Recovering Inner Slices of Layered Translucent Objects by Multi-frequency Illumination

Kenichiro Tanaka, *Student Member, IEEE,* Yasuhiro Mukaigawa, *Member, IEEE,*
Hiroyuki Kubo, Yasuyuki Matsushita, *Senior Member, IEEE,* and Yasushi Yagi, *Member, IEEE*

Abstract—This paper describes a method for recovering appearance of inner slices of translucent objects. The appearance of a layered translucent object is the summed appearance of all layers, where each layer is blurred by a depth-dependent point spread function (PSF). By exploiting the difference of low-pass characteristics of depth-dependent PSFs, we develop a multi-frequency illumination method for obtaining the appearance of individual inner slices. Specifically, by observing the target object with varying the spatial frequency of checker-pattern illumination, our method recovers the appearance of inner slices via computation. We study the effect of non-uniform transmission due to inhomogeneity of translucent objects and develop a method for recovering clear inner slices based on the pixel-wise PSF estimates under the assumption of spatial smoothness of inner slice appearances. We quantitatively evaluate the accuracy of the proposed method by simulations and qualitatively show faithful recovery using real-world scenes.

Index Terms—Descattering, layer separation, image restoration, projector-camera system.

◆

1 INTRODUCTION

SEEING inside a translucent object is a difficult task because its appearance is superposition of light rays emitted from every inner layers that are blurred due to subsurface scattering. Because observing the internal appearance of objects is of broad interest in medical, art, and industrial inspections, various imaging techniques have been developed for achieving this goal in the past. In particular, since the translucency effect becomes significant for many materials in near infrared (NIR) wavelengths, infrared photography is used as one of common techniques for this purpose. For example, it has been used for observing inner layers of oil paintings that tell us a lot about the drawing technique, history, and/or authenticity of old-age painters.

One of the major difficulties in observing inner layers of translucent objects is to separate individual inner appearances with properly dealing with scattering. To overcome this difficulty, we develop a *multi-frequency illumination* method, which can recover sharp appearance of inner slices at a desired depth with explicitly removing scattering blur. Compared with conventional techniques that aim at a similar goal, our method is faster and safer than the X-ray fluorescence technique [1], and sharper results can be obtained differently than infrared reflectography [2].

Our method exploits the fact that the spread of light due to scattering has dependency on the depth of inner layer from which light rays are emitted. By modeling the light spreads as depth-dependent point spread functions (PSFs) and exploiting their depth-dependent low-pass characteristics, we develop a method for recovering inner layer appearances from a set of images taken under a variant of high-frequency illumination [3]. Specifically, our method uses a spatial pattern projection with varying its pattern pitch – we call it *multi-frequency illumination*. Our multi-frequency illumination method allows us to separate direct (high-frequency) and global (low-frequency) components in a similar manner to [3], yet at various frequency levels that define high- and low-frequencies. The observed direction components are then used for recovering the appearance of inner slices, which are related to the direct components via depth-dependent PSFs.

Furthermore, to deal with inhomogeneous translucent objects that exhibit spatial PSF variations due to its inhomogeneity, we develop an extended method for determining the pixel-wise PSFs with an assumption of spatial smoothness of the target depth. In addition, to reduce the visible shadowing effect that appears in the inner layers, which is due to occlusion by their upper layers, we develop a shadow detection method so that the shadowed regions can be later post-processed for better visual quality.

The key contributions of this paper are threefold. First, we describe the relationship between depth inside a translucent object and its PSF by a physically motivated scattering model. Second, based on the relationship, we develop a method for recovering the appearance of inner slices using varying pitch pattern projection. Third, we develop a method for recovering clear inner slices of inhomogeneous objects based on the analysis of spatially non-uniform transmittance. We apply our method to the real-world scenes using a coaxial projector-camera system, and show the effectiveness of the proposed method using oil painting, mural, and paper scenes.

This paper extends its preliminary version [4] with the following major differences: (1) We provide more discussions and details about the PSF of subsurface scattering

---

- *K. Tanaka and Y. Yagi are with the Institute of Scientific and Industrial Research, Osaka University, Japan,*
  E-mail: {tanaka, yagi}@am.sanken.osaka-u.ac.jp,
- *Y. Mukaigawa and H. Kubo are with the Graduate School of Information Science, Nara Institute of Science and Technology, Japan,*
  E-mail: {mukaigawa, hkubo}@is.naist.jp,
- *Y. Matsushita is with the Graduate School of Information Science and Technology, Osaka University, Japan,*
  E-mail: yasumat@ist.osaka-u.ac.jp

Manuscript submitted November 30, 2015.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2631625, IEEE Transactions on Pattern Analysis and Machine Intelligence

blur and its relationship to depth inside the media and its translucency. (2) The proposed technique is further generalized for handling spatially inhomogeneous translucent objects. (3) A post-processing method is developed for detecting the shadowing effect due to occlusions.

The rest of the paper is organized as follows. Section 2 reviews related prior works. Section 3 describes the image formation model for translucent objects and the depth-dependent PSF model that represents blur effects due to subsurface scattering. Based on the model, we describe the proposed method for recovering inner slices of translucent objects in Sec. 4. Section 5 shows results for both simulation and real-world data. Finally, we conclude the paper in 6.

2 RELATED WORK

Our work aims at recovering the appearance of inner slices from superposed observation. There are threads of works for layer decomposition from super-imposed images [5], [6], [7]. Since these methods are designed for semi-transparent layers that do not exhibit multiple scattering, the image formation model is represented by simple alpha blending [8]. Our method, on the other hand, is developed for dealing with translucent scenes where significant multiple scattering is observed, where alpha blending cannot describe the image formation.

In computer graphics area, various image formation models for representing subsurface scattering have been proposed. Donner *et al.* [9] propose a multi-pole model for rendering translucent layered objects such as human skin by extending dipole [10] and Kubelka-Munk theory [11]. d'Eon *et al.* [12] propose an accurate image formation model for translucent objects by decoupling single and multiple scatterings and quantizing the diffusion Green's function [13]. Based on [12], Jakob *et al.* [14] propose a framework of rendering arbitrarily layered materials. These models enable high-quality rendering of translucent objects with multiple scattering. Since they are designed for the forward rendering process, it is not straightforward to directly adopt them in our backward inverse problem, *i.e.*, recovering appearance of inner layers of translucent objects, due to the computational complexity. In this work, we use a simplified version of Kubelka-Munk theory [11] for the image formation model, which makes the inverse estimation problem tractable while retaining the expression ability.

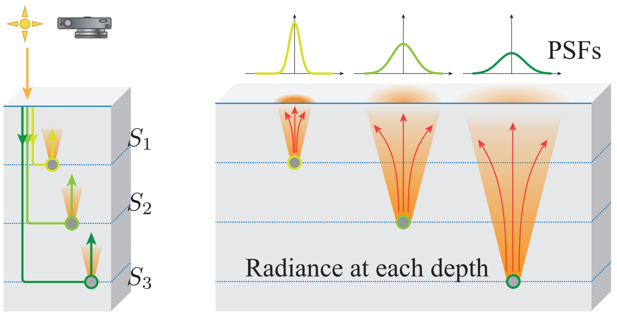
Our work is also related to imaging through scattering/occlusion methods in computational photography. Levoy *et al.* [15] combine a synthetic aperture technique [16] in remote sensing and a confocal imaging technique [17] in microscopy to reconstruct images of scenes behind occlusion or inside scattering medium. Fuchs *et al.* [18] also use the confocal imaging for the purpose of recording a solid target in scattering media. Because these methods are based on confocal imaging, for which both the camera and projector need exact focusing on the same depth plane, it is needed to have a priori knowledge about the target depth. For imaging through scattering media, Narasimhan *et al.* [19]'s and Gu *et al.* [20]'s methods sharpen images of a target scene in muddy liquid by precisely modeling single scattering. Their methods work well for those scenes that do not exhibit multiple scattering. Differently from these works, our method recovers images of inner slices of a translucent object, where significant multiple scattering is observed and the optical thickness of the target is unknown.

Our method can be grouped in a class of active sensing techniques that use high-frequency pattern projection. The original Nayar *et al.* [3]'s method and its extended methods [21], [22], [23], [24], [25], [26] separate direct and global components by projecting multiple high-frequency patterns. Reddy *et al.* [27] separate light transport into direct, near-indirect, and far-indirect rays by frequency-domain modeling and analysis. Our method is also based upon Nayar's method [3] and we use a relationship among direct components, the size of scattering blur, and the pitch of the projection pattern to separate depth layers. There are other pattern projection techniques to decompose light transport. Gupta *et al.* [28] acquire scene depths with direct-global separation by modeling both projector's defocus and global light transport. O'Toole *et al.* [29], [30], [31] illuminate a scene by a pattern while masking the camera by the complementary pattern to spatially probing the light transport of the target scene. Our method also uses a pattern projection technique not only for separating scattering effects, but also for specifically recovering appearance of multiple inner slices.

Time-domain coding for analyzing light transport is another approach to recover images that are not directly measurable. Heide *et al.* [32] sweep the modulation frequency and phase of their customized Time-of-Flight (ToF) camera to recover the light propagation inside scattering medium. Kadambi *et al.* [33] build a coded-illumination ToF camera with a deconvolution technique and use it for recovering a sharp image by observing through a diffuser. O'Toole *et al.* [34] combine spatial probing and ToF imaging to separate direct and indirect light-in-flight images. Tadano *et al.* [35] propose an imaging system that is capable to select a target depth using a coded ToF camera. While these methods effectively recover light transport, they require carefully tailored ToF cameras. Contrary, our method uses a simple projector-camera system and spatial pattern coding to analyze light transport inside translucent objects.

Methods for measuring transparent or translucent media, such as smoke scanning, are also related to our work. Morris *et al.* [36] recover the shape of clear transparent objects that refract light by recording light rays from different viewpoints. Hawkins *et al.* [37] acquire the density distribution of participating media, such as smoke, by laser scanning. Ihrke and Magnor [38] reconstruct the volume of dynamic semi-transparent media by a visual-hull based method. While related, our goal is to recover slices inside translucent objects instead of reconstructing 3-D shape of an object's surface.

There have been independent developments of technologies for imaging internal structures of target objects for special purposes. In art analysis, several techniques have been developed for imaging hidden layers of paintings. Infrared reflectography [2] and X-ray transmission radiography [1] have been used for visualizing internal layers of paintings, although the surface texture cannot be separated. X-ray fluorescence technique [1] uses spectroscopic information measured over a couple of days and estimates the metallic atom distribution for determining colored appearance of



(a) (b)

Fig. 1: Illustration of the image formation model for translucent objects. (a) Recorded intensity is the summation of all layer's appearance. (b) Spatial spread of light varies with depth.

inner layers. Tera-hertz imaging [39] is another technique that can see inner layers of paints. In the medical imaging and its related areas, optical coherence tomography [40], [41] techniques are widely used, especially for visualizing retina. These techniques enable to observe inside translucent objects based on interference of coherent light at the cost of expensive wavelength-order optics and mechanics. In contrast, our method uses a commodity camera and projector for recovering slices inside translucent objects, which allows low-cost implementation. In microscopy, there are methods that use pattern projection for visualizing inside substances, or reconstructing 3-D shape of a small sample, such as protein structure [42], [43], [44]. They sharpen the microscopic image by precisely taking into account both micro-scale optical blur and scattering blur. Our aim is to develop a technique that is applicable to a standard scale, where scattering blur becomes more significant.

3 APPEARANCE OF TRANSLUCENT OBJECTS

When an image of a translucent object is recorded, the observed intensity $L_o(c)$ at camera pixel $c \in \mathbb{Z}^2$ is a summation of all possible paths $\rho \in \mathcal{P}_t$ from depth t:

$$L_o(c) = \int_0^\tau \int_{\mathcal{P}_t} \iota(\rho) d\rho dt, \tag{1}$$

where τ is the maximum thickness of the object, and $\iota(\rho)$ is the intensity of the light path ρ. It can be modeled as a summation of the appearance of all depth slices as illustrated in Fig. 1(a), which can be described as

$$L_o(c) = \int_0^\tau S_t(c) dt, \tag{2}$$

where S_t is the appearance slice at depth t.

The appearance slice S_t is generally blurry due to the scattering effect inside the medium. The spread of radiance at a scene point inside a translucent object varies depending on its depth from the object surface [45]. In general, the spatial spread of light can be expressed using PSFs. Let us consider light rays emitted (or returned) from a specific depth inside a translucent object. When the depth is shallower, the PSF becomes sharper. It gradually wider spreads as the depth t becomes deeper inside the medium as illustrated in Fig. 1(b). In this manner, there is a close relationship between the PSF and depth. By denoting h_t as a PSF at depth t, the appearance slice S_t can be expressed as

$$S_t(c) = (R_t * h_t)(c), \tag{3}$$

where R_t is the sharp appearance slice that we are interested in estimating, which we call a radiance slice, and $*$ denotes a convolution operator.

Since the appearance of the translucent object under normal illumination is a superposition of radiance of multiple slices as Eq. (2), the observation L_o can be re-written as

$$L_o(c) = \int_0^\tau (R_t * h_t)(c) dt. \tag{4}$$

Specifically, we are interested in recovering a few planar parallel layers inside the medium; therefore, the image formation model can be discretized as a multiple layered model as

$$L_o(c) = \sum_d (R_d * h_d)(c), \tag{5}$$

where R_d and h_d are the radiance slice and depth-dependent PSF of d-th layer, respectively. Our goal is recovering radiance slices R_d convolved by unknown depth-dependent PSFs h_d from the composite observation L_o. Before introducing the solution method, we describe a model of depth-dependent PSFs h_d.

Depth-dependent PSFs

Our depth-dependent PSF model is motivated by the work of [46], which is a physically motivated scattering model designed for the PSF of scattering medium, named the *radiative transfer equation* (RTE) [47], [48]. In this work, we represent the depth-dependent PSFs h_d by the RTE model. With the RTE model, the intensity of an arbitrary light ray in a homogeneous scattering medium can be iteratively calculated [46] as

$$I(d, \theta) = \sum_{k=0}^{\infty} \left(g_k(T) + g_{k+1}(T)\right) L_k(\cos\theta), \tag{6}$$

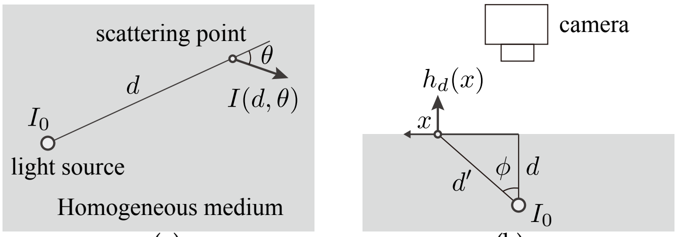
where L_k is the Legendre polynomial of order k, $g_0 = 0$, and

$$g_k(T) = I_0 \exp\left(-\frac{2k+1}{k}(1 - q^{k-1})T - (k+1)\log T\right).$$

The parameter T $(= \sigma d)$ represents optical thickness, which is the product of scattering coefficient σ and distance d between the point light source and the scattering point. The forward scattering parameter q controls how light rays spread; $q = 0$ corresponds to isotropic, while positive and negative values of q indicate forward and back scattering, respectively. θ is the angle between the depth axis and light ray's direction as depicted in Fig. 2(a). k is the number of light bounces, and I_0 is the intensity of the point light source.

In our setting, a camera is placed outside the media, where light rays do not scatter. In this case, the depth-dependent PSF applied to a point light source inside scattering media corresponds to the intensity of the light rays emitted from the surface, if the camera is placed sufficiently far from the object. When we consider that the x-axis lies

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2631625, IEEE Transactions on Pattern Analysis and Machine Intelligence



(a) (b)

Fig. 2: Illustration of RTE model described in [46]. (a) Solid arrow represents a scattering light ray, whose intensity depends on d and θ. d is the distance from the point light source I_0 to the point inside the scattering medium, and θ is the radial direction of the light ray. (b) Depth-dependent PSF h_d can be expressed using RTE model. We consider the x-axis lies along the tangent plane of the surface of translucent object and assume the direction of light rays emitted from the surface becomes parallel to the depth axis.

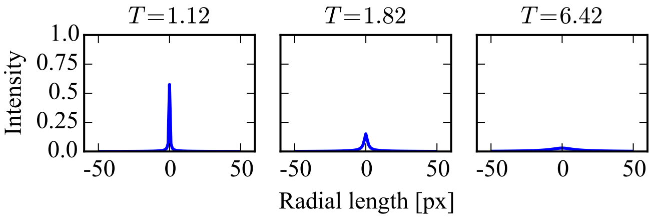


Fig. 3: Simulated PSF variations using the RTE model with varying optical thickness T

along the tangent plane of the surface of translucent object as illustrated in Fig. 2(b), the depth-dependent PSF $h_d(x)$ can be written as

$$h_d(x) = I\left(d', \phi\right), \tag{7}$$

where $d' = \sqrt{x^2 + d^2}$ and $\phi = \tan^{-1}(x/d)$. We assume that the camera is placed sufficiently far from the object compared to both d and x so that the direction of emitting light rays from the surface becomes parallel to the depth axis. In addition, we ignore the refraction of the surface for simplicity. Figure 3 shows simulated PSFs by changing the optical thickness T. When T is small (thin scattering media), a sharply pointed PSF is generated, and as T becomes larger (thicker scattering media), the PSF spatially spreads wider.

These PSFs only represent the spread of light but dot not consider the absorption by the medium. When the absorption is homogeneous, it only makes the deeper layers darker and has no effect on the sharpness of the layer recovery except for the siganl-to-noise ratio (SNR), hence it can be practically neglected. However, if the absorption is inhomogeneous, which is the case in inhomogeneous translucent objects, the radiance slice of inner layer exhibits artifacts if the spatial variation of PSFs is ignored. In this work, we explicitly take into account the spatially-varying PSFs in Sec. 4.3 for dealing with inhomogeneous translucent objects.

4 PROPOSED METHOD

We are interested in recovering radiance slices R_d from the mixed observation L_o. To achieve this goal, we develop a *multi-frequency illumination* measurement method, which is built upon the high-frequency illumination (HFI) method proposed by Nayar *et al.* [3]. To begin with, we briefly review the original HFI method.

High-frequency illumination method [3]

The HFI method separates direct and global components by projecting small pitch checker patterns. When the phase of the projection pattern changes slightly, the direct component $D(c)$ varies accordingly, but the global component $G(c)$ remains stable. Based on this observation, their method computes direct and global components using the maximum $L_{\max}(c)$ and minimum $L_{\min}(c)$ intensities that are obtained by shifting the projector pattern as

$$\begin{cases} D(c) = L_{\max}(c) - L_{\min}(c), \\ G(c) = 2L_{\min}(c). \end{cases} \tag{8}$$

The direct component $D(c)$ contains high-frequency components, while the global component $G(c)$ contains only lower frequency components than the frequency of projection pattern. Therefore, the HFI method can be viewed as a separation technique for high- and low-frequency components.

Pattern pitch of HFI

In our case, when a translucent object is measured under HFI with pattern pitch p, we can obtain direct component $D(p,c)$ and global component $G(p,c)$ as

$$\begin{cases} D(p,c) &= \sum_d D_d(p,c) \\ G(p,c) &= \sum_d G_d(p,c) \\ S_d(c) &= D_d(p,c) + G_d(p,c), \end{cases} \tag{9}$$

where $D_d(p,c)$ and $G_d(p,c)$ are the direct (high-frequency) and global (low-frequency) components at depth d, respectively, and the sum of direct and global components for each depth becomes the radiance slice. As mentioned in [3], the pattern pitch p must be sufficiently smaller than the scene texture for a faithful separation.

We have observed that the direct and global components vary when the projected pattern pitch p changes. Indeed, there is a tight relationship between the pitch size and strength of direct component. Suppose we measure the scene with two distinct pattern pitches p_i and p_j ($p_i < p_j$) independently. When the pattern pitch becomes larger, the wider spread of light rays are included in the direct component, hence the direct components have the following relationship:

$$D_d(p_i, c) < D_d(p_j, c), \quad p_i < p_j. \tag{10}$$

This relationship indicates that the separation frequency varies with the pattern pitch – wider frequency band is regarded as the direct component as the pattern pitch becomes larger. As low-pass characteristics of depth-dependent PSFs are also depth-dependent, the difference of direct components $D_d(p_j,c) - D_d(p_i,c)$ also varies with depth d. Our method exploits these differences for recovering the appearance of each inner slice by changing the projection pattern pitch.

4.1 Multi-frequency illumination

By measuring the target object using multi-frequency patterns, multiple of corresponding direct components are obtained. Unfortunately, increasing the number of measurements does not make the problem easier as it also increases the number of variables to solve for. To make the problem tractable, we assume that the texture of direct components does not vary drastically when the pattern frequency is high enough and the pitch variation $(p_j - p_i)$ is sufficiently small. These direct components D_d at a certain depth d are supposed to have a similar texture with the original radiance; therefore, we can expect the following relationship:

$$D_d(p, c) \approx \alpha(h_d, p) R_d(c), \tag{11}$$

where $\alpha(h_d, p)$ is the relative brightness of $D_d(p, c)$ to $R_d(c)$. We call $\alpha(h_d, p)$ the *direct component ratio* that represents the ratio of direct component's mean intensity to the radiance $R_d(c)$'s mean intensity. Hence, Eq. (9) can be rewritten as

$$D(p, c) = \sum_d \alpha(h_d, p) R_d(c). \tag{12}$$

These assumptions are based on the fact that the diffuse reflection and subsurface scattering can be regarded as the same physical phenomena [49], [50]; the light scatters on or beneath the surface and eventually bounces off of the material in random directions. Direct components D_d represent total intensities of lights from all the points inside the object whose distance from the incident point on the surface is smaller than the pattern pitch p. Hence the pattern pitch p controls the *scale* of the separation of scattered lights in the direct-global separation scheme, and thus controls the intensity of direct components. Furthermore, because p is sufficiently smaller than the scene texture, the texture of direct components and the original texture are largely similar. Based on these observations, we obtain the original texture at different brightnesses by changing p.

With these assumptions, a set of direct component images $D(p, c)$ taken under the multi-frequency illumination of m pitch variations $(p = p_1, p_2, \ldots, p_m)$ can be written in a matrix form as

$$\mathbf{D}(c) = \mathbf{A}\mathbf{R}(c), \tag{13}$$

where

$$\begin{aligned}
\mathbf{D}(c) &= \begin{bmatrix} D(p_1, c) & D(p_2, c) & \cdots & D(p_m, c) \end{bmatrix}^T, \\
\mathbf{A} &= \begin{bmatrix} \alpha(h_{d_1}, p_1) & \cdots & \alpha(h_{d_n}, p_1) \\ \vdots & \ddots & \vdots \\ \alpha(h_{d_1}, p_m) & \cdots & \alpha(h_{d_n}, p_m) \end{bmatrix}, \\
\mathbf{R}(c) &= \begin{bmatrix} R_{d_1}(c) & R_{d_2}(c) & \cdots & R_{d_n}(c) \end{bmatrix}^T.
\end{aligned}$$

Here, $\mathbf{D} \in \mathbb{R}^m$ is a vector of direct components measured under m variations of the pattern pitches at pixel c, $\mathbf{R} \in \mathbb{R}^n$ is a vector of n layers of radiance slices, and $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a matrix containing direct component ratios computed from the projected pattern pitch and the depth-dependent PSF.

When the number of projected patterns m is no less than the number of depth layers n $(m \geq n)$ and $\mathrm{rank}(\mathbf{A}) = n$, the radiance slices $\mathbf{R}$ can be obtained by a norm approximation of the residual vector, *i.e.*, $\mathbf{D}(c) - \mathbf{A}\mathbf{R}(c)$. For example, with a least-squares approximation, the radiance slices $\mathbf{R}(c)$ can be determined using the pseudo-inverse $\mathbf{A}^+$ as

$$\mathbf{R}(c) = \mathbf{A}^+ \mathbf{D}(c). \tag{14}$$

Computation of direct component ratio

The direct component ratio $\alpha(h_d, p)$ can be derived from the depth-dependent PSF h_d and the projected pattern pitch p. When a checker pattern is projected to a translucent object, it reaches the depth d with some blur effect and returns to the surface with the additional blur effect. To obtain the direct component ratio, we consider the difference between maximum and minimum intensities in a similar manner to the original HFI as :

$$\begin{aligned}\alpha(h_d, p) = &\max\left((l_p * h_d) * h_d\right) \\ &- \min\left((l_p * h_d) * h_d\right),\end{aligned} \tag{15}$$

where l_p is normalized projection patterns, whose pitch is p. The normalized illumination is defined in the range between 0 and 1; 0 being black and 1 being white. The operators $\max$ and $\min$ return the maximum and minimum value from all pixels, respectively.

4.2 Estimation of informative slices

Once we know the target depths (or PSFs) to inspect, we can set up a matrix $\mathbf{A}$ using Eq. (15), and thus can recover slices corresponding to the depths using Eq. (14). However, such a prior knowledge is difficult to obtain before measurement; therefore, automatically selecting a good set of depths becomes important for recovering *informative* slices. For example, if an arbitrary depth is chosen, it has a chance to correspond to the middle of distinct texture layers. To recover informative slices, we use a two-step approach. The first step is the estimation of a set of informative depths via optimization. This is equivalent to selecting a small number of useful PSFs from many other possible PSFs. The second step is the recovery of slices at the informative depths determined by the earlier step. The overall procedure of this strategy is illustrated in Fig. 4. Now, we explain the details of each step.

Step 1: Estimation of informative depths

Elements of matrix $\mathbf{A}$ depend on depth-dependent PSFs; hence estimation of informative depths corresponds to estimating the shape of PSFs. The candidate relative depth T, which determines the shape of the PSFs, are set in the range between $1 + \epsilon$ and $10 + \epsilon$ with a step size 0.05. We add the offset ϵ $(= 0.01)$ because Eq. (6) does not converge if $T \leq 1$. Initially, we set up the matrix $\mathbf{A}$ using all candidate parameters T (thus $m < n$) in order to estimate informative slices.

Frequently, there are only a small number of informative slices inside translucent objects. Hence we can regard such radiance slices exist sparsely along depth. Our method uses this sparsity to determine the informative slices by solving a l_1 regularized problem (as known as the lasso [51]) with a non-negative constraint about $\mathbf{R}$:

$$\begin{aligned}\hat{\mathbf{R}}(c) = &\operatorname*{argmin}_{\mathbf{R}(c)} \left\|\mathbf{A}\mathbf{R}(c) - \mathbf{D}(c)\right\|_2^2 + \lambda \left\|\mathbf{R}(c)\right\|_1 \\ &\text{subject to } \mathbf{R} \succeq 0.\end{aligned} \tag{16}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2631625, IEEE Transactions on Pattern Analysis and Machine Intelligence

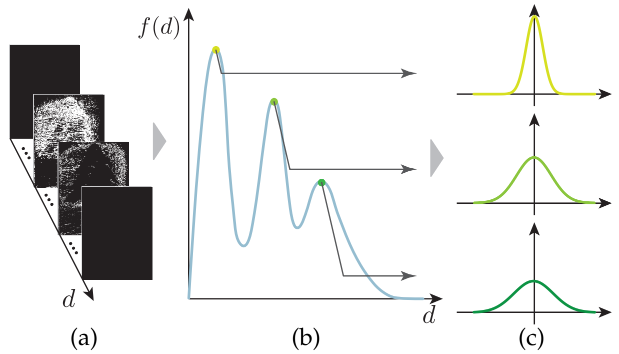


Fig. 4: Selection of informative PSFs (corresponding to certain depths). (a) Estimates $\hat{R}_d$ obtained via optimization (Eq. (16)). Non-zero pixels indicate the informative regions. (b) We find the local maxima of the non-zero pixel counts (corresponding to l_0-norm of $\hat{R}_d$) for all pixel coordinates across depth. (c) Selected PSFs corresponding to the local maximas

We can regard the depth d $(= \frac{T}{\sigma})$, where $\hat{R}_d(c)$ has a non-zero value, is informative while others are not. Equation (16) becomes a quadratic programming (QP) problem and thus can be efficiently solved in a polynomial time. We solve the optimization in a per-pixel manner. Solving a similar problem for the entire image at a time instead of computing in a pixel-wise manner is also a viable option; however, we have observed that they do not make much difference because of the following Step 2. Therefore, for efficient parallelization, we choose the per-pixel implementation.

Step 2: Informative slice recovery

This step determines informative depth slices of the whole image by consolidating all the pixel-wise selections. The informative depths $\hat{\mathbf{d}}$ are local maximas of the sum of l_0 norm of $\hat{R}_d(c)$ for all pixels. The evaluation function $f(d)$ is defined as

$$f(d) = \sum_{c} \|\hat{R}_d(c)\|_0, \tag{17}$$

and we find all local maxima of $f(d)$ along d in the rage of interest as shown in Fig. 4(b). Once the depths of interest are selected, we can set up a small matrix $\hat{\mathbf{A}}$. Finally, the appearances of informative slices are recovered using the matrix $\hat{\mathbf{A}}$ using the least-squares approximation as

$$\mathbf{R}(c) = \hat{\mathbf{A}}^{+}\mathbf{D}(c) \tag{18}$$

in the same manner to Eq. (14). We call the method described so far a *baseline method*.

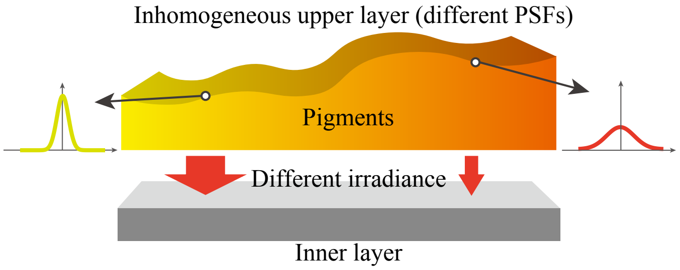


Fig. 5: Scene of inhomogeneous upper layer. PSFs and irradiance of inner layer vary depending on the upper layer's optical property.

4.3 Recovering inner slices inside inhomogeneous upper layer

Now we describe an extension of the baseline method for dealing with inhomogeneous translucent objects. When the upper layer is inhomogeneous, due to either different materials or thicknesses, or even both, recovered inner slices by the baseline method are affected by the upper layer's non-uniform scattering and transmission. Because of the spatially non-uniform scattering and absorption property, both PSF and irradiance of the inner layer vary even if the inner layer is homogeneous as illustrated in Fig. 5. Under this situation, the baseline method suffers from artifacts because it assumes a single PSF per slice for the entire image region.

To deal with the inhomogeneous upper layer, we employ a pixel-wise PSF selection approach. By applying Step 2 of the baseline method in a sliding window manner, a set of PSFs can be obtained for each pixel. We apply a 2D Gaussian-weighted filter for the sliding window for suppressing high-frequency observation noise. The evaluation function $f_c(d)$ for pixel location c is defined as

$$f_c(d) = \sum_{c_n \in W_c} \mathcal{N}_{\sigma_f}(\|c_n - c\|_2) \left\| \hat{R}_d(c_n) \right\|_0 , \tag{19}$$

where $\mathcal{N}_{\sigma_f}$ is a zero-mean Gaussian distribution with standard deviation σ_f centered at c, and c_n is a pixel location in the window W_c. By selecting the informative depths for each pixel via finding the local maxima of f_c, we can recover pixel-wise informative slices in a similar manner to Eq. (18) as

$$\mathbf{R}(c) = \hat{\mathbf{A}}^{+}(c)\mathbf{D}(c), \tag{20}$$

where $\hat{\mathbf{A}}^{+}(c)$ is the pseudo-inverse of a small matrix containing direct component ratios computed for pixel c using the selected PSFs.

The recovered inner slices may still contain the brightness discontinuity due to the non-uniform transmission, because PSFs considered so far neglect absorption by the medium. Inhomogeneous transmittance causes inhomogeneous irradiance on the inner layer, hence the recovered inner layer may have brightness variations. Indeed, this brightness discontinuity stands out at the boarder of distinct PSFs. In this work, we adopt a gradient-based image repairing approach to reduce the brightness discontinuity. Specifically, we suppress the gradient of edges at the brightness discontinuity in the inner with retaining other edges so that a cleaner appearance of the inner slice can be obtained. Such edges $\mathcal{E}$ can be identified by observing the change of PSF estimates because the discontinuity of transmission occurs at the boarder of different materials or thicknesses in the upper layer. Hence, the brightness corrected slice R'_d can be recovered by setting the gradient of such edges to zero as

$$\nabla R'_d(c) = \begin{cases} \mathbf{0} & \text{if } \ c \in \mathcal{E} \\ \nabla R_d(c) & \text{otherwise}, \end{cases} \tag{21}$$

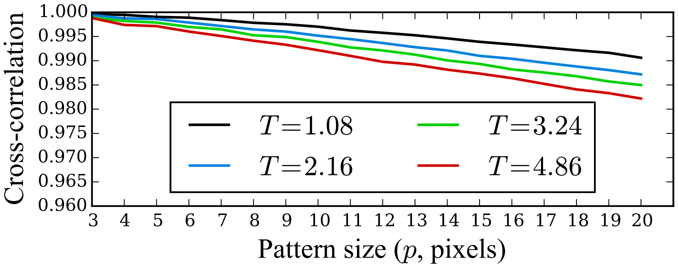


Fig. 6: Evaluation of the approximation in Eq. (11). We calculate the cross correlation between direct component slice $D_d(p)$ and the ground truth slice R_d. It shows high correlation even for the worst case.

where ∇ is the gradient operator and $\mathcal{E}$ is the set of edges corresponding to the material (or thickness) borders. The integration is computed by a standard sparse linear system solver as done in the work of Poisson image editing [52].

5 EXPERIMENTS

We evaluate our method numerically by simulation and also show qualitative results using real-world scenes.

5.1 Evaluation by simulation

We first assess the appropriateness of the approximation in Eq. (11), and then evaluate the accuracy of the slice recovery.

Validity of approximation in Eq. (11)

In this simulation, we change the pattern pitch p from 3 to 20 pixels, the depth d from 4 to 18 [mm] (in optical depths T, which correspond to 4σ to 18σ). The scene is one-layered for the purpose of assessing the approximation in Eq. (11). The coefficient σ is set to 0.27 [mm^{-1}], and forward scattering parameter q is set to 0.9. These parameters are chosen according to [53]. We generate the appearance slices $S^{\varphi}_{d,p}$ using checker pattern l^{φ}_{p} as

$$S^{\varphi}_{d,p} = ((l^{\varphi}_{p} * h_d) \circ R_d) * h_d. \tag{22}$$

p and φ are the pitch and phase of the pattern, and $\circ$ is the Hadamard (element-wise) product operator. We compute direct component $D_d(p)$ from these synthetic images with changing the phase φ.

We assess correlations between the recovered direct components $D_d(p)$ and the ground truth slices R_d using zero-mean cross correlation (ZNCC). The ZNCC value falls in the range from -1 to 1, where a negative value indicates negative correlation, and a greater value indicates higher similarity. The evaluation results are summarized in Fig. 6. The plots correspond to the average of ZNCC scores obtained from 15 different textures R_d. ZNCC values decreases as the depth becomes deeper (greater T) and pattern pitch becomes larger. However, they are all consistently highly correlated (minimum ZNCC value is 0.982 and the mean value is 0.992), and it shows the accuracy of the approximation.

Layers	ZNCC values		
	Max.	Mean	Min.
Top layer	0.99	0.93	0.89
Bottom layer	0.92	0.84	0.67

TABLE 1: Two layers recovery result for 20 sets of scenes. We compare recovered slices with the ground truth slices by ZNCC. Higher ZNCC scores indicate more accurate recovery.

Evaluation of slice recovery

Here we evaluate the overall accuracy of our method via simulation. We generate images L^{φ}_{p} under projection patterns l^{φ}_{p} as

$$L^{\varphi}_{p} = \sum_{d} ((l^{\varphi}_{p} * h_d) \circ R_d) * h_d. \tag{23}$$

We use 20 different two-layered scenes that have distinct textures R_5 and R_{15} at two depth ranges $5 \leq d < 15$ and $15 \leq d$. The textures are randomly paired from the ones that are used in the previous experiment. We change the pitch p of checker pattern and shift the pattern with φ. From the generated images, we compute $D(p, c)$ for each pattern pitch p, and apply our method to recover slices at $d = 5$ and 15.

We again use ZNCC values between the recovered and ground truth slices. The experimental results are summarized in Table 1. In all data sets, the ZNCC values of upper layer is higher than the lower layer as expected. The result of this simulation indicates that our method can recover slices of various textures with high accuracy.

Figure 7 shows a synthetic example of three-layer recovery, where slices at at $d = 1$, 5, and 15 are recovered. The ZNCC scores for the recovery results are 0.98, 0.83, and 0.51, respectively. While the result is generally consistent with the ground truth, negative intensities and ringing artifacts are observed due to the discrete pitch variations and convolution.

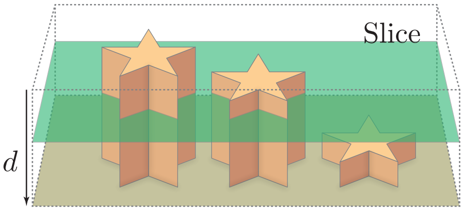
5.2 Real-world experiment

We develop a coaxial projector-camera setup for realizing the measurement setup as shown in Fig. 8. The coaxial setup has a favorable property; correspondence between projector and camera pixels becomes invariant with respect to depths. Unlike non-coaxial settings, with which a illumination ray inside the translucent object forms a line in the image coordinates [24], the coaxial setting allows us to easily separate the direct rays. We use a LightCommander projector, which is a DMD projector development kit by Texas Instruments, and use near infrared (NIR) light for measurements. The lenses of both camera and projector are set equivalent (Ai Micro-Nikkor 105mm f/2.8S) for making the alignment easy. In the experiment, we use 18 variations of checker patterns (3px to 20px with 1px interval), and shift the pattern for one-third of square size for each pattern.

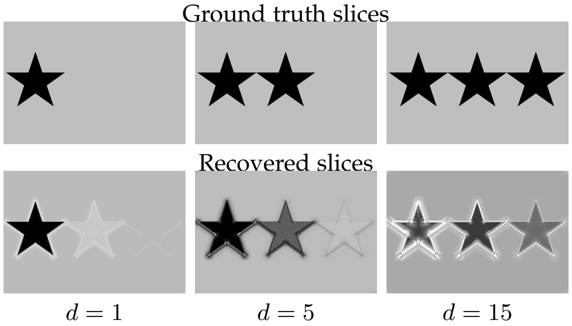
Experimental results

First, we use an oil painting as a target scene as shown in Fig. 9(a), which has draft and surface pigment layers as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2631625, IEEE Transactions on Pattern Analysis and Machine Intelligence



(a) Simulated scene. three star-shaped pillars with different heights are placed in scattering media. An example slice is illustrated by the green plane.



(b) Comparison of the recovered and ground truth slices

Fig. 7: (a) Target scene (b) Result of three-layer recovery at $d = 1$, 5, and 15. The ground truth radiances slices (upper) and recovered slices (lower) are shown. ZNCC scores are 0.98, 0.83, and 0.51, respectively.

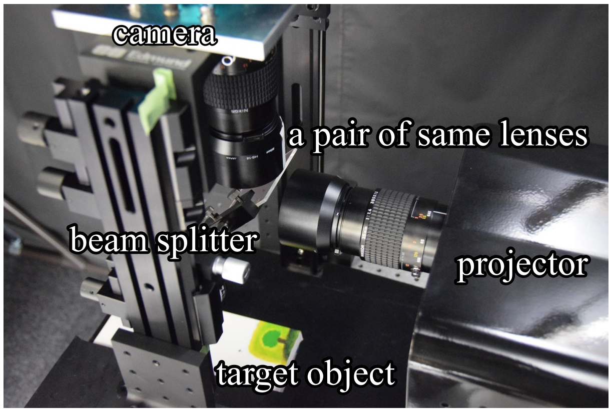
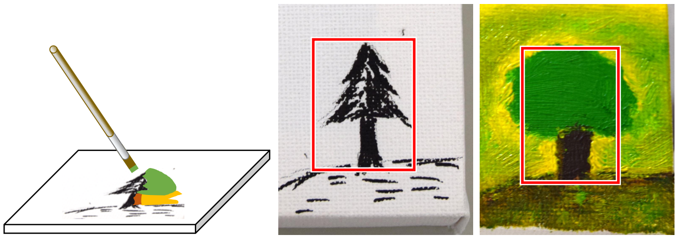
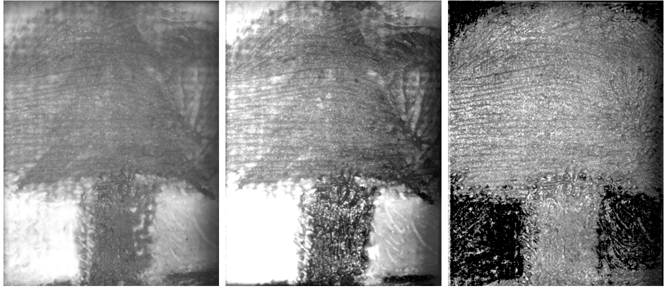


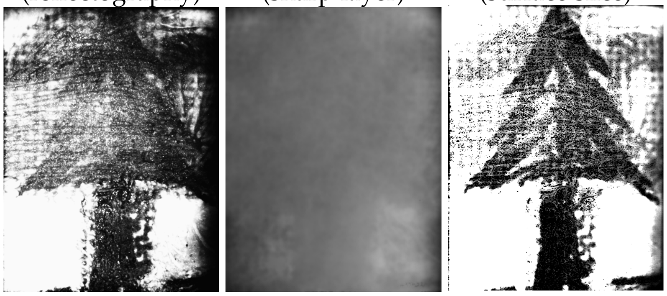
Fig. 8: Measurement setup. The coaxial system allows us to maintain the correspondences between projector and camera pixels.



(a) The scene (b) Draft layer (c) Oil painted



(d) NIR photo (reflectography) (f) Li and Brown [7] (sharp layer) (h) Ours (surface slice)

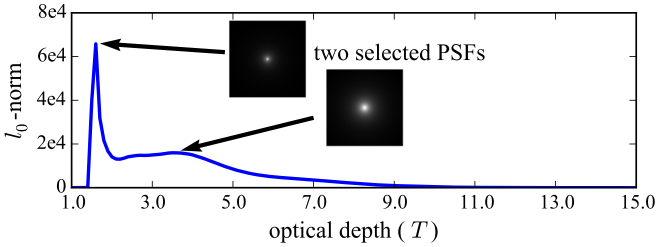


(e) Enhanced (g) Li and Brown [7] (blurred layer) (i) Ours (inner slice)

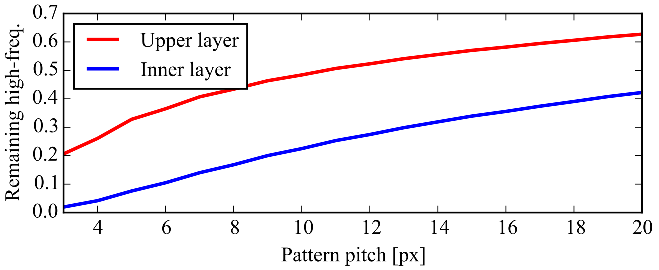
Fig. 9: Experimental result of oil painting using the baseline method. (a) Target scene. We draw a colored round tree on top of the draft of spiny tree. (b) Inner layer (draft) of the painting. (c) Painted scene. Red rectangle region is measured. (d) Normal photo using infrared light. (e) Enhanced image. Intensity range and contrast of (d) are manually adjusted. (f, g) Layer separation results of Li and Brown [7]. Because their method separates sharp and blurred layers, it suffers from global components. (h, i) Results of our method. Layer of surface texture and hidden drawing, respectively. Range of the intensities is adjusted for visualization.

depicted in Figs. 9(b) and 9(c). By taking a standard photograph under the near infrared light (as done in infrared reflectography [2] in the art field), we can only vaguely observe the draft layer as shown in Fig. 9(d). Since it is the superposition of draft and surface pigment layers, it naturally results in a blurry image. Even with a manual contrast adjustment, it is difficult to clearly observe the shape of draft tree as depicted in Fig. 9(e). A simple layer decomposition still suffers from the blur as shown in Figs. 9(f) and 9(g) . By applying our method to this scene, two PSFs are estimated as depicted in Fig. 10(a) and two slices are recovered as in Figs. 9(h) and 9(i). The upper surface layer corresponds to the surface pigment layer. Because the yellow pigment is almost transparent in the infrared wavelength, the corresponding painting regions become dark in the surface slice. The lower layer shows the inner layer, where the texture of the tree is clearly observed. The variation of the remaining high-frequency components of each layer, which can be computed from its PSF using Eq. (15), along with the pattern pitch is shown in Fig. 10(b). It shows that the high-frequency components of the upper layer remain more significantly than that of the inner layer, and it indicates the separability of the layers.

The second target scene is stacked translucent sheets that consist of layered translucent sheets and printed texture films as shown in Fig. 11(a). With a conventional NIR photograph, we can observe mixed textures as shown in Fig. 11(b), where textures are blurry and appearances across depths are superposed. Our method correctly selects two

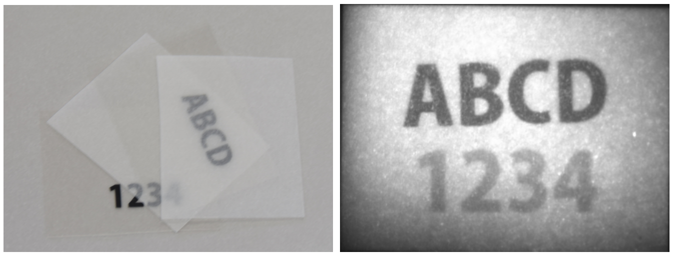


(a) Selected PSFs corresponding to local maxima

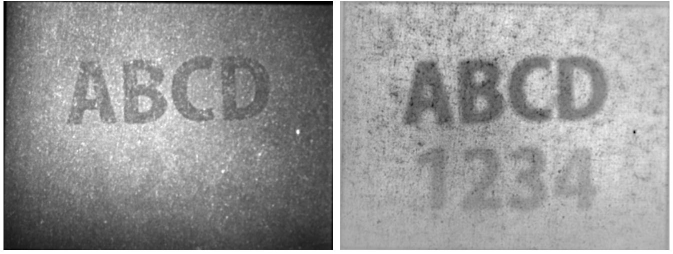


(b) Remaining high-frequency components for each layer

Fig. 10: Selected PSFs and remaining high-frequency components for each layer. (a) Selected PSFs. There are two peaks in the plot of Eq. (17), hence two corresponding PSFs are selected to recover. (b) Remaining high-frequency components for each layer. Upper layer remains more high-frequency components than that of the inner layer.



(a) The scene (b) NIR photo



(c) Upper slice (d) Lower slice

Fig. 11: Experimental result of layered scene using the baseline method. (a) Scene is a composite object of texture and translucent sheets (tracing paper). (b) Normal photo of the scene. Textures inside the object can be seen, although blurry. (c) Recovered slice of upper layer. (d) Recovered slice of lower layer. Textures in sub-millimeter gap can be separated. Further analysis of shadowed region is shown in Fig. 14

informative depths and recovers their radiance slices as shown in Figs. 11(c) and 11(d). In the upper slice, only 'ABCD' texture is visible, and '1234' texture appears in the lower slice. Due to the shadowing effect caused by the opaque material on the upper layer (the texture of 'ABCD'), the lower slice contains the 'ABCD' texture as an unobserved shadowed region.

Additional results are shown in Fig. 12. The top row is a piece of painting, where the painter's signature is hidden under pigment. Our method clearly recovers surface texture and inner signature slices. The middle row is a mural painting covered by white mold. In this example, we used RGB light sources instead of NIR, and record each color channel separately. Our method can recover the slices in this example as well, and the inner appearance is clearly visible in the result of Fig. 12(d). The bottom row is a double-sided ancient document with RGB light sources. While an ordinary digital scanner suffers from the back-side texture, our method can faithfully separate surface and back-side textures.

Figure 13 shows the result of the scene with an inhomogeneous upper layer. We use a 21×21 pixels Gaussian window for selecting PSFs. The target object is an oil painting, where the upper layer consists of a wide variety of pigments in their thicknesses and materials. The result of the baseline method shown in Fig. 13(b) exhibits brightness discontinuity at the middle of image region due to inhomogeneity of the upper pigments. By selecting PSFs per pixel using the Gaussian-window, the inner layer is recovered in a pixel-wise manner as shown in Fig. 13(f), although it still suffers from brightness discontinuity. Finally, with the gradient-domain filtering to suppress the brightness discontinuity, the inner layer that has uniform brightness can be recovered as shown in Fig. 13(h). It shows improvement in the visual quality, especially in the left part of the image and overall contrast.

Shadow detection from recovered inner layers

The recovered appearance of inner layers suffers from shadows caused by their upper layers. If the upper layer's transmittance is low, it casts shadow to the inner layer, *e.g.*, as 'ABCD' scene shown in Fig. 11(d). Here we describe a simple technique for dealing with shadows by post-processing. To simplify the discussion, we assume that there are only two layers in the scene, *i.e.*, upper and inner layers.

The shadowed regions have two common properties: (1) The shape of shadows observed in the inner slice is similar to the texture of upper slice. (2) The shadowed region becomes darker due to low irradiance. Based on these observations, we define a shadow likelihood measure P_s for identifying shadowed regions.

The similarity S of the texture shapes in the recovered slices R_1 and R_2 can be obtained using the absolute value of their cross correlation as

$$S(c) = \left| C_{w,c}(R_1, R_2) \right|, \tag{24}$$

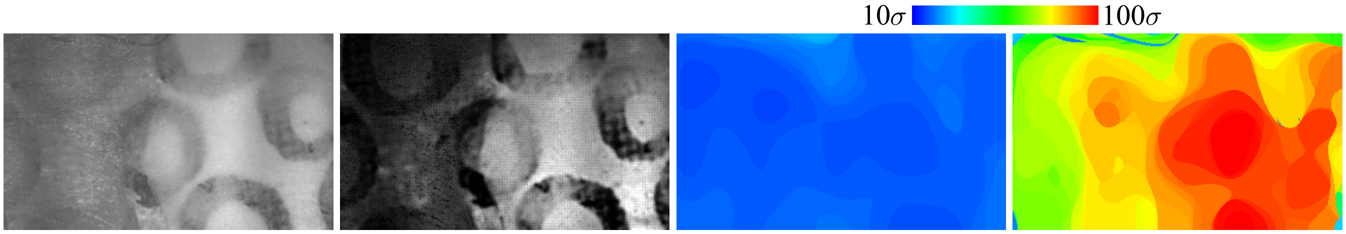
where $C_{w,c}$ is the cross-correlation within a small window w centered at pixel location c, and values in R_1 and R_2 are normalized in the range of $[0, 1]$. Using this similarity, the shadow likelihood P_s is defined using the darkness of the shadowed region in R_2 as

$$P_s(c) = S(c)(1 - R_2(c)). \tag{25}$$

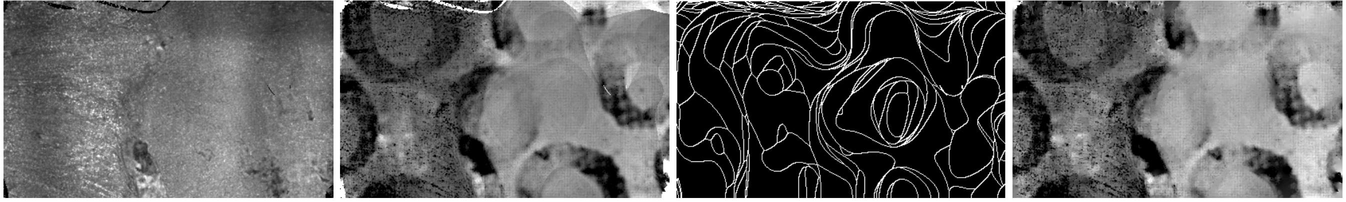
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2631625, IEEE Transactions on Pattern Analysis and Machine Intelligence

(a) Target scene (b) Normal photo (c) Surface slice (d) Inner slice

Fig. 12: Additional results for a painting, a mural, and an ancient document using the baseline method. (a) Target scenes. Top: author's signature covered by red pigment. Middle: A mural painting covered by white mold. Bottom: An ancient document of double side print. Rectangle regions are measured. (b) Normal photograph. Both upper and lower slices, and global components are composed. (c) Recovered slices of surface. (d) Recovered inner slices.



(a) Normal IR photo (b) Inner slice (baseline) (c) Upper layer's depth (d) Inner layer's depth



(e) Upper slice (pixel-wise) (f) Inner slice (pixel-wise) (g) Gradient mask (h) Inner slice (post-processed)

Fig. 13: Result of a scene with an inhomogeneous upper layer using the extended method. (a) Captured image using near infrared lights. (b) Recovered inner slice by the baseline method. (c, d) Selected depths for upper and inner layers, respectively. (e) Recovered upper slice by pixel-wise recovery. (f) Recovered inner slice by pixel-wise recovery. (g) Gradient mask computed from (d) that is used for gradient-domain filtering. (h) Recovered inner slice from (f) and (g) by image enhancement in the gradient domain.

(a) Inner slice including shadow (b) Likelihood of shadow (c) Shadow region mask (d) Inpainted inner layer

Fig. 14: Shadow detection and inpainting results. The scene are the same as Fig. 11 and Fig. 12. (a) The inner layer including shadowed region. (b) Shadow likelihood from recovered slices. White pixels indicate higher likelihood of being in shadow. (c) Shadow region is indicated by red by thresholding the likelihood. (d) Shadow region is recovered by image inpainting [54]. Shadow artifact caused by the upper layer is suppressed.

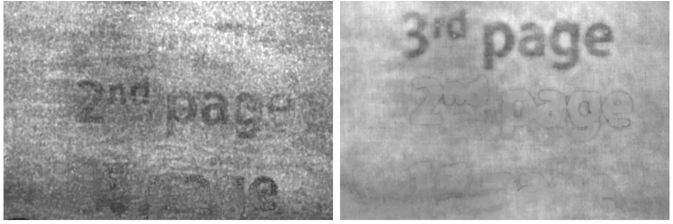
It yields a likelihood score for each pixel being in shadow, and by a simple thresholding, a shadowed region can be determined.

Once the shadow regions are identified, we can use a shadow removal technique. For example, the lost information within the shadow region can be filled in by arbitrary image inpainting methods. In this paper, we use a patch-match based image inpainting method [54].

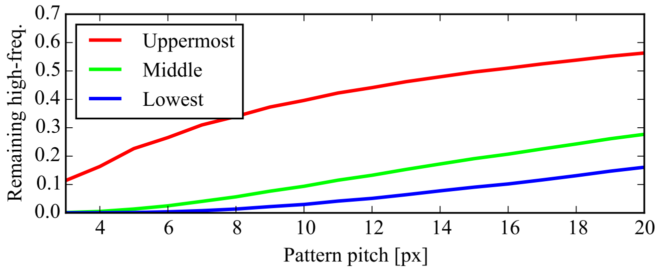
The result of shadow detection and removal for the 'ABCD' and ancient document scenes are shown in Fig. 14. For this experiment, we use 21×21 window size for computing the similarity measure. The shadow likelihood shown in Fig. 14(b) qualitatively corresponds to the shadowed regions in the inner layer. The shadow mask generated by thresholding the shadow likelihood is shown in Fig. 14(c). Figure 14(d) shows the image inpainting result using the detected shadow mask and the method of [54]. It shows improvement of visual quality of the inner layer.

Finally, we show a recovery and shadow removal result for a three-layer scene. Figure 15(a) shows the target scene, where three printed papers are overlaid. By applying our baseline method, three slices with shadows are recovered. With shadow detection and inpainting, we obtain the recovery of three layers as shown in Figs. 15(b)-15(d) For the shadow detection of the third slice, we compute the correlation with the original second slice which contains the shadows casted by the top slice. Figure 15(e) shows how high-frequency components remain along with varying pattern pitches. The lowest layer only retains a limited amount of high-frequency components hence its recovery becomes slightly blurry. However, it shows the different characteristics among layers, which indicates the separability of the layers.

(a) The scene (b) Uppermost slice



(c) Middle slice (d) Lowest slice



(e) Remaining high-frequency components

Fig. 15: The result of seeing through pages. (a) The target object. Three printed thin papers are superposed. (b) Recovered uppermost slice. Only the first page can be seen. (c) Recovered second page. Shadows from the upper layer is inpainted. (d) Recovered third page. Shadows from the first and second layers are inpainted. (e) Remaining high-freqency components for each pixel.

6 CONCLUSION

This paper described a method for recovering inner slices of translucent objects based on multi-frequency pattern projection. The proposed method is built upon the observation that the PSF inside translucent objects varies according to the depth of slices. Based on that, we have shown that inner radiance slices can be recovered by estimating PSFs using varying pitches of projection patterns. We also developed a method for automatically selecting informative slices via a sparse representation, *i.e.*, determining sparse coefficients that corresponds to radiance slices. We further extended the method for dealing with inhomogeneous translucent objects based on a combination of pixel-wise appearance recovery and gradient-based image repairing. The effectiveness of the proposed method are shown by several experiments on simulation and real-world translucent objects.

REFERENCES

[1] J. Dik, K. Janssens, G. Van der Snickt, L. van der Loeff, K. Rickers, and M. Cotte, "Visualization of a lost painting by Vincent van Gogh using synchrotron radiation based X-ray fluorescence elemental mapping." *Analytical chemistry*, vol. 80, no. 16, pp. 6436–42, Aug. 2008.

[2] D. Gavrilov, E. Maeva, O. Grube, X. Maldague, and R. Maev, "Infrared Methods in Noninvasive Inspection of Artwork," in *International Conference on NDT of Art*, no. May, 2008, pp. 1–5.

[3] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast Separation of Direct and Global Components of a Scene using High Frequency Illumination," in *Proc. SIGGRAPH*, 2006, pp. 935–944.

[4] K. Tanaka, Y. Mukaigawa, H. Kubo, Y. Matsushita, and Y. Yagi, "Recovering Inner Slices of Translucent Objects by Multi-frequency Illumination," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 5464–5472.

[5] R. Szeliski, S. Avidan, and P. Anandan, "Layer Extraction from Multiple Images Containing Reflections and Transparency," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2000, pp. 246–253.

[6] B. Sarel and M. Irani, "Separating Transparent Layers through Layer Information Exchange," in *Proc. European Conference on Computer Vision (ECCV)*, 2004.

[7] Y. Li and M. S. Brown, "Single Image Layer Separation using Relative Smoothness," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[8] T. Poter and T. Duff, "Compositing Digital Images," *Computer Graphics*, vol. 18, no. 3, pp. 253–259, 1984.

[9] C. Donner and H. W. Jensen, "Light Diffusion in Multi-Layered Translucent Materials," *ACM Transactions on Graphics (ToG)*, vol. 24, no. 3, p. 1032, jul 2005.

[10] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan, "A Practical Model for Subsurface Light Transport," in *Proc. SIGGRAPH*. ACM, 2001, pp. 511–518.

[11] P. Kubelka and F. Munk, "Ein Beitrag Zur Optik Der Farbanstriche," *Zeitschrift fr Technische Physik*, vol. 12, pp. 593–601, 1931.

[12] E. d'Eon and G. Irving, "A Quantized-Diffusion Model for Rendering Translucent Materials," in *Proc. SIGGRAPH*. ACM Press, 2011, p. 1.

[13] "Green Function," in *Encyclopedia of Mathematics*, M. Hazewinkel, Ed. Springer, 2001.

[14] W. Jakob, E. d'Eon, O. Jakob, and S. Marschner, "A Comprehensive Framework for Rendering Layered Materials," *ACM Transactions on Graphics (ToG)*, vol. 33, no. 4, pp. 1–14, jul 2014.

[15] M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas, "Synthetic Aperture Confocal Imaging," *ACM Transactions on Graphics (ToG)*, vol. 23, no. 3, pp. 825–834, 2004.

[16] M. Levoy and P. Hanrahan, "Light Field Rendering," in *Proc. SIGGRAPH*, 1996.

[17] T. R. Corle and G. S. Kino, Eds., *Confocal Scanning Optical Microscopy and Related Imaging Systems*. Academic Press, 1996.

[18] C. Fuchs, M. Heinz, M. Levoy, H.-P. Seidel, and H. P. A. Lensch, "Combining Confocal Imaging and Descattering," *Computer Graphics Forum, Special Issue for the Eurographics Symposium on Rendering 2008*, vol. 27, no. 4, pp. 1245–1253, 6 2008.

[19] S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal, "Structured Light in Scattering Media," in *Proc. International Conference on Computer Vision (ICCV)*, 2005, pp. 420–427.

[20] J. Gu, S. K. Nayar, E. Grinspun, P. Belhumeur, and R. Ramamoorthi, "Compressive Structured Light for Recovering Inhomogeneous Participating Media," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 35, no. 3, pp. 555–567, 2013.

[21] B. Lamond, P. Peers, and P. Debevec, "Fast Image-based Separation of Diffuse and Specular Reflections," in *Proc. ACM SIGGRAPH sketches*, 2007.

[22] Y. Mukaigawa, R. Raskar, and Y. Yagi, "Analysis of Light Transport in Scattering Media," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 153–160.

[23] ——, "Analysis of Scattering Light Transport in Translucent Media," *IPSJ Transactions on Computer Vision and Applications*, vol. 3, pp. 122–133, 2011.

[24] K. Tanaka, Y. Mukaigawa, Y. Matsushita, and Y. Yagi, "Descattering of Transmissive Observation using Parallel High-Frequency Illumination," in *Proc. International Conference on Computational Photography*, 2013.

[25] S. Achar, S. T. Nuske, and S. G. Narasimhan, "Compensating for Motion during Direct-Global Separation," in *Proc. International Conference on Computer Vision (ICCV)*, 2013.

[26] S. Achar and S. G. Narasimhan, "Multi Focus Structured Light for Recovering Scene Shape and Global Illumination," in *Proc. European Conference on Computer Vision (ECCV)*, 2014.

[27] D. Reddy, R. Ramamoorthi, and B. Curless, "Frequency-space Decomposition and Acquisition of Light Transport under Spatially Varying Illumination," in *Proc. European Conference on Computer Vision (ECCV)*, 2012, pp. 596–610.

[28] M. Gupta, Y. Tian, S. G. Narasimhan, and L. Zhang, "A Combined Theory of Defocused Illumination and Global Light Transport," *International Journal of Computer Vision (IJCV)*, vol. 98, no. 2, pp. 146–167, Oct. 2011.

[29] M. O'Toole, R. Raskar, and K. N. Kutulakos, "Primal-Dual Coding to Probe Light Transport," *ACM Transactions on Graphics (ToG)*, vol. 31, no. 4, pp. 1–11, jul 2012.

[30] M. O'Toole, J. Mather, and K. N. Kutulakos, "3D Shape and Indirect Appearance by Structured Light Transport," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2014, pp. 3246–3253.

[31] M. O'Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos, "Homogeneous Codes for Energy-Efficient Illumination and Imaging," *ACM Transactions on Graphics (ToG)*, vol. 34, no. 4, pp. 35:1–35:13, jul 2015.

[32] F. Heide, L. Xiao, A. Kolb, M. B. Hullin, and W. Heidrich, "Imaging in Scattering Media using Correlation Image Sensors and Sparse Convolutional Coding." *Optics express*, vol. 22, no. 21, pp. 26 338–50, oct 2014.

[33] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded Time of Flight Cameras: Sparse Deconvolution to Address Multipath Interference and Recover Time Profiles," *ACM Transactions on Graphics (ToG)*, vol. 32, no. 6, pp. 1–10, nov 2013.

[34] M. O'Toole, F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos, "Temporal Frequency Probing for 5D Transient Analysis of Global Light Transport," *ACM Transactions on Graphics (ToG)*, vol. 33, no. 4, pp. 1–11, jul 2014.

[35] R. Tadano, A. Kumar Pediredla, and A. Veeraraghavan, "Depth Selective Camera: A Direct, On-Chip, Programmable Technique for Depth Selectivity in Photography," in *Proc. International Conference on Computer Vision (ICCV)*, 2015, pp. 3595–3603.

[36] N. J. W. Morris and K. N. Kutulakos, "Reconstructing the Surface of Inhomogeneous Transparent Scenes by Scatter Trace Photography," in *Proc. International Conference on Computer Vision (ICCV)*, 2007.

[37] T. Hawkins, P. Einarsson, and P. Debevec, "Acquisition of Time-varying Participating Media," *ACM Transactions on Graphics (ToG)*, vol. 24, no. 3, pp. 812–815, Jul. 2005.

[38] I. Ihrke and M. Magnor, "Adaptive grid optical tomography," *Graphical Models*, vol. 68, no. 5, pp. 484 – 495, 2006, special Issue on the Vision, Video and Graphics Conference 2005.

[39] A. J. L. Adam, P. C. M. Planken, S. Meloni, and J. Dik, "TeraHertz imaging of hidden paint layers on canvas," *Optics Express*, vol. 17, no. 5, p. 3407, Feb. 2009.

[40] D. Huang, E. Swanson, C. Lin, J. Schuman, W. Stinson, W. Chang, M. Hee, T. Flotte, K. Gregory, C. Puliafito, and A. Et, "Optical

coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, Nov. 1991.

[41] A. F. Fercher, "Optical coherence tomography," *Journal of Biomedical Optics*, vol. 1, no. 2, pp. 157–173, 1996.

[42] M. G. L. Gustafsson, L. Shao, P. M. Carlton, C. J. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," *Biophysical journal*, vol. 94, no. 12, pp. 4957–70, Jun. 2008.

[43] P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, and M. G. L. Gustafsson, "Super-resolution video microscopy of live cells by structured illumination," *Nature methods*, vol. 6, no. 5, pp. 339–42, May 2009.

[44] O. D. Therrien, B. Aubé, S. Pagès, P. D. Koninck, and D. Côté, "Wide-field multiphoton imaging of cellular dynamics in thick tissue by temporal focusing and patterned illumination," *Biomedical optics express*, vol. 2, no. 3, pp. 696–704, Jan. 2011.

[45] K. Shimizu, K. Tochio, and Y. Kato, "Improvement of transcutaneous fluorescent images with a depth-dependent point-spread function," *Applied Optics*, vol. 44, no. 11, pp. 2154–2161, 2005.

[46] S. Narasimhan and S. Nayar, "Shedding Light on the Weather," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003, pp. 665–672.

[47] S. Chandrasekhar, *Radiative Transfer*. Dover Publications, Inc., 1960.

[48] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. IEEE Press, 1997.

[49] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan, "A Practical Model for Subsurface Light Transport," in *Proc. SIGGRAPH*. ACM Press, aug 2001, pp. 511–518.

[50] P. Hanrahan and W. Krueger, "Reflection from Layered Surfaces due to Subsurface Scattering," in *Proc. SIGGRAPH*. ACM Press, sep 1993, pp. 165–174.

[51] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.

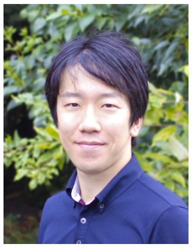
[52] P. Pérez, M. Gangnet, and A. Blake, "Poisson Image Editing," *ACM Transactions on Graphics (ToG)*, vol. 22, no. 3, pp. 313–318, Jul. 2003.

[53] S. G. Narasimhan, M. Gupta, C. Donner, R. Ramamoorthi, S. K. Nayar, and H. W. Jensen, "Acquiring Scattering Properties of Participating Media by Dilution," *ACM Transactions on Graphics (ToG)*, vol. 25, pp. 1003–1012, 2006.

[54] N. Kawai and N. Yokoya, "Image Inpainting Considering Symmetric Patterns," in *Proc. International Conference on Pattern Recognition*, 2012, pp. 2744–2747.

Kenichiro Tanaka received his B.S. and M.S. degrees in CS from Osaka University in 2012 and 2014, respectively. He is in pursuit of Ph.D. degree under supervision of Prof. Yagi, Prof. Mukaigawa, and Prof. Matsushita at Osaka University. He is a research fellow (DC1) of Japan Society for the Promotion of Science (JSPS). His research interests include computer vision, computational photography, and computational illumination. He is a member of CVF and a student member of IEEE.

Yasuhiro Mukaigawa received his M.E. and Ph.D. degrees from University of Tsukuba in 1994 and 1997, respectively. He became a research associate at Okayama University in 1997, an assistant professor at University of Tsukuba in 2003, an associate professor at Osaka University in 2004, and a professor at Nara Institute of Science and Technology (NAIST) in 2014. His current research interests include photometric analysis and computational photography. He is a member of IEEE.



Hiroyuki Kubo is an assistant professor at Nara Institute of Science and Technology (NAIST), Japan since 2014. His research interests include computer graphics and computer animation. He received the MS and Ph.D degrees from Waseda University, in 2008 and 2012, respectively. He is a member of ACM.

Yasuyuki Matsushita received his B.S., M.S. and Ph.D. degrees in EECS from the University of Tokyo in 1998, 2000, and 2003, respectively. From April 2003 to March 2015, he was with Visual Computing group at Microsoft Research Asia. In April 2015, he joined Osaka University as a professor. His research area includes computer vision, machine learning and optimization. He is on the editorial board of IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), International Journal of Computer Vision (IJCV), IPSJ Journal of Computer Vision and Applications (CVA), The Visual Computer Journal, and Encyclopedia of Computer Vision. He served/is serving as a Program Co-Chair of PSIVT 2010, 3DIMPVT 2011, ACCV 2012, ICCV 2017, and a General Co-Chair for ACCV 2014. He is a senior member of IEEE.

Yasushi Yagi is the Executive Vice President of Osaka university. He received his Ph.D. degrees from Osaka University in 1991. In 1985, he joined the Product Development Laboratory, Mitsubishi Electric Corporation, where he worked on robotics and inspections. He became a Research Associate in 1990, a Lecturer in 1993, an Associate Professor in 1996, and a Professor in 2003 at Osaka University. He was also Director of the Institute of Scientific and Industrial Research, Osaka university from 2012 to 2015. International conferences for which he has served as Chair include: FG1998 (Financial Chair), OMINVIS2003 (Organizing chair), ROBIO2006 (Program co-chair), ACCV2007 (Program chair), PSVIT2009 (Financial chair), ICRA2009 (Technical Visit Chair), ACCV2009 (General chair), ACPR2011 (Program co-chair) and ACPR2013 (General chair). He has also served as the Editor of IEEE ICRA Conference Editorial Board (2007–2011). He is the Editorial member of IJCV and the Editor-in-Chief of IPSJ Transactions on Computer Vision & Applications. He was awarded ACM VRST2003 Honorable Mention Award, IEEE ROBIO2006 Finalist of T.J. Tan Best Paper in Robotics, IEEE ICRA2008 Finalist for Best Vision Paper, MIRU2008 Nagao Award, and PSIVT2010 Best Paper Award. His research interests are computer vision, medical engineering and robotics. He is a fellow of IPSJ and a member of IEICE, RSJ, and IEEE.