Handheld reflectance acquisition of paintings

Niranjan Thanikachalam, Loïc Baboulaz, Damien Firmenich, Sabine Süsstrunk, *Fellow, IEEE*,
Martin Vetterli, *Fellow, IEEE*

Abstract—Relightable photographs are alternatives to traditional photographs as they provide a richer viewing experience. However, the complex acquisition systems of existing techniques have restricted its usage to specialized setups. We introduce an easy-to-use and affordable solution for using smartphones to acquire the reflectance of paintings and similar almost-planar objects like tablets, engravings and textile. Our goal is to enable interactive relighting of such artifacts by everyone. In our approach, we non-uniformly sample the reflectance functions by moving the LED light of a smartphone and simultaneously tracking the position of the smartphone by using its camera. We then propose a compressive-sensing based approach for reconstructing the light transport matrix from the non-uniformly sampled data. As shown with experiments, we accurately reconstruct the light transport matrix that can then be used to create relightable photographs.

Index Terms—Light transport matrix, computational relighting, image based relighting, compressive sensing, non-uniform sampling, mobile imaging.

I. Introduction

ARTWORKS such as paintings, iconographs, and tablets posses minute geometric details, on an almost planar surface, which add texture and visual richness to the observer. One way to preserve these characteristics in the digital replicas of such artifacts is to acquire their reflectance, thus enabling the creation of relightable photographs. More than a decade ago, the lighting sensitive display [1] was first proposed as a medium for displaying photographs that respond to the viewing environment by dynamically changing their incident illumination. Such relightable photographs are created by acquiring the scene's reflectance field, so that the scene can be rendered under novel illumination conditions. Acquiring the reflectance function of a real-world scene is called inverse rendering, and a large number of acquisition frameworks focus on fast and accurate acquisition of the reflectance field [2] [3] [4]. However, the complexity and the scale of most of the existing acquisition frameworks prevent widescale use beyond specialized studios. In this paper, we present a novel, affordable, portable and easy-to-use framework for creating relightable photographs of 2.5D objects such as paintings, which are characterized by fine geometric details as well as a specular reflectance field.

We build our acquisition system entirely with two smartphones as shown in Figure 1. In our proposed framework, a user samples the reflectance field of the scene by moving a smartphone (mobile $\mathcal{A}$) with its LED turned on, facing the

The authors are with the School of Computer and Communication Sciences, Ecole Polytechnique Federale de Lausanne, Lausanne 1015, Switzerland (e-mail: thanikachalam.niranjan@gmail.com; l.baboulaz@artmyn.com; d.firmenich@artmyn.com; sabine.susstrunk@epfl.ch; martin.vetterli@epfl.ch)

scene. Simultaneously, a second smartphone or any video-capable camera ($\mathcal{B}$) records the corresponding scene response. The user interface on the screen of phone $\mathcal{A}$ guides the user by displaying the path already taken by the user. (See Figure 2 for an illustration.) From the proposed handheld acquisition system, we obtain a sparse, non-uniform sampling of the reflectance function.

To obtain a meaningful, uniformly sampled light transport matrix (LTM) that we use for inverse rendering, we introduce a compressive-sensing based reconstruction strategy. As shown in Figure 2, the LTM can be defined as an array of arrays, where each subarray describes how the corresponding surface point behaves under varying illumination. Originally proposed as environment matting [5], the LTM has since grown into a powerful tool for image-based relighting, due to the versatility in its representational power. A major difference in our approach, compared with conventional methods that also acquire the LTM is the use of an LED that approximates a point-light source, instead of the more common structured-light (projector) setups. This enables us to generalize the incident illumination to 4D light fields, whereas structured-light sources only allow for spatially varying light.

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

II. Related Work

The most common setups for acquiring scene reflectance involve complex, specially constructed light stages [6] [7] [8] [2]. The object is photographed under varying illuminations; these photographs are then used to recover a parametric reflectance model of the object. An alternative strategy for scene relighting involves a structured-light source in a projector-camera setup that acquires the scene's LTM. Earlier work, [5] [9] [10] [11] focussed primarily on characterization of translucent and transparent objects. More recently, it has been extended to general scenes where compressive-sensing [3], low-rank matrix recovery [4], [12] and a combination of spatial-domain and frequency-domain sampling [13] have been proposed. However, in these latter methods, the usage of structured-light sources limits the angular range of the reflectance captured. An adaptive sampling approach [14] that uses indirect light from spotlight projectors has also been proposed. In general, such elaborate, specialized setups severely limit their usability outside of specialized studios. A simple, easily available hardware setup with an intuitive, fast, and non-restrictive usage scenario is therefore highly desirable.

(a)

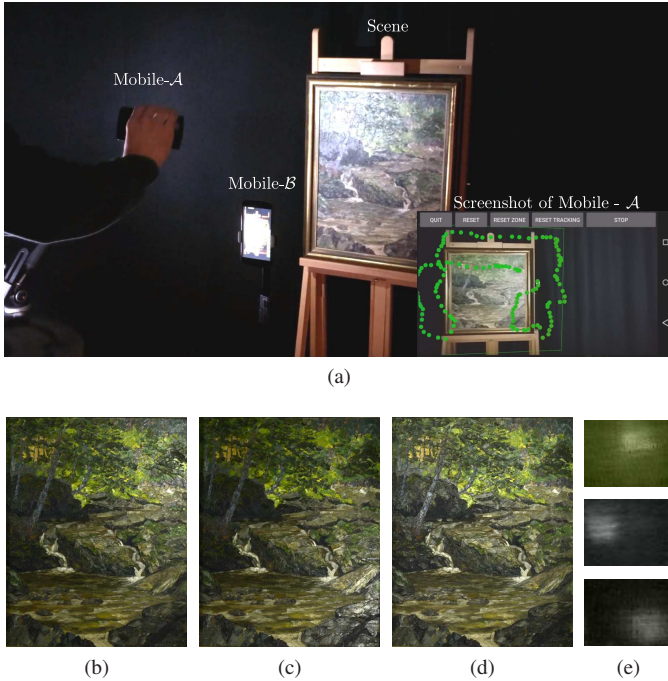(b)   (c)   (d)   (e)

Fig. 1: (a) Our acquisition system in action: Mobile $\mathcal{A}$ functions as a moving light source while mobile $\mathcal{B}$ records a video of the scene. A screenshot of mobile $\mathcal{A}$ displaying the user's trajectory is shown in inset. (b)-(d) Relighted images of an oil painting by Hohmann under three novel light positions. Image plane resolution was 555×437. (e) Randomly selected light transport functions exhibiting specular structures. Illumination plane resolution was 75×57.

Polynomial texture mapping [7] has a simple setup, as it can be used with a simple light probe and a user-controlled light source. However, the acquired per-pixel reflectance data is projected into the space of biquadratic polynomials, which results in the loss of specularities. As a compromise, the authors propose a simulated specular reflectance that uses parametric models along with the acquired normal maps. A more recent example is pocket reflectometry [15]. It uses a handheld linear-light source in combination with a reference bidirectional reflectance distribution function (BRDF) chart to acquire the spatially varying BRDF parameters of the scene. This parametric spatially varying BRDF is then used for rendering and is mainly aimed at scenes with convex surface geometry. The free-form light stage [16] uses a free-moving handheld light source in conjunction with four spheres in order to create basis images. A weighted combination of these basis images is then used for image based relighting of the acquired scene. We compare our reconstruction strategy with other reconstruction strategies in Section IV. Moreover, unlike these approaches, our approach does not require any special instrumentation apart from the now common smartphone. More recently, a method has been proposed to compute the SV-BRDF of a scene from two photographs, one with a mobile LED on and one without [17]. Another recent method proposes to use neural networks to model the light transport of complex

objects, from a few sparse samples [18].

A setup is proposed to capture light fields using a moving mobile camera in [19]. Although we capture the LTM instead of the light field, the two approaches take a similar smartphone-based scanning approach. However, [19] use bicubic interpolation for their reconstruction, hence a dense sampling of the light field is required during the acquisition stage. In contrast, as we operate in the frequency-domain, our approach allows us to trade-off the density of the trajectory with the resolution of the reflectance function.

A rigorous treatment of reconstruction of 2D spatial fields from samples along a trajectory is presented in [20]. However samples are obtained along well-defined curves, hence the approach is unsuitable for handheld sampling. Our acquisition system does not require the user to follow any specific trajectory in order to scan the object, thus making the acquisition more user-friendly.

In the following section, we introduce our mobile-based acquisition strategy, its sampling model and the reconstruction methodology for non-uniformly sampled light-transport. We discuss experimental results and future extensions of the current setup in Section IV.

III. Light-Transport Acquisition System

In this section, we first describe the continuous-space light-transport and introduce notations. Then we present the proposed light transport acquisition framework.

A. Continuous-Space Light-Transport

For a fixed viewing angle, scene appearance is entirely characterized by the light rays incident on the scene. Thus, any incident illumination can be represented as an incident 4D light field, which can be parametrized into a two-plane parametrization. Without loss of generality, the first plane with Cartesian co-ordinates $[x, y] \in \mathbb{R}^2$ is defined to coincide with the focal plane of the camera. As the camera projects this plane to its image plane (sensor or film), we refer to it as the image plane.

Let us define the second plane with Cartesian co-ordinates $[r, s] \in \mathbb{R}^2$ to be parallel to the image plane and located at a distance h from it, in front of the scene. Thus for a fixed viewing angle, image formation is a function of the incident light field that is solely parametrized by planes $[r, s]$ and $[x, y]$; this function, represented by $\mathcal{T}(r, s, x, y)$ is the continuous-space light-transport field. $\mathcal{T}(r, s, x, y)$ can be obtained by moving a point-light source through the plane defined by $[r, s]$. Since the light source is located on the plane $[r, s]$ in this acquisition scenario, we refer to it as the illumination plane. As moving the point-light along infinitely large planes is not possible, we restrict the illumination plane $[r, s]$ to a finite rectangle R and the image plane $[x, y]$ to a finite rectangle X.

The camera $\mathcal{B}$ discretizes the image plane $[x, y]$, as described by,

$$\tau_{ij}(r, s) \equiv \left\langle \mathcal{T}(r, s, x, y), \phi_{xy}\left(\frac{x}{\Delta_{xy}} - i, \frac{y}{\Delta_{xy}} - j\right)\right\rangle, \quad (1)$$
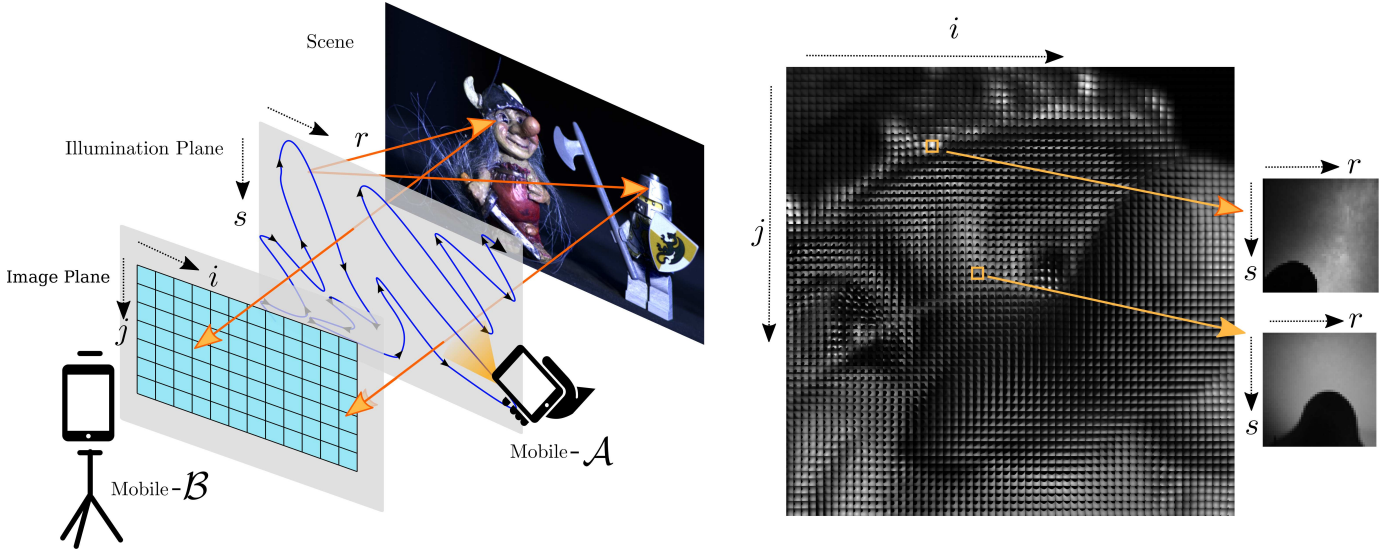
Fig. 2: Our acquisition system: Mobile $\mathcal{B}$ is placed in a stationary position facing the object. Its camera is turned on and captures videos. The user then scans a slice of the reflectance function by moving mobile $\mathcal{A}$, with its LED on, in a random trajectory along a plane. On the right, we show the array of array representation of a small part of the LTM from the Toys Dataset. Here, each sub-array is the light-transport function of the corresponding pixel. In inset, we show two of these sub-arrays.

where Δ_{xy} and ϕ_{xy} are the sampling rate and spatially invariant sampling kernel of the camera, respectively. $\tau_{ij}(r,s)$ denotes the continuous-space light-transport function (abbreviated as LTF) at the surface point corresponding to pixel (i,j) on the camera $\mathcal{B}$. As the LTFs at all 'J' camera pixels undergo the same sampling operation when the user scans with the mobile LED, we focus only on the sampling and reconstruction of a general LTF, $\tau(r,s)$. Let $\mathbf{t}(m,n)$ denote the discretized LTF, obtained by uniformly sampling $\tau(r,s)$ with the sampling rate Δ_{rs} and a box sampling kernel. In the rest of this section, we show how to obtain an approximation $\hat{\mathbf{t}}(m,n)$ of $\mathbf{t}(m.n)$. Our inverse rendering framework is composed of three stages : (B) handheld acquisition (C) resampling step and (D) LTM reconstruction.

B. Handheld Acquisition

The acquisition begins with the user moving the mobile $\mathcal{A}$ along an arbitrary trajectory γ on a chosen illumination plane, with its LED turned on and facing the object to be scanned. The stationary camera $\mathcal{B}$ then observes $\tau(\gamma)$. While lighting the scene, mobile $\mathcal{A}$ also tracks its own position using its camera. Similar to [19], to perform this tracking we use a mobile implementation of simultaneous localization and mapping (SLAM) technique called parallel tracking and mapping (PTAM) [21].

To summarize, the acquisition stage involves the following steps (see Figure 1-(a)):

- Camera $\mathcal{B}$ is placed in a stationary position, with the object of interest in its field of view.
- The user holds mobile $\mathcal{A}$ and marks the object of interest. A bounding box corresponding to the illumination plane to be sampled is shown to the user.
- The user now starts moving mobile $\mathcal{A}$ with its LED on, inside the depicted bounding box. A synchronisation

signal sent to mobile $\mathcal{B}$ starts recording a video. The trajectory that the user takes is tracked using the algorithm PTAM and is displayed on the bounding box in real-time. The user continues until the entire bounding box is spanned.
- Mobile $\mathcal{B}$ outputs a video, each frame corresponding to an observed image. Mobile $\mathcal{A}$ outputs the set of co-ordinates corresponding to the location of the light source on the illumination plane for each frame of the video.

In our proposed acquisition system, the trajectory γ and therefore $\tau(\gamma)$ can only be sampled at non-uniform intervals as a result of the non-uniform rate at which the positions of mobile $\mathcal{A}$ are computed, as determined by its framerate and the PTAM algorithm. The uniform sampling at mobile $\mathcal{B}$ as defined by its video framerate and the varying velocity with which the user moves mobile $\mathcal{A}$ also contribute to the non-uniform sampling pattern. Thus at the end of the proposed acquisition, only samples $\tau(\gamma_r)$ along a non-uniformly sampled trajectory γ_r defined on $[r,s]$ are obtained.

As discussed later, our reconstruction algorithm requires these samples to be distributed uniformly at random across the desired illumination plane. Hence, the user must move on arbitrary trajectories that span the entire illumination plane. To ease this requirement, a user interface displayed on mobile $\mathcal{A}$ indicates the trajectory on the simulated plane in real-time (see inset in Figure 1-(a)). Tracking mobile $\mathcal{A}$ is the only computation that is required at this stage.

C. Resampling Step

After acquisition, the light-transport was non-uniformly sampled, with the samples located in $\mathbb{R}$ on a grid with arbitrary precision, instead of a lattice defined on $\mathbb{Z}$. Since PTAM yields samples at a high rate, we can get an arbitrarily large number
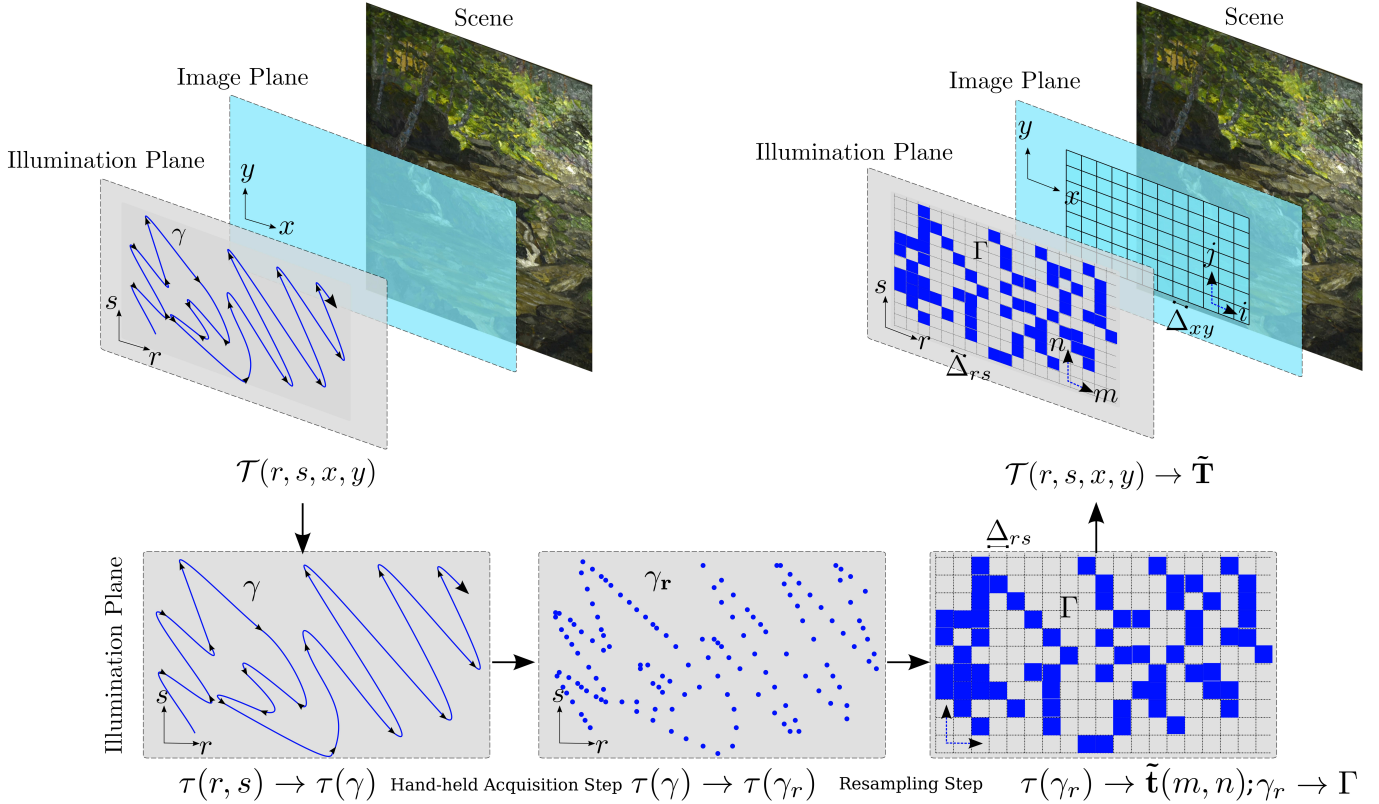
Fig. 3: Our non-uniform sampling model: The light-transport field is sampled by the user by moving mobile $\mathcal{A}$ along a random trajectory γ. This trajectory is discretized by PTAM a discrete trajectory γ_r. This is then resampled into the sampling lattice of the illumination plane to obtain a sparsely sampled LTM
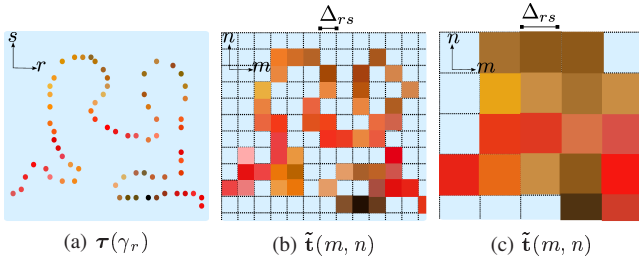

Fig. 4: An illustration of the resampling step. (a) A non-uniformly sampled LTF, with samples in $\mathbb{R}$. (b) and (c) The light-transport function is resampled with different sampling periods Δ_{rs} by binning. Note that a larger Δ_{rs} (c) results in a more densely sampled lattice, albeit at a loss of resolution, compared with the smaller Δ_{rs} (b).

of samples within a small spatial extent, thus making the reconstruction problem quickly untractable. We therefore use the resampling step to project these samples into a discrete lattice. We now resample $\boldsymbol{\tau}(\gamma_r)$ by projecting it onto a lattice defined by a user determined parameter - the sampling period Δ_{rs}. This projection is defined by the binning operation:

$$\tilde{\mathbf{t}}(m, n) = \mathcal{M}(\boldsymbol{\tau}(\gamma_r)W(r - m, s - n)), \qquad (2)$$

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

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

$\mathcal{M}(\cdot)$ is the averaging operator and m, n are indices on the discrete illumination plane. Similarly, the trajectory γ_r can also be projected onto the lattice to obtain an indicator function $\boldsymbol{\Gamma}$ that describes the set of elements in the uniform lattice for which samples have been observed. Let K be the number of elements for which samples were observed and M be the total number of elements in the lattice. As illustrated in the Figure 4, for a given trajectory, the choice of Δ_{rs} determines the sampling ratio $\frac{K}{M}$, as it affects both K and M. A second factor that determines the sampling ratio $\frac{K}{M}$ is the path density of the trajectory, a metric introduced in [20] as the length of continuous-space trajectory γ per unit area. We study how the choice of Δ_{rs} affects reconstruction and inverse rendering in section IV-D.

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

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

D. LTM Reconstruction

Shi et al [22] present a method to reconstruct light fields from a sparse set of samples by exploiting sparsity directly in the continuous frequency domain. In their work, they note that as a result of convolution with the sampling kernel, in the discrete frequency domain a given signal is only approximately sparse (due to sinc tails introduced by the convolution, a large number of co-efficients are close to zero, but not exactly zero). They therefore present a reconstruction method in the continuous domain, by solving a highly non-linear problem using an alternating minimization approach, which depends on a good initialization. While a similar approach can be used to reconstruct the non-uniformly sampled LTFs, we present a method where we trade-off for the approximate-sparsity of the discrete frequency domain, by solving a simple linear problem with guaranteed convergence.

Thus, our goal is to now obtain an approximation, $\hat{\mathbf{t}} \in \mathbb{R}^M$ from $K < M$ random Euclidean measurements, which can be studied under the sparse sampling framework.

It has been shown [23, Theorem 3.2] that when a signal is compressible (approximately sparse) in the transform domain, and when the measurement matrix satisfies Restricted Isometry, compressive sensing recovers the $S-$largest coefficients in the transform domain. Such a reconstruction can be seen as an orthogonal projection of the original signal into the S-bandlimited subspace, thus resulting in a Nyquist-like sampling theorem. Now, if $\mathbf{f}$ denotes the discrete Fourier transform of $\mathbf{t}$ such that

$$\mathbf{f} = \mathbf{H}\mathbf{t}, \tag{3}$$

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

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f}} \|\mathbf{f}\|_1 \quad \text{subject to} \quad \tilde{\mathbf{t}}_K = \mathbf{H}_K^T\mathbf{f}, \tag{4}$$

where $\mathbf{H}_K$ is obtained by sampling the columns of the transform matrix $\mathbf{H}$ using the indicator $\mathbf{\Gamma}$ and $\tilde{\mathbf{t}}_K$ are the K sparse samples from $\tilde{\mathbf{t}}$. It has been shown [23] that when the number of observations K is such that

$$K \geq C \cdot S \cdot \log M, \tag{5}$$

then, Equation (4) yields the S-sparse approximation of $\mathbf{f}_n$. Here C is a small constant independent of M. The discrete LTF can be reconstructed by the inverse transform

$$\hat{\mathbf{t}} = \mathbf{H}^T\hat{\mathbf{f}}. \tag{6}$$

We optimize the $l_1 - l_2$- minimization in Equation (4) by solving for the Least Angles Regression problem [24] using the Sparse Decomposition Toolbox [25]. We reconstruct $\hat{\mathbf{t}}_{ij}$ at all J pixels of the camera. The LTM can now be constructed by first vectorizing the image plane and stacking the discretized LTFs of all J pixels of the camera such that,

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

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

$$\mathbf{y} = \hat{\mathbf{T}}\mathbf{l},$$

where $\mathbf{y}$ is the vectorized image and $\mathbf{l}$ is the vectorized illumination configuration.

IV. EXPERIMENTS AND ANALYSIS

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

A. Evaluation of the Tracking Algorithm

We implemented the acquisition system presented in Section III-B on Android smartphones. As part of our reconstruction framework, we require the PTAM algorithm to give accurate estimates of the discrete trajectory $\gamma_{\mathbf{r}}$ of the light source positions in space. To measure the accuracy of PTAM, we compared its estimated coordinates with ground truth data by placing mobile $\mathcal{A}$ on a computer controlled X-Y table that enables precise displacements on a 2D plane. PTAM accurately and consistently predicted the location of the light source accurately with an acceptable root-mean-square error of less than 1mm (0.31mm). In practice, most inaccuracies are caused by the user moving out of the illumination plane while moving the light source. Both the rotation of the light source and its translation along the z-axis can also induce errors. However, PTAM enables us to track these errors and discard unusable data by using an error margin.

B. Experimental Validation against Ground Truth

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

- Dataset, *Notre Dame-a* is a small but specular region of an oil painting without any downsampling on the image plane. (Figure 5-a.)
- Dataset *Notre Dame-b* is the complete version of the same oil painting, downsampled to 1024×1024 on the image plane. (Figure 5-b.)
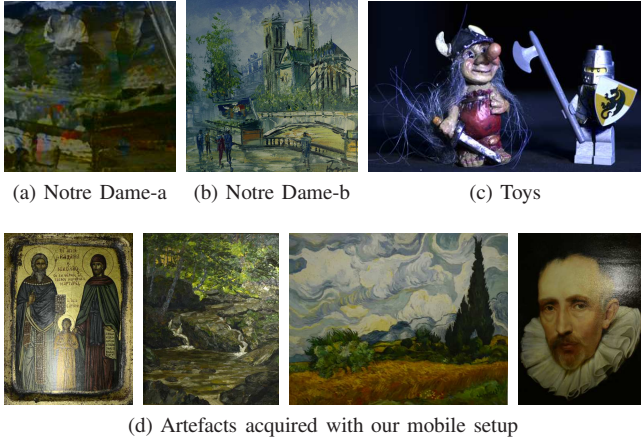- Dataset *Toys* is that of two toys with intricate 3D geometry, containing some specular surfaces. (Figure 5-c.)

(a) Notre Dame-a (b) Notre Dame-b (c) Toys

(d) Artefacts acquired with our mobile setup

Fig. 5: A list of objects acquired with our proposed framework

In our implementation, we chose to move to the YCbCr colorspace and assume that the Cb and Cr channels are illumination invariant. The SNR obtained for each dataset is tabulated in Table I. Here, SNR_Y denotes the signal-to-noise ratio on the luminance channel and is defined as $20\log_{10}\frac{\|\mathbf{T}\|_2}{\|\mathbf{T}-\hat{\mathbf{T}}\|_2}$. It can be seen that our recovery algorithm performs consistently well in all three datasets. In this experiment, the sampling ratio $\frac{K}{M}$ in all three datasets was 0.29. Relighting results are shown in Figure 14 and 15.

Dataset	SNR_Y
Notre Dame-a	29.09 dB
Notre Dame-b	31.88 dB
Toys	20.99 dB

TABLE I: SNR from ground-truth experiments on the three datasets. Sampling ratio $\frac{K}{M} = 0.29$.

C. Effect of Path Density

With 'approximately sparse' signals, compressive-sensing reconstruction recovers only the largest Fourier domain coefficients by truncating their Fourier transform. This results in ringing artifacts around these discontinuities due to the Gibbs phenomenon. The variance of these ringing artifacts can, however, be reduced by increasing the number of recovered components, as is known from classic Fourier theory. In Figure 6, we demonstrate this phenomenon for a single LTF (Figure 6-(a)) that has discontinuities due to both specularity (in the center) and hard shadows. In Figure 6-(b), we show the absolute error in reconstruction for a sampling ratio of 0.04. When this value is increased to 0.08, the reconstruction substantially improves, as seen in Figure 6-(c). Most of the error energy is now concentrated around the specularity and the border of the shadow. Further increasing the sampling ratio to 0.1 and 0.12 results in higher SNR and lower error energies around the discontinuities, as shown in Figures 6-(d) and (e).

The dataset *Toys* consists of large regions with both specularities and hard shadows. Figure 7-(a) shows the per-pixel mean of the absolute error on the luminance channel of the *Toys* dataset, given by $\frac{\langle|\mathbf{T}-\hat{\mathbf{T}}|, \mathbf{1}\rangle}{M}$ obtained for three
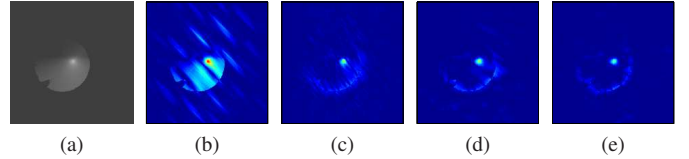

(a) (b) (c) (d) (e)

Fig. 6: The decaying error caused by truncation in the Fourier domain with compressive-sensing. (a) Ground truth LTF exhibiting both specularity and hard shadows. (b)-(e) Absolute error decreases for sampling ratios 0.04, 0.08, 0.1, 0.12, with SNR 14.08dB, 22.07dB, 22.96dB and 25.33dB, respectively. Note that when path density is increased, resulting in a higher sampling ratio, the reconstruction error around the discontinuities decreases.

different simulated trajectories with sampling ratios $\frac{K}{M}$ to be 0.2, 0.29, 0.4. Figure 7-(b) shows the per-pixel variance of the absolute error $|\mathbf{T}-\hat{\mathbf{T}}|$ for the three simulated trajectories. Both the per-pixel mean and variance of the absolute error decrease with increasing path density (and so, increasing $\frac{K}{M}$). The regions in the scene with a large error variance all have a large frequency support in their light-transport function; for example, the sword of the troll and the helmet of the soldier are specular. The knee of the soldier has hard shadows 'cast' upon it. The horn of the troll and the shield of the soldier have hard shadows 'attached' to them. The SNR_y of the entire LTM are 17.899dB, 20.99dB and 21.99dB, respectively, thus indicating a better reconstruction with increasing path densities.

D. Effect of Sampling Period Δ_r and Path Density

The projection of $\tau(\gamma_r)$ onto a uniform lattice to obtain the sparsely sampled $\tilde{\mathbf{t}}(m,n)$ as defined in Equation (2) is an important step in our framework as the choice of Δ_{rs} determines both the sampling ratio $\left(\frac{K}{M}\right)$ and the sparsity (S) of the signal in the transform domain due to asymptotic sparsity [26]. In Figure 8, we show three discrete trajectories with different path densities projected onto uniform lattices of increasing sampling periods. Increasing the sampling period Δ_{rs} results in an increase [1] in the sampling ratio $\frac{K}{M}$.

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

- the Restricted Isometry Principle is satisfied;
- the signal has a sparse representation or is compressible in the selected transform domain;
- the number of samples K (and therefore, the sampling ratio $\frac{K}{M}$) is such that equation (5) is satisfied.

Hence we expect the reconstruction to succeed when K is larger. In Figure 9, the reconstruction accuracy for trajectories with increasing path densities is shown for varying sampling periods. Note that the SNR grows with an increase in both

---

[1]Note that, even when the sampling ratio $\frac{K}{M} = 1$, as long as we sample along trajectories, the projection $\tilde{\mathbf{t}}(m,n)$ differs from a uniformly sampled LTF $\mathbf{t}(m,n)$. When $\frac{K}{M} = 1$, $\tilde{\mathbf{t}}(m,n)$ approaches $\mathbf{t}(m,n)$ as the path density approaches 1 (The trajectory is dense).
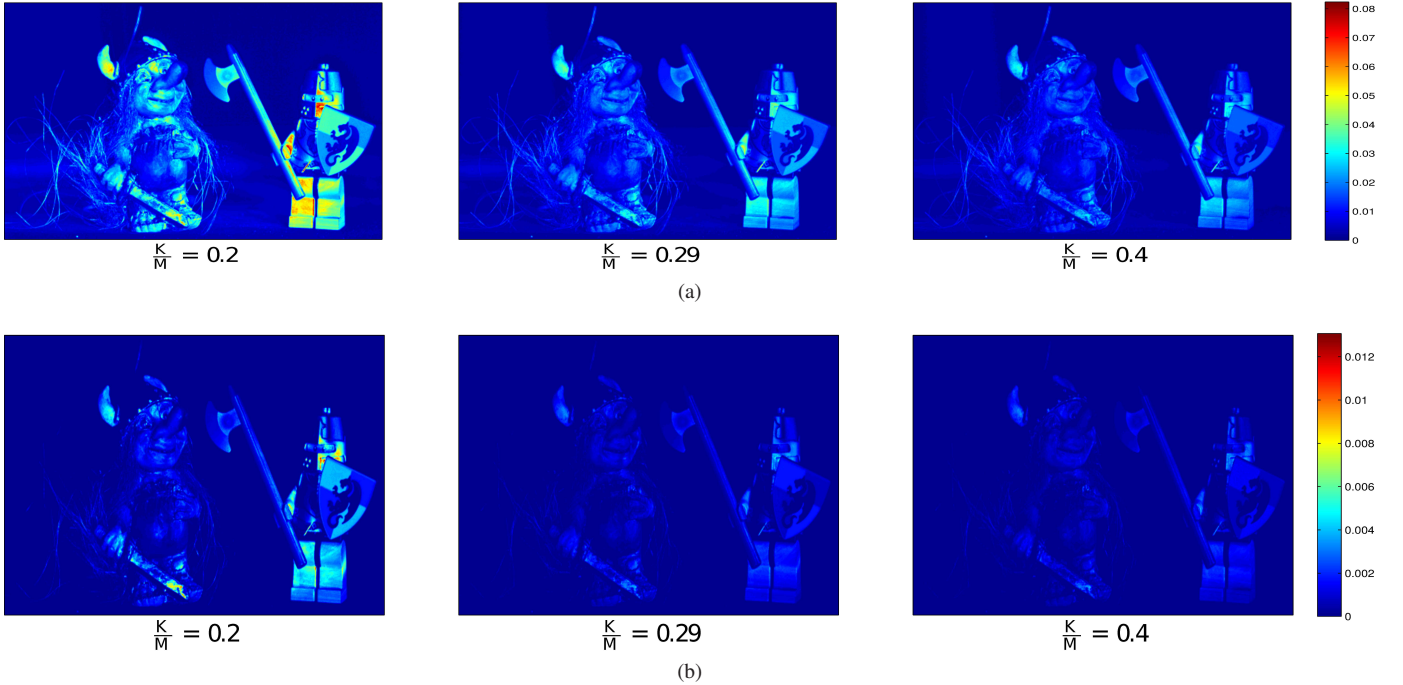
Fig. 7: (a) The per-pixel mean of the absolute errors from the *Toys* dataset is shown for various trajectories. It can be seen that the mean error decreases with increasing sampling ratios. (b) The per-pixel variance of the absolute errors from the *Toys* dataset is shown. Note that the largest error variance occurs around regions exhibiting discontinuities in the form of specularities or hard shadows in their light-transport.

path density and the sampling period, as both tend to increase the sampling ratio.

Although better reconstructions can be obtained by choosing larger Δ_{rs}, a larger sampling interval also results in a smaller resolution in the illumination plane. As can be seen in Figure 9, the reconstruction SNR is higher than $20dB$, even at small Δ_{rs}, as long as the path density is not too small. In general, if the user chooses a trajectory with a very low path density, we could increase reconstruction accuracy by trading off signal resolution, whereas a high resolution reconstruction with reasonable accuracy can be acquired when the user chooses a trajectory with higher path density. The user interface can be modified to indicate this trade-off, enabling the user to choose the trajectory accordingly.

E. Comparison with Other Methods

We also study the improvement in reconstruction with increasing path densities and compare the presented approach with other interpolation strategies such as Bicubic interpolation used in [19], Kernel Regression [27], and Low Rank Matrix Completion [28], as used in [12] [4], albeit in adaptive frameworks. To compare the performance of each reconstruction strategy, we first rendered a scene with considerable surface geometry using PovRay, a ray tracing software. We thus obtained the uniformly sampled LTM **T**. To simulate real-world sampling, we chose various trajectories of increasing path density. Shown in Figure 10 is the average performance of the various approaches for 100 randomly chosen illumination functions from the ray-traced data.

Note that when the trajectory is sparse, the spatial-interpolation operators fail completely, whereas the compressive-sensing based approach consistently yields a relatively good performance. Note that these performances are for sampling along discrete trajectories, whereas for the purely theoretical case of 'uniformly random' samples, all presented algorithms perform significantly better. When the user samples along a dense trajectory, we get good reconstructions irrespective of the choice of the reconstruction algorithm.

In Figure 11, we compare the ray-traced scene under a novel illumination position for the different approaches at the sampling ratio $\frac{K}{M} = 0.12$. As described earlier, our approach has projected the hard shadows present in the scene into a soft shadow. Bicubic interpolation, has resulted in a wrong hard shadow, which inturn will cause visible jumps when rendered under a moving light source. Kernel regression produced significant artifacts in the location of hard shadows.

F. 6-dimensional Light-Transport Matrices

We initially assume a fixed viewing direction, although interactive rendering can involve a change in both changing viewpoints and dynamic illumination. Our reconstruction strategy can be directly extended to incorporate a moving camera and a light source. We simulated an experiment where both mobiles $\mathcal{A}$ and $\mathcal{B}$ are moved to capture the 6D-light transfer, $\mathcal{T}(r, s, x, y, p, q)$, where $[p, q]$ denotes the plane of movement of mobile $\mathcal{B}$. Discretizing $\mathcal{T}(r, s, x, y, p, q)$ by extending the framework presented in this paper, we obtain the

(a) Trajectory 1   (b) $\frac{K}{M} = 0.034$ (c) $\frac{K}{M} = 0.153$ (d) $\frac{K}{M} = 0.609$

(e) Trajectory 2   (f) $\frac{K}{M} = 0.098$ (g) $\frac{K}{M} = 0.431$ (h) $\frac{K}{M} = 1$

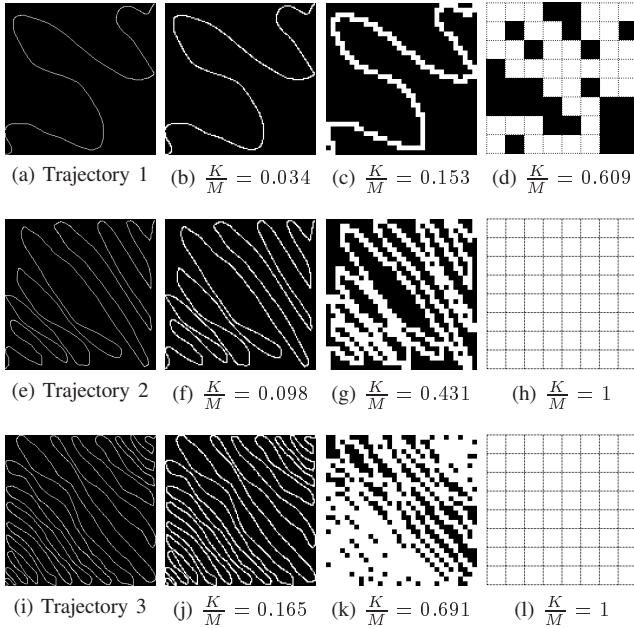(i) Trajectory 3   (j) $\frac{K}{M} = 0.165$ (k) $\frac{K}{M} = 0.691$ (l) $\frac{K}{M} = 1$

Fig. 8: Effect of path density of γ_r and Δ_r on the sampling ratio : $\frac{K}{M}$. The first column in this figure shows three different discrete trajectories γ_r with increasing path densities : (a) 0.012, (e) 0.028 and (i) 0.06 respectively. The next three columns represent projections of these trajectories into uniform lattices of increasing sampling periods : (b), (f), (j) have $2\Delta_r$, (c), (g), (k) have $8\Delta_r$ while (d), (h), (l) have $32\Delta_r$. The figures in columns 2, 3, 4 also show the corresponding sampling ratio.
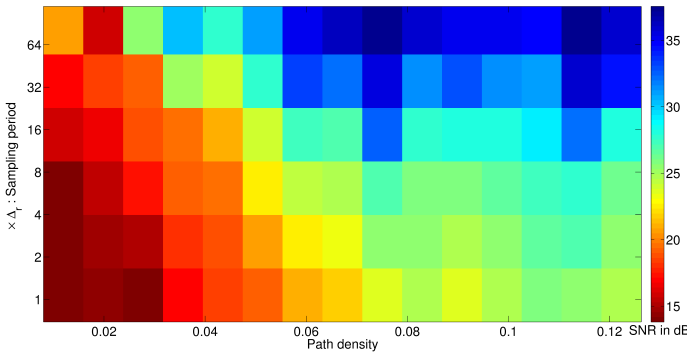

Fig. 9: Choice of sampling period. We plot the SNR of compressive-sensing based reconstructions for various path densities (along the x-axis) and increasing sampling periods (along the y-axis). As expected, the reconstructions are better for increasing sampling periods and path densities. The results shown are the average SNR of 50 experiments.

6D LTM, which allows for image-based rendering with both viewpoints and illumination configurations. This 6D signal can be interpreted either as an array of light fields, with each light field observed under a different illumination, or as an array of LTMs, with each LTM observed under a new viewpoint. With our compressive-sensing based reconstruction, for a sampling ratio of 0.1656, we obtain an SNR of $20.8dB$ for the entire
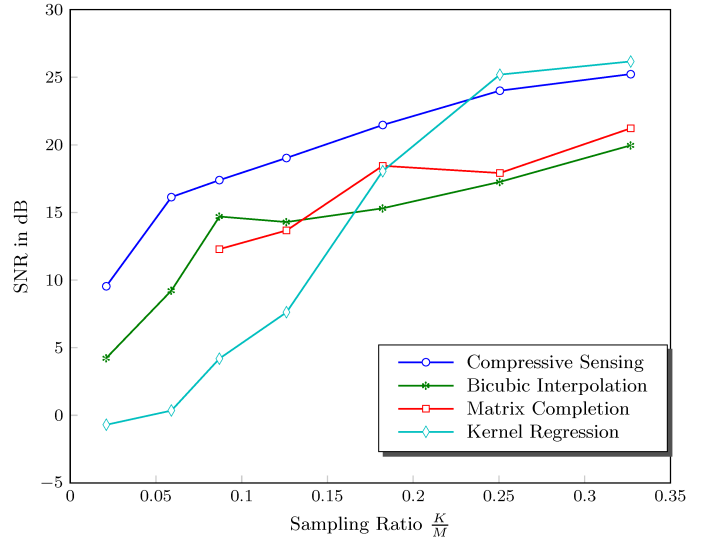

Fig. 10: Comparison of various reconstruction strategies for increasing path densities.

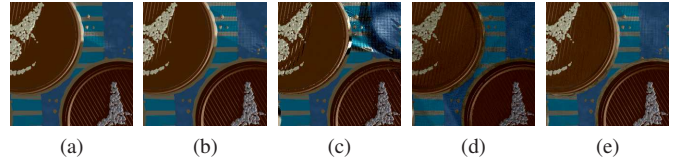(a)   (b)   (c)   (d)   (e)

Fig. 11: Relit images of the raytraced scene used in generating Figure 10, at a sampling ratio of 0.12. (a) Ground Truth image. (b) Bicubic interpolation. (c) Kernel Regression. (d) Matrix Completion. (e) Our Solution. Note that while kernel regression and matrix completion fail completely, bicubic interpolation has produced a hard shadow different from that in the ground truth image. This results in visible jumps in the shadows when rendered under a moving light source. Our solution on the other hand has projected the hard shadow as a soft shadow, resulting in a smoother rendering under a moving light source.

scene, compared to ray traced ground truth data. In Figure 12, we show the per-pixel light-transfer functions, where the x-axis denotes the light positions and the y-axis the camera positions. In Figure 13, we show images rendered from the ray-traced data, along side the ground truth data. Note that both view and illumination changes in the rendered images. The illumination plane resolution was fixed at 25×25 while the resolution of the plane along which the viewer moves was fixed at 16×16. This results in a light transfer function of size 625×256 at each pixel in the image plane.

G. Mobile Based Acquisition

We also tested our reconstruction algorithm using data acquired from our handheld acquisition system implemented on Android smartphones. We obtain the LTMs of three oil paintings and a very specular gold-plated iconograph. We did our acquisition in a room with an ambient light source.
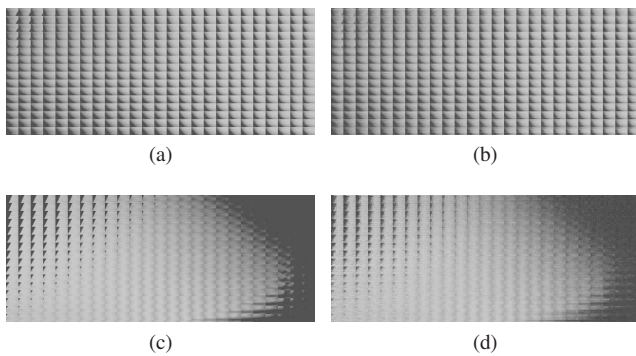
Fig. 12: Per pixel light transfer where the x-axis denotes light positions and the y-axis camera positions. (a) and (c) Ray Traced signals (Ground Truth). (b) and (d) Reconstruction results obtained with compressive-sensing. Per-pixel SNR = $23.1dB$ and $24.2dB$, respectively
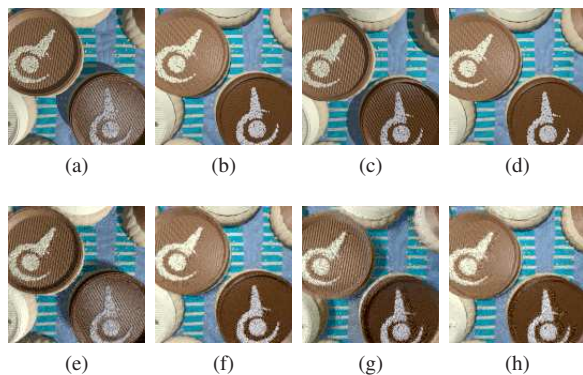

Fig. 13: Images rendered under varying viewpoints and light directions. (a)-(d) Ground truth images. (e)-(h) Reconstructed images. The illumination plane resolution was 25×25, image plane resolution was 128×128 and the viewing plane resolution was 16×16. At each pixel in the image plane, the size of the reconstructed light transfer function is thus 625×256. The overall sampling ratio was 0.165. Note that we chose a low resolution configuration on the illumination plane, in this example to keep the problem computationally tractable.

The presence of the ambient light source aids in a robust tracking system within PTAM. Each acquisition lasted less than five minutes, from setup and initialization of PTAM to completion of sparse sampling. During acquisition, since the user receives visual feedback on Mobile $\mathcal{A}$ about the position of the mobile relative to the illumination plane, less than 1% of acquired samples were discarded due to the user moving out of the illumination plane. Note that both the spectral and intensity profile of the LED varies depending on the mobile phone being used for acquisition. However, we observed that for a given mobile phone, the spectral and intensity profiles do not change over time, and samples acquired in a given acquisition are consistent. The examples presented in this section contain large amounts of glossy-specularities and our algorithm consistently recovers them. Experimental results are shown in Figure 16 to 18 and in Figure 1-(b) to 1-(e). We also present our results in a web-based gallery [29].

H. Discussion

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

Our acquisition stage is fast and requires only a few minutes, but the current implementation of our reconstruction algorithm is computationally expensive. However, since we essentially solve for Equation (4) for J pixels, the reconstruction algorithm is a very good candidate for a parallel implementation on modern General Purpose GPU platforms.

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

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

V. CONCLUSION

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

REFERENCES

[1] S. K. Nayar, P. N. Belhumeur, and T. E. Boult, "Lighting sensitive display," *ACM Trans. Graph.*, vol. 23, no. 4, pp. 963–979, Oct. 2004. [Online]. Available: http://doi.acm.org/10.1145/1027411.1027414

[2] B. Tunwattanapong, G. Fyffe, P. Graham, J. Busch, X. Yu, A. Ghosh, and P. Debevec, "Acquiring reflectance and shape from continuous spherical harmonic illumination," *ACM Trans. Graph.*, vol. 32, no. 4, pp. 109:1–109:12, Jul. 2013. [Online]. Available: http://doi.acm.org/10.1145/2461912.2461944

[3] P. Peers, D. K. Mahajan, B. Lamond, A. Ghosh, W. Matusik, R. Ramamoorthi, and P. Debevec, "Compressive light transport sensing," *ACM Trans. Graph.*, vol. 28, no. 1, Feb. 2009.

[4] M. O'Toole and K. N. Kutulakos, "Optical computing for fast light transport analysis," *ACM Trans. Graph.*, 2010.

[5] D. E. Zongker, D. M. Werner, B. Curless, and D. H. Salesin, "Environment matting and compositing," in *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '99. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1999, pp. 205–214. [Online]. Available: http://dx.doi.org/10.1145/311535.311558

[6] T. Hawkins, J. Cohen, and P. Debevec, "A photometric approach to digitizing cultural artifacts," in *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, ser. VAST '01, 2001.
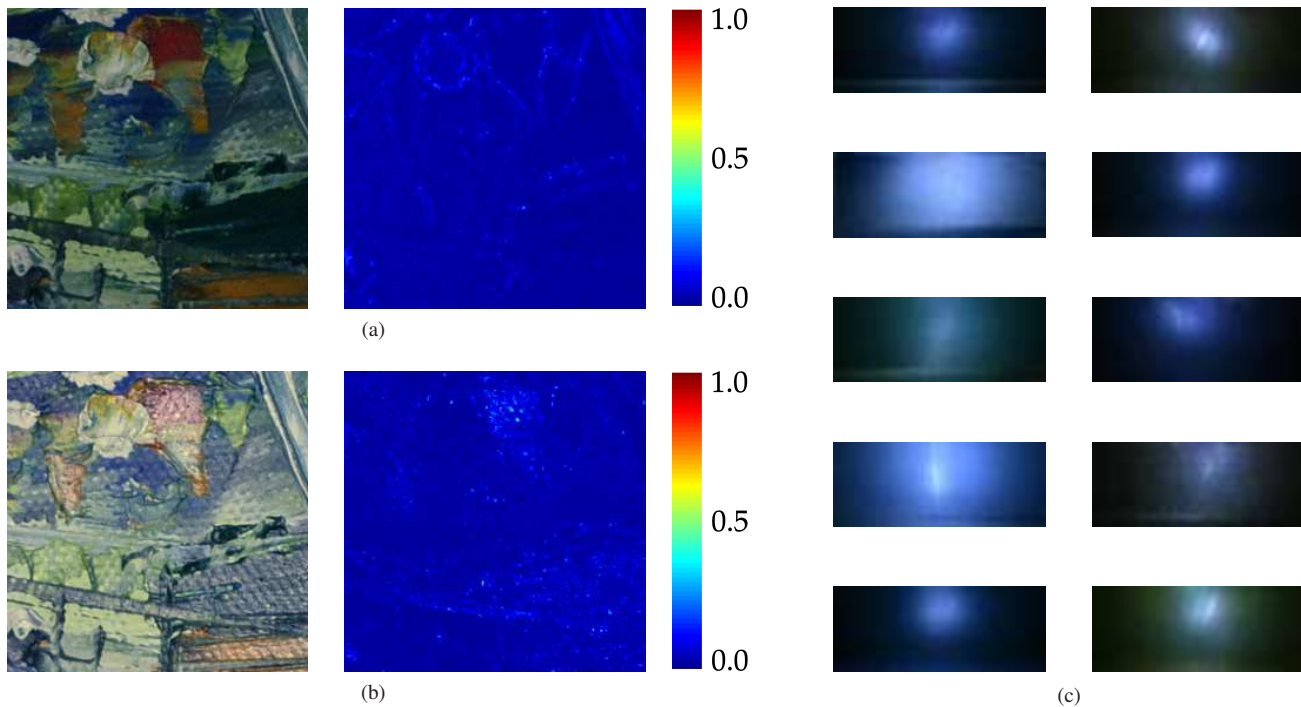
Fig. 14: Dataset *Notre Dame-a* Illumination plane resolution was 32×80. The illumination plane was not resampled. The sampling ratio $\frac{K}{M}$ was 0.29. (a) and (b) Rendered images under novel light source positions and the absolute error images are shown side by side. (c) Some recovered LTFs exhibiting interesting specularities are shown.
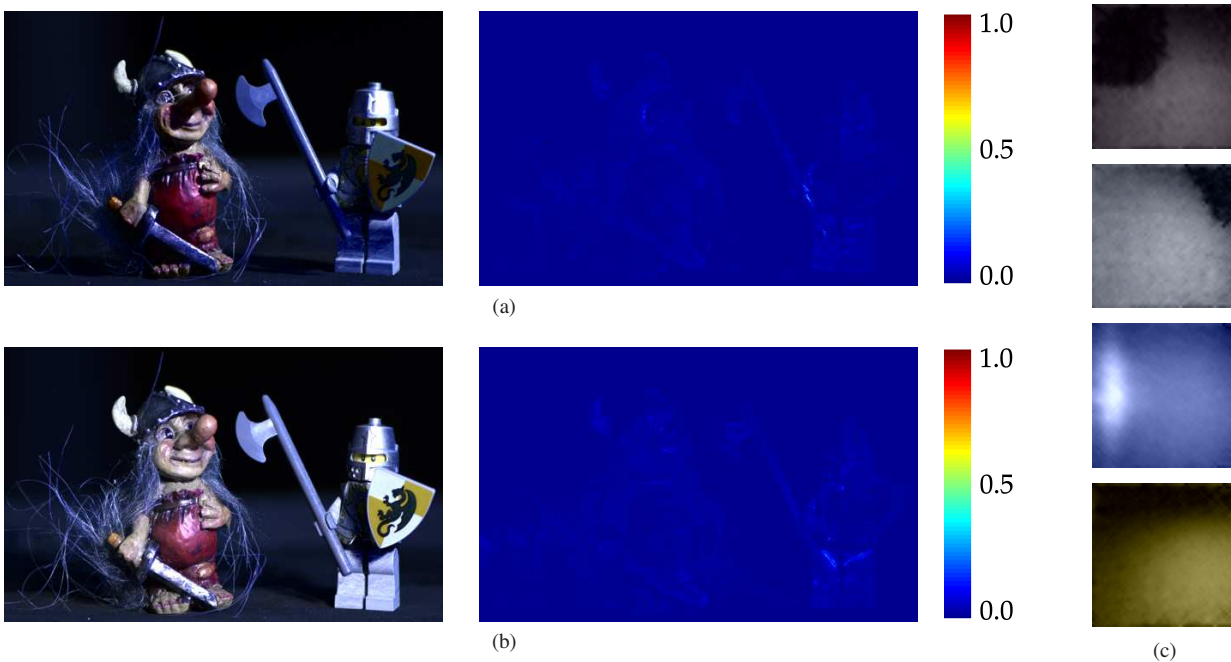

Fig. 15: Dataset *Toys* Image plane has a resolution of 320×512. Illumination plane has a resolution of 50×50. The illumination plane was not resampled. The sampling ratio $\frac{K}{M}$ was 0.29. (a) and (b) Rendered images under novel light source positions and the absolute error images are shown side by side. (c) Some recovered LTFs exhibiting specularities and hard shadows are shown.

Fig. 16: We acquired a highly specular iconograph that has a gold plating, using our hand held acquisition strategy.Image plane has a resolution of 561×421. Illumination plane has a resolution of 75×57. (a)-(c) shows relighted images for three different light positions. (d) shows some randomly selected LTFs exhibiting specular structures.
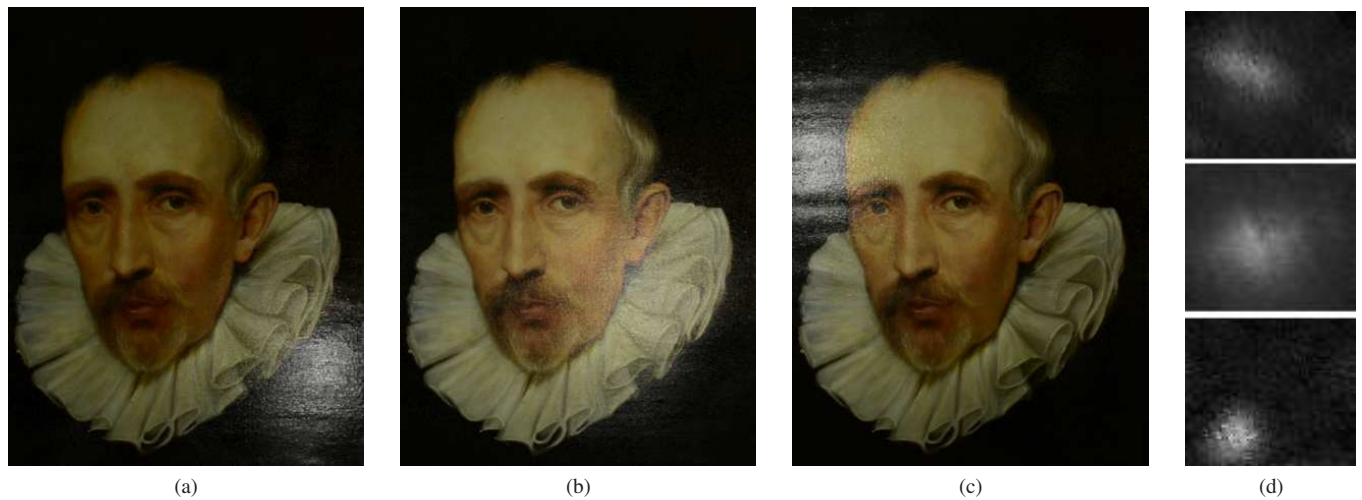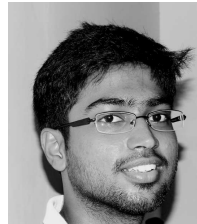

Fig. 17: We acquired a specular oil painting replica of the potrait of Cornelis van der Geest, using our hand held acquisition strategy. Image plane has a resolution of 451×350. Illumination plane has a resolution of 75×65. (a)-(c) shows relighted images for three different light positions. (d) shows some randomly selected LTFs exhibiting specular structures.

[7] T. Malzbender, D. Gelb, and H. Wolters, "Polynomial texture maps," in *In Computer Graphics, SIGGRAPH 2001 Proceedings*, 2001, pp. 519–528.

[8] A. Gardner, C. Tchou, A. Wenger, P. Debevec, and T. Hawkins, "Postproduction re-illumination of live action using interleaved lighting," in *ACM SIGGRAPH 2004 Posters*, ser. SIGGRAPH '04, 2004.

[9] W. Matusik, H. Pfister, R. Ziegler, A. Ngan, and L. McMillan, "Acquisition and rendering of transparent and refractive objects," in *Proceedings of the 13th Eurographics Workshop on Rendering*, ser. EGRW '02. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002, pp. 267–278. [Online]. Available: http://dl.acm.org/citation.cfm?id=581896.581931

[10] Y. Wexler, A. W. Fitzgibbon, and A. Zisserman, "Image-based environment matting," in *Proceedings of the 13th Eurographics Workshop on Rendering*, ser. EGRW '02. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002, pp. 279–290. [Online]. Available: http://dl.acm.org/citation.cfm?id=581896.581932

[11] P. Peers and P. Dutré, "Wavelet environment matting," in *Proceedings of the 14th Eurographics Workshop on Rendering*, ser. EGRW '03. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2003, pp.

157–166. [Online]. Available: http://dl.acm.org/citation.cfm?id=882404.882427

[12] J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo, "Kernel nystrom method for light transport," in *ACM SIGGRAPH 2009 papers*, ser. SIGGRAPH '09. New York, NY, USA: ACM, 2009, pp. 29:1–29:10. [Online]. Available: http://doi.acm.org/10.1145/1576246.1531335

[13] D. Reddy, R. Ramamoorthi, and B. Curless, "Frequency-space decomposition and acquisition of light transport under spatially varying illumination," in *ECCV (6)*, 2012, pp. 596–610.

[14] M. Fuchs, V. Blanz, H. P. Lensch, and H.-P. Seidel, "Adaptive sampling of reflectance fields," *ACM Trans. Graph.*, vol. 26, no. 2, Jun. 2007. [Online]. Available: http://doi.acm.org/10.1145/1243980.1243984

[15] P. Ren, J. Wang, J. Snyder, X. Tong, and B. Guo, "Pocket reflectometry," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 45:1–45:10, Jul. 2011. [Online]. Available: http://doi.acm.org/10.1145/2010324.1964940

[16] V. Masselus, P. Dutré, and F. Anrys, "The free-form light stage," in *Proceedings of the 13th Eurographics Workshop on Rendering*, ser. EGRW '02. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002, pp. 247–256. [Online]. Available: http://dl.acm.org/citation.cfm?id=581896.581928

Fig. 18: We acquired a replica of Van Gogh's 'Wheat Field with Cypresses', using our hand held acquisition strategy. Image plane has a resolution of 427×544. Illumination plane has a resolution of 75×49. (a),(b) shows relighted images for three different light positions. (c) shows some randomly selected largely diffuse LTFs.

[17] M. Aittala, T. Weyrich, and J. Lehtinen, "Two-shot svbrdf capture for stationary materials," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 110:1–110:13, Jul. 2015. [Online]. Available: http://doi.acm.org/10.1145/2766967

[18] P. Ren, Y. Dong, S. Lin, X. Tong, and B. Guo, "Image based relighting using neural networks," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 111:1–111:12, Jul. 2015. [Online]. Available: http://doi.acm.org/10.1145/2766899

[19] A. Davis, M. Levoy, and F. Durand, "Unstructured light fields," *Comp. Graph. Forum*, vol. 31, no. 2pt1, pp. 305–314, May 2012. [Online]. Available: http://dx.doi.org/10.1111/j.1467-8659.2012.03009.x

[20] J. Unnikrishnan and M. Vetterli, "Sampling and Reconstruction of Spatial Fields using Mobile Sensors," *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2328–2340, 2013. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6463464

[21] G. Klein and D. Murray, "Parallel tracking and mapping for small ar workspaces," in *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ser. ISMAR '07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 1–10. [Online]. Available: http://dx.doi.org/10.1109/ISMAR.2007.4538852

[22] L. Shi, H. Hassanieh, A. Davis, D. Katabi, and F. Durand, "Light field reconstruction using sparsity in the continuous fourier domain," *ACM Trans. Graph.*, vol. 34, no. 1, pp. 12:1–12:13, Dec. 2014. [Online]. Available: http://doi.acm.org/10.1145/2682631

[23] E. J. Candès, "Compressive sampling," in *Proceedings oh the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures*, 2006, pp. 1433–1452.

[24] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, pp. 407–499, 2004.

[25] "Spams: a sparse modeling software," (Date last accessed 25-Mar-2017). [Online]. Available: http://spams-devel.gforge.inria.fr/

[26] B. Adcock, P. Univ, A. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing," Tech. Rep., 2013.

[27] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel Regression for Image Processing and Reconstruction," *IEEE Transactions on Image Processing*, vol. 16, pp. 349–366, 2007.

[28] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. on Optimization*, vol. 20, no. 4, pp. 1956–1982, Mar. 2010. [Online]. Available: http://dx.doi.org/10.1137/080738970

[29] "A web-based galley for relightable photographs," (Date last accessed 15-Dec-2015). [Online]. Available: http://go.epfl.ch/relightables

Niranjan Thanikachalam received his Bachelor's degree in Electronics and Communication Engineering from Madras Institute of Technology, Anna University in 2011 and his Ph.D degree in Computer and Communication Sciences from École Polytechnique Fédérale de Lausanne in 2016. During his Ph.D, he worked on computational methods for the digitization of artworks and was part of the eFacsimile project. He is currently a researcher at Rayform, a company specializing in computational caustics. His research interests include inverse problems, computational photography and computational geometry.

Loïc Baboulaz received the M.Sc. degree in 2004 in communication systems from the École Polytechnique Fédérale de Lausanne and the Ph.D degree in electrical engineering from Imperial College London in 2008. He was a research intern at NEC Multimedia Research Laboratories in Japan in 2002 and at Sony Europe Research Laboratories in Germany in 2004. From 2008 to 2011, he was a research geophysicist in multicomponent seismic imaging at CGG UK. In 2012, he led the Google eFacsimile project on Art imaging at the Laboratory of Audiovisual Communications of École Polytechnique Fédérale de Lausanne. Since 2016, Loïc Baboulaz is CTO and co-founder of Artmyn, a company specialised in digitising artworks for interactive and 5D web rendering. His interests include computational imaging, sampling theory and data visualisation.
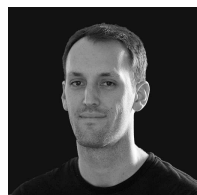
Damien Firmenich specializes in computational photography, more specifically on structure from shading, light-field cameras and near-infrared image processing. He collaborated on the eFacsimile project on light-field acquisition, as well as with the Montreux Jazz Foundation for the digitization of concert videos. Damien holds a MS in communication systems from EPFL and is currently working at Artmyn on artwork digitization.

Sabine Süsstrunk leads the Images and Visual Representation Lab (IVRL) in the School of Computer and Communication Sciences at EPFL since 1999 and is the Director of the Digital Humanities Institute (DHI), College of the Humanities (CdH) since May 2015. Her main research areas are in computational photography, color image processing and computer vision, multimedia, and image quality. She has a BS in Scientific Photography from ETH Zürich, Switzerland, a MS in Electronic Publishing from the Rochester Institute of Technology (RIT), Rochester, NY, USA, and a PhD on computing chromatic adaptation from the School of Computing Sciences, University of East Anglia (UEA) in Norwich, UK.

From 2003-2004, Sabine was a Visiting Scholar in the Computational Color Reproduction Group at Hewlett-Packard Labs in Palo Alto, CA, USA. From 1995-1999, she was the Principle Imaging Researcher at Corbis Corporation in Seattle, WA, USA. From 1991-1995, she was a Visiting Assistant Professor in the School of Photographic Arts and Sciences at the Rochester Institute of Technology (RIT).

Sabine has authored and co-authored over 150 publications, of which 7 have received best paper/demo awards, and holds 10 patents. She served as chair or committee member in many international conferences on color imaging, digital photography, and image systems engineering (General Chair for the IS&T/SPIE Annual Symposium on Electronic Imaging in 2011, Area Chair for IEEE CVPR 2011 and 2016, Area Chair for IEEE ICIP 2008, 2009, 2016, 2017, etc.). She was an Associate Editor for the IEEE Transactions on Image Processing from 2007-2011, the Technical Secretary of the International Color Consortium (ICC) from 2000-2001, Vice-President of IS&T, the Society of Imaging Science and Technology, from 1999-2003 and Conference Vice-President from 2011-2014, Director of CIE Division 8 (Commission Internationale de l'Eclairage, Imaging Technology) from 2007-2011, and member of the ACM E.L. Lawler Award Committee for Humanitarian Contributions within Computer Science and Informatics from 2004-2014.

Sabine is President and member of the board of the EPFL-WISH (Women in Science and Humanities) Foundation. She received the IS&T/SPIE 2013 Electronic Imaging Scientist of the Year Award for her contributions to color imaging, computational photography, and image quality. Sabine is a Fellow of IEEE and IS&T, senior member of ACM, and a member of SPIE and OSA.

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. After his dissertation, he was an Assistant and Associate Professor of Electrical Engineering with Columbia University, New York, 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. From 2013 to 2016 he was President of the National Research Council of the Swiss National Science Foundation and since January 2017 he is President of École Polytechnique Fédérale de Lausanne.

He works in the areas of 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, as well as 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.

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. 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 (20002004), is an ISI Highly Cited Researcher in engineering and a foreign member of the NAE.