

Multiplexed Illumination for Scene Recovery in the Presence of Global Illumination

Jinwei Gu¹, Toshihiro Kobayashi², Mohit Gupta³, Shree K. Nayar³
¹Rochester Institute of Technology, ²Canon Inc., ³Columbia University

jwgu@cis.rit.edu, kobayashi.toshihiro558@canon.co.jp, {mohitg, nayar}@cs.columbia.edu

Abstract

Global illumination effects such as inter-reflections and subsurface scattering result in systematic, and often significant errors in scene recovery using active illumination. Recently, it was shown that the direct and global components could be separated efficiently for a scene illuminated with a single light source. In this paper, we study the problem of direct-global separation for multiple light sources. We derive a theoretical lower bound for the number of required images, and propose a multiplexed illumination scheme which achieves this lower bound. We analyze the signal-to-noise ratio (SNR) characteristics of the proposed illumination multiplexing method in the context of direct-global separation. We apply our method to several scene recovery techniques requiring multiple light sources, including shape from shading, structured light 3D scanning, photometric stereo, and reflectance estimation. Both simulation and experimental results show that the proposed method can accurately recover scene information with fewer images compared to sequentially separating direct-global components for each light source.

1. Introduction

Most classical photometry based scene recovery techniques assume that scene points are illuminated only directly by the light source(s). Consequently, global illumination effects, such as inter-reflections, subsurface and volumetric scattering cause systematic biases in recovered scene properties. Because of the complex nature of these effects, separating the direct and global components of illumination has traditionally been considered a hard problem. Recently, Nayar et al. [11] presented simple and efficient techniques for performing the separation for a single light source. They showed that, in theory, separation can be performed by projecting only two high-frequency illumination patterns with ideal step edges. The method has subsequently been adapted in many scene recovery applications [2, 8, 7].

The goal of this paper is to analyze the problem of extracting the direct components for multiple light sources. Many scene recovery techniques require direct components under varying illumination conditions, such as intensity ra-

Application	Lights Needed	Number of Images Needed			
		Previous Work [11]		Our Method	
		Ideal	Practical	Ideal	Practical
Shape From Shading	1	2	3	2	3
Intensity Ratio (Structured Light)	2	4	6	3	5
Phase Shifting (Structured Light)	3	6	9	4	7
Photometric Stereo	3	6	9	4	7
BRDF/Normal Estimation	N	$2N$	$3N$	$N+1$	$2N+1$

Table 1. Number of images needed for scene recovery in the presence of global illumination. The “ideal” case refers to light sources that can project perfect step edges. The “practical” case refers to physically realizable light sources, and uses sinusoidal patterns.

tio structured light [3], phase shifting, photometric stereo and reflectance estimation. One straight-forward method is to use the technique in [11] sequentially for each light source. For N light sources, this would require $2N$ images. Can we do better, and what is the minimum number of images required? There are two pertinent observations. First, only the direct components are required for scene recovery - the global components can be considered as *noise*. Second, by carefully designing the illumination patterns, the collective contribution of all the global components can be made constant across all the captured images. Based on these observations, we establish a lower bound — at least $N + 1$ images are required to extract the N direct components. Furthermore, we design a separation method which actually achieves this lower bound.

In practice, due to various optical effects (*e.g.*, defocus, color bleeding, screen-door), light sources cannot project perfect step edges. Nayar et al. [11] showed that sinusoidal patterns can be used to mitigate this problem. Using three shifted sinusoidal patterns, the direct and global components can be recovered. Applying this technique sequentially for N sources would require capturing $3N$ images.

In this paper, we show that multiplexed illumination can be used to recover the N direct components corresponding to N different light sources by capturing $2N + 1$ images. Since global components remain constant for high-frequency illumination, our key idea is to modulate each

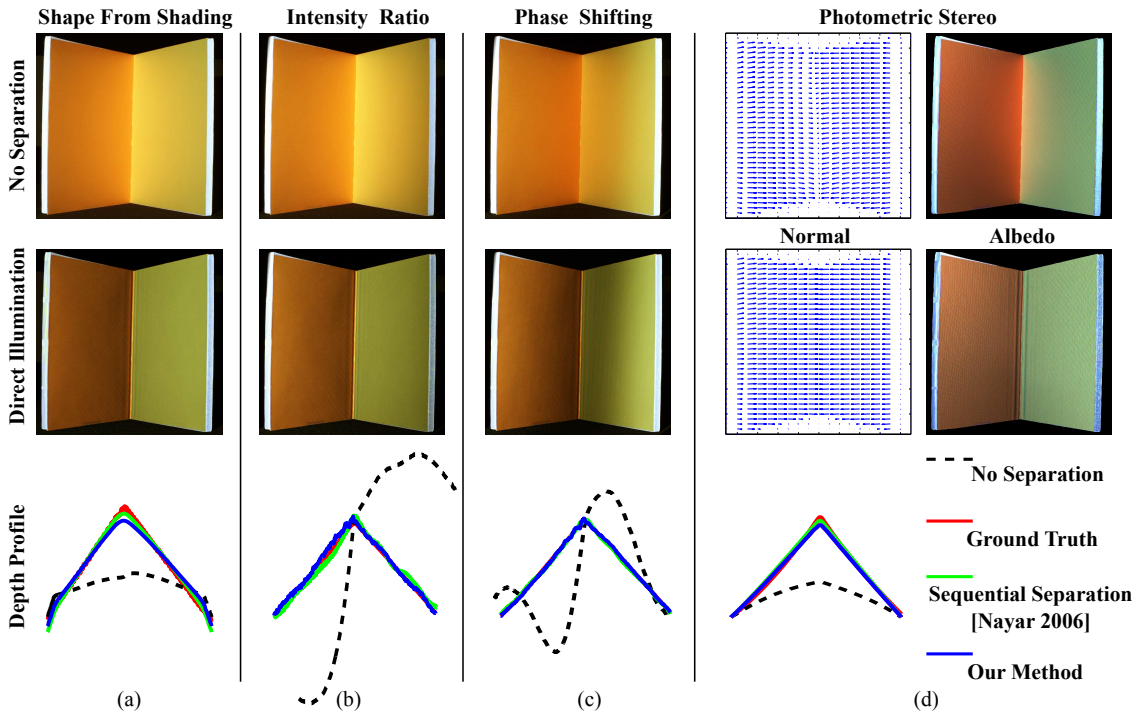


Figure 1. **Scene recovery results for a v-groove:** (a) shape from shading (one source); (b) intensity ratio (two sources); (c) phase shifting (three sources); and (d) photometric stereo (three sources). **Row 1:** One of the captured images without direct-global separation. **Row 2:** The separated direct component using our method. **Row 3:** Recovered depth profiles. In (d), we also show the recovered surface normals (as needle maps) and albedo maps obtained with and without direct-global separation. Our method faithfully recovers scene information, while requiring fewer images than applying the separation method [11] sequentially.

light source with a unique high frequency sinusoidal pattern shifted over time. The proposed multiplexed illumination method has implications for a wide range of scene recovery techniques, as shown in Table 1.

We perform signal-to-noise ratio (SNR) analysis of the proposed multiplexed illumination method and find that it also has SNR benefits compared to the sequential separation method, especially under low light conditions. Our conclusion is similar to that of conventional multiplexed illumination [14, 15] without direct-global separation. In summary, the key contributions of this paper are three-fold:

- We analyze the problem of simultaneously extracting the direct components for multiple light sources and provide a theoretical lower bound for the number of acquired images.
- We propose a multiplexed illumination method and analyze its SNR characteristics in the context of direct-global separation.
- We show that our techniques can be easily integrated with a variety of scene recovery techniques. This enables accurate recovery even in the presence of strong global illumination effects, while requiring fewer images as compared to previous approaches. Examples are shown in Figures 1, 5, 6 and 7.

2. Related Work

Scene Recovery under Global Illumination: One of the first attempts at scene recovery under global illumination was looking at shape from inter-reflections for recovering faceted, Lambertian scenes [10]. More recently, Nayar et al. [11] proposed using high frequency illumination patterns to separate direct and global illumination for more general scenes. Gupta et al. [5] studied the relation between illumination defocus and global light transport. Chen et al. [2] used modulated structure light patterns with high-frequency patterns to mitigate the effects of global illumination. Lamond et al. [8] used high frequency light patterns to separate the diffuse and specular components of BRDF. Holroyd et al. [7] constructed a high-accuracy imaging system for measuring surface shape and BRDF. All these techniques are sequential in that they assume that the global illumination in each of the acquired images is caused by a single source. In contrast, we consider *simultaneous* separation of direct components for multiple light sources.

Multiplexing: Multiplexing has a long history in imaging [6]. Schechner et al. [14] proposed using multiplexed illumination to increase SNR for image relighting. The multiplexing benefit has been analyzed under various noise sources [13, 15]. In addition to illumination, multiplexing has also been widely used in several other domains. Liang

et al. [9] proposed multiplexing camera aperture patterns for light field acquisition. Park et al. [12] used multiplexed LEDs for spectral imaging. Agrawal et al. [1] proposed multiplexing camera exposure for high speed imaging. In our work, we analyze illumination multiplexing in the context of direct-global separation.

3. Direct-Global Separation: Single Source

Let us first briefly review direct-global separation for a single light source [11]. Suppose the scene is illuminated with a high-frequency black-white checkerboard pattern such that half the pixels are on (white). If a scene point is illuminated (*i.e.*, in white squares), its radiance is the sum of the direct illumination and half the global illumination, *i.e.*, $L_{max} = L_d + \frac{1}{2}L_g$. Next, suppose that the scene is illuminated with the inverted checkerboard pattern. Now, the point is not directly illuminated by the light source, and its radiance contains only half the global illumination, $L_{min} = \frac{1}{2}L_g$. This provides two linear equations in the two unknowns, L_d and L_g . Thus, in theory, two images are sufficient to solve for L_d and L_g .¹

Because real-world light sources cannot produce perfect step edges, Nayar et al. [11] suggested using a high frequency sinusoid pattern with three different phases to modulate the light source. The one extra image provides the phase information of the scene with respect to the light source. Suppose the light source has an amplitude of A , and is modulated with a sinusoidal pattern of frequency f cycles/pixel, which is shifted three times with a speed of v pixels/second. The scene irradiance at time t is

$$L(t) = A \cdot (1 + \sin(\omega t + \phi)) / 2, \quad (1)$$

where $\omega = 2\pi f v$ is the modulation frequency and ϕ is the phase with respect to the source. As the sinusoidal pattern is shifted, only the direct component is modulated differently; the fraction of the global component in the captured images remains the same at $1/2$. If the reflectance of the scene point is R , its radiance at time t is

$$\begin{aligned} I(t) &= R \cdot L(t) + \frac{L_g}{2} = \frac{L_d}{2} \sin(\omega t + \phi) + \frac{L_d + L_g}{2} \\ &= \alpha \cdot \sin(\omega t) + \beta \cdot \cos(\omega t) + \frac{1}{\sqrt{2}}\gamma, \end{aligned}$$

where $L_d = AR$ and L_g are the direct and global illumination, $\alpha = L_d \cos(\phi)/2$, $\beta = L_d \sin(\phi)/2$, and $\gamma = (L_d + L_g)/\sqrt{2}$. The above equation is linear in the three unknowns, α , β , γ , and thus three images are sufficient. The direct and global components are given by

$$L_d = 2\sqrt{\alpha^2 + \beta^2}, \quad L_g = \sqrt{2}\gamma - L_d. \quad (2)$$

¹As shown in [11], the key assumption in this argument (and also in our proposed method) is that global light transport acts as a low pass filter, which is true for most scenes. The assumption, however, does not hold for highly specular materials, such as mirrors.

While three sinusoid patterns are theoretically enough, in practice the results are prone to image artifacts due to noise and quantization. In our experiments, we capture 25 images with a shifting checkerboard to obtain high-quality direct/global separation as ground truth for comparison with the methods proposed in this paper. Refer to the supplementary document for details.

4. Direct-Global Separation: Multiple Sources

What is the minimum number of images required to separate the direct illumination components for N light sources? A straight forward answer is $N \cdot K$ where K is the number of images needed for direct-global separation for a single light (two and three for ideal and practical light source, respectively). Is it possible to separate the direct components with fewer images? If so, what is the least number of images required?

4.1. A Theoretical Lower Bound

Lemma 4.1. *To separate the direct components for N light sources, at least $N + 1$ images are required. If the light sources can project perfect step edges, $N + 1$ images are sufficient.*

Necessary: Similar to a single light source, for multiple light sources, if all the projected patterns are high frequency and the average pattern intensity is half the maximum pattern intensity, the contribution of the N global components will remain the same at $1/2$, for any combinations of shifts of the patterns. The sum of N global components can thus be treated as one unknown. Together with the N direct components, we have $N + 1$ unknowns, requiring at least $N + 1$ images to recover.

Sufficient: We now show that $N + 1$ images are sufficient, using a constructive argument. Let $L_d^{(i)}$ and $L_g^{(i)}$ denote the direct and global illumination components corresponding to the i -th light source, $i = 1, \dots, N$. First, we turn on all N lights at half brightness and capture an image I_0 :

$$I_0 = \sum_{i=1}^N (L_d^{(i)} + L_g^{(i)}) / 2.$$

Next, we set the i -th light source to be a checkerboard pattern (half the pixels on) while keeping all the other $N - 1$ light sources at half brightness. In the captured image I_i , only the coefficient of $L_d^{(i)}$ changes. Thus

$$I_i = \begin{cases} I_0 - L_d^{(i)}/2, & \text{for points in black squares} \\ I_0 + L_d^{(i)}/2, & \text{for points in white squares.} \end{cases} \quad (3)$$

Therefore, the N direct components can be computed as $L_d^{(i)} = 2|I_i - I_0|$, $i = 1, \dots, N$, from $N + 1$ images.

The above algorithm provides only a theoretical lower bound, since real-world light sources cannot produce perfect step edges, as discussed earlier. The patterns were chosen to make the lower bound argument. Below, we propose a practical algorithm that uses sinusoidal patterns and requires $2N + 1$ images.

4.2. Frequency Modulated Multiplexing

In this section, we propose a practical algorithm, *frequency modulated multiplexing*, that uses high frequency sinusoids as modulation patterns. While the sequential separation with sinusoid patterns [11] requires $3N$ images for N light sources, the proposed method requires only $2N + 1$ images. Compared to the lower bound given in Lemma 4.1, the extra N images are required to compute the N projections (phase values) of the scene-point into the image plane of the N light sources.

The core idea of frequency modulated multiplexing is to modulate the N light sources with high frequency, shifting sinusoid patterns, with a single unique frequency per light source. Suppose the i -th light source has an amplitude of A_i and a modulation frequency ω_i , $i = 1, \dots, N$. The total irradiance for a scene point at time t is $L(t) = \sum_{i=1}^N A_i \cdot (1 + \sin(\omega_i t + \phi_i)) / 2$. The captured image $I(t)$ includes the N modulated direct components and half the sum of all N global components, which is given by

$$\begin{aligned} I(t) &= R \cdot L(t) + \frac{1}{2} \sum_{i=1}^N L_g^{(i)} \\ &= \sum_{i=1}^N (\alpha_i \sin(\omega_i t) + \beta_i \cos(\omega_i t)) + \frac{1}{\sqrt{2}} \gamma, \end{aligned} \quad (4)$$

where $\alpha_i = L_d^{(i)} \cos(\phi_i) / 2$, $\beta_i = L_d^{(i)} \sin(\phi_i) / 2$, and $\gamma = \sum_{i=1}^N (L_d^{(i)} + L_g^{(i)}) / \sqrt{2}$, for $i = 1, \dots, N$. Equation (4) is linear with respect to the $2N + 1$ unknowns, $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N, \gamma$. By shifting the sinusoidal patterns, we capture $2N + 1$ images $I(t_1), \dots, I(t_{2N+1})$ and obtain $2N + 1$ linear equations, which can be written as:

$$\mathbf{F} \cdot \mathbf{x} = \mathbf{b}, \quad (5)$$

where $\mathbf{x} = [\alpha_1, \beta_1, \dots, \alpha_N, \beta_N, \gamma]^T$ is the unknown, $\mathbf{b} = [I(t_1), \dots, I(t_{2N+1})]^T$ is the measurement, and the multiplexing matrix \mathbf{F} is

$$\mathbf{F} = \left[\mathbf{c}_1, \mathbf{s}_1, \dots, \mathbf{c}_N, \mathbf{s}_N, \frac{1}{\sqrt{2}} \mathbf{1} \right], \quad (6)$$

where the column vectors \mathbf{c}_i and \mathbf{s}_i are defined as $\mathbf{c}_i = [\cos(\omega_i t_1), \dots, \cos(\omega_i t_{2N+1})]^T$ and $\mathbf{s}_i = [\sin(\omega_i t_1), \dots, \sin(\omega_i t_{2N+1})]^T$, $i = 1, \dots, N$. Once \mathbf{x} is solved via matrix inverse, the N direct components are computed as $L_d^{(i)} = 2\sqrt{\alpha_i^2 + \beta_i^2}$ for $i = 1, \dots, N$.²

²We can also compute the N phase maps, $\phi_i = \tan^{-1}(\beta_i / \alpha_i)$, which are related to scene depth.

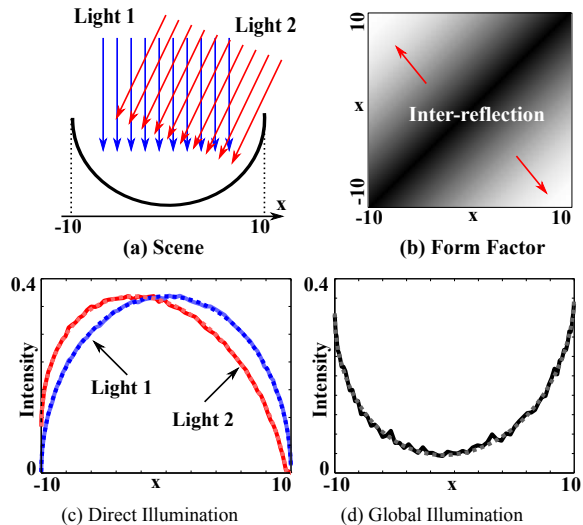


Figure 2. **Simulation results for frequency modulated multiplexing.** (a) The scene is a 2D Lambertian half circle, illuminated by two directional light sources. (b) The form factor matrix for the scene, used to simulate inter-reflections. $2 \times 2 + 1 = 5$ images (with 0.5% Gaussian additive noise) are simulated and used for direct-global separation. (c) The two estimated direct components and (d) the estimated sum of the two global components (solid lines) accurately match the ground truth (dotted lines).

Optimal Modulation Frequencies and Timings: The robustness of the estimates of the direct components depends on the condition number of the matrix \mathbf{F} , which is a function of the modulation frequencies ω_i and timings t_j , $i = 1, \dots, N$, $j = 1, \dots, 2N + 1$. Our goal is to obtain a non-singular matrix \mathbf{F} with the smallest condition number.

Theorem 4.2. *For the frequency modulated multiplexing method, if the frequencies ω_i , $i = 1, \dots, N$, and timings t_j , $j = 1, \dots, 2N + 1$ satisfy*

$$t_j = j, \quad \omega_i = \frac{2\pi k}{2N + 1}, \quad k, j = 1, \dots, 2N + 1, \quad (7)$$

where k can take any of the N values (without repeating) from $1, \dots, 2N + 1$, the matrix \mathbf{F} is an orthogonal matrix with the minimum condition number 1.

The proof for the above theorem is given in the appendix. We can implement the modulation frequency ω_i by setting either the spatial frequency f_i of the sinusoid pattern or the shifting speed v_i , or both, since $\omega_i = 2\pi f_i v_i$.

Figure 2 shows a simulation result to verify the accuracy of the proposed method. The scene is a 2D Lambertian half circle, illuminated by two directional light sources. The ground truth of the direct components and the global components (*i.e.*, inter-reflection) are simulated with the radiosity method [10]. We simulate $2 \times 2 + 1 = 5$ images under the modulated light patterns (with 0.5% additive noise) as input for the proposed method. As shown in Figs. 2(c) and (d), the estimated direct components and the sum of the global component accurately match the ground truth.

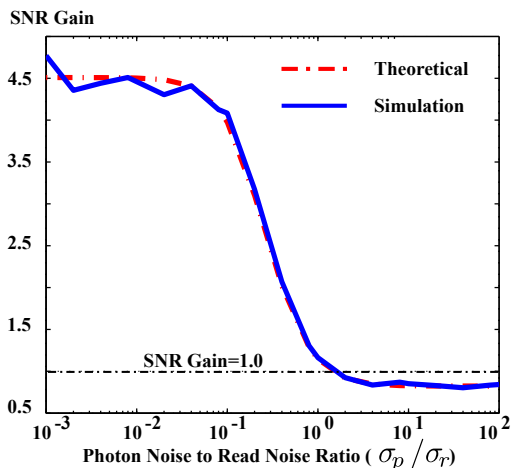


Figure 3. **Signal-to-Noise Ratio (SNR) characteristics of the proposed method.** We assume a Gaussian model for both the photon noise and the read noise. The x -axis is the ratio between the standard deviation of the photon noise (σ_p) and that of the read noise (σ_r). The y -axis is the SNR gain of the proposed method with respect to the sequential separation method [11]. **Red dot-dash line:** the theoretical result. **Blue solid line:** the simulation result (for $N = 30$ lights).

4.3. Signal-to-Noise Ratio (SNR) Analysis

Schechner et al. [14, 15] analyzed the SNR characteristics for conventional light multiplexing. We extend this analysis to the problem of direct-global separation. Our results are summarized below. Derivations and details are given in the appendix and the supplementary document.

- For dim scenes, if the imaging system is **read noise limited**, the proposed method increases SNR by $\sqrt{(2N+1)}/3$ with respect to the sequential separation method. If the imaging system is **photon noise limited**, the proposed method has a slightly lower SNR ($\sqrt{(2N+1)}/(3N) \approx \sqrt{2/3} \approx 0.83$).
- For bright scenes where saturation is an issue, the proposed method will shorten exposure time and thus reduce SNR by a factor of $1/\sqrt{N}$.
- One can also perform conventional multiplexing (with Hadamard code) [14] of N light sources for direct-global separation — by multiplexing N images for every modulation light pattern.³ It requires the same number of images ($3N$). Compared to this Hadamard-based method, our proposed method has an SNR gain of $\sqrt{4N(2N+1)}/(3(N+1)^2) \approx \sqrt{8/3} \approx 1.6$, while requiring fewer images ($2N+1$).

Figure 3 shows the SNR gain with respect to the sequential separation [11] for a variety of photo noise to read noise ratios. The red dot-dash line is the theoretical result, and the blue solid line is the simulation result (for $N = 30$ light sources). As expected, the SNR gain $\approx \sqrt{2N/3} \approx 4.47$ if

³Refer to the supplementary document for a detailed explanation.

the read noise dominates, and it reduces as the photon noise increases, approaching the asymptotic value of ≈ 0.83 .

5. Applications to Scene Recovery

We applied the proposed method to a variety of scene recovery applications that require multiple light sources:⁴ shape from shading ($N = 1$), intensity ratio structured light scanning ($N = 2$), phase shifting structured light scanning ($N = 3$), photometric stereo ($N = 3$), and BRDF and surface normal estimation ($N = 9$). We used a PointGrey Grasshopper camera (5.0 mega pixels) and an Optoma DLP projector (1280×720 pixels) to implement our method. Both the camera and the projector are geometrically and radiometrically calibrated before image acquisition.

For intensity ratio and phase shifting, each structured light pattern is treated as a separate (but collocated) light source. For photometric stereo and BRDF estimation, which use multiple light sources in different directions, we implemented one of two strategies: (1) use mirrors to create multiple lights using a single projector (an example is given in the supplementary document), or (2) move the projector and sum the captured images to get the multiplexed illumination images.

Recovery of a V-Groove: We applied several scene recovery techniques to a v-groove made of two color panels. The groove has strong inter-reflection. As an example, Figure 4 shows the projected light patterns and the captured images of phase shifting. The amplitudes of the three (collocated) light sources A_1, A_2, A_3 are shown in Fig. 4(a). For direct-global separation, we modulate A_1, A_2, A_3 with high frequency sinusoids shifting over time (Fig. 4(b)) and project the modulated light patterns.⁵ Figure 4(c) shows the corresponding captured images, which are the input for the proposed method. Figure 1 shows the results. Our multiplexing approach recovers accurate scene information using fewer images than the traditional approach.

Surface Normal Estimation for a Banana: We performed photometric stereo for a banana. As shown in Figs. 5(a) and (b), comparing the images with and without direct-global separation, the banana exhibits significant subsurface scattering. We obtained the ground truth depth map (Fig. 5(c)) by using the sequential separation method with a shifting checkerboard pattern [11]. As shown in Figs. 5 (d)~(i), the proposed method yields accurate depth recovery with $2 \times 3 + 1 = 7$ images. In contrast, applying the separation method [11] sequentially using sinusoids will require 9 images.

Surface Normal/BRDF Estimation of a Cake Mold: Figure 6 shows an example to recover the surface normal and BRDF of a shiny cake mold. The cake mold is made

⁴Project web page: www.cis.rit.edu/jwgu/research/demux.

⁵Since the three light sources are collocated, we use a single projector to project the *sum* of the three modulated light patterns.

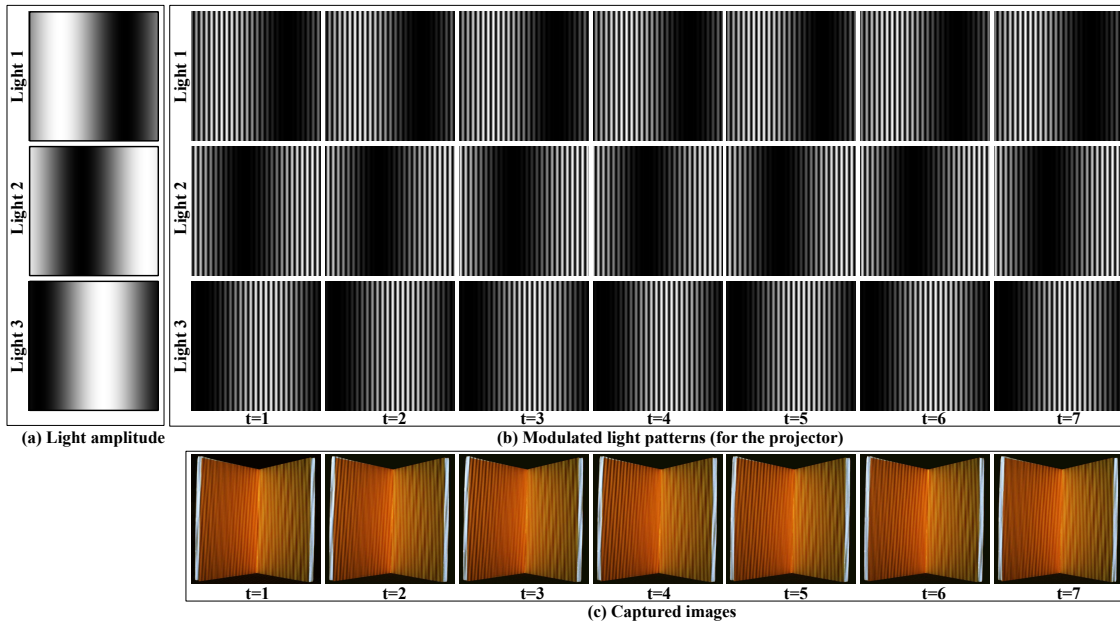


Figure 4. **Projected light patterns and captured images for phase shifting on a v-groove.** (a) The amplitudes A_1, A_2, A_3 for the three (collocated) light sources, implemented with a low frequency (1 cycle/image width) to avoid unwrapping. (b) We modulate the three light sources with high frequency sinusoids shifting over time and simultaneously project the modulated light patterns. (c) The corresponding captured input images for the proposed method. Depth estimation results are given in Fig.1.

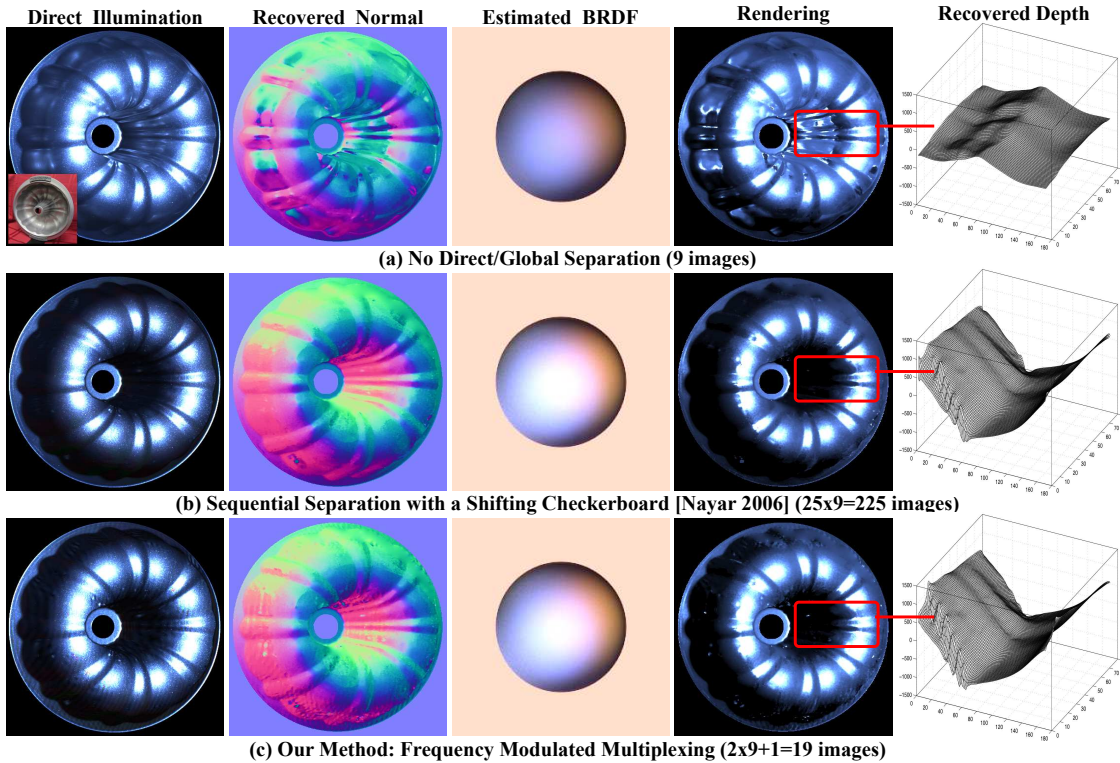


Figure 6. **BRDF and surface normal estimation of a shiny cake mold.** We used $N = 9$ lights and compared three methods: no direct-global separation, the conventional method (*i.e.*, sequential separation with a shifting checkerboard) [11], and the proposed method. Sequential separation using a sinusoid requires $9 \times 3 = 27$ images in this case. **Column 1:** One of the direct components (for no separation, it is one of the captured image). **Column 2:** Recovered surface normal map (color coded). **Column 3:** Estimated BRDF (rendered as a sphere under natural environment lighting). **Column 4:** Rendered images with the estimated BRDF and surface normals. **Column 5:** Recovered depth for the selected region (red rectangle).

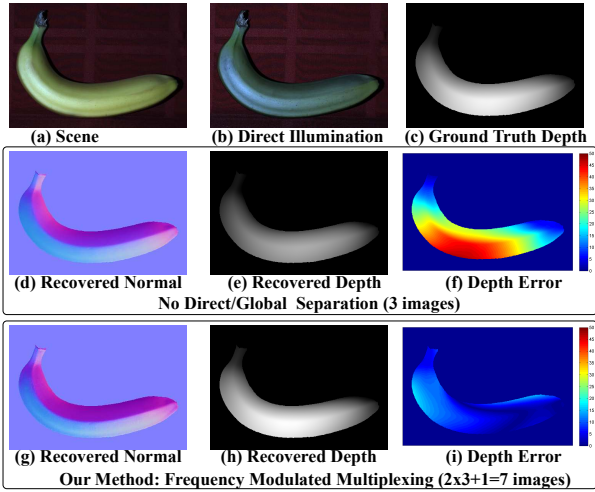


Figure 5. **Recovery of surface normal and depth of a banana using photometric stereo** ($N = 3$). (a) One of the three captured images without direct-global separation. (b) The corresponding direct illumination separated with the proposed method. (c) Ground truth depth map estimated by the sequential separation with a shifting checkerboard pattern [11] ($3 \times 25 = 75$ images). **Row 2:** Results without direct-global separation — (d) recovered normals, (e) estimated depth map, and (f) depth error ((e)-(c)). **Row 3:** Results of the proposed method ($2 \times 3 + 1 = 7$ images), where (i) depth error is (h)-(c). Without separation, there is an average of 19% error in the recovered depth; with our method, it’s only 4%.

of cast iron with a shiny coating, with a concave sink between the boundary and the center, as shown in the inset in Fig. 6(a). We used $N = 9$ light sources from different directions to normal and BRDF estimation [4]. As shown in Fig. 6, our proposed method obtains much more faithful scene recovery (compared to no separation) while requiring fewer images as compared to the conventional sequential separation method.

Depth Recovery of a Room In this example, we recover the depth of a room in a pop-up book with phase shifting ($N = 3$). As shown in Figs. 7(a) and (b), there are strong inter-reflection around the corners and the ceiling. We obtained the ground truth depth by scanning the scene with a single stripe of light (to minimize inter-reflection). As shown in Fig. 7, the proposed method accurately recovered scene depth with $2 \times 3 + 1 = 7$ images. In contrast, sequential separation using shifting sinusoids requires 9 images.

6. Discussion and Limitations

In this paper, we studied the problem of extracting direct illumination for multiple light sources. We derived a theoretical lower bound on the number of required images and proposed a multiplexed illumination method for direct-global separation which achieves this bound. We analyzed the signal-to-noise ratio characteristics of the proposed method. Experimental results show that the proposed method can accurately recover scene information in

the presence of global illumination, with fewer images compared to previous approaches.

There are several limitations of the proposed method. Similar to conventional illumination multiplexing, for bright scenes where saturation is an issue or where photon noise dominates read noise, multiplexing will reduce SNR, and thus it is not recommended. An interesting avenue of future work is to determine the optimal illumination scheme for these situations [13]. Another interesting direction is to exploit the recovered phase maps for scene recovery. As discussed in Section 4.2, the phase maps for N light sources are recovered as by-products of the proposed method. These phase maps are related to scene structure, and can be used to further constrain shape recovery. This may result in a further reduction of the required images.

Acknowledgments: This research was supported in parts by Canon Inc., NSF (grant number IIS 09-64429) and ONR (grant number N00014-08-1-0638).

Appendix

Proof of Theorem 4.2 We prove that if the condition $t_j = j$, $\omega_i = \frac{2\pi k}{2N+1}$, $k, j = 1, \dots, 2N + 1$, is satisfied, the matrix \mathbf{F} is an orthogonal matrix, and it has the smallest condition number 1. Recall that $\mathbf{F} = \left[\mathbf{c}_1, \mathbf{s}_1, \dots, \mathbf{c}_N, \mathbf{s}_N, \frac{1}{\sqrt{2}} \mathbf{1} \right]$. Under the above condition, we have $\mathbf{s}_i = \sin(\omega_i) = \sin \frac{2\pi k j}{2N+1}$ and $\mathbf{c}_i = \cos(\omega_i) = \cos \frac{2\pi k j}{2N+1}$, for $i = 1, \dots, N$ and $k, j = 1, \dots, 2N + 1$. We note that

$$\mathbf{s}_i^T \mathbf{1} = 0, \quad \mathbf{c}_i^T \mathbf{1} = 0, \quad \mathbf{s}_{i_1}^T \mathbf{c}_{i_2} = 0, \quad \mathbf{1}^T \mathbf{1} = 2N + 1,$$

and

$$\mathbf{s}_{i_1}^T \mathbf{s}_{i_2} = \begin{cases} 0 & i_1 \neq i_2, \\ \frac{2N+1}{2} & i_1 = i_2, \end{cases} \quad \mathbf{c}_{i_1}^T \mathbf{c}_{i_2} = \begin{cases} 0 & i_1 \neq i_2, \\ \frac{2N+1}{2} & i_1 = i_2, \end{cases}$$

Therefore, $\mathbf{F}^T \mathbf{F} = \text{diag} \left(\frac{2N+1}{2}, \dots, \frac{2N+1}{2}, \frac{2N+1}{2} \right) = \frac{2N+1}{2} \mathbf{I}$. Thus, the matrix \mathbf{F} is an orthogonal matrix, and its condition number reaches the minimum value 1. \square

SNR Gain Derivation Here we briefly summarize the derivation of SNR gain for the proposed method. A detailed derivation is given in the supplementary document. We recall that $L_d^{(i)} = 2\sqrt{\alpha_i^2 + \beta_i^2}$, and thus we have $(\Delta L_d^{(i)})^2 = \left(\frac{\partial L_d^{(i)}}{\partial \alpha_i} \right)^2 \Delta \alpha_i^2 + \left(\frac{\partial L_d^{(i)}}{\partial \beta_i} \right)^2 \Delta \beta_i^2$. Thus, assuming N lights have the same brightness, the noise after de-multiplexing (averaged across an entire image) is $\overline{(\Delta L_d^{(i)})^2} = 2(\Delta \alpha_i^2 + \Delta \beta_i^2)$. The mean square error (MSE) of the N direct components is $\overline{\Delta L_d^2} = \sum_{i=1}^N \overline{(\Delta L_d^{(i)})^2} / N$. Compared to the sequential separation method (*i.e.*, non-multiplexing, $N = 1$), the SNR gain G can be derived as

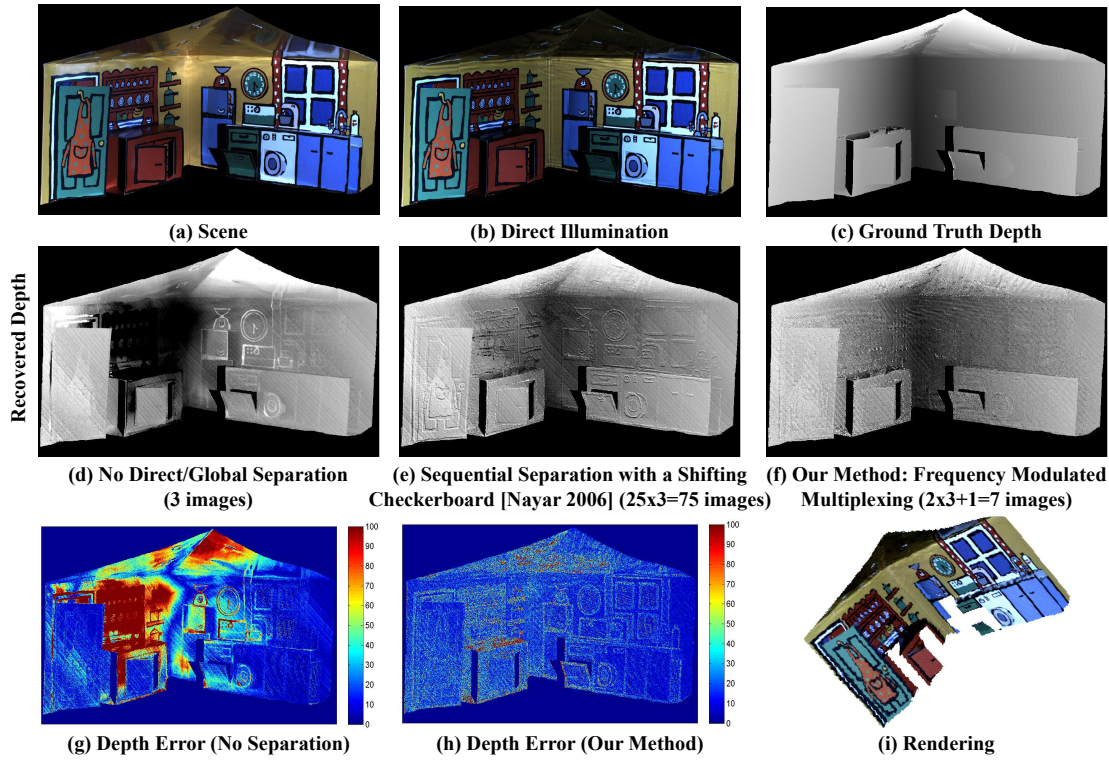


Figure 7. **Depth recovery of a room in a pop-up book using phase shifting** ($N = 3$). (a) The scene exhibits strong inter-reflections. (b) The corresponding direct component, separated with the proposed method. (c) Ground truth depth measured by scanning a single stripe of light. (d)(e)(f) Recovered depth maps for three methods: no direct-global separation, the sequential separation method [11], and the proposed method. Note that performing direct-global separation sequentially using sinusoids requires 9 images. (g)(h) Depth error maps computed using the ground truth. (i) Rendering of (f) for a different view.

$$G = \frac{\text{SNR}_{\text{multiplexed}}}{\text{SNR}_{\text{sequential}}} = \frac{\overline{\Delta L_d}_{\text{sequential}}}{\overline{\Delta L_d}_{\text{multiplexed}}} = (\sigma_0/\sigma) \cdot \sqrt{(2N+1)/3}.$$

where σ_0 is the noise level for the captured image under a single light source and σ is the noise level in each of the captured images under multiple light sources. If the imaging system is read-noise limited, $\sigma \approx \sigma_0$, and thus $G = \sqrt{(2N+1)/3}$; if it is photon-noise limited, $\sigma \approx \sqrt{N}\sigma_0$, and thus $G = \sqrt{(2N+1)/(3N)} \approx 0.83$.

References

- [1] A. Agrawal, M. Gupta, A. Veeraraghavan, and S. Narasimhan. Optimal coded sampling for temporal super-resolution. In *CVPR*, 2010. 3
- [2] T. Chen, H.-P. Seidel, and H. P. A. Lensch. Modulated phase-shifting for 3D scanning. In *CVPR*, 2008. 1, 2
- [3] B. Carrihilla and R. Hummel. Experiments with the intensity ratio depth sensor. In *CVGIP*, 32(3):337–358, 1985. 1
- [4] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying BRDFs from photometric stereo. *IEEE TPAMI*, 32(6):1060–1071, 2010. 7
- [5] M. Gupta, Y. Tian, S. G. Narasimhan, and L. Zhang. (De) Focusing on global light transport for active scene recovery. In *CVPR*, 2009. 2
- [6] M. Harwit and N. J. A. Sloane. *Hadamard Transform Optics*. Academic Press, 1979. 2
- [7] M. Holroyd, J. Lawrence, and T. Zickler. A coaxial optical scanner for synchronous acquisition of 3D geometry and surface reflectance. *SIGGRAPH*, 29(4):99, July 2010. 1, 2
- [8] B. Lamond, P. Peers, A. Ghosh, and P. Debevec. Image-based separation of diffuse and specular reflection using environmental structural illumination. In *ICCP*, 2009. 1, 2
- [9] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen. Programmable aperture photography: Multiplexed light field acquisition. *SIGGRAPH*, 27(3):55:1–55:10, 2008. 3
- [10] S. K. Nayar, K. Ikeuchi, and T. Kanade. Shape from inter-reflection. *IJCV*, 6(3):173–195, 1991. 2, 4
- [11] 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. *SIGGRAPH*, 25(3):935–944, 2006. 1, 2, 3, 4, 5, 6, 7, 8
- [12] J.-I. Park, M.-H. Lee, M. D. Grossberg, and S. K. Nayar. Multispectral imaging using multiplexed illumination. In *ICCV*, 2007. 3
- [13] N. Ratner and Y. Y. Schechner. Illumination multiplexing within fundamental limits. In *CVPR*, 2007. 2, 7
- [14] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. A theory of multiplexed illumination. In *ICCV*, 2003. 2, 5
- [15] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. Multiplexing for optimal lighting. *IEEE TPAMI*, 29:1339–1354, 2007. 2, 5