An Analysis of Using High-Frequency Sinusoidal Illumination to Measure the 3D Shape of Translucent Objects

Michael Holroyd
University of Virginia

Jason Lawrence
University of Virginia

Abstract

Using optical triangulation methods to measure the shape of translucent objects is difficult because subsurface scattering contaminates measurements of the "direct" reflection at the surface. A number of recent papers have shown that high-frequency sinusoidal illumination patterns allow isolating this direct component [16], which in turn enables accurate estimation of the shape of translucent objects [4]. Despite these encouraging results, there is currently no rigorous mathematical analysis of the expected error in the measured surface as it relates to the parameters of these systems: the frequency of the projected sinusoid, the geometric configuration of the source and camera, and the optical properties of the target object. We present such an analysis, which confirms earlier empirical results and provides a much needed tool for designing 3D scanners for translucent objects.

1. Introduction

Measuring the 3D shape of real-world scenes is a primary goal of many computer vision systems. This problem is particularly challenging for scenes that contain translucent objects. Unlike opaque objects, which reflect incident light back into the environment at their surface, translucent objects exhibit subsurface scattering wherein a significant portion of incident illumination scatters inside the object before being re-emitted at a different surface location. Although it is sometimes possible to first coat these objects with a diffuse powder and then apply standard optical triangulation techniques, there are many applications such as scanning archaeological artifacts for which this is not practical.

It has recently been shown that by imaging translucent surfaces under a series of shifted high-frequency illumination patterns, scattering at the surface (direct reflection) can be reliably isolated from multiple scattering below the surface (indirect reflection), a process known as *optical descattering* [16]. In the particular case where the illumination is a sinusoid, it is possible to incorporate traditional phase profilometry techniques [21] to estimate the object's surface geometry based on the direct component alone [3, 4]. This has proven to be a reliable way of estimating the 3D shape of translucent objects.

Despite the promise of this approach, however, there is currently no theory that connects the nature and magnitude of the error in the estimated surface to the main parameters in these systems: the spatial frequency of the projected sinusoid, the geometric setup of the source and camera, and the optical properties of the target object. We derive a closed-form analytic expression that relates these three components. This allows one to compute the expected error for a particular experimental setup or estimate a lower-bound on the frequency of the projected sinusoid necessary to achieve a desired level of accuracy.

2. Related work

Many techniques have been developed for measuring the shape of *opaque* objects, whose appearance is characterized by the bidirectional reflectance distribution functions (BRDF) [17]. In reality, all nonmetals exhibit some degree of subsurface scattering, and using techniques that assume a perfectly opaque response in these cases will produce errors in the recovered geometry. For example, laser stripes and other active lighting patterns are effectively blurred by subsurface scattering and therefore the maximum response observed by a camera can be below the true surface [6, 3] as illustrated in figure 1. The following sections review methods designed specifically for measuring the 3D shape of translucent scenes.

2.1. Inverse rendering

One approach for scanning translucent objects is to perform inverse rendering and explicitly account for subsurface scattering while optimizing for the geometry and material parameters that best match a set of input images. However, this is an ill-posed problem in general, and a system capable of capturing high-resolution geometry of complex translucent objects has not yet been demonstrated [18, 20]. A related method is that of Gu et al. [7], which uses com-

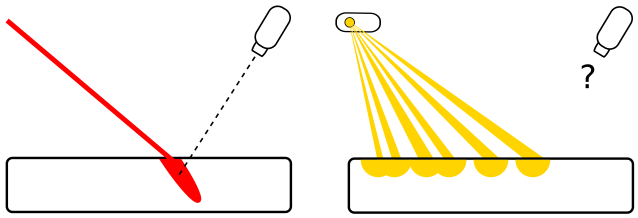

Figure 1. Translucent objects present difficulties for traditional optical triangulation methods. (left) The maximum response of a laser beam striking a translucent surface as seen by a camera is often below the actual surface. (right) Active lighting patterns are also blurred by the material, making it difficult to establish a correct correspondence between rays in a projector and camera.

pressive sensing to recover the volumetric density of semi-transparent inhomogeneous participating media, but it is not clear how this technique could be extended to optically dense translucent objects with well-defined surfaces.

2.2. Diffuse and specular separation

The most common approach for extending optical triangulation methods to handle translucent objects is to first isolate the direct reflection at the object surface before recovering geometry. One set of techniques is based on the idea that the specular component of the surface response is due to light that reflects directly off of the object surface while the diffuse component is produced by subsurface scattering. Chen et al. [2] apply a simple threshold to a histogram of the measured intensities in order to identify specular highlights, which are subsequently used to estimate a high-resolution normal field. A similar set of methods rely on color heuristics to estimate normal fields of human faces [23] and repair errors in a structured light scan [1]. These methods generally suffer in the presence of glossy highlights (as opposed to sharp specular highlights), objects with a small specular response, or objects whose diffuse color is similar to that of the illuminant. Additionally, they require a very dense sampling of light directions in order to observe the specular highlight everywhere over the object surface.

2.3. Polarization difference imaging

If the light striking a surface is polarized, the portion that is directly reflected at the surface typically retains this polarization, whereas subsurface scattering acts to depolarize the light [22]. Under this assumption, the direct component of the scene can be isolated by computing the difference between images taken behind parallel and perpendicular polarization filters. This approach has been used in combination with sinusoidal illumination for recovering the geometry of translucent objects [3]. Ma et al. [14] extend this idea to use circularly polarized spherical gradient illumination in order to recover dense normal fields of translucent objects. However, a limitation of these approaches is that, for certain materials, light that has been scattered multiple times may still retain a significant amount of polarization and other materials are known to depolarize light that is reflected at the surface [3].

2.4. Optical descattering

Another approach to the separation of direct and indirect illumination uses high-frequency spatially modulated light to remove the subsurface scattering in a scene [16], a process known as optical descattering. Light that scatters multiple times within the material is significantly diffused whereas light that scatters only once retains the projected pattern and can be isolated. Specifically, a modulated light source (e.g., a projector) illuminates a scene with a high-frequency pattern and a series of images are captured while this pattern is translated within the focal plane of the projector.

Nayar et al. [16] note that this descattering technique can be integrated with a number of active scanning methods that also use high-frequency lighting patterns. Chen et al. [4] present a method that uses sinusoidal patterns both for the purpose of isolating the direct surface reflection and for recovering depth using a standard phase unwrapping technique, a special case of a broader class of phase profilometry methods [21]. Holroyd et al. [10] demonstrate a similar system that uses high-frequency sinusoidal patterns in a multi-view stereo algorithm. Gupta et al. [8] analyze the effect of defocus on this descattering process and show how to both correct for it and leverage it for the purpose of depth recovery. More recent work uses the light transport equation to estimate each component of the indirect response through a recursive procedure [15].

We build on this prior work by providing a radiometric analysis in the specific case of sinusoidal illumination. This analysis demonstrates the nature and degree of error in the estimated surface for systems based on sinusoidal patterns. The next section reviews the key ideas behind these systems in greater detail.

3. Background: Depth estimation using active sinusoidal illumination

Active stereo triangulation systems that use sinusoidal lighting capture a sequence of images of a scene while it is illuminated by a high-frequency sinusoidal pattern. Between each exposure, the sinusoid is translated within the lightsource's focal plane by a known amount. A key observation about this approach is that summing together multiple sinusoids with the same frequency produces a sinusoid that also has that frequency:

$$\sum_i A_i \cos(ft + \theta_i) + G_i = A \cos(ft + \theta) + G. \quad (1)$$

Regardless of the properties of the scene, due to the linearity of light transport and the closure of sinusoids under addition, the intensities $\{I_i | i \in [1, n]\}$ measured at each camera pixel will trace a time-varying sinusoid, $I_i = A\cos(t_i + \phi_0) + G$, where t_i can be calculated from the spatial frequency of the projected sinusoid and the magnitude of the translation in the i^{th} image. This fact makes sinusoidal lighting especially desirable for scanning translucent objects: *the signal measured at the camera has a simple relationship to the signal produced by the source without having to explicitly account for global illumination in the scene*. The per-pixel amplitude, phase, and gain can be computed by solving:

$$\begin{bmatrix} \cos(t_1) & -\sin(t_1) & 1 \\ & \vdots & \\ \cos(t_n) & -\sin(t_n) & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} I_1 \\ \vdots \\ I_n \end{bmatrix} \tag{2}$$

$$A = \sqrt{a_1^2 + a_2^2} \qquad \phi_0 = \arctan(a_2/a_1) \qquad G = a_3$$

The initial phase offset at each pixel ϕ_0 can be used to identify corresponding rays between the source and camera and estimate geometry in a process known as phase profilometry [21]. However, many pixels along each epipolar line will observe the same phase value. Resolving this "phase ambiguity" is a key problem and a number of solutions have been developed (a good review is provided by Salvi et. al [19]). In the following, we will not address this aspect of phase profilometry and assume the ambiguity can be resolved using one of the available methods. Instead, our goal is to analyze the effect of subsurface scattering on the recovered phase value at each pixel and how this in turn biases the estimated depth.

4. Analysis

As illustrated in Figure 2, we will assume the scene is composed of a homogeneous flat translucent surface, which ignores effects due to local curvature or thin shells. The translucent medium is characterized by the absorption coefficient σ_a, scattering coefficient σ_s, and extinction coefficient $\sigma_t = \sigma_a + \sigma_s$, as well as the medium's index of refraction η and phase function $p(\cdot, \cdot)$. A camera images the scene while the only source of illumination is a light source modulated to produce a spatially-varying sinusoid (a "projector"). Translating this projected pattern produces time-varying sinusoids along individual rays as they leave the projector and strike the surface. Both the projector and camera are assumed to be orthographic.

Paths of light can reach the camera along the view direction v either due to surface reflection (green), single scattering within the material (red), or multiple scattering (blue). We ignore surface reflection by assuming the camera is not

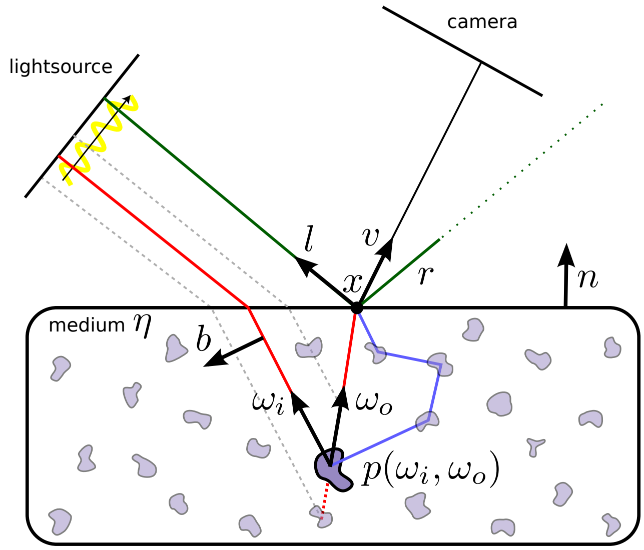

Figure 2. Setup and notation used in this paper. The vector b is perpendicular to ω_i and in the direction of the sinusoid projected into the scene. The function $p(\omega_i, \omega_o)$ is the material's phase-function.

positioned along the mirror direction r. At the surface, the Fresnel equations predict the amount of light that is transmitted into the medium and Snell's law predicts the direction it travels due to refraction, $-\omega_i$. Light that enters the object is eventually either absorbed or scattered until it exits the medium. Scattering occurs according to the material's phase function $p(\omega_i, \omega_o)$, which characterizes the angular distribution of light scattered when striking a particle in the medium.

Some portion of the incident light scatters exactly once in the direction ω_o and exits the medium at point x traveling in the view direction v towards the camera. Another portion will scatter multiple times before exiting at x in this same direction. The key assumption in optical descattering [16, 3, 4] is that the contribution made by multiple scattering to the exiting light traveling along v is unaffected (constant) under translations of the projected sinusoid. Therefore, it only affects the DC gain of the time-varying sinusoid measured along the camera ray and can be ignored (we describe a simulation we performed to validate this important assumption in Section 6.1). Building on this prior work, we will assume that only single scattering contributes to the amplitude and phase of the time-varying sinusoid measured along the camera ray.

Let L_o represent the response measured by the camera along this ray. An analytic expression for L_o can be obtained by integrating the contribution due to single scattering along the camera ray as it travels through the medium. Let F_η represent the percentage of light lost at the surface due to Fresnel effects while entering and exiting the medium, $f_{\sin}$ represent the spatial frequency of the sinusoid, and ϕ_0 represent the initial phase offset of the time-varying sinusoid at point x. We integrate along s, the total distance light travels through the medium – from the point of entrance, to the scattering event, and to the point of exit. Then, L_o can then be written as

$$F_\eta p(\omega_i, \omega_o) \int_{s=0}^{\infty} e^{-\sigma_t s} \cos(f_{\sin} \frac{(n \cdot l)(b \cdot \omega_o)(n \cdot \omega_i)}{(n \cdot \omega_i) + (n \cdot \omega_o)} s + \phi_0)$$

The rate at which the projected sinusoid oscillates *along the direction* ω_o depends both on the spatial frequency of the sinusoid $f_{\sin}$ as well as the sinusoid's direction of variation within the medium b, which is always perpendicular to ω_i.

We use the relationship

$$\int_{s=0}^{\infty} e^{-\sigma s} \cos(As + \theta) = \frac{\cos(\theta + \arctan(A/\sigma))}{\sqrt{A^2 + \sigma^2}},$$

to derive a closed-form expression for L_o:

$$\boxed{L_o = \frac{F_\eta p(\omega_i, \omega_o)}{\sqrt{A^2 + \sigma_t^2}} \cos(\phi_0 + \arctan(A/\sigma_t))} \tag{3}$$

where

$$\boxed{A = f_{\sin} \frac{(n \cdot l)(b \cdot \omega_o)(n \cdot \omega_i)}{(n \cdot \omega_i) + (n \cdot \omega_o)}}.$$

Note that the "phase error" $\Delta\phi = \arctan(A/\sigma_t)$. If $\Delta\phi = 0$ then the same phase produced along the projector ray l will be precisely measured along the camera ray v that intersects it at point x. When this quantity is not zero, this triangulation will intersect at some point away from x, introducing an error in the estimated depth. Naturally, the phase error depends on the geometric setup of the projector and camera (through A), the extinction coefficient σ_t, and indirectly on the index of refraction η (which determines ω_i and ω_o).

5. Discussion

The phase error in equation 3 has a number of important implications for scanning translucent objects. The phase error will impact recovered geometry differently based on the specific experimental setup and reconstruction algorithm that is used. Perhaps the simplest and most common setup to consider uses one calibrated projector and one calibrated camera, often referred to as "structured lighting". In this case, the geometric error depends on the phase offset scaled by the period of the sinusoid along the camera ray v. Specifically, the geometric error ϵ is equal to

$$\epsilon = \frac{\Delta\phi}{2\pi} \frac{(1 - v \cdot l)}{f_{\sin}}. \tag{4}$$

Figure 3. Legend for the different materials we consider in section 5. Colors correspond to: $\sigma_t = 0.1\text{mm}^{-1}$ (ocean water), 0.2mm^{-1} (chicken broth), 0.5mm^{-1} (potato), 1.0mm^{-1} (skin), 2.0mm^{-1} (marble). These are approximate values based on measurements published by Jensen et al. [13].

This relationship allows one to establish upper bounds on ϵ as a function of the material properties and system parameters. For example, we can answer the question: how is ϵ affected by the spatial frequency of the sinusoid? Figure 4 visualizes this relationship by plotting ϵ, measured in millimeters, for five different materials over a range of sinusoid frequencies (the graph actually shows the period). Since the maximum error is bounded by one period of the sinusoid, and assuming the phase ambiguity can be resolved perfectly, ϵ tends to 0 as the frequency of the sinusoid approaches infinity. Of course, as the frequency increases the ability of the camera to resolve it also decreases (discussed in section 5.2). Additionally, note that for optically dense materials such as marble, lower-frequency sinusoids have less of an impact on ϵ, because light rays are not able to penetrate as far into the medium.

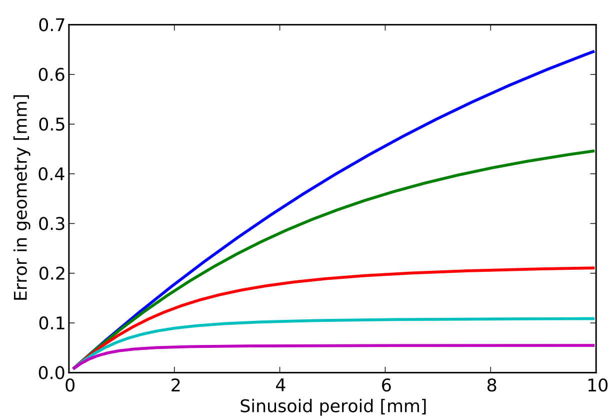

Figure 4. Geometric error ϵ for a structured lighting setup. Light and view are at $\pm 45^\circ$ across the surface normal. The colored lines correspond to the materials in figure 3. Ocean water ($\sigma_t = 0.1\text{mm}^{-1}$) is the most translucent and has the largest phase offset error (blue).

Another interesting question is: for a given projector location, where should the camera be placed to minimize ϵ? This could be used, for example, to optimize a specific setup or to drive a view planning algorithm. Figure 5 shows upper bounds on ϵ derived from equation 3 for a fixed projector located 45° off the surface normal, and view directions that vary in the plane formed by the surface normal and direction to the projector (i.e., l in figure 2). Note that as the camera and projector become coaxial, each camera ray integrates along a matching projector ray so that the phase error is zero; however, the ability to robustly triangulate two rays diminishes rapidly as the baseline becomes very small.

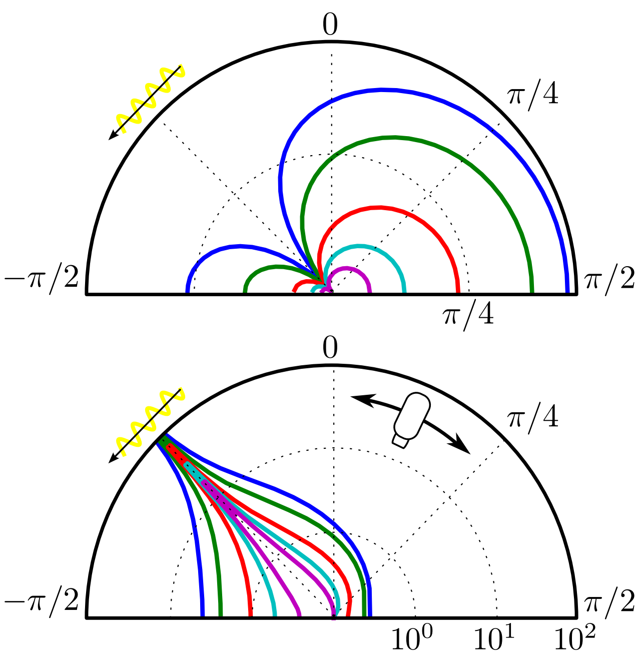

Figure 5. Phase error (top) and geometric error (bottom, in mm, log-scale) for a light at 45° elevation angle and view directions from -90° to 90°. The index of refraction $\eta = 1.55$, and the sinusoid frequency is $2\pi/5.0\,\mathrm{mm}^{-1}$, modulated spatially in the plane formed by the light and camera (worst case). The colored lines correspond to the materials in figure 3.

Finally, we visualize the manner in which ϵ depends on the material properties. Figure 6 plots ϵ for a specific camera/projector configuration and a sinusoidal frequency of $2\pi/5.0\,\mathrm{mm}^{-1}$. Other frequencies follow a similar trend, although the absolute error varies as seen in figure 4.

5.1. Relation to phase unwrapping

A common method for solving the phase ambiguity problem is to project progressively lower-frequency sinusoidal patterns into the scene, until each point receives a unique phase offset [5]. For translucent objects, figure 4 gives some insight into when this "phase unwrapping" may fail. Note that modifying the sinusoid's frequency can easily produce a large difference in the phase error in certain cases. If care is not taken in choosing these progressively lower-frequency sinusoids, each iteration could make the problem of localizing the correct period more difficult. We expect equation 3 will help strengthen these techniques when they are applied to translucent surfaces.

5.2. Amplitude loss

Although it is possible to decrease the phase error arbitrarily by increasing the frequency of the sinusoid (figure 4), at some point this process breaks down. Each camera pixel integrates over a finite region of the surface. This can be modeled as a convolution of the scene radiance with a compact kernel, which has the effect of diminishing the measured amplitude of the time-varying sinusoid. At some point it is no longer possible to reliably measure the amplitude and, consequently, the phase. This effect was recently analyzed and termed "amplitude loss" [10]. Combined with this previous result, our analysis provides a tool for choosing the optimal frequency in these types of systems: one that is as small as possible to reduce $\Delta\phi$ while not becoming impractical due to amplitude loss.

6. Validation

We validated equation 3 using a volumetric raytracer [11] to simulate a structured light setup (one projector and one camera). This simulation included both single and multiple scattering, and was used to verify our analysis over a wide range of parameters including various indices of refraction η and phase functions $p(\omega_i, \omega_o)$. Figure 6 compares our simulated data to the predictions made by equation 3 for a range of extinction coefficients.

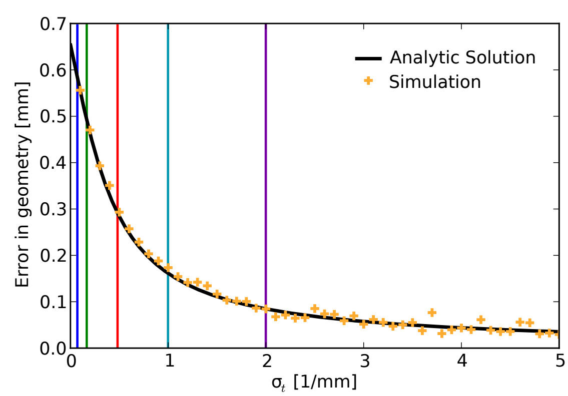

Figure 6. Geometric error ϵ as a function of the material extinction coefficient for a light -30° and view 45° off of the surface normal and a sinusoidal frequency of $2\pi/5.0\,\mathrm{mm}^{-1}$. The analytic solution is derived from equation 3 and is compared to a simulation of this setup using a physically-based volumetric ray-tracer [11].

6.1. Multiple scattering

Although light that is scattered multiple times within a translucent object contributes significantly to its overall *appearance*, for many materials it has a negligible effect on the *phase error*. We performed an additional series of simulations to validate this assumption. We first computed the phase at one camera pixel based on single scattering only, followed by a full simulation that includes multiple scattering. The projector and camera were both located $\pm 45°$ off the surface normal, and the phase function was assumed to be constant (i.e., the parameter $g = 0$ in the Henyey-Greenstein function). We ran these simulations over a range of extinction coefficients $\sigma_t \in [0.1, 2.0]$ and spatial frequencies $f_{\sin} \in [0.04, 2.0]$. In all of these cases, the difference between the phases produced for single scattering only and single+multiple scattering were on average 0.007 radians apart with a standard deviation of 0.005. This offers further support for our choice to assume that the phase is predominantly affected by single scattering.

7. Conclusion

We have presented a mathematical analysis of stereo triangulation systems that use high-frequency sinusoids to estimate the geometry of translucent surfaces. Despite recent empirical results that indicate the promise of this approach, no theoretical study of the nature and magnitude of the errors in the estimated surface had been performed. Our analysis assumes that single scattering within the object dominates the response measured by the camera. Under this assumption, we derived a closed-form expression that predicts the error in the estimated phase as a function of the geometric arrangement of a projector and camera and the properties of the material (extinction coefficient, phase function, and index of refraction). Equation 3 confirms a number of earlier empirical observations and offers a tool for making precise predictions about the operating tolerances and performance of actual scanners. Further, it indicates the limitations of phase unwrapping methods, which are commonly used to resolve the phase ambiguity problem, in the case of translucent scenes.

We expect this work will lead to further improvements in active stereo techniques for translucent surfaces. This includes the possibility of estimating material properties, or making more precise depth measurements, based on observations of how the phase measured at the camera is affected by *adjusting* the frequency of the projected sinusoid. We also plan to use this result to study more complex setups including those that incorporate polarized filters [3] or combine multiple views of the scene [9] to help identify corresponding rays. Finally, more research is warranted on how multiple scattering affects these systems. While our simulations indicate that it has only a very minor effect for a wide range of common materials (section 6 and figure 6), it is possible that this may not be the case for more exotic, strongly anisotropic, materials [12].

Acknowledgements

We wish to thank Todd Zickler for his very helpful input and the anonymous reviewers for their constructive feedback. This research was supported by NSF grant CCF-0811493.

References

[1] S. Barsky and M. Petrou. Colour photometric stereo: Simultaneous reconstruction of local gradient and colour of rough textured surfaces. 2:600–605, 2001. 2986

[2] T. Chen, M. Goesele, and H.-P. Seidel. Mesostructure from Specularity. *CVPR*, 2006. 2986

[3] T. Chen, H. Lensch, C. Fuchs, and H.-P. Seidel. Polarization and Phase-Shifting for 3D Scanning of Translucent Objects. In *CVPR*, 2007. 2985, 2986, 2987, 2990

[4] T. Chen, H.-P. Seidel, and H. Lensch. Modulated phase-shifting for 3D scanning. In *CVPR*, 2008. 2985, 2986, 2987

[5] D. C. Ghiglia and M. D. Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Wiley-Interscience, 1998. 2989

[6] G. Godin, M. Rioux, J.-A. Beraldin, M. Levoy, L. Cournoyer, and F. Blais. An Assessment of Laser Range Measurement of Marble Surfaces. *Proceedings of the 5th Conference on Optical 3-D Measurement Techniques, Vienna, Austria*, pages 49–56, 2001. 2985

[7] J. Gu, S. Nayar, E. Grinspun, P. Belhumeur, and R. Compressive structured light for recovering inhomogeneous participating media. *ECCV*, 2008. 2985

[8] M. Gupta, Y. Tian, S. Narasimhan, and L. Zhang. (De)Focusing on Global Light Transport for Active Scene Recovery. *CVPR*, 2009. 2986

[9] M. Holroyd, J. Lawrence, and T. Zickler. A coaxial optical scanner for synchronous acquisition of 3D geometry and surface reflectance. *ACM SIGGRAPH 2010*, pages 1–12, 2010. 2990

[10] M. Holroyd, J. Lawrence, and T. Zickler. A radiometric analysis of projected sinusoidal illumination for opaque surfaces. Technical report, University of Virginia, Computer Science, Charlottesville, VA, 2010. 2986, 2989

[11] W. Jakob. http://mitsuba-renderer.org, 2010. 2989

[12] W. Jakob, A. Arbree, J. Moon, and K. Bala. A radiative transfer framework for rendering materials with anisotropic structure. *ACM SIGGRAPH*, 2010. 2990

[13] H. Jensen, S. Marschner, and M. Levoy. A practical model for subsurface light transport. *ACM SIGGRAPH*, 2001. 2988

[14] W.-C. Ma, T. Hawkins, P. Peers, C.-F. Chabert, M. Weiss, and P. Debevec. Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient Illumination. In *Rendering Techniques*, pages 183–194, 2007. 2986

[15] Y. Mukaigawa, Y. Yagi, and R. Raskar. Analysis of light transport in scattering media. *CVPR*, 2010. 2986

[16] S. Nayar, G. Krishnan, M. Grossberg, and R. Raskar. Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination. *ACM Transactions on Graphics*, 25(3):935–944, July 2006. 2985, 2986, 2987

[17] F. Nicodemus, J. Richmond, J. Hsia, I. Ginsberg, and T. Limperis. Geometrical considerations and nomenclature for reflectance. *NBS monograph*, 160(October):201–231, 1977. 2985

[18] G. Patow and X. Pueyo. A survey of inverse rendering problems. In *Computer graphics forum*, volume 22, pages 663–687, 2003. 2985

[19] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, Aug. 2010. 2987

[20] S. Seitz and Y. Matsushita. A theory of inverse light transport. *ICCV*, 2005. 2985

[21] V. Srinivasan, H. Liu, and M. Halioua. Automated phase-measuring profilometry: a phase mapping approach. *Applied Optics*, 24(2):185–188, 1985. 2985, 2986, 2987

[22] J. Tyo, M. Rowe, E. Pugh Jr, and N. Engheta. Target detection in optically scattering media by polarization-difference imaging. *Applied Optics*, 35(11):1855–1870, 1996. 2986

[23] T. Weyrich, W. Matusik, H. Pfister, B. Bickel, C. Donner, C. Tu, J. McAndless, J. Lee, A. Ngan, H. Jensen, and M. Gross. Analysis of Human Faces Using a Measurement-based Skin Reflectance Model. *ACM Transactions on Graphics*, 25(3):1013–1024, 2006. 2986