods for various values of $P < \nu$ is shown in Figure 10(b) and it can be seen that in this under-determined case, the dictionary helps in a more accurate reconstruction.

We present similar comparisons on a full LTM by first acquiring the light transport matrix of resolution 262144×16384 ($N \times M$ of a $15cm \times 15cm$ stained glass slab using equation (7) by projecting illumination vectors drawn from a Gaussian random ensemble $\mathbf{L}$. A dictionary $D \in \mathbb{R}^{1024\times 81}$ shown in Figure 11, was then learnt on the light transport matrix using Algorithm 1.

On a second stained glass slab made of similar material type and the same thickness as the first slab, we project Gaussian illumination vectors and reconstruct the light transport matrix $\hat{\mathbf{T}}$ using equation (11) for various values of $P < \nu$. We also obtain the light transport matrix $\mathbf{T}$ of the same slab under the same configuration using equation (7). Figure 13 shows how the SNR improves for light transport

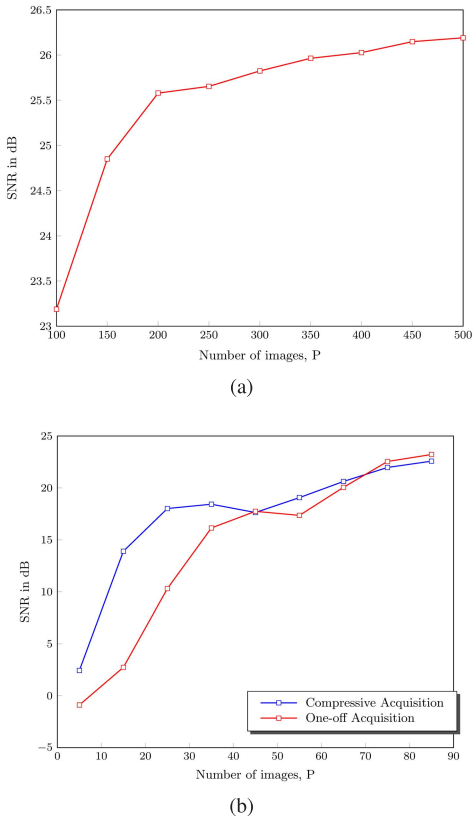


Fig. 10. Improvement of SNR for increasing number of input images (a) Over-determined scenario ($P > \nu$): One-off acquisition without a dictionary. (b) Under-determined scenario ($P < \nu$, $\nu = 81$): Comparison of one-off acquisition and compressive acquisition. It can be seen that for $P < \nu$, the compressive acquisition system performs better.

Algorithm 1 Learning priors

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

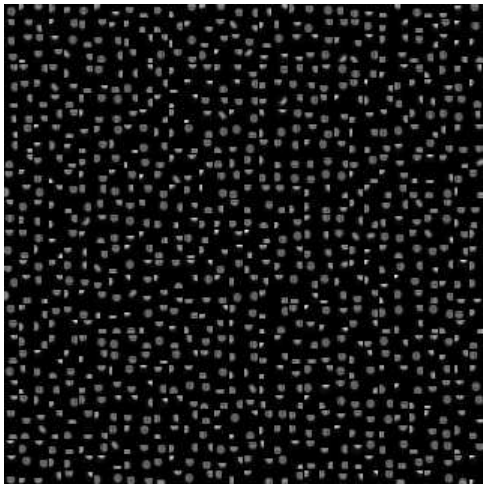
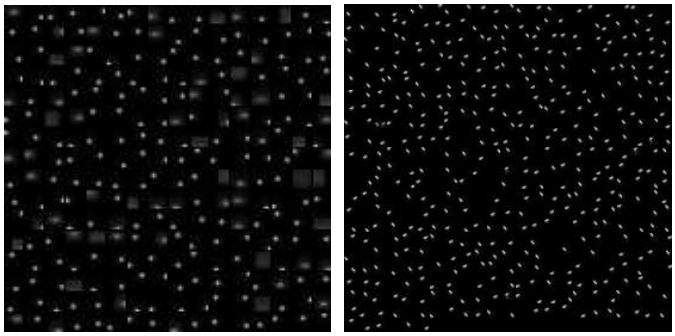


Fig. 11. An overcomplete basis $\mathbf{D} \in \mathbb{R}^{81\times1024}$ with 1024 atoms was learnt from the translated versions of a total of 262144 LTFs. Each basis vector has been re-organized as a 9×9 discretized 2D light transport function.



(a) (b)

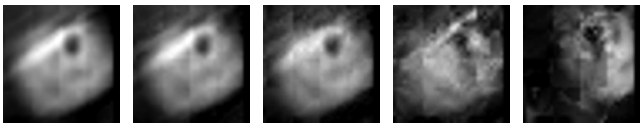
Fig. 12. Material dictionaries learnt from light transport matrices of different types of glass slabs where individual basis vectors have been re-organized as 2D light transport functions. (a) $\mathbf{D} \in \mathbb{R}^{81\times256}$ (b) $\mathbf{D} \in \mathbb{R}^{81\times512}$.

recovery under controlled illumination when the number of illumination vectors used is increased. Here the SNR is defined as $20\log\frac{|\mathbf{T}\boldsymbol{\ell}|_2}{|\mathbf{T}\boldsymbol{\ell}-\hat{\mathbf{T}}\boldsymbol{\ell}|_2}$, where $\mathbf{T}$ is the reference light transport matrix and $\hat{\mathbf{T}}$ is the recovered LTM. Figure 14 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 for SNR. Such an approach will be extremely useful, when a large number of objects made of a similar material are digitized.

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 $\mathbf{Y}$, 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 15 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 16 we

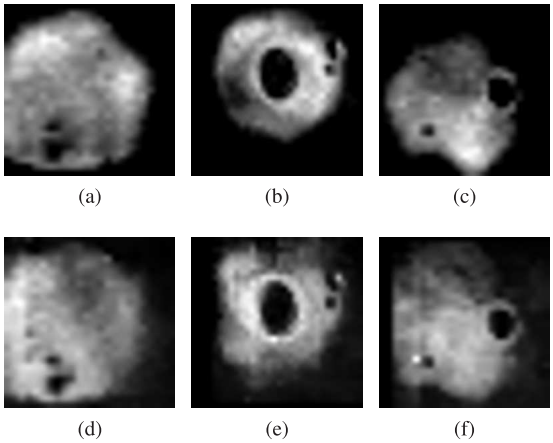
Fig. 13. Improvement of SNR for recovered LTMs of a piece of stained glass, as a function of the number of controlled (Gaussian) illumination vectors used.

(a) Ground Truth (b) 30.1dB (c) 29.7dB (d) 29.6dB (e) 29.4dB



(f) 27.8dB (g) 19.5dB (h) 14.9dB (i) 9.1dB (j) 5.5dB

Fig. 14. 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.



(a) (b) (c)

(d) (e) (f)

Fig. 15. Some point scattering functions (blue channel) of a piece of blue stained glass: Figures (a),(b) and (c) are extracted from an LTM obtained under controlled illuminations. Figures (d),(e) and (f) are the same point scattering functions, when the LTM was recovered under uncontrolled illumination.

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. In Figure 17, we compare a single light transport vector reconstructed under uncontrolled illumination, using nonnegative leastsquares, compressive sensing without a dictionary, a block sparse nonnegative solution and our proposed solution. Note

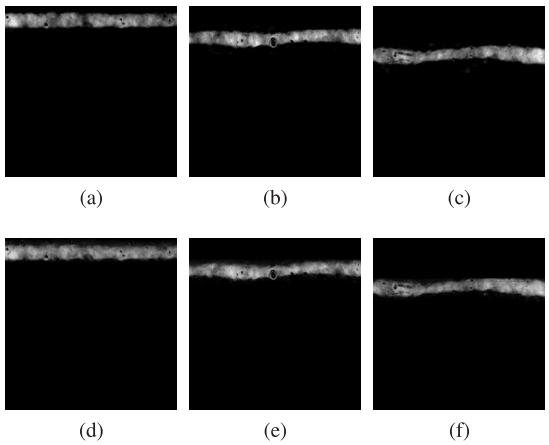


Fig. 16. Rendered images of a piece of stained glass observed under a light source projecting a line on the glass to demonstrate volume scattering. Figures (a),(b) and (c) were obtained by solving the linear system (11) under controlled illumination patterns (Compressive sensing). Figures (d),(e) and (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 (d) 28.95dB ,(e) 32.28dB and (f) 32.95dB.

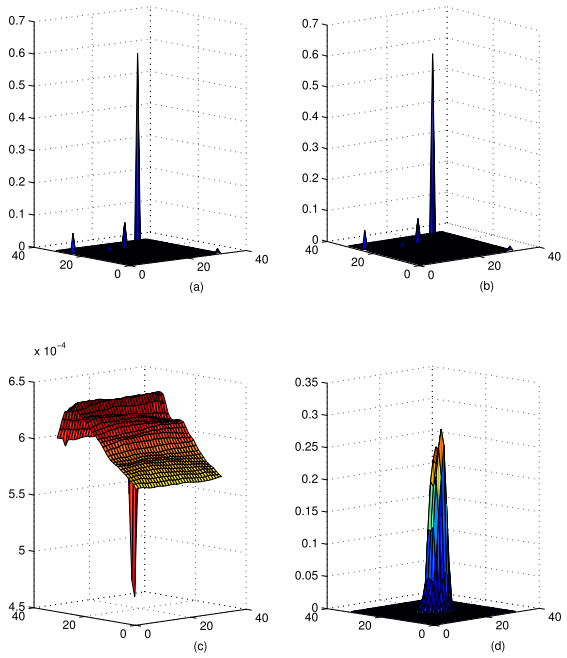


Fig. 17. The same light transport vector, $\mathbf{t}_n$ recovered from data obtained in an outdoor illumination, using (a) ℓ_2 minimization with non-negativity (b) $\ell_1 - \ell_2$ minimization, with non-negativity (c) block sparse solution under non-negativity and (d) Our solution obtained under equation (11). 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 where preserved in any of these three cases.

that only our method preserves the compact support of the light transport vector.

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 18 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.

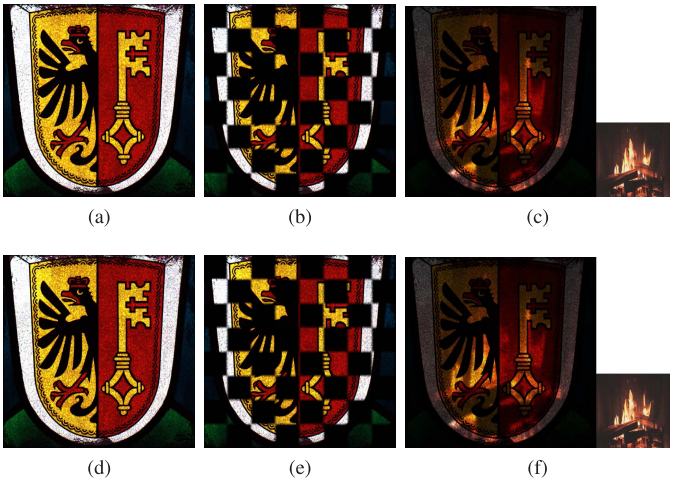


Fig. 18. Relighting results obtained with the Geneva Flag experiment. The top row contains images rendered from an LTM obtained under controlled illuminations. The bottom row contains images rendered from an LTM obtained under uncontrolled illuminations using Method-1. Figures (a) and (d) contain floodlit scenes; PSNR $= 21.97dB$. (b) and (e) contain images generated when a checkerboard pattern is projected upon the glass artifact; PSNR $= 24.26dB$. Figures (c) and (f) are generated by the illumination pattern shown in bottom-right; PSNR $= 26.32dB$.

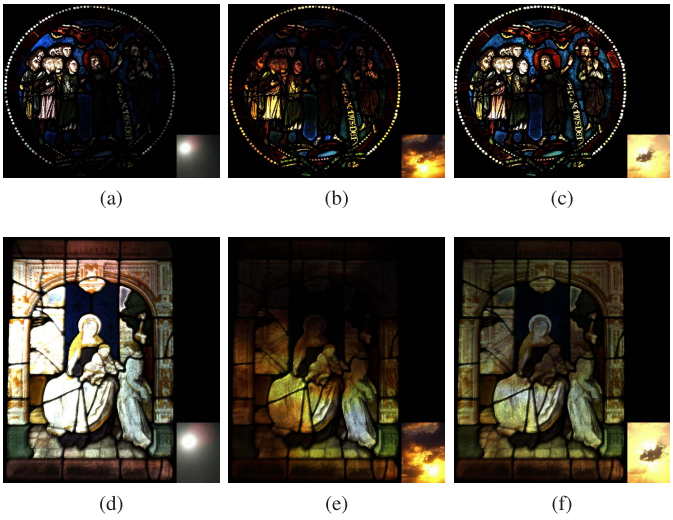


Fig. 19. Relighting results obtained with exhibits at Romont Vitrocentre. (a)-(c) A medieval window that formed part of the Rose Window of the Cathedral of Lausanne. (d)-(f) A window belonging to the Renaissance era.

B. Controlled Acquisition at Romont

We used our acquisition method described by equation (7) to obtain the light transport matrices of two exhibits at the Vitromusée, Romont, Switzerland from a total of 1024 images with each light transport matrix having a resolution of $\simeq 1M \times 2.6 * 10^5$. One was that of a medieval stained glass window dating from ca. 1200 AD, depicting 'Ecce Agnus Dei' that was once located at the cathedral of Lausanne. The second window belongs to the Rennaisance era. Shown in Figure 19 are some relighting results obtained with the acquired light transport matrices.

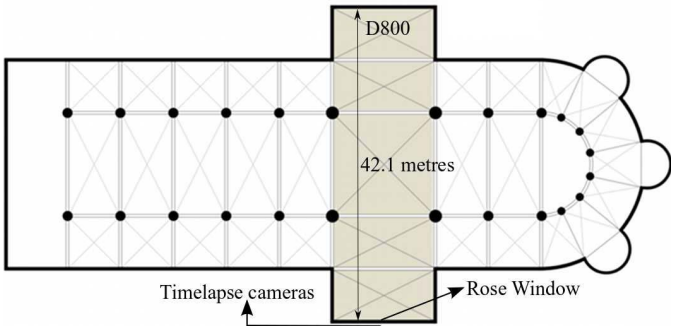
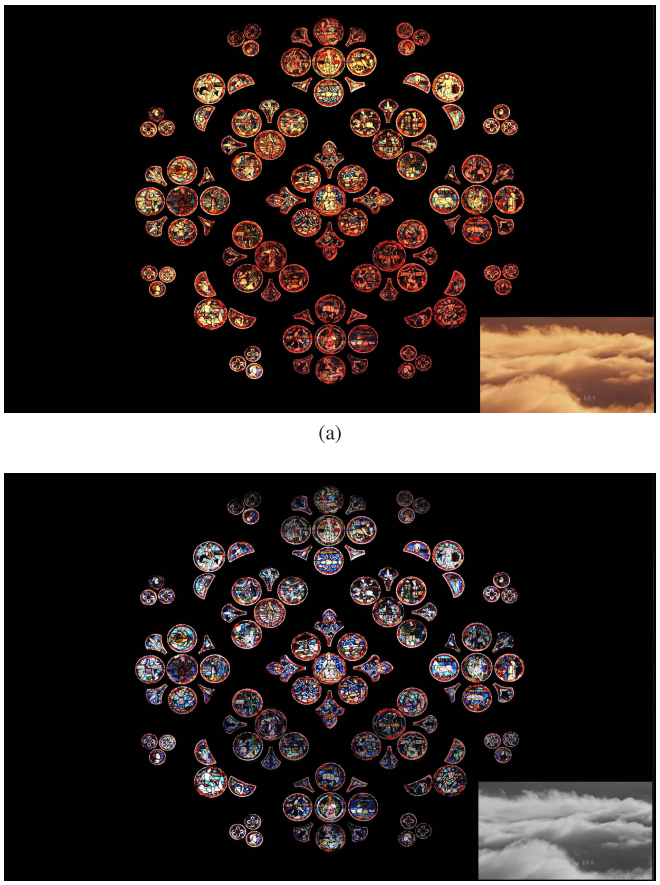


Fig. 20. The Rose Window acquisition setup. On the left is a schematic of the cathedral. On right, the Nikon D800 focused at the rose window and the time lapse cameras on the exterior are shown.



(a)

(b)

Fig. 21. Relighting results obtained with the Rose window experiment. Figures (a) shows the relighting result obtained by using LTM recovered with equations (12). Figures (b) shows the relighting result obtained by using LTM recovered with equations (11). The illumination pattern used for relighting is shown in inset.

C. In-Situ Acquisition of Rose Window, Lausanne

In this section, we describe an experiment carried out at a 12^{th} century Cathedral to digitize its rose window. The window is circular with a diameter of 8 meters. Figure 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

Fig. 22. Parts of results from Figure 21, shown in full resolution to compare both methods described in this paper for LTM recovery under uncontrolled illuminations. The left column includes relighting results obtained by using the LTM recovered by solving for equations (11). The right column includes relighting results obtained by using the LTM recovered by solving for equations (12). Compare the flushed texture of the images on the left with the sharp texture on the right.

the incident light in a synchronized manner with the D800. Both camera setups captured an image once in 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. 138 image pairs were available for analysis. A dictionary was learnt from the light transport matrices of three separate glass slabs. 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 ($\sim 36.1M \times 1.89M : N \times M$). 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 maximum matrix size.

We recovered the approximate LTM by solving for equations (11). Since the spectral composition of outdoor illumination is non-uniform, different color channels of incident light

are differently conditioned. As seen in Figure 22, the rendered images have a washed out appearance at some parts of the stained glass. When the LTM was recovered on the same dataset by solving for the system of equations (12), we observe a remarkable improvement in the preserved details, as shown in Figure 22. 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.

VIII. Conclusion

We have presented and validated by experiments a practical workflow for digitizing stained glass, that can be readily utilized in large scale acquisitions. Based on methods presented in this paper, 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 will encounter.

Appendix

Algorithmic Issues

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

Handling Surface Irregularities

The convex problem in (10) provides the exact light transport vector when the object has an exactly planar surface geometry. $\boldsymbol{\gamma}_n$ for $n = [1, 2, 3..N]$ are exactly known *i.e*, $\boldsymbol{\tau}_n$ 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, $\boldsymbol{\tau}_n$ occur with a slight offset about the leading diagonal. $\boldsymbol{\gamma}_n$ is only approximately known. We account for this by defining a series of $\boldsymbol{\gamma}_n^r$, $r \in [1, 2, ..\delta]$ 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 $\boldsymbol{\tau}_n^r$. We now redefine $\mathbf{L}_n^r = \mathbf{L}[\boldsymbol{\gamma}_n^r, :]$. The light transport matrix at pixel n is then calculated as

$$\mathbf{T}[n, :] = \frac{1}{\delta} \sum_r \mathbf{t}_n^r$$

where,

$$\begin{aligned} \mathbf{t}_n^r[\boldsymbol{\gamma}_n^r] &= \boldsymbol{\tau}_n^r, \qquad \forall r \in [1, 2, ..\delta] \\ \boldsymbol{\tau}_n^r &= \boldsymbol{\alpha}_n^r \mathbf{D} \end{aligned}$$

where,

$$\begin{aligned} \boldsymbol{\alpha}_n^r = \arg\min_{\boldsymbol{\alpha}} & ||\mathbf{Y}[n, :] - \boldsymbol{\alpha}_n^r \mathbf{D} \mathbf{L}_n^r||_2 \\ \text{s.t} \;\; & ||\boldsymbol{\alpha}_n^r||_1 < \mu, \quad \boldsymbol{\alpha}_n^r \geq 0. \end{aligned} \tag{11}$$

As we include translated versions of sample $\boldsymbol{\tau}_n$ while learning $\mathbf{D}$, each $\boldsymbol{\tau}_n^r$ will infact 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.

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 matrix 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.

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 $\mathbf{T}[n, :] \in \mathbb{R}^{1 \times 3\,M}$, 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 $\mathbf{L}_n^{rgb} = \begin{bmatrix} \mathbf{L}_n^r \\ \mathbf{L}_n^g \\ \mathbf{L}_n^b \end{bmatrix} \in \mathbb{R}^{3S \times P}$. Then, for each channel 'c',

$$\begin{aligned} \boldsymbol{\tau}_n &= [\boldsymbol{\alpha}_{n1}\mathbf{D} \; \boldsymbol{\alpha}_{n2}\mathbf{D} \; \boldsymbol{\alpha}_{n3}\mathbf{D}] \quad \boldsymbol{\tau}_n \in \mathbb{R}^{1 \times 3S}, \\ \{\boldsymbol{\alpha}\}_{n1,n2,n3} &= \arg\min_{\boldsymbol{\alpha}_{n1},\boldsymbol{\alpha}_{n2},\boldsymbol{\alpha}_{n3}} ||\mathbf{Y}[n, :] - [\boldsymbol{\alpha}_{n1}\mathbf{D} \; \boldsymbol{\alpha}_{n2}\mathbf{D} \; \boldsymbol{\alpha}_{n3}\mathbf{D}]\mathbf{L}_n^{rgb}||_2, \\ & \qquad \text{subject to, } ||\boldsymbol{\alpha}_{ni}||_1 \leq \mu, \; i \in [1, 2, 3], \\ \boldsymbol{\alpha}_{ni} &\geq 0. \end{aligned} \tag{12}$$

Image formation is now described by

$$\mathbf{y}[n] = [\boldsymbol{\alpha}_{n1}\mathbf{D} \; \boldsymbol{\alpha}_{n2}\mathbf{D} \; \boldsymbol{\alpha}_{n3}\mathbf{D}]\boldsymbol{\ell}_n^{rgb}.$$

Light transport recovery under ill-conditioned illumination is better posed 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. While solving for (12) the light transport recovery is better conditioned, by trading off for colour in the incident light. If we solve for (11) instead, both incident and outgoing light are coloured, however each channel of light transport can have different accuracies depending on the particular illumination spectrum (e.g. different weather conditions).

REFERENCES

[1] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn.*, 2000, pp. 145–156.

[2] T. Hawkins, J. Cohen, and P. Debevec, "A photometric approach to digitizing cultural artifacts," in *Proc. Conf. Virtual Reality, Archeol., Cult. Heritage*, 2001, pp. 333–342.

[3] 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, no. 3, pp. 547–556, 2002.

[4] A. Gardner, C. Tchou, A. Wenger, P. Debevec, and T. Hawkins, "Postproduction re-illumination of live action using interleaved lighting," in *Proc. ACM SIGGRAPH Posters*, 2004, p. 47.

[5] T. Malzbender, D. Gelb, and H. Wolters, "Polynomial texture maps," in *Proc. 28th SIGGRAPH Comput. Graph.*, 2001, pp. 519–528.

[6] B. Tunwattanapong *et al.*, "Acquiring reflectance and shape from continuous spherical harmonic illumination," *ACM Trans. Graph.*, vol. 32, no. 4, Jul. 2013, Art. no. 109. [Online]. Available: http://doi.acm.org/10.1145/2461912.2461944

[7] 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

[8] D. E. Zongker, D. M. Werner, B. Curless, and D. H. Salesin, "Environment matting and compositing," in *Proc. 26th ACM SIGGRAPH*, Jul. 1999, pp. 205–214.

[9] W. Matusik, H. Pfister, A. Ngan, R. Ziegler, and L. McMillan, "Acquisition and rendering of transparent and refractive objects," in *Proc. Eurograph. Symp. Rendering (EGSR)*, 2002, pp. 277–288.

[10] P. Peers *et al.*, "Compressive light transport sensing," *ACM Trans. Graph.*, vol. 28, no. 1, Feb. 2009, Art. no. 3.

[11] Q. Duan, J. Cai, J. Zheng, and W. Lin, "Fast environment matting extraction using compressive sensing," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2011, pp. 1–6.

[12] M. O'Toole and K. N. Kutulakos, "Optical computing for fast light transport analysis," *ACM Trans. Graph.*, vol. 29, no. 6, Dec. 2010, Art. no. 164.

[13] J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo, "Kernel Nyström method for light transport," in *Proc. ACM SIGGRAPH*, New York, NY, USA, 2009, Art. no. 29. [Online]. Available: http://doi.acm.org/10.1145/1576246.1531335

[14] D. Reddy, R. Ramamoorthi, and B. Curless, "Frequency-space decomposition and acquisition of light transport under spatially varying illumination," in *Proc. ECCV*, 2012, pp. 596–610.

[15] P. Peers, T. Hawkins, and P. Debevec, "A reflective light stage," USC Inst. Creative Technol., Los Angeles, CA, USA, Tech. Rep. ICT-TR-04.2006, Dec. 2006.

[16] I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich, "Transparent and specular object reconstruction," *Comput. Graph. Forum*, vol. 29, no. 8, pp. 2400–2426, 2010. [Online]. Available: http://dx.doi.org/10.1111/j.1467-8659.2010.01753.x

[17] H. W. Jensen, S. R. Marschner, S. R. Marschner, M. Levoy, and P. Hanrahan, "A practical model for subsurface light transport," in *Proc. 28th Annu. Conf. Comput. Graph. Interact. Techn.*, 2001, pp. 511–518.

[18] I. Gkioulekas, S. Zhao, K. Bala, T. Zickler, and A. Levin, "Inverse volume rendering with material dictionaries," *ACM Trans. Graph.*, vol. 32, no. 6, Nov. 2013, Art. no. 162.

[19] J. Wang *et al.*, "Modeling and rendering of heterogeneous translucent materials using the diffusion equation," *ACM Trans. Graph.*, vol. 27, no. 1, Mar. 2008, Art. no. 9.

[20] P. Peers *et al.*, "A compact factored representation of heterogeneous subsurface scattering," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 746–753, 2006.

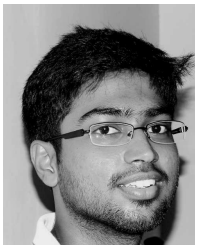
[21] C. Donner, T. Weyrich, E. d'Eon, R. Ramamoorthi, and S. Rusinkiewicz, "A layered, heterogeneous reflectance model for acquiring and rendering human skin," in *Proc. ACM SIGGRAPH Asia Papers*, 2008, Art. no. 140. [Online]. Available: http://doi.acm.org/10.1145/1457515.1409093

[22] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[23] 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, 2009. [Online]. Available: http://dx.doi.org/10.1137/060657704

[24] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible illumination conditions?" *Int. J. Comput. Vis.*, vol. 28, no. 3, pp. 245–260, 1998. [Online]. Available: http://dx.doi.org/10.1023/A:1008005721484

[25] R. Garg, H. Du, S. M. Seitz, and N. Snavely, "The dimensionality of scene appearance," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 1917–1924.



Niranjan Thanikachalam received the bachelor's degree in electronics and communication engineering from the Madras Institute of Technology, Anna University in 2011, and the Ph.D. degree in computer and communication sciences from the École Polytechnique Fédérale de Lausanne (EPFL) in 2016. He is currently a Post-Doctoral Researcher with EPFL. His research interests are image based rendering, inverse problems, and low level computer vision.

Loïc Baboulaz received the M.Sc. degree in communication systems from the École Polytechnique Fédérale de Lausanne in 2004, and the Ph.D. degree in electrical engineering from Imperial College London in 2008. He was a Research Intern with NEC Multimedia Research laboratories, Japan, in 2002, and with Sony Europe Research laboratories, Germany, in 2004. From 2008 to 2011, he was a Research Geophysicist in multicomponent seismic imaging with CGG UK. Since 2012, he has been leading the eFacsimile project on art imaging with the Laboratory of Audiovisual Communications, École Polytechnique Fédérale de Lausanne. In 2016, He Co-Founded Artmyn, a company specialised in digitising artworks for interactive and 3-D web rendering. His interests include image processing, sampling theory and data visualisation.

Paolo Prandoni is currently a Lecturer with École Polytechnique Fédérale de Lausanne, Switzerland. He has co-authored the textbook entitled *Signal Processing for Communication* and designed the first DSP MOOC on Coursera in 2102, now in its fifth edtion. He is the Co-Founder and current CSO of Quividi, the leading video analytics company for out-of-home audience measurement.

Stefan Trümpler received the Ph.D. degree in art history, classical archaeology and history of architecture from the University of Berne, Switzerland, in 1986. In 1988, he became the Director of Vitrocentre Romont, Swiss Research Centre for Stained Glass and Glass Art. In 1991, he was also appointed as a Chief Curator and the Director of Vitromusée Romont, Swiss Museum for Stained Glass and Glass Art. Stefan Trümpler specialized in the history, technology and conservation of stained glass. His research focuses on the technology of stained glass and its interrelation with artistic concepts and the perception of these works of art. His recent research study of preparative designs and cold paint on late medieval and post medieval stained glass and the preventive conservation of historical stained glass. He is also interested in the museology of glass art.

Sophie Wolf received the Degree in geology, palaeontology, and mineralogy from the Universities of Bonn and Cologne in 1995, and the Ph.D. degree from the Department of Geosciences, University of Fribourg, Switzerland, in 1996. She was a Research Fellow with the Research Laboratory for Archaeology and the History of Art, Oxford, U.K., from 2000 to 2001. She was a Lecturer and a Researcher with Fribourg University, the Swiss Gemmological Institute, Basel and the ETH Zurich. Since 2011, she has been with Vitrocentre Romont. Her main research interests are the history, production technology and conservation of glass, stained glass and ceramics.

Martin Vetterli (S'86–M'86–SM'90–F'95) received the Dipl.El.-Ing. degree from Eidgenossische Technische Hochschule, in 1981, the M.Sc. degree from Stanford University, in 1982, and the D.Sc. degree from the École Polytechnique Fédérale de Lausanne (EPFL), in 1986. He was an Assistant and an Associate Professor of Electrical Engineering with Columbia University, NY, and became an Associate Professor and then a Full Professor with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in 1993. In 1995, he joined EPFL as a Full Professor. He held several positions with EPFL, including the Chair of Communication Systems and the Founding Director of the National Competence Center in Research on Mobile Information and Communication systems. From 2004 to 2011, he was the Vice President of EPFL for international affairs, and the Dean of the School of Computer and Communications Sciences from 2011 to 2012. Since 2013, he has been the President of the National Research Council of the Swiss National Science Foundation.

He is currently involved in electrical engineering, computer sciences and applied mathematics. His work covers wavelet theory and applications, image and video compression, self-organized communication systems and sensor networks, and fast algorithms, and has led to about 150 journals papers, and about 30 patents that led to technology transfer to high-tech companies and the creation of several start-ups.

He has co-authored three textbooks entitled *Wavelets and Subband Coding* (Prentice-Hall, 1995) with J. Kovacevic, *Signal Processing for Communications* (EPFL Press, 2008) with P. Prandoni, and *Foundations of Signal Processing* (Cambridge University Press, 2014) with J. Kovacevic and V. Goyal. These books are available in open access, and his research group follows the reproducible research philosophy.

He is a fellow of ACM and EURASIP and a Foreign Member of the U.S. National Academy of Engineering, was a member of the Swiss Council on Science and Technology (2000-2004), and is an ISI Highly Cited Researcher in engineering. His work won him numerous prizes, such as best paper awards from EURASIP in 1984 and the IEEE Signal Processing Society in 1991, 1996, and 2006, the Swiss National Latsis Prize in 1996, the SPIE Presidential Award in 1999, the IEEE Signal Processing Technical Achievement Award in 2001, and the IEEE Signal Processing Society Award in 2010.