Micro Phase Shifting

Mohit Gupta
Columbia University
New York, NY 10027
mohitg@cs.columbia.edu

Shree K. Nayar
Columbia University
New York, NY 10027
nayar@cs.columbia.edu

Abstract

We consider the problem of shape recovery for real world scenes, where a variety of global illumination (interreflections, subsurface scattering, etc.) and illumination defocus effects are present. These effects introduce systematic and often significant errors in the recovered shape. We introduce a structured light technique called Micro Phase Shifting, which overcomes these problems. The key idea is to project sinusoidal patterns with frequencies limited to a narrow, high-frequency band. These patterns produce a set of images over which global illumination and defocus effects remain constant for each point in the scene. This enables high quality reconstructions of scenes which have traditionally been considered hard, using only a small number of images. We also derive theoretical lower bounds on the number of input images needed for phase shifting and show that Micro PS achieves the bound.

1. Introduction

Phase shifting is one of the most widely used shape measurement techniques [16]. Because of its precision and low cost, it is applied in surgery, factory automation and digitization of cultural heritage. Like all active scene recovery techniques, phase shifting assumes that scene points are only directly illuminated by the light source. As a result, it recovers erroneous shapes for scenes that have global illumination due to interreflections and subsurface scattering. Furthermore, conventional phase shifting algorithms assume that illumination comes from a perfect point source with infinite depth of field. In practice, all sources have a limited depth of field, resulting in defocus. In order to account for defocus, existing techniques need to capture a large number of input images. It is worth noting that global illumination and defocus effects are ubiquitous - they arise in virtually any real world scene.

In this paper, we introduce a shape recovery technique called *Micro Phase Shifting (Micro PS)* which addresses the problems of global illumination and illumination defocus [1]. While these problems have seen a lot of research activity in the last few years [7, 4, 8, 19, 5], for the first time, we present a technique which is fast, accurate and widely applicable. The key idea is to project sinusoidal patterns with spatial frequencies limited to a narrow high-frequency band. The word *Micro* refers to the small values for both the bandwidth and periods of the patterns. The bandwidth as well as the periods are chosen to be *small enough* so that for each scene point, both global illumination and defocus effects remain constant over all the input images. In contrast, conventional phase shifting has patterns with a broad range of spatial frequencies. Figure 1 shows comparisons between conventional and Micro PS.

Getting around global illumination: Nayar *et al*. showed that high-frequency patterns can be used to separate global illumination from the direct component [14]. Several subsequent phase shifting techniques [4, 7] have built upon this method. These techniques modulate the low frequency patterns with high-frequency sinusoids to remove the global illumination component. In contrast, Micro PS *avoids* the problem of global illumination. Since all the patterns are high frequency sinusoids, global illumination for any given scene point remains constant over all the captured images and hence direct-global separation is not required. As we show in our results, this significantly improves the reconstruction quality.

Invariance to defocus: Illumination defocus manifests as a low-pass filter on the projected images. If the images are sinusoids, as in phase shifting, defocus reduces their amplitude. The amplitude varies significantly with the sinusoid frequency, as well as across the scene due to varying amount of defocus, as shown in the second row of Figure 1. In conventional PS, the amplitudes for different frequencies need to be treated as separate unknowns. In Micro PS, since all the frequencies are in a narrow band, the amplitudes for all the frequencies are approximately the same, and can be treated as a single unknown. As a result, the number of input images required for Micro PS is considerably reduced. We derive a lower bound on the number of input images required for phase-shifting and show that Micro PS achieves this bound.

Resolving depth ambiguities: In conventional PS, high frequency sinusoids provide high-resolution phase (depth) information, albeit in a small depth range.

[1] We do not consider the effects of camera defocus. Camera defocus results in blurring of captured images, resulting in incorrect depths, specially at depth edges.

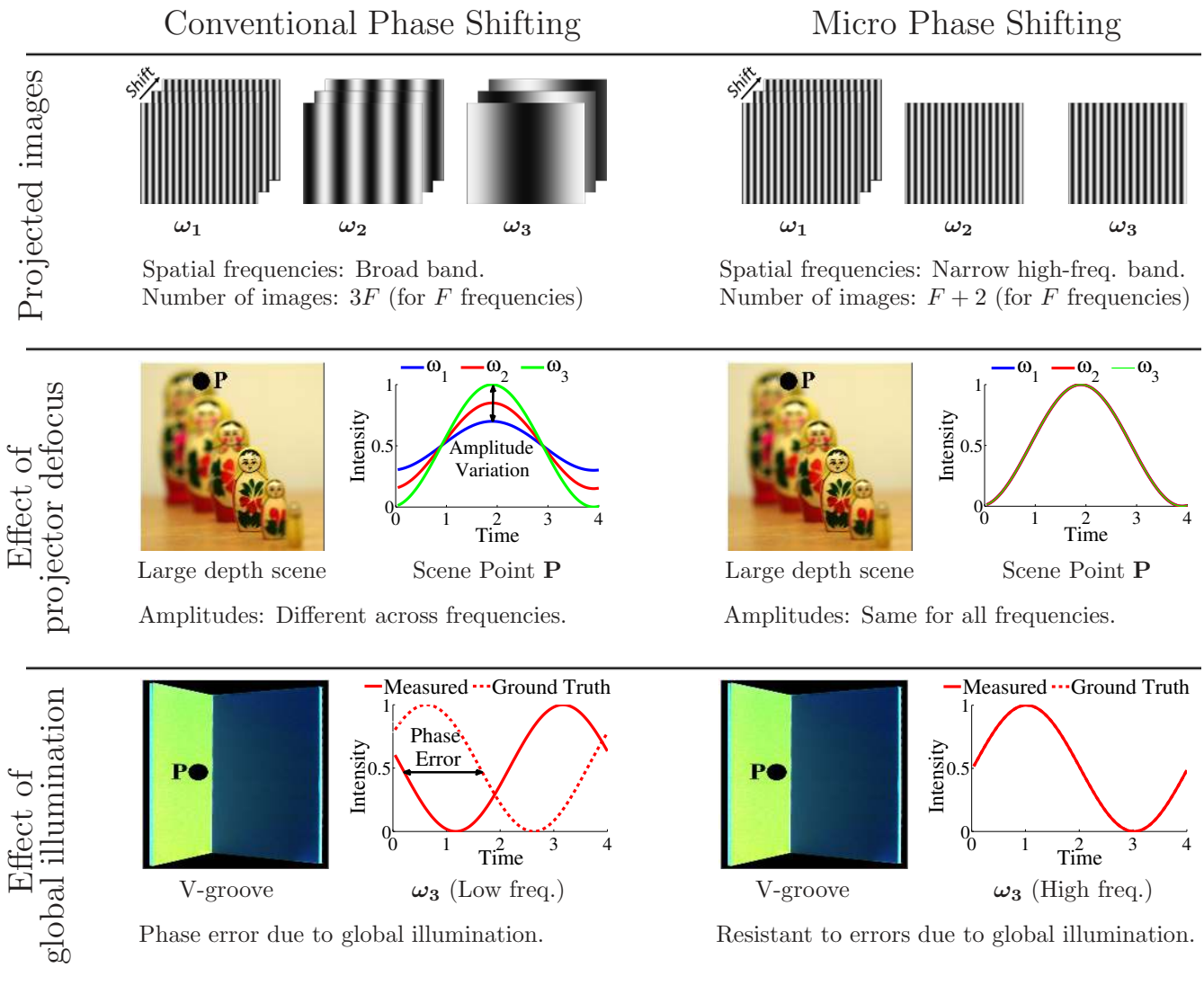


Figure 1. **Conventional versus Micro phase shifting.** The top row shows projected images. The second row shows the effect of projector (source) defocus. Scene point P receives defocused illumination due to the large depth of the scene. The plots show the temporal intensity profiles at **P** as the projected patterns are shifted. For conventional PS, the profiles have different amplitudes for different frequencies. For Micro PS, since all the frequencies are in a narrow band, all the amplitudes are similar. The third row shows the effect of global illumination. The scene consists of a V-groove. **P** receives global illumination due to interreflections. In conventional PS, interreflections result in a large phase error for low-frequency sinusoids. Micro PS is resistant to such errors since all the frequencies are high.

The low-frequency sinusoids are used to disambiguate the phase information over a larger range. This process is called phase unwrapping [20]. One might wonder: How can we perform phase unwrapping with Micro PS, where only high frequencies are used? Fortunately, it is possible to emulate a low-frequency sinusoid with a period equal to the product of the periods of several high-frequency sinusoids. There are algorithms in interferometry literature which combine several high-frequency phase values into a single, unambiguous phase [10, 17]. While originally proposed for interferometry, these algorithms are easily adopted for phase shifting. An example is shown in Figure 2.

Practical implications: With Micro PS, it is now possible to achieve high quality reconstructions of scenes which have traditionally been considered hard, while requiring only a small number of input images. For example, scenes exhibiting a variety of global illumination effects (scattering, interreflections), or combinations of multiple effects - can be handled. Micro PS differs from existing techniques in only the projected patterns. As a result, it can be readily integrated with existing structured light systems.

1.1. Related Work

In 1991, Nayar *et al.* [13] addressed the problem of interreflections in the context of shape recovery. The problem has gained renewed interest in the last

few years. Several methods have been proposed for handling defocus and global illumination effects, e.g., Chandraker *et al.* [2] and Liao *et al.* [12] for photometric stereo, and Gupta *et al.* [9] for shape from illumination defocus.

For structured light triangulation, techniques for handling global illumination have been proposed too. However, most current methods require several tens or hundreds of images [8, 6, 19, 4], require moving cameras or light sources [11, 5, 15], and are often limited to scenarios where only one global illumination effect is dominant [8, 19]. Chen *et al.* [3] use polarizers to remove subsurface scattering. This method does not address the problem of interreflections. In comparison, Micro PS requires a small number $(5-7)$ of input images, does not require any moving parts and can handle scenes where multiple global illumination and defocus effects are present at the same time.

1.2. Background

Phase shifting belongs to the class of active stereo triangulation techniques. The simplest setup for these techniques consists of a projector and a camera. Depth of a scene point is computed by intersecting rays from the corresponding projector and camera pixels. A projector and a camera pixel correspond if the projector pixel directly illuminates the scene point imaged at the camera pixel. Correspondences between camera and projector pixels are established by projecting coded intensity patterns on the scene. In phase-shifting, the projected patterns are spatial sinusoids with different frequencies $\{\omega_1, \omega_2, \ldots, \omega_f, \ldots, \omega_F\}$ [2] [3]. For each frequency, the sinusoid is spatially shifted $N (N \geq 3)$ times, and an image is captured for each shift.

Consider an idealized model of image formation, i.e., scene points receive perfectly focused illumination and there are no global illumination effects. Then, $R_{nf}(c)$, the intensity at a camera pixel c for the n^{th} shift of frequency ω_f is given as:

$$R_{nf}(c) = O_f(c) + A_f(c) * cos\left(\phi_f(p) + \frac{2\pi n}{N}\right), \quad (1)$$

where p is the projector pixel that illuminates the scene point imaged at c. For each frequency ω_f, the set of intensities $R_{nf}(c)$ form a sinusoid as a function of the time instant n. The sinusoid has three unknown parameters - the offset $O_f(c)$, the amplitude $A_f(c)$ and the phase $\phi_f(p)$. The amplitude $A_f(c)$ encapsulates the BRDF of the scene point, the intensity fall-off from the projector, foreshortening and the camera intensity response. The offset $O_f(c)$ includes the contribution of ambient illumination due to other (temporally stable) light sources illuminating the scene. The phase provides the correspondence information. Thus, the goal is to recover the phases $\{\phi_1, \phi_2, \ldots, \phi_f, \ldots, \phi_F\}$. The recovered phases are then combined into a single unambiguous phase ϕ using temporal phase unwrapping [20]. Equation 1 can be written as a system of linear equations:

$$\mathbf{R_f} \quad = \quad \mathbf{M} * \mathbf{U_f}\ , \quad (2)$$

where $\mathbf{R_f}$ is the vector of recorded intensities for the frequency ω_f. $\mathbf{M}$ is an $N \times 3$ matrix, the n^{th} row of which is given as $[1 \;\; cos(\frac{2\pi n}{N}) \;\; -sin(\frac{2\pi n}{N})]$. The unknown vector $\mathbf{U_f} = [O_f(c) \;\; A_f(c) * cos(\phi_f(p)) \;\; A_f(c) * sin(\phi_f(p))]$ is solved using linear least-squares. The phase information ϕ_f is obtained as:

$$A_f(c) \quad = \quad \sqrt{\mathbf{U_f}(2)^2 + \mathbf{U_f}(3)^2}\,, \quad (3)$$

$$\phi_f(p) \quad = \quad acos\left(\frac{\mathbf{U_f}(2)}{A_f(c)}\right). \quad (4)$$

2. Global illumination and Defocus

In this section, we analyze the effects of global illumination and illumination defocus on phase shifting. The intensity at camera pixel c in the presence of global illumination can be written as:

$$R_n(c) = \sum_q \alpha_{qc}\left(O(c) + A(c) * cos\left(\phi(q) + \frac{2\pi n}{N}\right)\right), \quad (5)$$

where α_{qc} is the light transport coefficient between projector pixel q and camera pixel c. For brevity, we drop the subscript f. Intuitively, because of global illumination, the scene point imaged at c receives illumination originating from multiple projector pixels. After simplification, we get:

$$R_n(c) = O'(c) + A'(c) * cos\left(\phi'(c) + \frac{2\pi n}{N}\right), \text{ where} \quad (6)$$

$$O'(c) = \sum_q \alpha_{qc}\, O(c), \quad (7)$$

$$A'(c) = A(c)\sqrt{P(c)^2 + Q(c)^2}, \quad (8)$$

$$\phi'(c) = atan\left(\frac{Q(c)}{P(c)}\right), \quad (9)$$

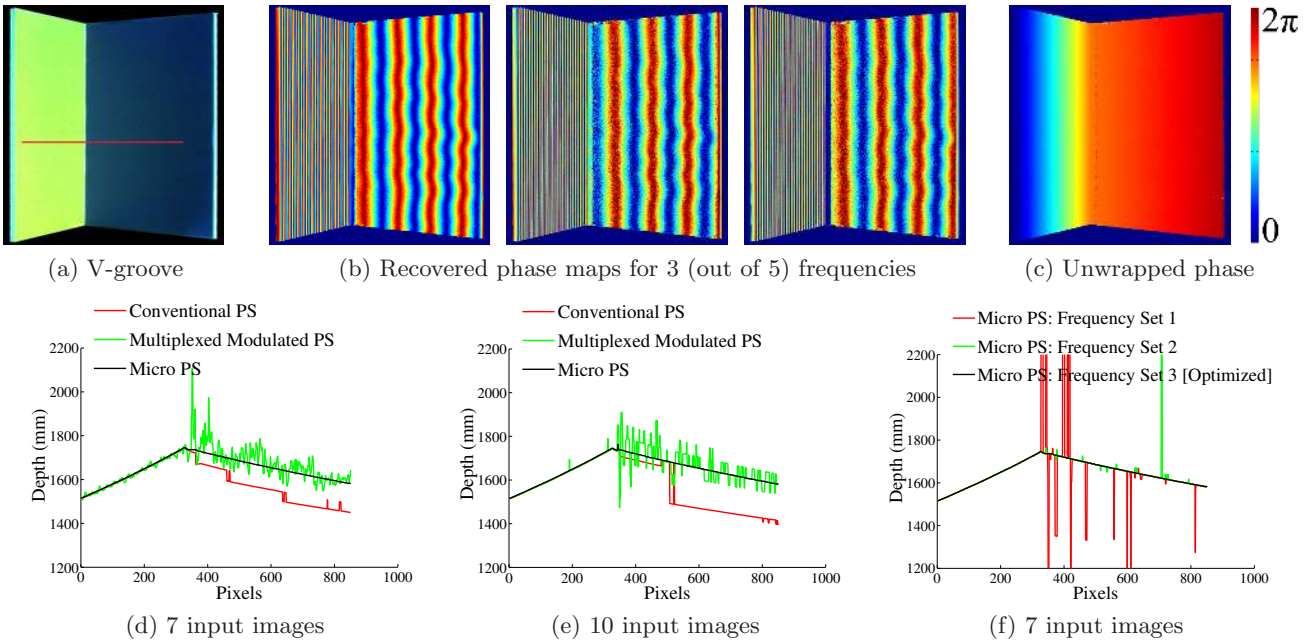
$$P(c) = \sum_q \alpha_{qc}\, cos\left(\phi(q)\right), \quad (10)$$

$$Q(c) = \sum_q \alpha_{qc}\, sin\left(\phi(q)\right). \quad (11)$$

Note that the above model is valid for all forms of global light transport - multi-bounce interreflections, subsurface scattering, etc. Due to global illumination, the phase $\phi'(c)$ of the sinusoid (Eq. 9) is changed. The phase error $\nabla\phi = |\phi(p) - \phi'(c)|$ results in systematic errors in the recovered shape [8, 7]. An example is shown in Figures 2 (d,f).

[2] The number of frequencies F is determined by the desired accuracy and acquisition speed. More the frequencies used (more input images), higher the reconstruction quality.

[3] In this paper, frequencies are represented by the period of the sinusoid, in projector pixel units.



(a) V-groove (b) Recovered phase maps for 3 (out of 5) frequencies (c) Unwrapped phase

(d) 7 input images (e) 10 input images (f) 7 input images

Figure 2. **Micro Phase Shifting in the presence of interreflections.** (a) V-groove scene. (b) Recovered phase maps for 3 out of 5 frequencies. Since all the frequencies belong to a narrow high-frequency band, the phase maps have ambiguities. (c) A single unambiguous phase map is computed using the G-S algorithm [10]. Depth comparisons for different schemes using (d) 7 images and (e) 10 images, along the red line in (a). Conventional PS results in large systematic errors. Modulated PS reduces the systematic bias due to interreflections, but suffers from low SNR. Depth using Micro PS is nearly error free. (f) Comparisons for different frequency sets for Micro PS. The frequency selection procedure (Section 3.2) is used to determine the optimal frequency set.

Illumination defocus manifests as a local blur on the projected, and hence, observed sinusoids, thus lowering their amplitude. The reduced amplitude is a function of both the frequency of the sinusoid and the amount of defocus. As a result, the observed amplitudes are different for different scene points. An example is shown in Figure 3 (b). Points receiving global and defocused illumination show larger variations in the amplitude with respect to the projected frequency. Because of this, in conventional phase shifting, the amplitudes for different frequencies need to be treated as separate unknowns. The offsets at each scene point (Eq. 7) are independent of the frequency, and can be treated as a single unknown, as shown in Figure 3 (c).

3. Micro Phase Shifting

We now present our technique, which addresses the problems of global illumination and defocus. The *key idea* is simple - use patterns so that all the spatial frequencies are within a narrow high-frequency band. Let the frequency band be characterized by the mean frequency ω_m, and the width of the band δ. All the frequencies $\Omega = \{\omega_1, \ldots, \omega_F\}$ lie within the band $\left[\omega_m - \frac{\delta}{2},\ \omega + \frac{\delta}{2}\right]$. For Micro PS, the frequency set Ω must meet the following two conditions: (a) ω_m is sufficiently high (period is small) so that global illumination does not introduce errors in the recovered phase, and (b) the bandwidth δ is sufficiently small so that the amplitudes for all the frequencies are approximately the same, i.e., $A_1 \approx A_2 \approx \ldots \approx A_F \approx A$.

There is a tradeoff while selecting the mean frequency ω_m. Theoretically, the higher ω_m is, the more resistant it is to global illumination. However, due to optical aberrations, projectors cannot project arbitrarily high frequencies reliably. For typical currently available projectors with a resolution of 1024×768 pixels, we found that the mean frequency $\omega_m = 16$ pixels period is sufficiently high to prevent global illumination errors for a large collection of scenes. Similar frequencies were chosen in [14] to separate the direct and global components of an image. When projectors with small optical aberrations are available, patterns with higher frequencies can be used. For an analytical expression for the sinusoid frequencies for which global illumination does not influence the phase, please see the technical report available at the following location [1].

Similarly, there is a tradeoff while selecting the bandwidth δ. Theoretically, the narrower the bandwidth, the fewer the unknowns (due to invariance to defocus). However, due to finite spatial resolution of the projectors and finite intensity resolution of the projectors and cameras, two frequencies that are very close cannot be distinguished reliably. This imposes a lower-bound on δ. If ϵ is the minimum difference between two frequencies so that they can be distinguished reliably, $\delta \geq (F-1)\epsilon$.

We compute δ empirically by measuring the amplitudes for different frequencies around the mean fre-

quency ω_m. The amplitudes are averaged over several scene points receiving different amounts of global illumination and defocus. Then, δ is chosen to be the largest value so that the amplitudes for all the frequencies in the resulting frequency band are approximately the same (the maximum difference in the amplitudes between any pair of frequencies in the band is less than 1%). For our projector with a resolution of 1024×768 pixels, δ was computed to be approximately 3 pixels for $\omega_m = 16$ pixels. The resulting frequency band is $[14.5, 17.5]$ pixels.

Algorithm: Using these properties, we design a phase recovery algorithm which requires capturing only $F+2$ images for F frequencies. Three images are captured for the first frequency. Subsequently, one image is captured for each of the remaining $F-1$ frequencies:

$$R_n(c) = \begin{cases} O(c) + A(c) * cos\left(\phi_1(p) + (n-1)\frac{2\pi}{3}\right) \\ \text{if } 1 \leq n \leq 3, \\ \\ O(c) + A(c) * cos\left(\phi_{n-2}(p)\right) \\ \text{if } 4 \leq n \leq F+2. \end{cases} \tag{12}$$

The above system of equations can be solved jointly as a single linear system:

$$\mathbf{R_{micro}} = \mathbf{M_{micro}} * \mathbf{U_{micro}}\ , \tag{13}$$

where $\mathbf{R_{micro}}$ is the vector of recorded intensities. $\mathbf{M_{micro}}$ is a square matrix of size $F+2$, and is given as $\mathbf{M_{micro}} =$

$$\begin{bmatrix} 1 & a_1 & 0 & 0\dots0 \\ 1 & a_1 * cos(\frac{2\pi}{3}) & -a_1 * sin(\frac{2\pi}{3}) & 0\dots0 \\ 1 & a_1 * cos(\frac{4\pi}{3}) & -a_1 * sin(\frac{4\pi}{3}) & 0\dots0 \\ 1 & 0 & 0 & \\ \vdots & \vdots & \vdots & \mathbb{I}_{F-1} \\ 1 & 0 & 0 & \end{bmatrix}, \tag{14}$$

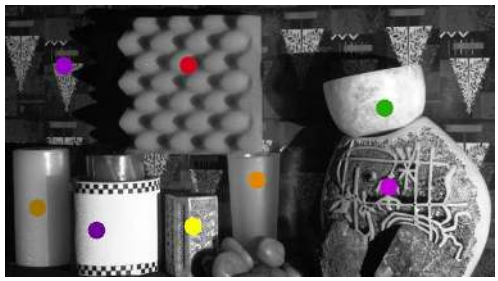
where $\mathbb{I}_{F-1}$ is an identity matrix of size $F-1 \times F-1$. The unknown vector,

$$\mathbf{U_{fact}} = \begin{bmatrix} O(c) \\ A(c) * cos(\phi_1(p)) \\ A(c) * sin(\phi_1(p)) \\ A(c) * cos(\phi_2(p)) \\ \vdots \\ A(c) * cos(\phi_F(p)) \end{bmatrix}, \tag{15}$$

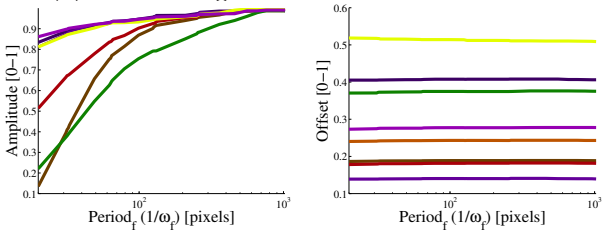
is computed by solving the linear system given in Eq. 13. The individual phases are computed by:

$$A(c) = \sqrt{\mathbf{U_{fact}}(2)^2 + \mathbf{U_{fact}}(3)^2}, \tag{16}$$

$$\phi_f(p) = \begin{cases} acos\left(\frac{\mathbf{U_{fact}}(f+1)}{A(c)}\right) & \text{if } f = 1, \\ acos\left(\frac{\mathbf{U_{fact}}(f+2)}{A(c)}\right) & \text{if } 2 \leq f \leq F. \end{cases} \tag{17}$$



(a) Scene with global illumination and defocus



(b) Normalized amplitudes (c) Offsets

Figure 3. **Variation in amplitude with sinusoid frequency (period).** (a) Scene with global illumination (subsurface scattering, diffusion) and defocus. (b) Amplitudes for points marked in (a). Because of global illumination and defocus, the variation in amplitudes is different for different scene points. (c) The offsets are invariant across frequencies, and can be treated as a single unknown.

3.1. Phase unwrapping

The set of frequencies used in Micro PS belong to a narrow high-frequency band. There is no unit frequency, which is conventionally used for disambiguating the high-frequency phases. Thus, conventional phase unwrapping algorithms are not applicable here. The problem of disambiguating the phases without the unit frequency has been studied in the field of interferometry [10, 17]. With the Gushov-Solodkin (G-S) algorithm [10], it is possible to combine several high-frequency phases into a single low-frequency phase. If the periods of the high-frequency sinusoids are pairwise co-prime (no common factors), then a low-frequency sinusoid with a period equal to the product of the periods of all the high-frequency sinusoids can be emulated. The details required for implementing the algorithm and the MATLAB code are given on [1]. Figure 2 shows an example with the wrapped high-frequency phases and the final unwrapped phase.

3.2. Frequency Selection for Micro Phase Shifting

How should the spatial frequencies of the projected patterns be selected for Micro PS? [4]. In this section, we outline our frequency selection algorithm for Micro PS. See [1] for more details and analysis.

The frequencies should be chosen so that depth errors due to incorrect phase computations (resulting

[4]The optimal set of spatial frequencies for conventional PS form a geometric series, with the smallest frequency (largest period) being the unit frequency [18].

from noise and global illumination) are minimized. For a camera pixel, suppose the correct correspondence is projector column p, but the computed correspondence is column q. The resulting depth error is proportional to the phase error $\triangle\phi = |p - q|$. For F-frequency Micro PS, each projector column is encoded with a unique $F + 2$ dimensional intensity vector. In order to minimize the probability of a phase error, the optimal set of frequencies should maximize the distance d_{pq} between vectors corresponding to distinct projector columns. For a given frequency set Ω, the average weighted distance between intensity vectors is:

$$E(\Omega) \quad = \quad \frac{1}{N^2} \sum_{p,q=1}^{N} |p - q| \; d_{pq}, \tag{18}$$

where N is the total number of projector columns. For d_{pq}, we used the norm-2 Euclidean distance. Then, the optimal set of frequencies in the frequency band $[\omega_{min}, \omega_{max}]$ is the one that minimizes $E(\Omega)$:

$$\Omega* \quad = \quad \arg\min_{\Omega} E(\Omega), \; \omega_f \in [\omega_{min}, \omega_{max}] \; \forall f \, . \tag{19}$$

This is a constrained F-dimensional optimization problem. We used the simplex search method implemented in the MATLAB optimization toolbox to solve this. For the frequency band of $[14.5, 17.5]$ pixels and $F = 5$, the above procedure returns the following frequency set: $[14.57, 16.09, 16.24, 16.47, 16.60]$ pixels. Figure 2 (e) shows a comparison of reconstructions of the optimized set versus two randomly selected frequency sets. Notice the spikes in the depth maps for the two unoptimized sets (first two). In contrast, the reconstruction for the optimized set is error-free.

3.3. Proof of Optimality

We now show that Micro Phase Shifting is theoretically optimal in terms of the number of input images.

Lemma 1 *The minimum number of images required for F-frequency phase-shifting is $F + 2$.*

Proof 1 *The proof involves counting the number of unknowns for each camera pixel. For a pixel, the unknowns are (a) the direct illumination component, (b) the ambient+global illumination component, and (c) a phase value for each frequency. Given F frequencies, the total number of unknowns is $F+2$. Since the system of equations is linear, the minimum number of equations (images) required for F-frequency phase-shifting is $F + 2$.*

Corollary 1 *Micro Phase Shifting is theoretically optimal in terms of the number of input images.*

This directly follows from Lemma 1 and the fact that Micro PS requires $F + 2$ images for F-frequency phase shifting [5].

[5] The lower bound of F for Micro PS is 2, since a minimum of 2 high-frequency sinusoids are required to emulate a low-frequency sinusoid. In practice, F depends on the complexity of the scene. In our experiments, $3 \leq F \leq 5$ was found to be sufficient.

3.4. Comparison with the State of the Art

Modulated phase shifting needs to explicitly separate global illumination by modulating the low-frequency patterns with high-frequency sinusoids [4, 7]. Each low-frequency pattern is modulated by at least 3 high-frequency shifting sinusoids. Since a minimum of 3 low-frequency patterns are required to compute the phase, at least 9 images are needed per low frequency. Recently, Gu *et al.* [7] showed that by using multiplexed illumination, the number of input images for each low frequency can be reduced to 7.

Let F_l be the number of low frequencies used, and F_h be the number of high frequencies. Since the high-frequency patterns do not require explicit separation [14], the minimum number of images required with the current state of the art is $7F_l + 3F_h$. For Micro PS, explicit direct-global separation is not required. As a result, Micro PS needs only $F_l + F_h + 2$ input images. For example, in order to measure phase for one high frequency and one low-frequency, modulated phase shifting would require $7+3 = 10$ images. In contrast, given a budget of 10 images, Micro PS can use 8 different frequencies, resulting in significantly higher quality reconstructions.

4. Experiments

For our experiments, we used a Sanyo PLC-XP18N projector. The number of projector columns is 1024. Hence, unit frequency patterns have a period of 1024 pixels. For modulated PS, at least 7 input images are needed (Section 3.4). Given a budget of seven input images, only a single unit frequency can be used. In case of 10 input images, one high frequency and one unit frequency can be captured. For conventional PS, a budget of 7 and 10 input images allows 2 and 3 frequencies, respectively. Camera defocus was minimized by capturing images with a small camera aperture.

Figure 2 shows depth recovery results for a V-groove with interreflections. As shown in Figures 2 (d,e), conventional PS (red curve) results in significant as well as systematic errors. Modulated PS reduces the systematic bias by separating the effects of interreflections. However, the resulting depth estimates suffer due to low SNR of the separated direct illumination component. In contrast, Micro PS produces nearly error-free results with the same number of input images.

Figures 4 (a-e) show shape recovery results for scenes with different global illumination and defocus effects. The ceramic bowl has strong interreflections. The lemon skin is translucent, resulting in subsurface scattering. The depth-range for the Russian dolls scene is large, resulting in illumination defocus. The wax bowl is challenging because it has *both strong interreflections and subsurface scattering.*

Conventional PS results in large and systematic errors due to interreflections and low SNR in low albedo regions (e.g., regions on the Russian dolls). Modulated PS requires a large number of images per frequency. With a budget of 7 images, only the unit frequency can be acquired, which is not sufficient for accurate reconstruction. Moreover, the explicit separation of direct component from the captured images further enhances the noise. The problem is especially severe for regions of low albedo and highly translucent materials (wax bowl and lemon), for which the direct component is a small fraction of the total radiance. Micro PS does not require explicit separation and captures many more spatial frequencies given the same image budget, thus resulting in high quality reconstructions.

The metal bowl illustrates a failure case. Due to high-frequency interreflections, the reconstructed shapes for all the schemes have large holes and errors. For more results, see [1].

5. Discussion

Scope and limitations: While Micro PS reduces the errors due to global illumination, it may not completely remove them. For example, in the presence of high-frequency light transport such as mirror interreflections, Micro PS is prone to errors (see Figure 4 (e) for an example). However, for scenes where the frequency of light transport is less than the frequency of the sinusoids used, Micro PS will mitigate the effects of global illumination.

Error and resolution characteristics: Although the error characteristics of Micro and Conventional PS are different due to different projected patterns and decoding algorithms, both can resolve correspondences with sub-pixel accuracy, as both belong to the class of continuous coding schemes. However, in case of low SNR, the resolution is often limited by light source and camera noise. There is a tradeoff between depth resolution and the number of input images. In the examples shown in the paper, a relatively small number of input images are used, resulting in a low SNR. If more images are used, noise ceases to be the limiting factor and sub-pixel resolution can be achieved.

Polarization and Micro PS: Polarization has been used to reduce the effect of sub-surface scattering in phase shifting [3]. This approach can also be used in conjunction with Micro PS by placing polarizers in front of the camera and the projector.

Frequency selection: The frequency selection algorithm (Section 3.2) does not necessarily yield the optimal frequency set. The optimal set of frequencies depends on the noise levels and resolutions of the projector and the camera, scene's albedos and light transport characteristics. While high frequencies are preferred for coping with global illumination, they are also more prone to amplitude loss, and hence, low SNR. In future, we envisage incorporating the projector and camera characteristics and coarse light transport characteristics of the scene (acquired with a few initial measurements) in the frequency selection algorithm.

Acknowledgments: This research was supported in parts by NSF (grant number IIS 09-64429) and ONR (grant number N00014-11-1-0285).

References

[1] Project webpage. *http://www.cs.columbia.edu/CAVE/projects/MicroPhaseShifting/*. 4, 5, 7

[2] M. Chandraker, F. Kahl, and D. Kriegman. Reflections on the generalized bas-relief ambiguity. In *CVPR*, 2005. 3

[3] T. Chen, H. P. A. Lensch, C. Fuchs, and H. peter Seidel. Polarization and phase-shifting for 3D scanning of translucent objects. In *CVPR*, 2007. 3, 7

[4] T. Chen, H.-P. Seidel, and H. P. A. Lensch. Modulated phase-shifting for 3D scanning. In *CVPR*, 2008. 1, 3, 6

[5] C.Hermans, Y.Francken, T.Cuypers, and P.Bekaert. Depth from sliding projections. In *CVPR*, 2009. 1, 3

[6] V. Couture, N. Martin, and S. Roy. Unstructured light scanning to overcome interreflections. In *ICCV*, 2011. 3

[7] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar. Multiplexed illumination for scene recovery in the presence of global illumination. In *ICCV*, 2011. 1, 3, 6

[8] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. Narasimhan. Structured light 3D scanning the presence of global illumination. In *CVPR*, 2011. 1, 3

[9] M. Gupta, Y. Tian, S. G. Narasimhan, and L. Zhang. A combined theory of defocused illumination and global light transport. In *IJCV*, To appear. 3

[10] V. I. Gushov and Y. N. Solodkin. Automatic processing of fringe patterns in integer interferometers. *Optics Lasers Engineering*, 14, 1991. 2, 4, 5

[11] M. Holroyd, J. Lawrence, and T. Zickler. A coaxial optical scanner for synchronous acquisition of 3D geometry and surface reflectance. *ACM Trans. Graph.*, 29(3), 2010. 3

[12] M. Liao, X. Huang, and R. Yang. Interreflection removal for photometric stereo by using spectrum-dependent albedo. In *CVPR*, 2011. 3

[13] S. K. Nayar, K. Ikeuchi, and T. Kanade. Shape from Interreflections. *IJCV*, 6(3), 1991. 2

[14] 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. *ACM Trans. Graph.*, 25(3), 2006. 1, 4, 6

[15] J. Park and A. C. Kak. 3D modeling of optically challenging objects. *IEEE TVCG*, 2008. 3

[16] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43, 2010. 1

[17] M. Takeda, Q. Gu, M. Kinoshita, H. Takai, and Y. Takahashi. Frequency-multiplex fourier-transform profilometry: A single-shot three-dimensional shape measurement of objects with large height discontinuities and/or surface isolations. *Applied Optics*, 36(22), 1997. 2, 5

[18] C. E. Towers, D. P. Towers, and J. D. C. Jones. Absolute fringe order calculation using optimised multi-frequency selection in full-field profilometry. *Optics and Lasers in Engineering*, 43, 2005. 5

[19] Y. Xu and D. Aliaga. An adaptive correspondence algorithm for modeling scenes with strong interreflections. *IEEE TVCG*, 2009. 1, 3

[20] H. Zhao, W. Chen, and Y. Tan. Phase-unwrapping algorithm for the measurement of three-dimensional object shapes. *Applied Optics*, 33(20), 1994. 2, 3

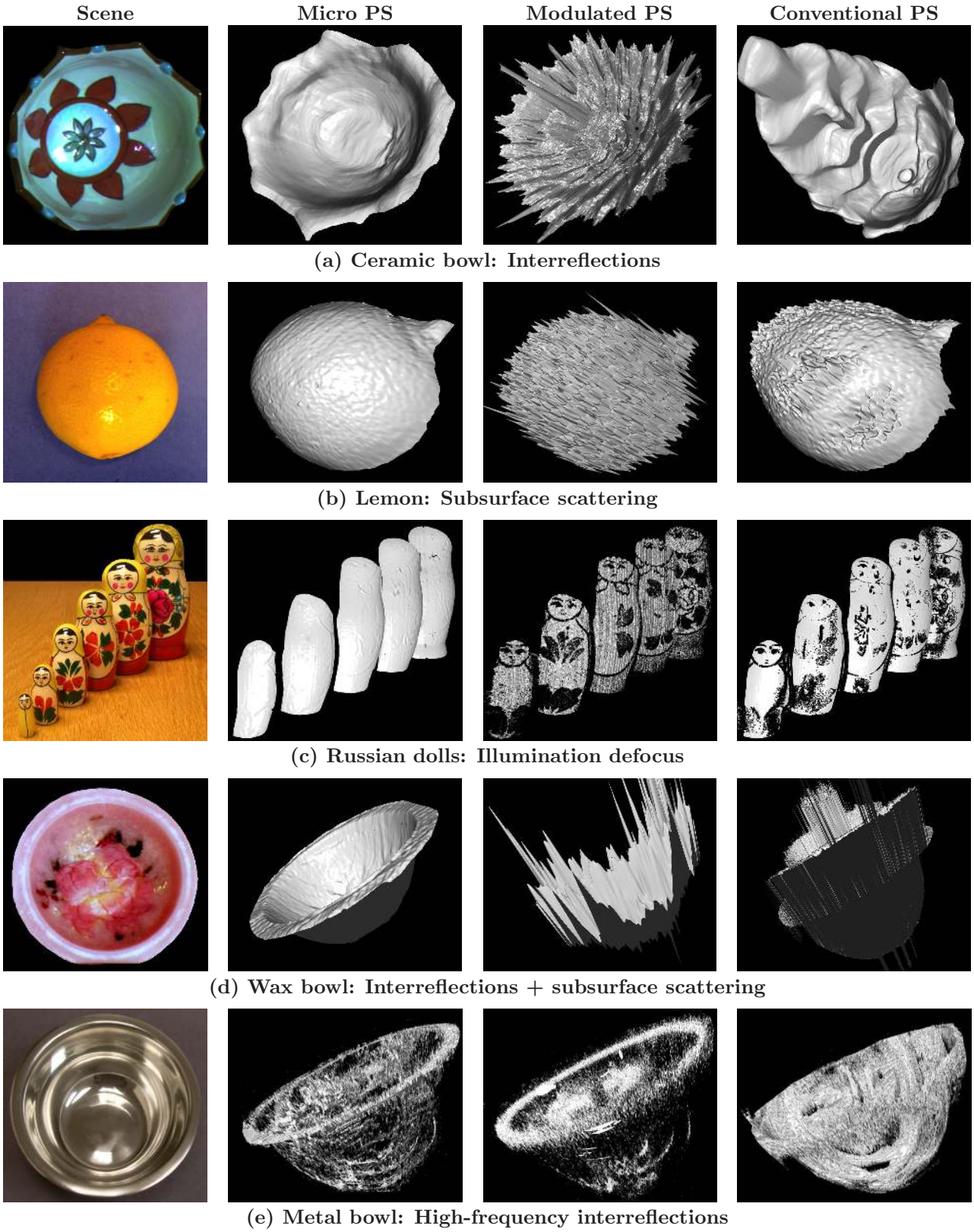


Figure 4. **Shape recovery for scenes with different global illumination and defocus effects.** For every scene, 7 input images were used for each scheme. (a) Ceramic bowl with interreflections. (b) Lemon with subsurface scattering. (c) Russian Dolls. Large depth-range of the scene results in illumination defocus. (d) Wax bowl with both interreflections and strong subsurface scattering. With only 7 input images, Micro PS achieves a high quality reconstruction. In contrast, conventional PS results in systematic errors due to interreflections, and both conventional and modulated PS suffer due to low SNR of the direct irradiance component in low albedo and highly translucent regions. (e) Metal bowl (failure case). Due to high-frequency specular interreflections, reconstructed shapes using all the schemes have large holes and errors.