Causes and Corrections for Bimodal Multi-path Scanning with Structured Light

Yu Zhang[1], Daniel L. Lau[2], Ying Yu[2]
[1] Nanjing University
[2] University of Kentucky

zhangyu606@gmail.com dllau@uky.edu ying.yu@uky.edu

Abstract

Structured light illumination is an active 3D scanning technique based on projecting/capturing a set of striped patterns and measuring the warping of the patterns as they reflect off a target object's surface. As designed, each pixel in the camera sees exactly one pixel from the projector; however, there are multi-path situations when the scanned surface has a complicated geometry with step edges and other discontinuities in depth or where the target surface has specularities that reflect light away from the camera. These situations are generally referred to multi-path where a camera pixel sees light from multiple projector positions. In the case of bimodal multi-path, the camera pixel receives light from exactly two positions which occurs along a step edge where the edge slices through a pixel so that the pixel sees both a foreground and background surface. In this paper, we present a general mathematical model to address the bimodal multi-path issue in a phase-measuring-profilometry scanner to measure the constructive and destructive interference between the two light paths, and by taking advantage of this interesting cue, separate the paths and make two decoupled phase measurements. We validate our algorithm with a number of challenging real-world scenarios, outperforming the state-of-the-art method.

1. Introduction

Structured light illumination (SLI) refers to a method of 3D scanning that uses a projector to project a series of light striped patterns such that a camera can reconstruct depth based on the warping of the pattern over the target object's surface [22, 11, 16, 14, 20, 13, 27, 26]. Example of SLI includes single pattern techniques which project a static pattern that is continuously projected and from which a 3D reconstruction can be made from a single snapshot [11, 2, 10, 8].

Multiple pattern SLI scanners, alternatively, project a series of patterns, trading temporal resolution for spatial resolution such that each pixel can be independently processed from its neighbors to produce a single point for each pixel in the camera. In Phase Measuring Profilometry (PMP), the row coordinates of each pixel are encoded through phase modulation [28, 20, 11, 3]. These PMP scanners are common for industrial metrology applications with resolutions that can be below 10 microns.

As an active imaging technique, structured light is susceptible to errors and distortions caused by the redirection of the projected light to form multiple paths from projector to camera besides the direct path of projector to target to camera [29]. It is a common problem and one of great interest to researchers because of the potentially catastrophic effects on scans. The same problem can be found in a range of 3D imaging modalities such as time-of-flight (ToF) where light will reflect off specular surfaces onto neighboring surface points before reflecting back to the camera.

Examples of how to deal with multi-path issues in ToF include Dorrington *et al.* [7] as well as Bhandari *et al.* [1] and Godbaz *et al.* [12] who take the common approach of making multiple depth measurements over many different modulation frequencies such that they derive a set of equations from which to fit the phase and magnitude of a multitude of possible component paths. Freedman *et al.* [9] assume sparsity in reflection and assume the problem is restricted to a small number of multi-path components, which restrict further extension to other scenarios.

Naik *et al.* [23] take the approach of deriving a light transport model [24] to combine the standard measurements from a ToF camera with information from direct and global light transport. By doing so, they separate the phase associated with the direct light path, placing all subsequent paths into a single indirect light component. O'Toole *et al.* [26, 25] employ the epipolar geometry constraint and re-design the optical system to separate the direct and indirect light paths. They modify the optical system and block the global component during the data capture procedure. Gupta *et al.* [15] study temporal illumination and report that global light transport vanishes at high frequencies. They propose a ToF based shape recovery technique and a method to separate direct and global light. Kadambi *et al.* [17] use

a coded illumination ToF camera to achieve light sweep imaging with multi-path correction.

Dedrick [6] identifies multi-path in SLI scans without presenting an effective algorithm for extracting the absolute paths from the collected scans. Courture *et al.* [4] design special pattern to overcome interreflections which is quite different from traditional phase shifting pattern. Nayar *et al.* [24] show that the radiance of a scene point is due to direct illumination of the point by the source and global illumination arising from diffuse interreflection, subsurface scattering, volumetric scattering and translucency. Gupta and Nayar [14] use this conclusion [24] and present a state-of-the-art approach using a narrow, high frequency band structured light pattern to separate direct and global illumination for shape recovery for real scenes. However, the separated direct component can still suffer from bimodal multi-path. Their method cannot address bimodal multi-path in the direct image and will cause severe artifacts in the reconstruction because they still use traditional phase shifting method to solve phase/depth in the direct component.

1.1. Contributions

While this paper limits its discussion to the two-path problem, we note that it can easily be extended to more paths; however, the two-path problem is particularly important because its presence is unavoidable in any scanning situation where a target surface includes a step edge, resulting in pixels of the camera collecting light from a fore and background surface. We further note that this problem has received little attention in the published literature even through extensive studies have been published on phase-shifting structured light.

This paper studies the causes and corrections for bimodal multi-path in terms of a structured light scanner and includes an intuitive construction that explains how paths interact as a function of the spatial frequency to produce standing waves of constructive and destructive interference. In doing so, we establish an equation for this interference such that we can visualize multi-path as a sinusoidal pattern plotted versus pattern frequency and varying as a function of the phase difference between component paths.

The experimental results that we present also deal with a problem unique to structured light, and that is the low-pass filtering effect of the component optics that cause high spatial frequency patterns to have a lower amplitude than low spatial frequencies. In traditional structured light, this is an issue that is largely ignored since the final phase is determined by the high spatial frequency, with lower frequencies used for unwrapping the high frequencies. This paper deals directly with the issue by establishing an envelope function during scanner calibration such that we can observe bimodal multi-path in the presence of a non-flat

spatial frequency response.

To the best of our knowledge, we are the first to report the interesting constructive and destructive cue for bimodal multi-path using signal processing theory and present a practical approach to simultaneously identify and extract the dominant and non-dominant phases/magnitudes by taking advantage of that cue in an intuitive way without any hardware modifications or additional requirements for customized patterns. As a result, it is easy to be integrated with existing structured light systems. Central to this separation, we propose the idea of a zero-frequency PMP patterns which projects a time-varying but spatially constant structured light patterns as a way to observe the modulated light component absent the multi-path interference that may otherwise partially cancel the modulated light.

2. Background

Three-dimensional surface scanning by means of structured light is performed using a series of striped patterns projected onto a target scene and captured by a digital camera, placed at a triangulation angle of the projector's line of sight. The pixels of the captured images are then processed to identify a unique projector row coordinate for which the subject camera pixel corresponds. Perhaps one of the simplest means of SLI is through the use of phase-shift keying where the component patterns are defined by the set, $\{I_n^p : n = 0, 1, \ldots, N-1\}$, according to:

$$I_n^p(x^p, y^p) = \frac{1}{2} + \frac{1}{2} \cos\left(2\pi(\frac{n}{N} - y^p)\right). \quad (1)$$

where (x^p, y^p) is the column and row coordinate of a pixel in the projector, I_n^p is the intensity of that pixel in a projector with dynamic range from 0 to 1, and n represents the phase-shift index over the N total patterns.

For reconstruction, a camera captures each image where the sine wave pattern is distorted by the scanned surface topology, resulting in the patterned images expressed as:

$$I_n^c(x^c, y^c) = A^c + B^c \cos\left(\frac{2\pi n}{N} - \theta\right). \quad (2)$$

where (x^c, y^c) is the coordinates of a pixel in the camera while $I_n^c(x^c, y^c)$ is the intensity of that pixel. θ represents the phase value of the captured sinusoidal pattern. The term A^c is the averaged pixel intensity across the pattern set that includes the ambient light component, which can be derived according to:

$$A^c = \frac{1}{N} \sum_{n=0}^{N-1} I_n^c(x^c, y^c). \quad (3)$$

Correspondingly, the term B^c is the intensity modulation of a given pixel and is derived from $I_n^c(x^c, y^c)$ in terms of real

and imaginary components where:

$$B_{\mathcal{R}}^c = \sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos\left(\frac{2\pi n}{N}\right) \qquad (4)$$

and

$$B_{\mathcal{I}}^c = \sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin\left(\frac{2\pi n}{N}\right) \qquad (5)$$

such that

$$B^c = \|B_{\mathcal{R}}^c + jB_{\mathcal{I}}^c\| = \left\{B_{\mathcal{R}}^{c\,2} + B_{\mathcal{I}}^{c\,2}\right\}^{\frac{1}{2}}, \qquad (6)$$

which is the amplitude of the observed sinusoid.

If $I_n^c(x^c, y^c)$ is constant or less affected by the projected sinusoid patterns, B^c will be close to zero. Thus B^c is employed as a shadow noise detector/filter [18] such that the shadow-noised regions, with small B^c values, are discarded from further processing. Of the reliable pixels with sufficiently large B^c, θ represents the phase value of the captured sinusoid pattern derived as:

$$\theta = \angle(B_{\mathcal{R}}^c + jB_{\mathcal{I}}^c) = \arctan\left\{\frac{B_{\mathcal{I}}^c}{B_{\mathcal{R}}^c}\right\}, \qquad (7)$$

which is used to derive the projector row according to $\theta = 2\pi y^p$.

Given that the reconstructed θ is affected by distortions in the projector/camera such as thermal noise [5] or gamma [21], Eq. (2) is commonly modified to include higher spatial frequencies according to:

$$I_n^p(x^p, y^p) = \frac{1}{2} + \frac{1}{2}\cos\left(2\pi(\frac{n}{N} - Ky^p)\right), \qquad (8)$$

where K is the number of sinusoidal wavelengths across the projector in any one frame. These higher frequency scans result in ambiguities in θ which are resolved by phase unwrapping via lower frequency Ks. For instance, one might use three separate scans with $K = 1$, 4, and 16 using the $K = 1$ scan to unwrap the $K = 4$ scan and then using that resulting scan to unwrap the $K = 16$ scan. This procedure results in a scan with $1/16^{th}$ the noise of the $K = 1$ scan where $y^p = \theta/(K2\pi)$.

In choosing K, an experienced operator knows that quantization noise in the projector requires that K be selected such that the corresponding wavelength of the spatial sinusoids corresponds to integer multiples of N pixels; otherwise, banding artifacts are visible in the reconstruction of θ. At the same time, larger values of N result in less thermal noise as well as in the elimination of gamma. So while a small N allows for higher spatial frequency K, it also results in high levels of Gaussian noise in θ while also making θ susceptible to gamma distortion. As such, we recommend an N no smaller than 8, meaning a VGA projector is limited to a maximum frequency of $K = 60$ with the sinusoid moving 1 pixel with each step in n.

3. Bimodal Multi-Path Model

In signal processing, it is often convenient to assume a sample of an analogue signal is its value at an infinitesimally small amount of time, but in fact, a sample is the average value of the signal over a fixed interval in time. In digital cameras, a pixel collects light over a fixed angle in the horizontal and vertical directions. As such, a camera pixel can collect light from both a fore and background surface. So a more accurate version of Eqs. (4) and (5) can be written as:

$$B_{\mathcal{R}}^c = B_{\mathcal{R}}^{c,f} + B_{\mathcal{R}}^{c,b} \qquad (9)$$

and

$$B_{\mathcal{I}}^c = B_{\mathcal{I}}^{c,f} + B_{\mathcal{I}}^{c,b} \qquad (10)$$

where we added the superscripts f and b to distinguish between the foreground and background components on $B_{\mathcal{R}}^c$ and $B_{\mathcal{I}}^c$.

Now notice that increasing the spatial frequency of the PMP patterns by a factor of K increases the phase term by an equal amount while keeping the amplitude of the sinusoid constant. In the case of multi-path, this frequency scaling has a far different effect as illustrated graphically in Fig. 1 where we show (left) the foreground and background components assuming unit frequency while (center) and (right) show the same components when $K = 8$ and $K = 12$. The red vectors in Fig. 1 show the complex vectors formed by $B_{\mathcal{R}}^{c,f}$ and $B_{\mathcal{I}}^{c,f}$ and $B_{\mathcal{R}}^{c,b}$ and $B_{\mathcal{I}}^{c,b}$, while the blue vector shows the superimposed vectors forming the single vector formed by $B_{\mathcal{R}}^c$ and $B_{\mathcal{I}}^c$.

By using a frequency scaling of K, we expect the direction or phase of the foreground and background vectors to scale by an equal amount. Graphically, this is depicted by a rotation of the vectors around the origin. Notice, though, that by rotating the vectors separately, it is quite likely that the phase of the combined vectors are not equal to the scaling of the phase term prior to frequency scaling. Likewise, the vectors may swing from constructively interfering where the magnitude of the combined vectors is equal to the sum of the individual magnitudes to destructively interfering where the magnitude of the combined vectors is equal to the difference of the individual magnitudes.

4. Bimodal Multi-Path Reconstruction

Mathematically, the magnitude and phase of the subject pixel can be defined according to vector $\vec{AB}$, with foreground vector $\vec{A}$ and background vector $\vec{B}$, such that:

$$|\vec{AB}|^2 = |\vec{A}|^2 + |\vec{B}|^2 + 2|\vec{A}||\vec{B}|cos(2\pi K(y_a^p - y_b^p)) \quad (11)$$

where y_a and y_b are the projector row coordinates for the two paths. This change in vector phase and magnitude, in the superimposed vectors, as a function of K is the prime means by which to detect multi-path in the scanned image.
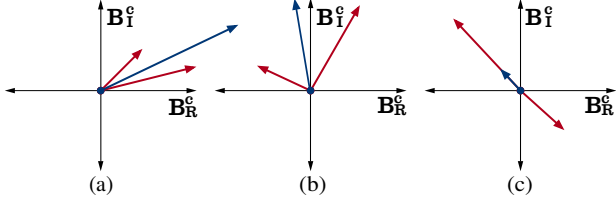
Figure 1: Illustration of the change in direction and magnitude, on the $(B^c_\mathcal{R}, B^c_\mathcal{I})$-axis, in the (blue) observed complex vector $B^c_\mathcal{R} + jB^c_\mathcal{I}$ created by the superposition of (red) complex vectors from multi-path fore and background objects for (a) $K = 1$, (b) $K = 8$, and (c) $K = 12$.
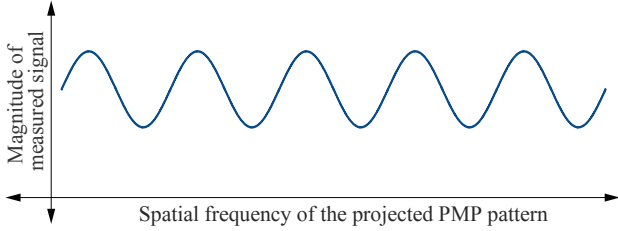

Figure 2: Illustration of the change in (vertical axis) magnitude of the observed complex vector, $B^c_\mathcal{R} + jB^c_\mathcal{I}$, as a function of the (horizontal axis) scaling factor K.

The goal is to separate the vectors $\vec{A} = |\vec{A}|e^{(iy_a)}$ and $\vec{B} = |\vec{B}|e^{(iy_b)}$ from the observation $\vec{AB}$. The search space of this task is determined by the four independent variables $|\vec{A}|, |\vec{B}|, y^p_a$, and y^p_a with computational complexity $O(n^4)$.

In order to reduce the complexity, we present a two-step procedure where the first step finds the parameters $|\vec{A}|, |\vec{B}|$, and $dy = y^p_a - y^p_b$ that minimize the mean-squared error given by:

$$(|\vec{A^*}|, |\vec{B^*}|, dy^*) = \underset{|\vec{A}|,|\vec{B}|,dy}{\mathrm{argmin}} \sum_K \{|\vec{AB}| - |\vec{A} + \vec{B}|\}^2. \quad (12)$$

And once this is solved, the second step is to obtain the absolute phases, y^p_a and y^p_b, by minimizing the mean squared error given by:

$$(y^{p^*}_a, y^{p^*}_b) = \underset{y^p_a, y^p_b}{\mathrm{argmin}} \sum_K \{\vec{AB} - (\vec{A} + \vec{B})\}^2 \quad (13)$$

such that $|\vec{A}| = |\vec{A^*}|, |\vec{B}| = |\vec{B^*}|$, and $y^p_a - y^p_b = dy^*$.

Specifically to reduce the search space in the first step from the three independent variables $|\vec{A}|, |\vec{B}|$, and dy to the two $|\vec{A}|$ and dy, we define a zero-frequency scan where $K = 0$ to obtain $\vec{AB}_0$ such that:

$$|\vec{AB}_0|^2 = |\vec{A}|^2 + |\vec{B}|^2 + 2|\vec{A}||\vec{B}|. \quad (14)$$

From this, we get the constraint:

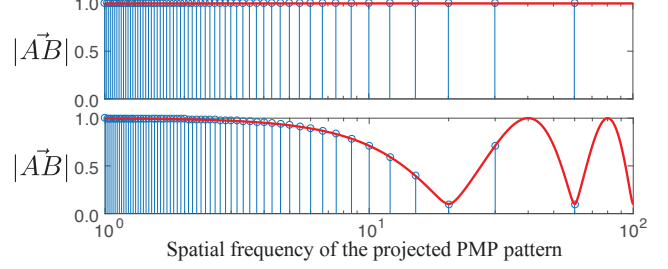$$|\vec{A}| + |\vec{B}| = |\vec{AB}_0| \quad (15)$$

Figure 3: Stem plots showing the normalized intensity, in the range [0,1], versus frequency, K, for noiseless $|\vec{AB}|$ of (top) the background/foreground pixel, and (bottom) the edge pixel where the red line illustrates $|\vec{AB}|$ over continuous frequency.

so that we can perform an exhaustive search over $|\vec{A}|, |\vec{B}|$, and the phase difference $y^p_a - y^p_b$ along the line $|\vec{B}| = |\vec{AB}_0| - |\vec{A}|$ to find the values that minimize the mean-squared error in Eq. (11) over all scanned values of K. In this way, we reduce the computing complexity in the first step from $O(n^3)$ to $O(n^2)$ while keep the second step as $O(n)$ since $y^p_a - y^p_b$ has already been determined from the first step. So the overall algorithm complexity is reduced from $O(n^4)$ to $O(n^2)$ based on the proposed two-step optimization.

As an illustration of the proposed algorithm, Fig. 3 shows plots of simulated $|\vec{AB}|$ over K for two pixels, separated in the projector by 12 pixels, with (top) the background/foreground pixel, and (bottom) a linear combination of 55% foreground and 45% background pixel. As will be the case for these stem plots in this paper, the frequency, K, ranges from 1 to 60 sinusoids, at wavelength intervals of 8 pixels, across the projector field of view and is plotted in Fig. 3 on the log scale. Also note that, the x-axis represents K, the y-axis is normalized by $|\vec{AB}_0|$ and will range from 0 to 1. Shown in red are plots of the resulting best-fit $\vec{A}$ and $\vec{B}$ vectors where $|\vec{AB}|$ is plotted over continuous K from 1 to 60 Hz.

5. Experimental Evaluations

In order to demonstrate the proposed de-coupling technique, we consider the case of scanning two layers of a half-inch, textureless, foam board where Fig. 4 (bottom) shows the variance in the magnitude in the observed phasors, $\vec{AB}$, over all K where the step edge is clearly visible as indicated by the bright vertical line. To illustrate this sinusoidal shape on $|\vec{AB}|$, Fig. 5 shows stem plots of $|\vec{AB}|$ versus K for the three pixels in Fig. 4, labeled **A**, **B**, and **AB** where **A** corresponds to the foreground surface to the right of the edge, **B** the background surface to the left of the edge, and **AB** a pixel on the edge of the surface.

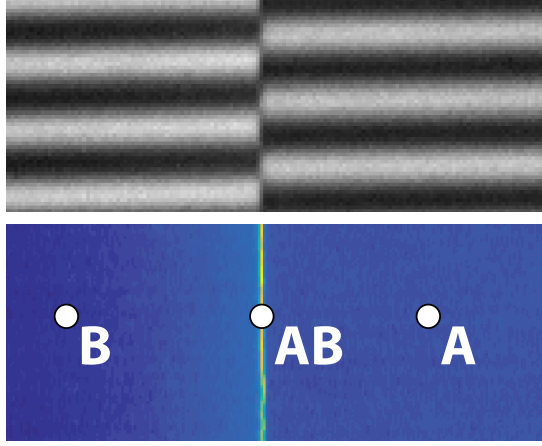Observing the stem plot in Fig. 5, one can see a consis-

Figure 4: Illustration of the variance in $|\vec{AB}|$ versus K (bottom) for a textureless surface (top) with a step edge composed of foreground pixel labeled **A**, background pixel labeled **B**, and edge pixel **AB**.
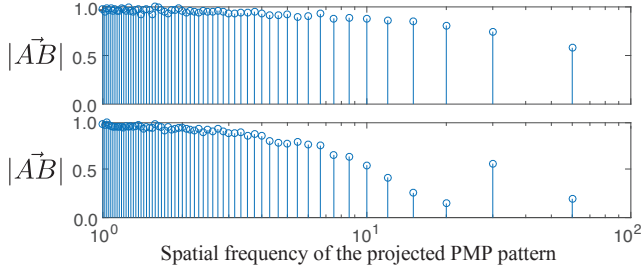

Figure 5: Stem plots showing the measured $|\vec{AB}|$ versus K for (top) the background/foreground pixel **B/A**, and (bottom) the edge pixel **AB**.
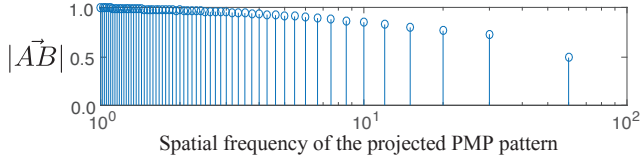

Figure 6: Stem plot of the measured $|\vec{AB}|$ versus K for a flat, textureless surface at the center of the scanners focal distance averaged over all pixels as an estimate of the systems modulation transfer function.

tent drop in magnitude at higher frequencies. This is caused by the low-pass nature of the projector and camera optics, blurring the peaks and valleys of the projected sinusoids. In order to account for the modulation transfer function of the projector/camera optics, we scan a white, textureless foam board at the center of our depth range and then average the value of $|\vec{AB}|$ over all pixels for all K to produce the stem plot in Fig. 6. This resulting vector is then used as a normalizing factor for all subsequent scans. Apply-
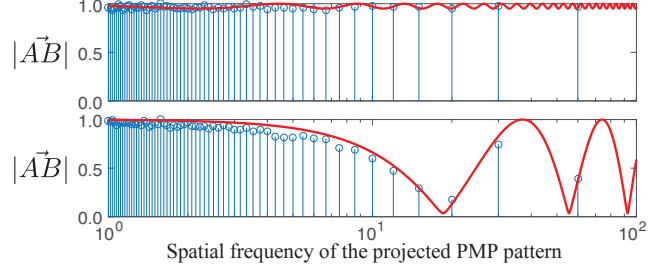

Figure 7: Plots showing the measured $|\vec{AB}|$ over K for (top) the background/foreground pixel **B/A**, (bottom) the edge pixel **AB** where the red line illustrates best-fit $|\vec{AB}|$ over continuous frequency.

ing this normalization to Fig. 5 produces the stem plots in Fig. 7 which now show the expected flat response to foreground and background pixels **A** and **B** and the distinctive sinusoidal shape for the edge pixel **AB**.

Using the proposed algorithm on the edge pixel **AB**, we obtain the normalized magnitudes of 0.5560 and 0.4440 and phases of 0.3647 and 0.3917 (projector row coordinates 188 and 175 or 13 pixels difference), respectively, where the actual pixels have phases of 0.3650 and 0.3916. If we also apply the algorithm to background pixel **B** under the assumption of multi-path, we extract magnitudes of 0.9770 and 0.0230 with phase values 0.3918 and 0.1698, resulting in the small sinusoidal curve. For the foreground pixel **A**, we extract magnitudes of 0.9680 and 0.0320 with phase values 0.3648 and 0.0998. We associate these weak, secondary multi-path signals with noise in the sensor and, ignoring these terms, focus on the edge pixel, **AB**, noting how close our estimated values are to the true phases derived through the traditional structured light phase processing.

Applying the exhaustive search over $|\vec{A}|$, $|\vec{B}|$, and the phase difference $y_a^p - y_b^p$ along the line $|\vec{A}|^2 + |\vec{B}|^2 = |\vec{AB_0}|^2$ for a small region of interest about the step edge, Fig. 8a shows the value of the magnitude of the (left) primary, the larger of $|\vec{A}|$ or $|\vec{B}|$, and the (right) secondary or smaller term. The corresponding primary and secondary phase terms are illustrated in Fig. 8b. Relying on the primary term for reconstructing depth, Fig. 9 illustrates the improved edge rendition sans bimodal multi-path.

For a demonstration of multi-path separation in a structured light system, Fig. 10 shows an experimental setup (left) where we scanned a white plaster owl figurine through a polyester lining mesh fabric and four challenging scan scenarios (right). Shown in Figs. 11 and 12 are the resulting phase reconstructions and point cloud showing the effects of using the proposed multi-path detection scheme. It should be evident that this phase unwrapping error is devastating to the 3D reconstruction which we illustrate in Fig. 12 where the reconstruction from the raw phase image is given
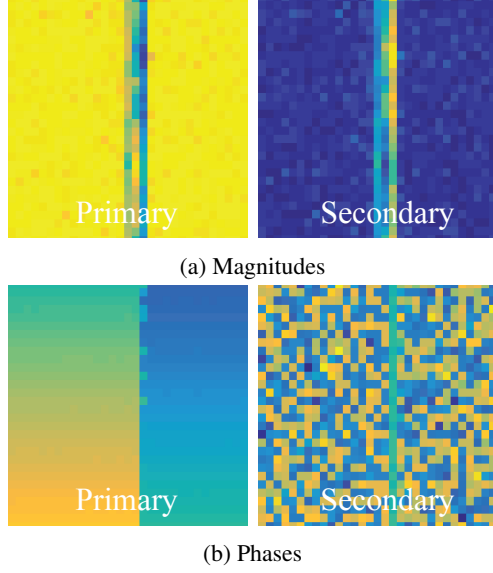
(a) Magnitudes

(b) Phases

Figure 8: Pseudo-color plot of the magnitudes (a) and the phases (b) of the primary (stronger) and secondary (weaker) bimodal path component along the step edge in Fig. 4.
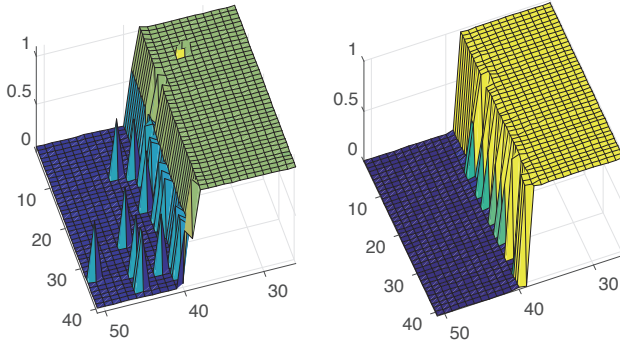

Figure 9: Surface plots of the phase images (left) before and (right) after applying the multi-path separation procedure where (left) shows the unprocessed phase image while (right) shows the phase image of the dominant component from Fig. 8b.

in Fig. 12 (top) while Fig. 12 (bottom) shows the reconstruction using the multi-path phase image. Again, the multi-path reconstructions are correct, the traditional phase unwrapping is incorrect in these figures. From visual inspection, the proposed technique is a clear improvement over the scan produced without the process.

As a demonstration of multi-path reconstruction on transparent objects, Fig. 13 shows an owl sitting inside of a semi-transparent plastic box. The 3D point cloud reconstruction (top) showing the raw phase reconstruction, (middle) and (bottom) showing the multi-path reconstruction without and with our proposed calibration approach. In
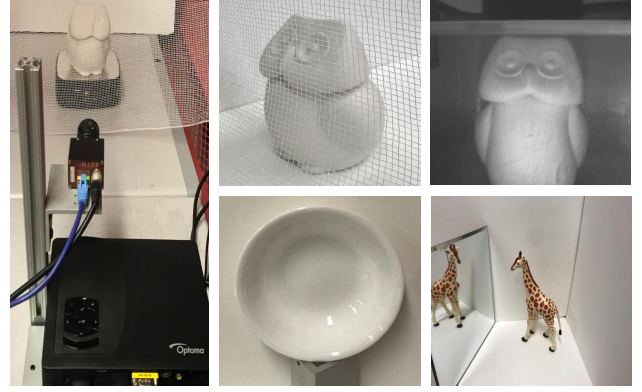

Figure 10: Experimental setup (left) scanning a white owl figurine through a polyester lining mesh fabric and four challenging scan scenarios (right).
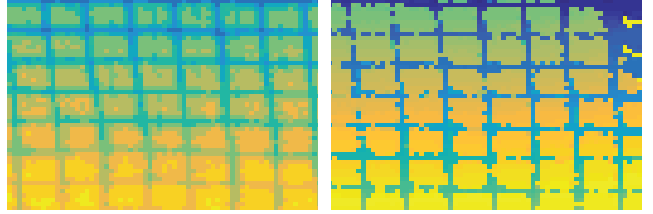

Figure 11: Pseudo-color plot of the phases of the (left) raw phase image of the owl figurine beak and eyes through the mesh and the (right) primary path component.
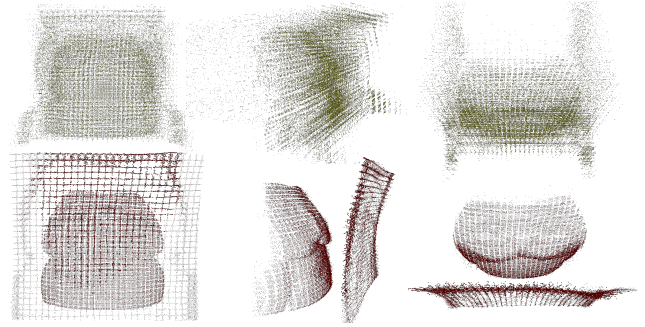

Figure 12: Point cloud reconstructions of the owl figurine using (top) raw phase, (bottom) multi-path processed phase. Left to right: front view, side view and top view.

this illustration, we note that without the multi-path algorithm, traditional SLI reconstruction will result in multiple layers of the transparent surface that appear at incorrect positions in front and over the owl. With our proposed algorithm, the ghost-layers disappear and result in an accurate reconstruction of the transparent surface in front of the owl. Figs. 13 (middle) and (bottom) also demonstrate the effectiveness of our proposed novel calibration approach with an incorporated curve estimation as part of routine scan-
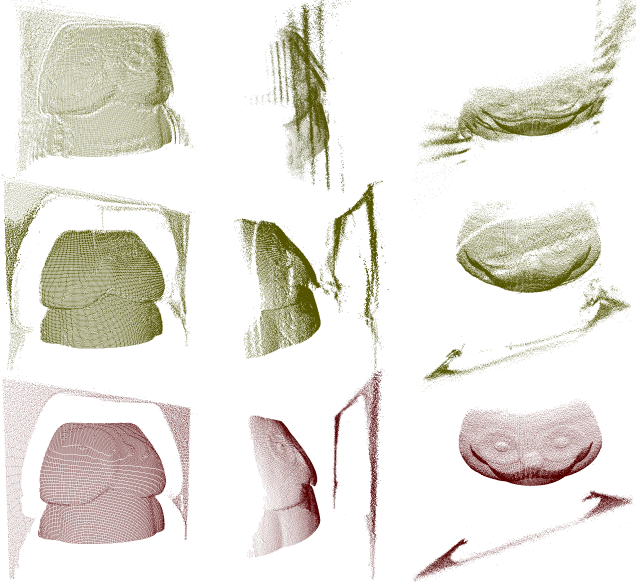
Figure 13: Point cloud reconstructions of the owl figurine in a semi-transparent box using raw phase (top), multi-path processed phase without (middle) and with (bottom) our proposed novel calibration approach. Left to right: front view, side view and top view.
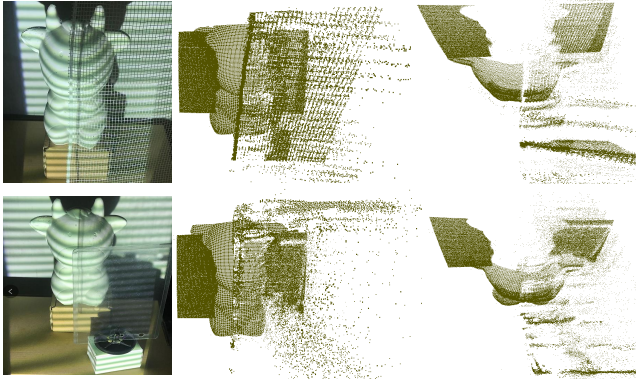

Figure 14: Scenarios and point cloud reconstruction rendered in top view of the multi-path cases in terms of sharp step edge and transparent surface with the state-of-the-art micro phase shifting [14]. Note the left side of each scan does not have the mesh or the transparent layer. So the left side represents the ground truth while the right side shows MicroPS failing.

ner calibration to model the projector defocus effect by estimating the shape of $|\vec{AB}(K)|$ for a flat, textureless surface from an in-the-wild scan by taking the median value of $|\vec{AB}(K)|/|\vec{AB}(0)|$ over all camera pixels.

In order to compare to the current state-of-the-art, we apply micro phase shifting [14] on our owl scan data. We have used Micro PS with the 7 optimized patterns from Gupta

and Nayar [14] and it produces obvious artifacts as shown in Fig. 14. We believe the reason is that they cannot handle the bimodel multi-path like the step edge in the separated direct image since the authors still use traditional phase shifting method to solve phase/depth in the direct component. Furthermore, the authors relate that they do not consider the camera defocus effect, resulting in incorrect depths especially at depth edges.

As a third demonstration of the multi-path technique, Fig. 15 shows the phase reconstructions comparing again the traditional phase unwrapping procedure versus our proposed multi-path procedure when the target image is the inside of a white, porcelain bowl. In this sample, specular reflections off the surface of the bowl create multi-paths, most evident at the top and the bottom of the bowl where the reflections stay within the epipolar geometry of the camera/projector lens alignment. While the new multi-path procedure is not completely immune to issues caused by specularities on the target surface, it is greatly improved over the board artifacts introduced through phase unwrapping, as indicated in Fig. 15.
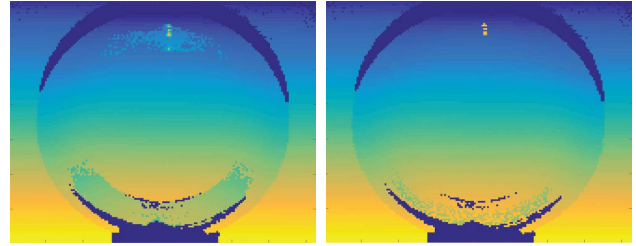

Figure 15: Bowl experiment. The (left) raw and (right) multi-path phase images of a porcelain bowl.

As a final demonstration, we used a mirror to reflect light from off to on target, a plastic giraffe figurine, as illustrated in the photograph in Fig. 10 (bottom right) with phase results in Fig. 16. Looking at the raw phase image versus multi-path reconstructed, there are substantial artifacts in the raw phase as indicated by posterization, most visible in the region of the giraffe's neck/chest facing the mirror and especially in the top-right corner of the background screen and on the right side wall. These posterization effects are also visible in the reflected image of the mirror.

Now of course the more frequencies that we use, the better the scan will be as we minimize the impact of sensor noise on our final path estimates. Under ideal or noise-less conditions, we can separate the two paths with no more than 9 frequencies. Note these results are for noiseless simulations. For an illustration of the number of frequencies used on real data, Figs. 17 and 18 show the evolution of our reconstructed scenario using an increasing number of frequencies. From visual inspection, the true data show the same result that 10 unique scan frequencies are able to sep-
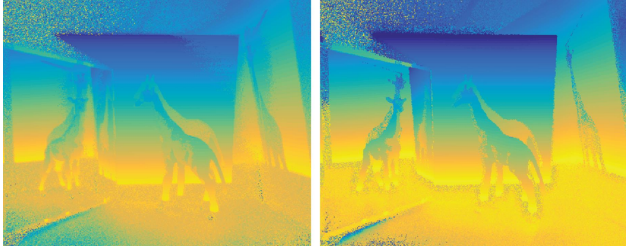
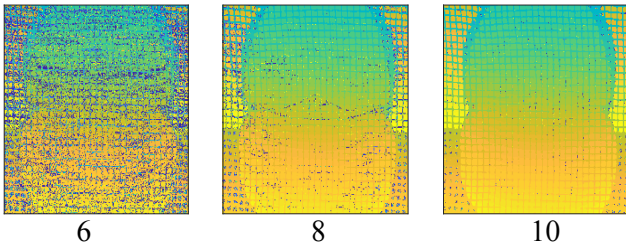Figure 16: Mirror experiment. The (left) raw and (right) multi-path phase images of a plastic giraffe.

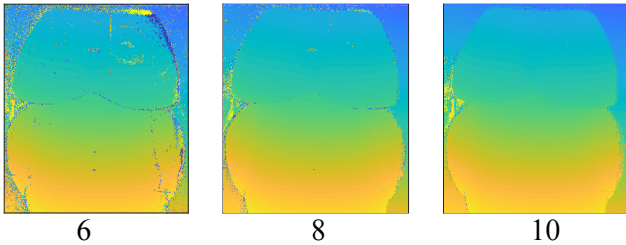Figure 17: The evolution of our owl and mesh scan using 6, 8, and 10 unique frequencies.

Figure 18: The evolution of our owl and transparent box scan using 6, 8, and 10 unique frequencies.

arate the majority of multi-path instances in the scan.

6. Conclusions and Future Work

In this paper, we introduced a novel procedure for extracting the bimodal multi-path phase terms for a PMP structured light scan based on modeling the change in magnitude of the observed phasors caused by modulating the spatial frequency of the projected PMP patterns. Furthermore, we introduced the first PMP scans to employ zero-frequency PMP patterns as a way to measure the magnitude in the observed phasors sans multi-path. As demonstrated here, the proposed technique is especially geared toward step edges and scanning through semi-transparent surfaces; however, the proposed derivation can be expanded to include more than two paths, although additional investigation is necessary to gauge how practical doing so is.

For the polyester mesh example, that is a contrived ex-

periment meant to maximize the likelihood of step edge multi-paths; however, it's not inconceivable in an industrial environment to have a screen of some kind between a machine vision system and a conveyor belt or robotic assembly line. For the semi-transparent experiment, we are mimicking a situation found in deep water SLI scanners where an off-the-shelf scanner is placed inside a pressured box with clear plastic window. Being inside the box, projected light internally reflects off the window, back at the camera. This is actually a benefit because we can measure the flexing of the plastic window when under heavy pressure at deep depths of the water.

For the bowl and mirror examples, these represent the kind of multi-path we see for inter-oral dental scanners where even though we may spray a white power onto target teeth to cut down on specular reflections, there is always some part of a tooth that is not adequately covered with powder. Or there is a gum line or the subject's tongue that has a specular/wet surface, and we absolutely have to scan down to the gum line. And because this is a dynamic environment, we cannot actively adapt the projector based on the scene since the patterns are pre-generated and stored inside the scanner electronics [19]. Our bowl and mirror test examples are just a way to simulate strong specularities, such as the inter-oral dental scanning case.

Although not considered here, the problem of multi-texture is very similar to the multi-path problem. Here, a single pixel sees a continuous smooth surface, but the surface texture has a discontinuity or step edge mid-way across the pixel's field of view. We can define the brighter side of the edge as the foreground surface and the darker side of the edge as the background surface. This means that the phase values inside the foreground surface will have a greater weight, per unit area, than the background surface. This has the effect of pushing the combined vector closer to the foreground phase than the background. While the change may not be as severe as the multi-path problem, the solution is the same by taking advantage of the presented interesting cue of measuring the constructive and destructive interference between the two light paths.

Acknowledgements

References

[1] Ayush Bhandari, Achuta Kadambi, Refael Whyte, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar. Multifrequency time of flight in the context of transient renderings. In *ACM SIGGRAPH 2013 Posters*, page 46, 2013. 1

[2] Kim L Boyer and Avinash C Kak. Color-encoded structured light for rapid active ranging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1):14–28, 1987. 1

[3] Tongbo Chen, Hans-Peter Seidel, and Hendrik PA Lensch. Modulated phase-shifting for 3d scanning. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008. 1

[4] Vincent Couture, Nicolas Martin, and Sebastien Roy. Unstructured light scanning to overcome interreflections. In *IEEE International Conference on Computer Vision*, pages 1895–1902, 2011. 2

[5] Raymond C. Daley and Laurence G. Hassebrook. Channel capacity model of binary encoded structured light-stripe illumination. *Applied Optics*, 37(17):3689–3696, 1998. 3

[6] Eric Dedrick. Improving sli performance in optically challenging environments. PhD thesis, University of Kentucky, 2011. 2

[7] Adrian A Dorrington, John Peter Godbaz, Michael J Cree, Andrew D Payne, and Lee V Streeter. Separating true range measurements from multi-path and scattering interference in commercial range cameras. In *Three-Dimensional Imaging, Interaction, and Measurement*. International Society for Optics and Photonics, 2011. 1

[8] Barak Freedman, Alexander Shpunt, Meir Machline, and Yoel Arieli. Depth mapping using projected patterns, 2012. US Patent 8,150,142. 1

[9] Daniel Freedman, Yoni Smolin, Eyal Krupka, Ido Leichter, and Mirko Schmidt. Sra: Fast removal of general multipath for tof sensors. In *European Conference on Computer Vision*, pages 234–249, 2014. 1

[10] Jason Geng. Rainbow three-dimensional camera: new concept of high-speed three-dimensional vision systems. *Optical Engineering*, 35(2):376–383, 1996. 1

[11] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Advances in Optics and Photonics*, 3(2):128–160, 2011. 1

[12] John P Godbaz, Michael J Cree, and Adrian A Dorrington. Closed-form inverses for the mixed pixel/multipath interference problem in amcw lidar. In *Computational Imaging X*. International Society for Optics and Photonics, 2012. 1

[13] Mohit Gupta, Amit Agrawal, Ashok Veeraraghavan, and Srinivasa G Narasimhan. Structured light 3d scanning in the presence of global illumination. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 713–720, 2011. 1

[14] Mohit Gupta and Shree K Nayar. Micro phase shifting. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 813–820, 2012. 1, 2, 7

[15] Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics*, 34(5):156, 2015. 1

[16] Mohit Gupta, Qi Yin, and Shree K Nayar. Structured light in sunlight. In *IEEE International Conference on Computer Vision*, pages 545–552, 2013. 1

[17] Achuta Kadambi, Refael Whyte, Ayush Bhandari, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics*, 32(6):167, 2013. 1

[18] Jielin Li, Hongjun Su, and Xianyu Su. Two-frequency grating used in phase-measuring profilometry. *Applied Optics*, 36(1):277–280, 1997. 3

[19] DLP LightCrafter. 4500 evaluation module users guide. *Texas Instruments Inc*, pages 560–590, 2015. 8

[20] Kai Liu, Yongchang Wang, Daniel L Lau, Qi Hao, and Laurence G Hassebrook. Dual-frequency pattern scheme for high-speed 3-d shape measurement. *Optics Express*, 18(5):5229–5244, 2010. 1

[21] Kai Liu, Yongchang Wang, Daniel L. Lau, Qi Hao, and Laurence G. Hassebrook. Gamma model and its analysis for phase measuring profilometry. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 27(3):553–562, 2010. 3

[22] Raymond A Morano, Cengizhan Ozturk, Robert Conn, Stephen Dubin, Stanley Zietz, and J Nissano. Structured light using pseudorandom codes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):322–327, 1998. 1

[23] Nikhil Naik, Achuta Kadambi, Christoph Rhemann, Shahram Izadi, Ramesh Raskar, and Sing Bing Kang. A light transport model for mitigating multipath interference in time-of-flight sensors. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 73–81, 2015. 1

[24] Shree K Nayar, Gurunandan Krishnan, Michael D Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics*, 25(3):935–944, 2006. 1, 2

[25] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos. Temporal frequency probing for 5d transient analysis of global light transport. *ACM Transactions on Graphics*, 33(4):87, 2014. 1

[26] Matthew O'Toole, John Mather, and Kiriakos N Kutulakos. 3d shape and indirect appearance by structured light transport. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3246–3253, 2014. 1

[27] Guy Rosman, Daniela Rus, and John W Fisher. Information-driven adaptive structured-light scanners. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 874–883, 2016. 1

[28] Venugopal Srinivasan, Hsinchu Liu, and Maurice Halioua. Automated phase-measuring profilometry of 3-d diffuse objects. *Applied Optics*, 23(18):3105–3108, 1984. 1

[29] Yongchang Wang, Kai Liu, Qi Hao, Xianwang Wang, Daniel L Lau, and Laurence G Hassebrook. Robust active stereo vision using kullback-leibler divergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(3):548–563, 2012. 1