

Light Transport Matrix recovery for nearly planar objects

Niranjan Thanikachalam, Loic Baboulaz, Paolo Prandoni and Martin Vetterli
École Polytechnique Fédérale de Lausanne, Switzerland

ABSTRACT

The light transport matrix has become a powerful tool for scene relighting, owing to the versatility of its representational power of various light transport phenomenon. We argue that scenes with an almost planar surface geometry, even with significant amounts of surface roughness, have a banded structure in the light transport matrix. In this paper, we propose a method that exploits this structure of the light transport matrix and provide significant savings in terms of both acquisition time and computation time, while retaining a high accuracy. We validate the proposed algorithm, by recovering the light transport of real objects that exhibit multiple scattering and with rendered scenes exhibiting inter-reflections.

Keywords: Light transport matrix, scene relighting, rendering artworks, sparsity, banded diagonal matrix

1. INTRODUCTION

With the advent of digital museums and virtual tourism, interactive rendering of real world objects has become increasingly relevant. Artworks like stained glass, paintings or tablets have a nearly planar surface as a common factor. Despite the almost planar structure, these objects exhibit complex micro-structures and some visible macro-structures resulting in intricate reflectance and scattering functions. Given the steady increase in the computational power of consumer electronics, efficient methods to recover the surface reflectance of real-world objects can make relightable-photographs an attractive alternative to conventional photographs.

Scene-relighting is a well studied problem, and is completely defined by the underlying 8D bidirectional scattering distribution function (BSDF), $S(\omega_o, \mathbf{x}_o, \omega_i, \mathbf{x}_i)$. The light transport matrix (LTM) is the discretized version of a 4D slice of the 8D bidirectional scattering distribution function, obtained by fixing the incident and outgoing directions. Let $\mathbf{p} \in \mathbb{R}^{1 \times M}$ be the vectorized illumination image projected on the scene by a programmable light source. Let $\mathbf{y} \in \mathbb{R}^{1 \times N}$ be the vectorized image captured by a camera. The LTM, $\mathbf{T} \in \mathbb{R}^{M \times N}$, then defines the transport of light between the \mathbf{p} plane and the \mathbf{y} plane such that,

$$\mathbf{y} = \mathbf{p}\mathbf{T}$$

Given the high spatial resolution of current projectors or CRTs and cameras, the LTM is a highly resolved discretization of the light transport function. Hence, it is a very powerful tool for image based relighting, as it captures specular and diffuse reflections or refraction, subsurface scattering, inter-reflections, caustics and other such phenomena in the scene, for a given incoming and outgoing angles. However, this also results in a prohibitively large matrix, which, if acquired by a brute force method, requires M images. For a projector or CRT with a spatial resolution of 1024×1024 , this amounts to ~ 1 million images.

Three distinct acquisition strategies have emerged as the current state of art. Peers et al¹ and Duan et al² assume that the transport matrix is compressible in a wavelet basis, and employ compressive sensing either out-of-the-box or in a hierarchical manner to exploit spatial redundancies in the camera plane. Another successful approach is to assume that the light transport matrix can be efficiently approximated by a low rank equivalent. Wang et al³ use the kernel-nystrom method for low rank recovery, while Toole et al⁴ use projections in the Krylov subspace, by first imposing symmetry constraints on the LTM using a co-axial configuration for the camera and the projector, for LTM recovery. More recently, there has been focus on decomposing the various components of light transport. Reddy et al⁵ were the first to study in detail the structure of the LTM for a general scene. They

Further author information: (Send correspondence to Niranjan Thanikachalam)
Niranjan Thanikachalam.: E-mail: niranjan.thanikachalam@epfl.ch

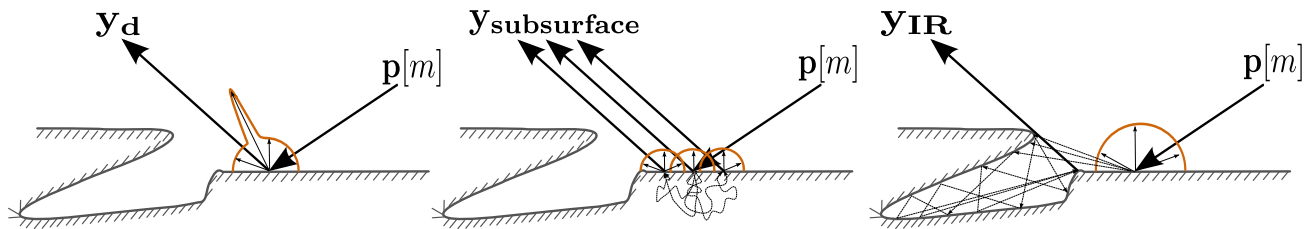
proposed a combination of frequency domain and spatial domain reconstructions by assuming a component-wise low frequency light transport.

In this paper, by considering only objects with a specific type of surface geometry, we propose a recovery algorithm which outperforms existing methods, for this class of objects.

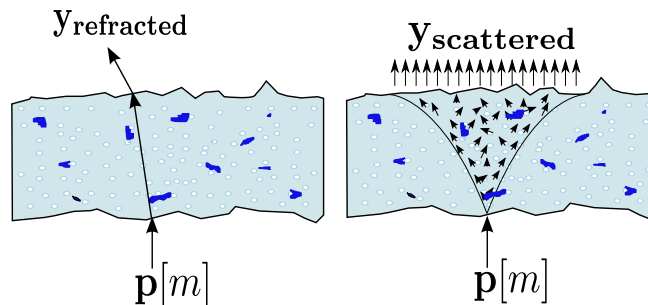
In the following section, we motivate the structure of the LTM for objects with an almost planar geometry. In Section 3, we pose LTM recovery as a well posed inverse problem, with a convex solution. Section 4 outlines our experimental setup and simulation results, that verify the effectiveness of the proposed algorithm. We conclude the article in section 5.

2. STRUCTURE OF THE LIGHT TRANSPORT MATRIX

Building on previous observations, the reflected or refracted light from a surface point is composed of a direct component and a scattered component. In case of opaque and translucent objects, the direct component corresponds to light that is reflected at the surface of the object while the scattered component can be the result of subsurface scattering and inter-reflections. In case of transparent objects and translucent objects, when the light is incident from behind the scene, the direct component is the singly refracted light ray, while the scattered component is from subsurface scattering and other phenomenon such as caustics. Due to additivity of light, it is straight forward to see that the light transport matrix can also be split into a direct and a scattered component, $\mathbf{T} = \mathbf{T}_d + \mathbf{T}_s$. Figure 1 illustrates various light transport phenomena for scenes with different scattering properties.



(a) Illustration of direct, subsurface scattering and inter-reflection components of light transport



(b) Illustration of refracted and scattered components of light transport in a translucent slab

Figure 1: Illustration of various light transport phenomena. $\mathbf{p}[m]$ denotes the m^{th} light element of the programmable light source, while \mathbf{y} denotes outgoing light.

For our analysis of the structure of the LTM, we note that it is the dual of the light field, i.e. it is obtained by the translations of a light source along a plane, while the camera remains fixed. This lets us derive two complementary views of the light transport matrix : one as an array of point scattering functions and another as an array of light transport functions. The i^{th} column of the LTM, \mathbf{t}_i is the vectorized light transport function for the i^{th} pixel in the image sensor plane. It describes the transport of light from the illumination plane \mathbf{p} to the i^{th} pixel of the image sensor plane. The j^{th} row of the LTM on the other hand is the vectorized point

scattering function, for the j^{th} light element in the illumination plane. It corresponds to the impulse response of the scene to the j^{th} illumination element. Without the loss of generality, consider a setup with equal image sensor and projector resolutions. The direct component of the j^{th} point scattering function is a scaled dirac-delta at the j^{th} sensor element, i.e $\mathbf{T}_d[j, :] = c\delta_{(i-j)}$ where $i \in [1, 2..N]$ c is a scalar and $j \in [1, 2, ..M]$. Vectorizing all $\mathbf{T}_d[j, :]$ and stacking them to form the LTM \mathbf{T}_d , yields a diagonal matrix. As is observed in Peers et al,⁶ the subsurface scattering component, for light entering a surface at a given point is a spatially varying filter with a finite support defined around the point of entry of light, where the support depends on the constituent material and the surface thickness. Thus, on the sensor plane, the subsurface scattering component has a finite support, which when vectorized becomes a group sparse vector and stacking them into a matrix yields a banded matrix. The inter-reflections and caustics components only occur in scenes with a non-convex surface geometry. The inter-reflections and caustics component for a given light element, in an arbitrary scene can span the entire scene, depending on the scene geometry. However, in a scene with nearly planar geometry, the inter-reflections component is localized around the point of entry of light, and its span is dependant on the maximum depth difference in the scene. Thus, the inter-reflections component also forms a banded matrix. Reddy et al⁵ also argue that the inter-reflections component occurs anywhere along the off diagonal elements and can be divided into a banded ‘near range’ and an unstructured ‘far range’ component. In almost planar scenes, only near range inter reflections are present. Thus, the LTM for nearly planar objects is always a band matrix, where the bandwidth \hat{S} is related to the surface regularity in case of a reflective object or material scattering properties in case of a refractive object and both in case of translucent objects.

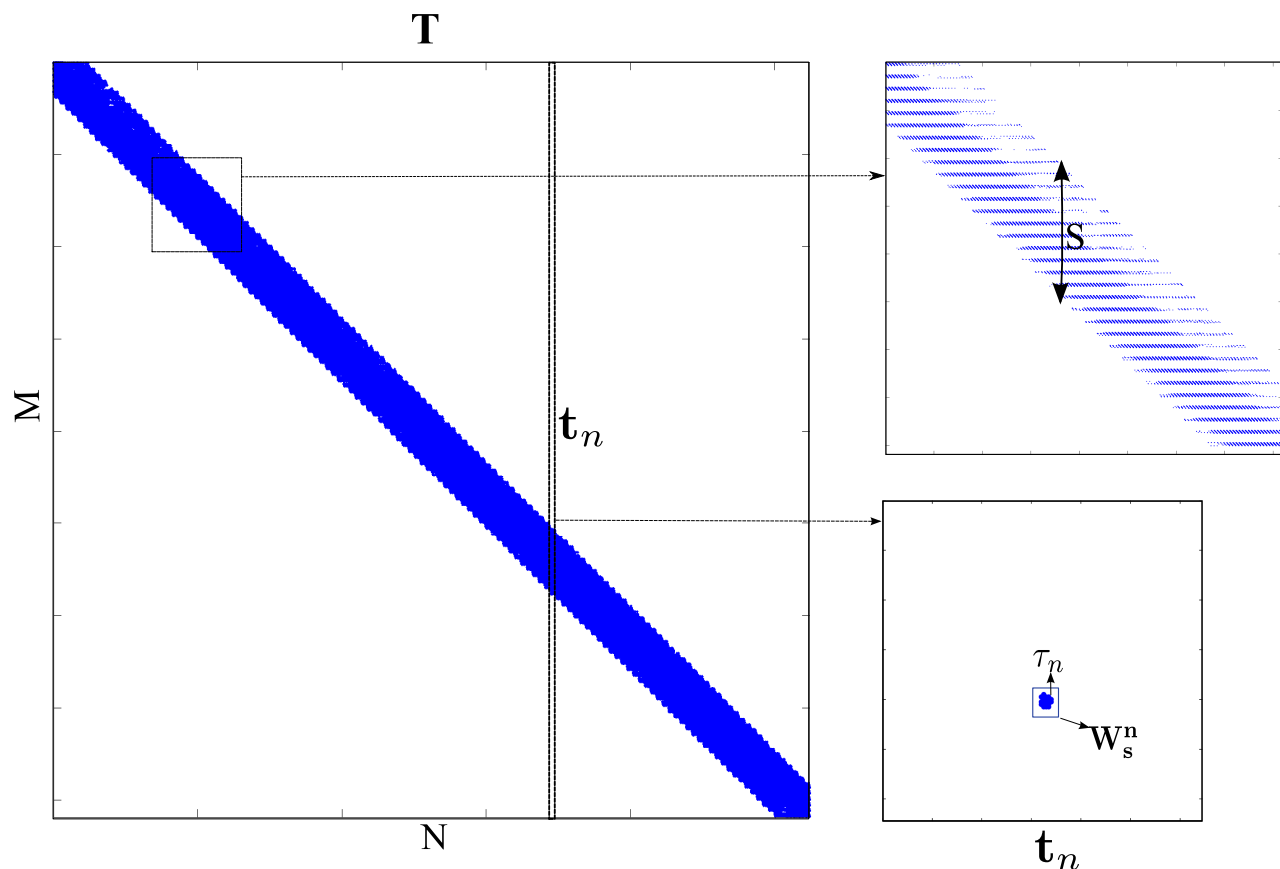


Figure 2: Shown in right is the sparsity profile of the LTM \mathbf{T} of a piece of blue stained glass. Shown in top right is the sparsity profile of a part of the LTM, at a higher resolution. In bottom right, the sparsity profile of a light transport function, \mathbf{t}_n is shown.

3. LIGHT TRANSPORT MATRIX RECOVERY

When represented as an array of LTFs, the banded nature of the LTM manifests itself as a light transport function with finite spatial support, \hat{S} , defined about pixel $n \in [1 \dots N]$. Let $\mathbf{Y} = [\mathbf{y}_1; \mathbf{y}_2; \dots; \mathbf{y}_L]$ and $\mathbf{P} = [\mathbf{p}_1; \mathbf{p}_2; \dots; \mathbf{p}_L]$ be the ensemble of observed image and illumination vectors respectively, while $m \in [1 \dots M]$ be the index of light elements in the illumination vector. Let \mathbf{W}_s^n be a square window of area S defined about the pixel n , such that $\hat{S} < S$. We then define an indicator function $\Gamma_n[\mathbf{m}]$ that describes which light elements from the illumination plane contribute to the pixel n ,

$$\Gamma_n[\mathbf{m}] = \begin{cases} 1 & \text{if } m \in \mathbf{W}_s^n \\ 0 & \text{otherwise} \end{cases}$$

The illumination vector, $\hat{\mathbf{P}}_n$, that contributes to any given pixel in the image plane can then be calculated as

$$\hat{\mathbf{P}}_n = \mathbf{P}[:, \Gamma_n]$$

Then, $\tau_n = \mathbf{t}_n[\Gamma_n]$ denotes the non-zero component of the light transport function.

Figure 2 shows the banded nature of the sparsity pattern of the LTM of a real stained glass piece. Also shown is the sparsity pattern of one LTF of the stained glass piece, with an illustration of the square window.

The light transport function at pixel n , \mathbf{t}_n , is then the solution of

$$\mathbf{y}_n = \hat{\mathbf{P}}_n \tau_n$$

where $\mathbf{y}_n \in \mathbb{R}^{L \times 1}$ is the set of L observations at pixel n . Note that, since $S \ll M$ for objects of interest, the above equation solves a much smaller system than the original system. We can now consider two cases.

Case-a: Brute force approach ($S < L$) When the number of observations is larger than the spatial support, S , of τ_n , the problem is well posed. It is a known prior that $\tau_n \geq 0$. Thus the LTM can be recovered by solving a convex problem that minimizes the l_2 -norm, while imposing a non-negativity constraint on τ_n .

$$\hat{\tau}_n = \arg \min_{\tau_n} \|\mathbf{y}_n - \hat{\mathbf{P}}_n \tau_n\|_2 \quad \text{subject to } \tau_n \geq 0 \quad \forall n \in [1 \dots N]$$

where $\mathbf{y}_n \in \mathbb{R}^L$ and $\hat{\mathbf{P}}_n \in \mathbb{R}^{L \times S}$. The above problem is well posed as long as $\hat{\mathbf{P}}_n$ is well conditioned. This simplified problem by itself provides considerable savings in both computational power and acquisition time, compared to out-of-the-box compressive sensing.

Case-b: Compressive Sensing ($L < S \ll M$) If the scene is composed of refractive / reflective objects, each light transport function, bounded by \mathbf{W}_s^n is sparse even inside the finite support of the window. If the scene is composed of translucent material, both subsurface scattering and reflection occur (ideal refraction can happen depending on the position of the illumination plane). Thus in this case each light transport function is a low frequency signal and hence is compressible. Let $\phi \in \mathbb{R}^{S \times S}$ be any basis that renders the bounded light transport sparse (or approximately sparse). Then, when the observation matrix $\hat{\mathbf{P}}_n \in \mathbb{R}^{L \times S}$ satisfies the Restricted Isometry Principle (RIP), the bounded light transport can be recovered by compressive sensing using fewer than S samples. We use non-negativity as an additional criterion for light transport recovery.

$$\hat{\tau}_n = \arg \min_{\tau_n} \|\phi^T \tau_n\|_1 \quad \text{subject to } \|\mathbf{y}_n - \hat{\mathbf{P}}_n \tau_n\|_2 \leq \epsilon \quad \text{and } \tau_n \geq 0 \quad \forall n \in [1 \dots N]$$

where $\mathbf{y}_n \in \mathbb{R}^L$ and $\hat{\mathbf{P}}_n \in \mathbb{R}^{L \times S}$.

The number of measurements required for successful recovery of a k -sparse signal depends upon the kind of sensing matrix used. When a partial Fourier ensemble is used the minimum number of measurements required for out of the box compressive sensing to succeed in signal recovery is

$$L \sim k(\log M)^6.$$

However, with our problem formulation, the problem size reduces to S from an initial size of M , while the sparsity of the signal remains the same. Thus with our problem formulation even fewer (L_1) measurements are required for the same signal, such that

$$L_1 \sim L \left(\frac{\log S}{\log M} \right)^6.$$

Note that in case of approximately sparse signals (compressible), the number of measurements can be traded for a better SNR. When Gaussian matrices are used as the sensing matrix $L \geq k \log S/L$ measurements are required for successful recovery. In our experiments, the sensing matrix is the ensemble of images such that each pixel value is independently drawn from the Gaussian distribution. Also note that since the modified problem is of size $S \ll M$ the computational time also significantly falls.

4. EXPERIMENTAL RESULTS

We validate our method for LTM acquisition with both real world data and with ray-traced models of nearly planar objects. Our experimental setup consists of a radio-metrically calibrated CRT monitor, that projects patterns on the scene of interest, and a NIKON D800 that captures the resulting scene. For a given camera-display setup, we establish the camera-display correspondence, by projecting a few calibration patterns, as is the standard practice. We now describe the performance of our algorithm in two experiments : the Geneva flag stained glass setup and the cookies - a ray traced scene.

Figure 4 shows relighting results obtained with the Geneva flag experiment. The object is a stained glass artifact, that exhibits considerable amounts of surface roughness, and subsurface-multiple scattering when the light source is located behind the object. The figure shows relit images with various illumination patterns : a flood light, a high frequency light (checker-board pattern), a multi-coloured light and with natural illumination (fire). The SNR variation for the LTM recovery of a stained glass artifact as a function of undersampling factor L/M is shown in figure 3-a. Here SNR is defined as $20 \log \frac{\|\mathbf{T}\|_2}{\|\mathbf{T} - \hat{\mathbf{T}}\|_2}$, where $\hat{\mathbf{T}}$ is the recovered LTM. Note that a SNR of 25dB is obtained for an undersampling factor of even 0.0011. 80 image pairs (L) were used in the recovery of LTM and M was 65536 and $N = 262144$. A supplemental video of the same scene relit with an illumination sequence of a fireplace has been made available.

For the cookies experiment, we used PovRay to create a synthetic scene made of opaque objects, with an almost flat, but non-convex surface geometry. With this setup, we repeated our recovery algorithm. Relighting results for various illumination patterns : a flood light, high frequency light (stripes) and a coloured illumination pattern are shown. The transport matrix is sparse even along the diagonal band (the previous scenes where approximately sparse in the bounded window). Thus ϕ in the reconstruction formula, is now an identity matrix, $\mathbf{I}_s \in \mathbb{R}^{s \times s}$. The SNR of reconstructed transport matrix is shown in Figure 3-b. Note that the undersampling factor is defined similar to the previous case.

5. CONCLUSION

We have presented an algorithm for light transport recovery that exploits the banded nature of the LTM of scenes with an almost planar surface geometry. We have verified the applicability of our algorithm for scenes with both reflective and translucent surfaces. The main contribution of the paper is, a new problem formulation that greatly reduces the size of the linear system of equations solved for per-pixel-light-transport recovery, thus enabling savings on both acquisition time and computational requirements. We have also provided a compressive sensing formulation of our modified problem that enables further reduction in acquisition time by a trade off with the accuracy of recovered light transport. Due to the gaining popularity of digital museums and computational photography, alternative representations of visual information are becoming increasingly attractive. Relightable photographs can be a viable alternative for traditional photographs if the acquisition and computational overhead is minimal. Towards this end, it is interesting to have a common class of light transport recovery algorithms for the large variety of paintings, engravings and stained glass artifacts all of which have a nearly planar (2.5D) geometric structure. Our approach for light transport recovery aids in high fidelity image based 2.5D rendering, while reducing both the computation load and the acquisition time.

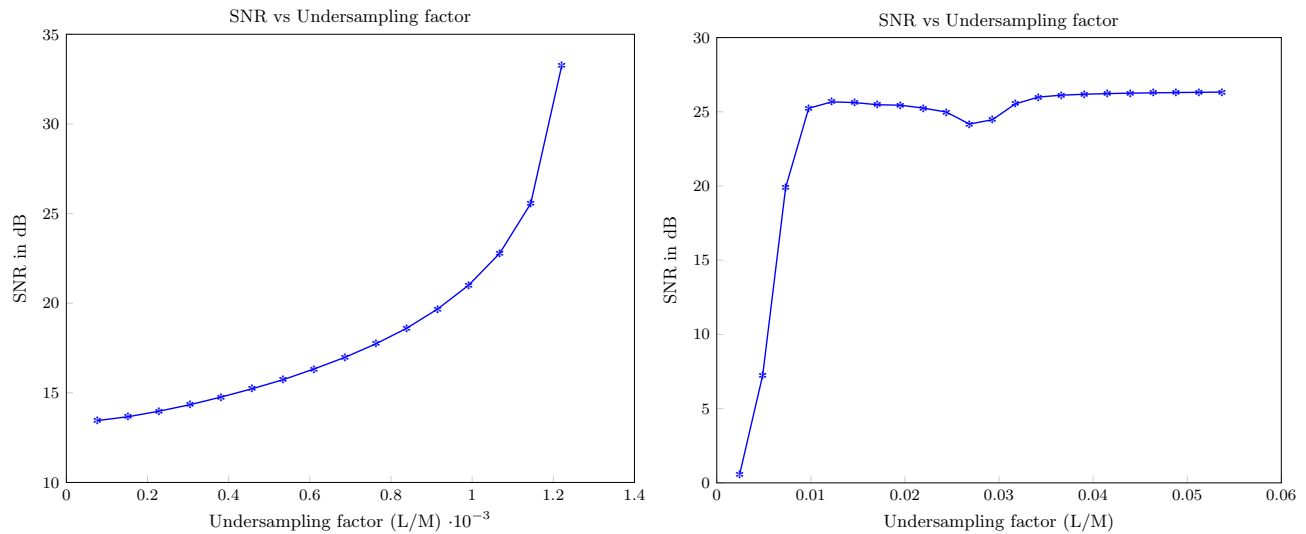


Figure 3: SNR variation for stained glass artifact, and rendered scene with non-convex-but almost flat surface, as a function of undersampling factor

ACKNOWLEDGMENTS

This research is being supported by Google Focus Award - eFacsimile research, SNF Grant and ERC Advanced Investigators Grant.

REFERENCES

- [1] Peers, P., Mahajan, D. K., Lamond, B., Ghosh, A., Matusik, W., Ramamoorthi, R., and Debevec, P., "Compressive light transport sensing," *ACM Trans. Graph.* **28**, 3:1–3:18 (Feb. 2009).
- [2] Duan, Q., Cai, J., Zheng, J., and Lin, W., "Fast environment matting extraction using compressive sensing," in [*Multimedia and Expo (ICME), 2011 IEEE International Conference on*], 1–6 (2011).
- [3] Wang, J., Dong, Y., Tong, X., Lin, Z., and Guo, B., "Kernel nystrom method for light transport," in [*ACM SIGGRAPH 2009 papers*], *SIGGRAPH '09*, 29:1–29:10, ACM, New York, NY, USA (2009).
- [4] O'Toole, M. and Kutulakos, K. N., "Optical computing for fast light transport analysis," *ACM Trans. Graph.* **29**, 164:1–164:12 (Dec. 2010).
- [5] Reddy, D., Ramamoorthi, R., and Curless, B., "Frequency-space decomposition and acquisition of light transport under spatially varying illumination," in [*Proceedings of the 12th European conference on Computer Vision - Volume Part VI*], *ECCV'12*, 596–610, Springer-Verlag, Berlin, Heidelberg (2012).
- [6] Peers, P., Berge, K., Matusik, W., Ramamoorthi, R., Lawrence, J., Rusinkiewicz, S., and Dutr, P., "A compact factored representation of heterogeneous subsurface scattering," *ACM Transactions on Graphics SIGGRAPH* (2006).



(a) Floodlit scene



(b) Checkerboard pattern



(c) Coloured illumination

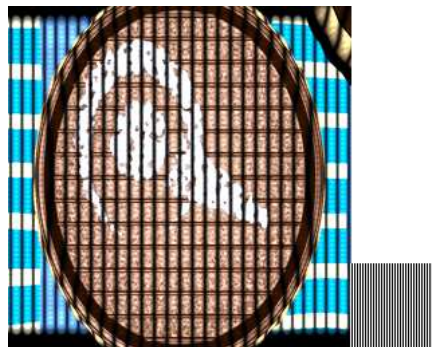


(d) Natural illumination : Fire

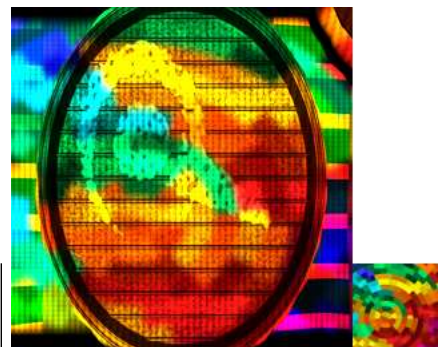
Figure 4: Relit images of a stained glass artifact under various illuminations. A video sequence showing this glass artifact rendered with illumination from a fireplace has been made available as supplemental material <http://dx.doi.org/10.1117/12.2038909.1>



(a)



(b)



(c)

Figure 5: Relit images of a rendered scene with non-convex surface : a) Floodlit scene b),c) Illuminated by image shown in inset