Dense Depth and Albedo from a Single-Shot Structured Light

Conference Paper · October 2015

DOI: 10.1109/3DV.2015.22

CITATIONS
5

READS
297

3 authors:

Hyowon Ha
Korea Advanced Institute of Science and Technology
20 PUBLICATIONS **226** CITATIONS
SEE PROFILE

Jaesik Park
Intel
38 PUBLICATIONS **1,462** CITATIONS
SEE PROFILE

Inso Kweon
Korea Advanced Institute of Science and Technology
206 PUBLICATIONS **2,952** CITATIONS
SEE PROFILE

Dense Depth and Albedo from a Single-shot Structured Light

Hyowon Ha Jaesik Park In So Kweon
Korea Advanced Institute of Science and Technology
Daejeon, Republic of Korea
[hwha,jspark]@rcv.kaist.ac.kr, iskweon@kaist.ac.kr

Abstract

Single-shot structured light scanning has been actively investigated as it can recover accurate geometric shapes even in dynamic scenes. Since many single-shot approaches focus on improving depth accuracy, recovering the intrinsic properties of the scene, such as albedo and shading, is also valuable. In this paper, we propose a novel method that reconstructs not only the metric depth but also the intrinsic properties from a single structured light image. We extend the conventional color structured light model to embrace the Lambertian shading model. By using a color phase-shifting pattern, we parameterize the captured image with only two variables, albedo and depth. For an initial solution, a simple but powerful method to decompose sinusoids from the input image is presented. We formulate a non-linear cost function and jointly optimize albedo and depth efficiently by calculating the analytic Jacobian. We demonstrate that our algorithm works reasonably on various real-world objects that exhibit challenging surface reflectance and albedo.

1. Introduction

Structured light is a method that measures depth by analyzing projected pattern(s) on the scene. It is known to be one of the most successful approaches for acquiring accurate geometric shapes and has been applied to a broad range of areas such as digital heritage, reverse engineering, and 3D printing.

Among various types of structured light patterns [15], the continuous sinusoidal function is widely used since it gives dense and precise correspondences with sub-pixel precision. Scharstein and Szeliski [16] utilize dozens of patterns consisting of various gray-codes and phase-shifted sinusoids. This method is used to obtain the ground-truth depth maps for stereo matching evaluation. Gupta *et al*. [6] introduce multiple high-frequency patterns which are robust to inter-reflection. Other works utilize multiple color channels to acquire depth from a single-shot structured light. Huang *et al*. [9] design a color phase-shifting pattern which contains three phase-shifted sinusoidal images for RGB color channels, which has been improved by considering the response curve between camera and projector [10].

Compared to these approaches focusing on accurate depth acquisition, we present a novel structured light method which produces not only the metric depth but also the intrinsic properties of the scene (e.g. albedo and shading). To our knowledge, the proposed method is the first attempt to compute depth, albedo and shading from a single-shot structured light image, jointly. Our method also differs from geometry refinement approaches utilizing shape-from-shading [4, 7] or photometric stereo [13] because we solely utilize an image of the scene lightened by a structured light.

Our method is inspired by the color structured light model [3] where the observed image intensities are formed as the multiplication of reflectance and pattern intensities. We extend this model to embrace the well-known Lambertian shading model by substituting the reflectance with albedo and shading, which are determined by the depth and the lighting direction. By adopting the sinusoidal-function-based pattern [9, 10], we parameterize the observed intensities with two variables, *i.e*. albedo and depth. Afterwards, the optimization that minimizes the residual error between the observed image and a synthetic image rendered by our model is done for estimating both albedo and depth.

From the perspective of estimating albedo and shading (obtained from depth), our method can be related to the *intrinsic image decomposition* problem which decomposes albedo and shading from a single natural image [17, 2]. Compared to these approaches, our method fully exploits the structured light constraint to decompose depth and intrinsic images efficiently. Based on the Lambertian shading model, we develop a non-linear optimization framework that jointly estimates the disparity map (depth map when the camera-projector rig is calibrated) as well as the albedo map. Since our energy function is efficiently parameterized, an analytic form of Jacobian of the energy function can be computed in a closed form, which makes the subsequent optimization step efficient.

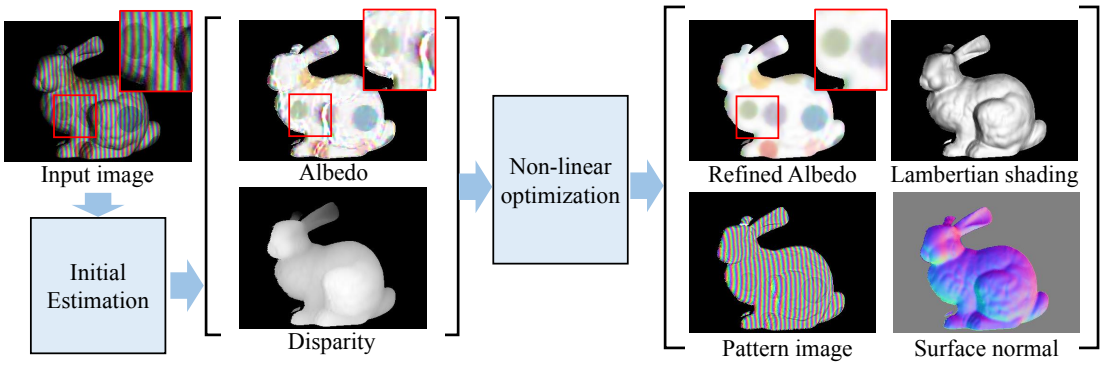


Figure 1. An overview of our algorithm. From a single color image illuminated by a structured light source, our method initially estimates surface albedo and a disparity map which gives a depth map (described in Sec. 2.2). Our optimization stage (described in Sec. 2.3) refines the estimation with our image formation model (described in Sec. 2.1). Note that all of the images in this figure are generated from the single input image.

Compared to previous approaches that capture color images and depth maps using two heterogeneous cameras, such as Kinect [1], our method is free from the burden of the alignment problem between the color image and the depth map. This relieved issue on the multi-modal sensor calibration is another surplus of the proposed scheme. In addition, our approach does not require a white pattern image for the intrinsic image decomposition. Our algorithm is applicable to high-frame-rate video capture by simply utilizing a high-frame-rate color camera.

2. Proposed Method

We describe our method to estimate depth and albedo from a single structured light. Throughout this paper, we assume that the camera is located on the right side of the projector. Also, the camera and the projector are calibrated and their image domains are rectified where all epipolar lines become horizontal. Therefore, any point correspondence between the two images can be described with a disparity d. For instance, a point $\mathbf{x} = (x, y)$ in the rectified camera image domain corresponds to $\mathbf{x}' = (x-d, y)$ in the rectified projector image domain.

The overview of the proposed method is shown in Fig. 1. From a single color image of a scene illuminated by our structured light, the proposed method initially estimates the surface albedo and the disparity map. Then our non-linear optimization refines the estimates based on the proposed image formation model.

This section is composed of four subsections. In Sec. 2.1, the proposed image formation model is introduced. In Sec. 2.2, we suggest a simple but effective method to obtain initial depth and albedo. A non-linear optimization for refining depth and albedo according to the proposed image formation is described in Sec. 2.3.

2.1. Image Formation Model

When a diffuse scene is illuminated by a color structured light pattern, we can regard the color image of the scene to be formed by the multiplication of surface reflectance and the intensities of the projected structured light pattern. For modeling this idea mathematically, we propose a linear image formation model for an image intensity I as

$$\begin{bmatrix} I_r(a_r, d) \\ I_g(a_g, d) \\ I_b(a_b, d) \end{bmatrix} = \mathbf{n}(d)^\intercal \mathbf{l}(d) \begin{bmatrix} a_r & 0 & 0 \\ 0 & a_g & 0 \\ 0 & 0 & a_b \end{bmatrix} \begin{bmatrix} S_r(d) \\ S_g(d) \\ S_b(d) \end{bmatrix}, \quad (1)$$

where d is the disparity, $\{\cdot\}^\intercal$ the vector transpose, $\mathbf{n} \in \mathbb{R}^3$ the surface normal vector, $\mathbf{l} \in \mathbb{R}^3$ the light direction, $a_{\{r,g,b\}}$ the color albedos, and $S_{\{r,g,b\}}(d)$ the sinusoidal intensities emitted from the projector for each color channel. In Eq. (1), we omit $\mathbf{x}$ from $I, \mathbf{n}, \mathbf{l}$ and d for simplicity.

In this model, it is assumed that the surface exhibits a Lambertian shading with multi-channel albedos and that the ambient light is negligible. Compared to the previous model proposed by Caspi *et al.* [3], our new model has the shading component, $\mathbf{n}^\intercal \mathbf{l}$ in Eq. (1). It is worth noting that the image can simply be parameterized with albedo and the disparity.

Among various continuous coding methods, we use phase-shifting. The phase-shifted sinusoidal pattern $S_{\{r,g,b\}}$ for each channel in our single pattern is defined as

$$S_n(\mathbf{x}, d) = (1-\alpha) + \alpha \sin\left(\frac{2\pi(x-d)}{T} - \frac{2n\pi}{3}\right), \quad (2)$$

where $n \in \{0, 1, 2\}$ stands for the red, green, and blue channels of the pattern, T is the period of the sinusoidal function in pixels, and $\alpha \in [0, 0.5]$ is the amplitude of the sinusoids bounding S_n to be $[0, 1]$. The sinusoidal pattern is invariant to the y coordinate since we are only interested in finding correspondences along the horizontal epipolar lines.

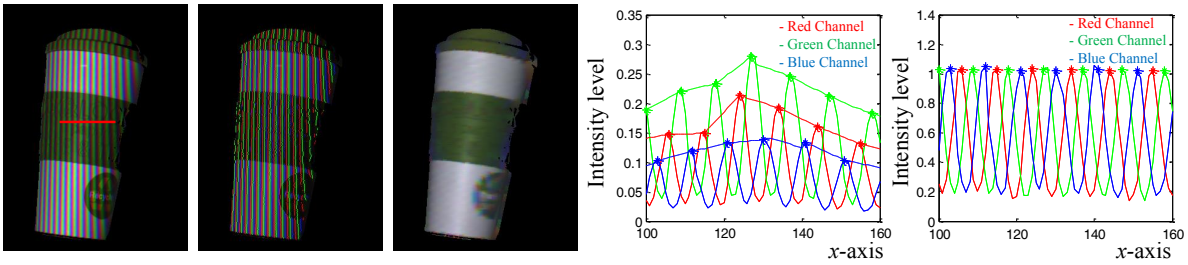


Figure 2. Initial estimation step. (a) Input image. (b) Local maximum intensities are marked in each color channel. (c) Estimated pattern-free image. We interpolate peak intensities of the input image to get a pattern-free image which barely exhibits any sinusoidal patterns. (d) Color intensity profiles correspond to the red horizontal line in (a). (e) Decomposed sinusoidal pattern profiles from the input image. Details are described in Sec. 2.2.

2.2. Initial Estimation

In this section, we describe how the initial guesses for depth and scene radiance can be obtained from the given image. Compared to previous literatures [9, 10] which assume uniform albedo over the scene, our method is designed to capture scenes with varying albedo.

For this, we first decompose the sinusoidal pattern from the input image and we estimate the initial depth from the decomposed pattern image. It is done based on the characteristics of the sinusoidal pattern. Since the pattern intensity approaches 1.0 at every maximum point, the local maxima in each color channel can be regarded as the scene radiance. Therefore, these maxima can be regarded as sparse samples of the scene radiance without the pattern. We upsample the dense scene radiance from these samples by using bi-linear interpolation. We name this scene radiance image as the *pattern-free image*. One example is shown in Fig. 2. Local maxima (peak points) are detected from the input image and the peak intensities are interpolated to generate a pattern-free image in Fig. 2 (c).

As we discussed in Sec. 2.1, images are formed by multiplying the Lambertian reflectance, multi-channel albedo and sinusoidal pattern together. Based on this, we can divide the original image by the pattern-free image to get a *pattern image* which only contains the pure components of the sinusoidal pattern. Ideally, the pattern image is equivalent to the image that we get when we project the pattern on an object having white albedo and maximum reflectance. Fig. 2 (e) shows the color intensity profiles of the decomposed pattern image from Fig. 2 (d).

The pattern image lets us use the conventional approach [9] of determining the phase $\phi \in \{0, 2\phi\}$ of the sinusoidal pattern. From the decomposed pattern image I', ϕ can be uniquely determined as

$$\phi = \frac{1}{2\pi}\text{atan}\left(\frac{2I'_r - I'_g - I'_b}{\sqrt{3}(I'_b - I'_g)}\right), \tag{3}$$

where $I'_{\{r,g,b\}}$ correspond to the red, green and blue intensities of I'.

Note that we cannot apply Eq. (3) directly on the input image since the scene has multi-channel albedo. This is because the multi-channel albedo scales the magnitudes of three colored sinusoids differently. Fig. 2 (d) shows a real-world example where the intensity profiles of different channels have different scales of sinusoids.

Since the sinusoidal pattern has a period T and the arctangent function gives a phase within only one period, ϕ cannot directly determine the disparity. Therefore, we utilize the recursive propagation method [10] to get an unwrapped phase ϕ'. Then, the disparity can be calculated as $d = x - \frac{T\phi'}{2\pi}$.

2.3. Refinement of Depth and Albedo

This section introduces the optimization step to refine the depth and albedo. To simplify the problem, we parameterize the shading image with the disparities. The parameterization and the non-linear optimization algorithm are described in this section.

2.3.1 Parameterization of Shading

We derive the normal vector $\mathbf{n}$ and the lighting vector $\mathbf{l}$ in terms of d. This allows us to describe the Lambertian shading with only local disparities so that the optimization problem becomes simpler.

As we mentioned at the beginning of Sec. 2, the camera and the projector are calibrated and their images are rectified. Since we can consider the projector to be a pinhole camera, we can apply the same camera model to the projector as follows:

$$\mathbf{K}_c = \begin{bmatrix} f_x & 0 & x_c \\ 0 & f_y & y_c \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{K}_p = \begin{bmatrix} f_x & 0 & x_p \\ 0 & f_y & y_p \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

where $f_{\{x,y\}}$ are the focal lengths for the rectified images,

and (x_c, y_c), (x_p, y_p) are the principal points of the camera and the projector, respectively. Since all epipolar lines are aligned horizontally, y_c and y_p are the same. Also the relative pose between two rectified sensors is a pure translation $[-b, 0, 0]^\mathsf{T}$ along the x-axis where b is the metric baseline between the two sensors. When a 2D point $\mathbf{x} = [x, y]^\mathsf{T}$ in the rectified camera image corresponds to a point $\mathbf{x}' = [x - d, y]^\mathsf{T}$ in the rectified projector image, the corresponding 3D point $\mathbf{X}$ is described as a function of $\mathbf{x}$ and d as:

$$\mathbf{X} = \left[\frac{b(x-x_c)}{d-x_c+x_p}, \frac{f_x b(y-y_c)}{f_y(d-x_c+x_p)}, \frac{f_x b}{d-x_c+x_p} \right]^\mathsf{T}. \quad (5)$$

Since the projector is the only light source, the lighting direction $\mathbf{l}$ for this 3D point can be approximated as the direction of the ray from the center of the projector to $\mathbf{X}$. Thus, $\mathbf{l}$ is computed as the unit vector of the normalized coordinates of $\mathbf{x}'$, which again is a function of $\mathbf{x}$ and d as:

$$\mathbf{l}(\mathbf{x}, d) \sim \left[\frac{x - d - x_p}{f_x}, \frac{y_p}{f_y}, 1 \right]^\mathsf{T}. \quad (6)$$

The surface normal $\mathbf{n}$ at $\mathbf{X}$ is given by the surface gradient vector which we approximate as the unit vector of the cross-product of two adjacent vectors as:

$$\mathbf{n}(\mathbf{x}, d) \sim (\mathbf{X} - \mathbf{X}_u) \times (\mathbf{X} - \mathbf{X}_r) \quad (7)$$

where $\mathbf{X}_u$ and $\mathbf{X}_r$ are the 3D points of the neighboring pixels above and to the right in the camera image. The Lambertian shading is then obtained as $\mathbf{n}^\mathsf{T}\mathbf{l}$. The initial albedo is estimated by dividing the pattern-free image by the Lambertian shading image.

2.3.2 Non-linear Optimization

In this section, we optimize the initial depth and the albedo to find the best albedo map $\mathbf{A}^*$ and the disparity map $\mathbf{D}^*$ by minimizing the following non-linear objective function:

$$\{\mathbf{A}^*, \mathbf{D}^*\} = \underset{\mathbf{A},\mathbf{D}}{\operatorname{argmin}} \sum_{p \in \mathcal{M}} (\mathbf{O}_p - \mathbf{I}_p(\mathbf{a}_p, d_p))^2 + \lambda_a \sum_{p \in \mathcal{M}} \sum_{q \in \mathcal{N}_p} (\mathbf{a}_p - \mathbf{a}_q)^2 + \lambda_d \sum_{p \in \mathcal{M}} \sum_{q \in \mathcal{N}_p} (d_p - d_q)^2, \quad (8)$$

where $p \in \mathcal{M}$ stands for every non-zero pixel in the image with albedo $\mathbf{a}_p \in \mathbf{A}$ and disparity $d_p \in \mathbf{D}$, $\mathbf{O}_p$ is the observed image intensities, $\mathbf{I}_p(\mathbf{a}_p, d_p)$ is the rendered image intensities, and $q \in \mathcal{N}_p$ represents the neighboring pixels of p.

Our optimization consists of three terms. The first term enforces that the rendered image should be similar to the observed image. The second and the third terms regulate smoothness of the albedo map and the disparity map, respectively. λ_a and λ_d are balancing weights that control the importance of the two smoothness terms. Similar to previous approaches for intrinsic image decomposition [2], we also assume that albedo and depth are spatially smooth.

To solve this non-linear problem, we utilize the Levenberg-Marquardt method implemented in Matlab using analytic Jacobian which is easily derived thanks to our parameterization. During this process, the parameters are iteratively updated to reduce the residual error in a gradient descent manner. In our experiment, this optimization took approximately one minute for $0.1M$ pixels of an image.

Fig. 3 shows the initial estimates and the optimization results. It is shown that the estimated albedo, shading and depth are noticeably improved by the proposed method. Also, Fig. 4 shows how the estimates are optimized along the iterations. At the third iteration, albedo becomes converged and shows similar texture to the ground-truth. The depth error has reduced and converged within seven iterations. For most of our real-world datasets, seven iterations had been found to be enough for convergence.

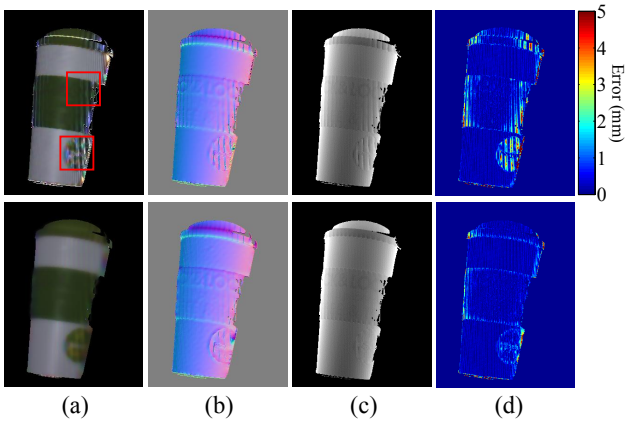


Figure 3. Refinement of the initial guess by using the proposed optimization. The first row shows the initial guess acquired from Sec. 2.2 and the second row shows the result of the optimization described in Sec. 2.3. (a) Estimated albedo map, (b) surface normal map, (c) shading map, (d) absolute depth error (mm).

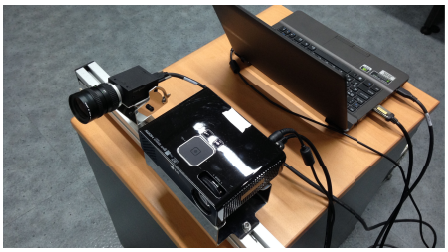


Figure 5. The camera-projector system used in this paper.

3. System Configuration

In this section, we introduce our camera-projector rig and explain our geometric/photometric calibration. Our

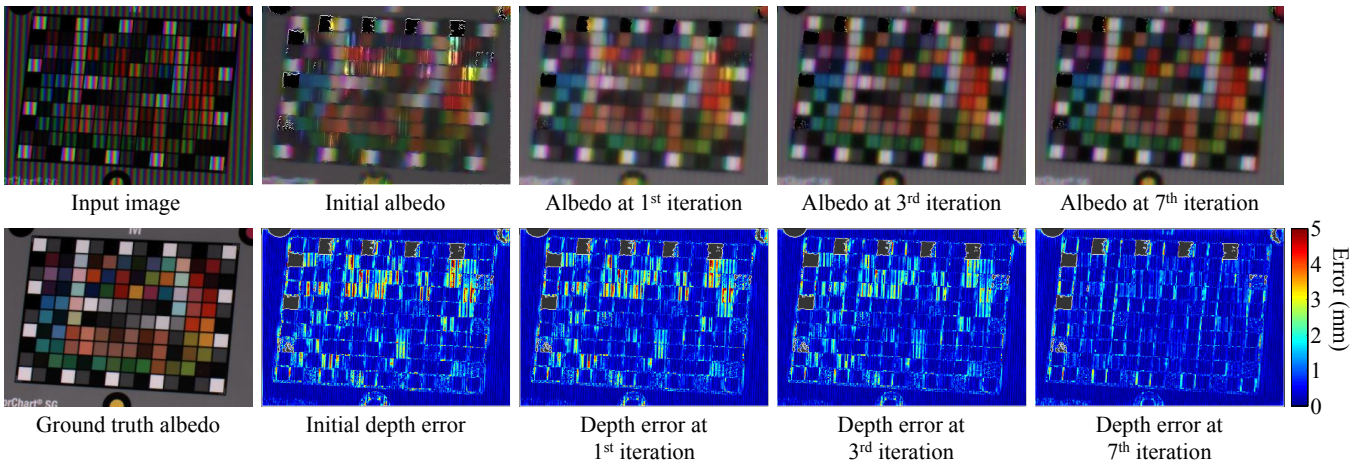


Figure 4. Intermediate images in our iterative optimization step. The first row shows the input image and albedo maps after each iteration. The initial estimation from Sec. 2.2 had obvious artifacts on the initial albedo. The iterative optimization decreases the albedo error and depth error until the seventh iteration. Note that the albedo and depth map are jointly refined using our refinement algorithm.

camera-projector rig consists of a Grasshopper3 camera by Point Grey, Inc. with 1280×800 image resolution and an LG color projector (See Fig. 5) with the same resolution and a refresh rate of 60Hz. The shutter speed of the camera is set to 33ms. The camera lens parameters are 12.5mm and f/5.6.

3.1. Geometric Calibration

We use a simple camera-projector calibration method proposed by Moreno *et al*. [12] Based on the estimated parameters, we rectify the camera and the projector domains [8]. Since the projector has some lens distortion, we intentionally generate an inversely distorted pattern to emit an undistorted pattern on the scene.

3.2. Photometric Calibration

In order to have image intensities which are linearly proportional to the scene radiance, we calibrate the photometric properties of the camera and the projector. We apply the rank-minimization approach to linearize the camera response function [11] and we apply the vignetting correction algorithm [5] to alleviate the intensity degradation. For the gamma correction of the projector, we adjust the intensities of the pattern image using a lookup table. Also, there was a cross correlation within red, green and blue channels. Therefore, we adopt the linear cross-correlation model $C \in \mathbb{R}^{3\times3}$ in [3]. In this model, C maps the emitted radiance (pure red, green and blue with varying intensity) from the projector to the radiometrically calibrated camera. We get C from pairs of emitted intensities from the projector and observed intensities from the camera. The effect of the cross correlation can be removed by multiplying C^{-1} to the color intensity vector.

4. Experimental Results

In this section, we evaluate our algorithm on various kinds of objects. One synthetic example is already shown in Fig. 1. To check the validity of our algorithm on challenging shading models, we focus on real-world experiments.

We conduct real-world experiments by using the camera-projector system in Fig. 5. For generating the sinusoidal pattern, we follow the idea suggested by Gupta *et al*. [6] which concludes that high-frequency patterns are robust against interreflection. Therefore, we set $T = 10$ in units of pixels which makes sinusoids of short periods for our system. For the amplitude α in the sinusoidal pattern, we observed that a small offset helps to get stable results because the projector is prone to emit low light. Therefore, we set $\alpha = 0.4$ to make S be bounded within $[0.2, 1]$. $\lambda_a = 0.5$ and $\lambda_d = 0.1$ are used for the optimization.

To evaluate the accuracy of our depth estimation result, we obtain the ground-truth depth by using the well-known structured light method proposed by Scharstein and Szeliski [16]. This method uses dozens of time-multiplexing patterns to achieve accurate depth measurement. Throughout this paper, we use the metric depth error compared to this method for evaluation.

4.1. Real-world Objects

For evaluating our algorithm on real-world objects, we take five objects with varying albedo and depth. Fig. 6 shows our results on the objects. From top to bottom, the datasets are named as TUMBLER, PIGGY-BANK, T-SHIRT, MUG, and HAND.

Each of these objects are made of different materials such as plastics (TUMBLER and PIGGY-BANK), fabric (T-SHIRT), ceramic (MUG), and human skin (HAND). How-

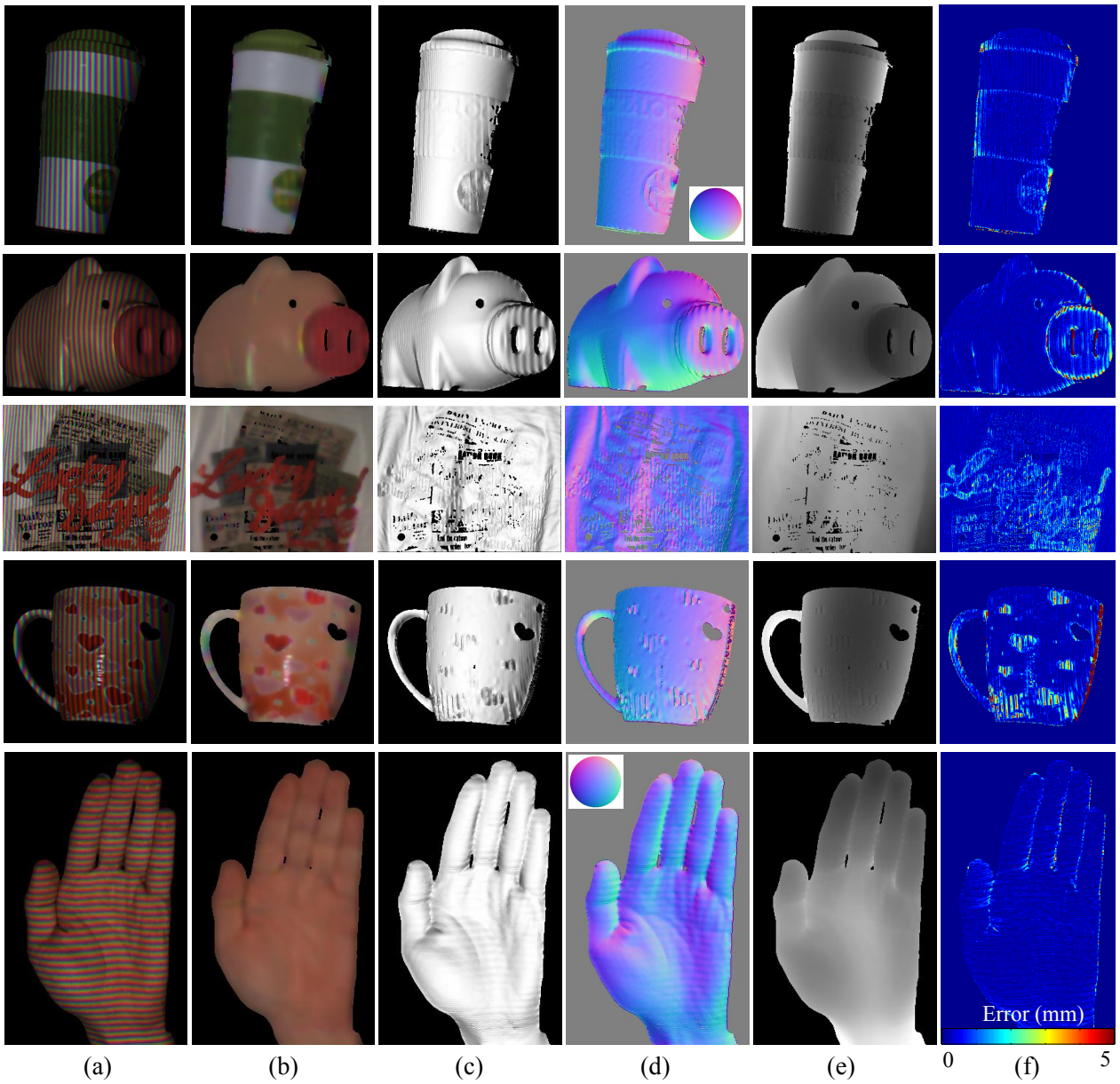
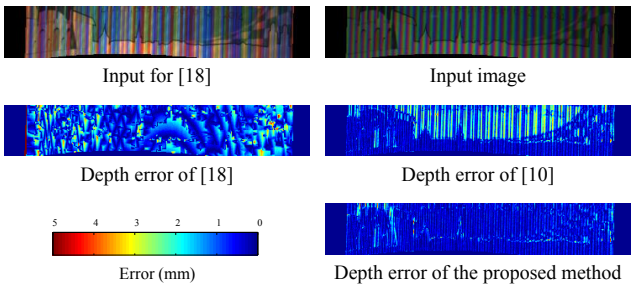


Figure 6. Intrinsic images and depths from real-world objects. (a) Scene illuminated by our structured light pattern which is a sole input of our pipeline. (b) Estimated albedo. (c) Estimated Lambertian shading. (d) Estimated surface normal. (e) Estimated depth map. (f) Depth error in millimeter unit. Images of HAND dataset are rotated 90 degrees counter-clock wise for better visualization.

ever, the promising results imply our method works well not only on the Lambertian-like surfaces. Even though the estimated albedo maps are smooth, we can observe that, except for the boundaries, they are homogeneous as expected. Most regions in the depth maps have errors less than $1mm$. For all of the datasets, the mean of absolute depth error is less than $2mm$.

In PIGGY-BANK and MUG dataset, there are strong specular regions which do not obey our Lambertian-shading-based model. However, if we study the estimated depth map, the specular regions do not severely distort the solution. This is mainly due to the constraint of the light condition (sinusoidal light pattern) which is used in the non-linear optimization. Although the surface does not obey Lambertian reflectance, the optimization step favors depths that generate similar shapes of sinusoids as those in the input image. Therefore, the optimization results in finding a solution whose albedo is adjusted to minimize the residual error.

The result also shows some issues to be improved in



Method	[18]	[10]	Proposed
Mean of error (mm)	0.4563	0.7987	**0.3196**
Variance of error (mm)	0.4894	0.9664	**0.0933**

Figure 7. Evaluating accuracy of estimated depth map. We compare our method with Zhou *et al*. [18] utilizing the De Bruijn sequence and Je *et al*. [10] utilizing a similar pattern as ours. The table shows the mean and the variance of the depth error compared to the ground-truth depth.

our algorithm. As we introduced in Sec. 2.2, the pattern-free image is generated from the interpolation of the local maximum intensities. However, due to the lack of samples around the thin structure (for example, the handle in MUG) it makes a wrong initial guess. This wrong guess cannot be recovered by our refinement algorithm. We also observe strong wiggling on the nose area in PIGGY-BANK. This usually happens on regions with very strong chroma. For those regions, we find that the projected sinusoidal pattern is observed as clipped sinusoids, hence the refinement stage does not work properly.

4.2. Comparison with Existing Methods

We compare our depth map with other methods proposed by Zhou *et al*. [18] and Je *et al*. [10] that use single-shot patterns. Zhou *et al*. use the De Bruijn sequence which makes unique codes between neighboring color stripes. Compared to the De Bruijn approach, our method can generate the depth map with sub-pixel accuracy. Since Je *et al*. [10] utilize a similar phase-shifting pattern, the method may be basically identical to ours, however, their method cannot be applied on colored objects because it assumes that the scene albedo is only white. Fig. 7 shows that our depth quality outperforms that of the others.

4.3. Application: Re-lighting of Human Face

As a possible application, from estimated albedo and depth, we can synthesize new images with certain light conditions at different viewpoints. We captured a human face using our system and obtained the intrinsic images as shown in Fig. 8. Since our method can handle varying albedo, we can capture a human face without white powder makeup which is conventionally used in previous methods to make a uniform surface albedo.

5. Discussion

In practice, our initial estimation algorithm may not be ideal because of the followings reasons: 1) Our initial guess is estimated by a step-by-step process, which implies that the error can be accumulated. 2) Estimating a pattern-free image is prone to depth discontinuity because the method cannot correctly interpolate peak intensities from unreliable or missing samples. 3) Due to the effects of cast shadow, specularity, highly textured albedo or non-Lambertian shading, the peak positions can be drifted. 4) If the texture of the scene is complex or the geometry has significant depth discontinuity, our approach produces wiggling depth as shown in Fig. 6 (e).

Another weakness of the proposed method may be the dependency on the object colors. Since the proposed method requires three channel observations to decode the phase-shifting pattern, objects having albedo close to the primary colors (red, green or blue) could not guarantee to give good results. The question of measuring the correctness of estimated surface albedo is another issue.

6. Conclusion

In this paper, we propose a method to reconstruct both the metric depth and the intrinsic properties *i.e*. albedo and shading, from a single structured light image. The proposed method extends the conventional color structured light model by adopting the Lambertian shading model. Based on the extended image formation, we first detect peak points with local maxima in each color channel and interpolate the samples to estimate the initial albedo and depth of the scene. The parameterization of the normal, the lighting direction and the sinusoidal pattern with the disparity makes the optimization step simple and efficient. The effectiveness of the proposed method is shown from the real experiments on various multi-albedo objects and it shows the feasibility of estimating both albedo and shape simultaneously from a single-shot structured light. In the future, we plan to improve our method by designing a robust energy function to handle outliers. Careful consideration on the non-isotropic characteristics of the structured light source such as [14] is another interesting direction.

Acknowledgments

This work was supported by the Development of Intra-Oral Scanner System Project funded by the FOURDRS Co., Ltd.

References

[1] J. T. Barron and J. Malik. Intrinsic scene properties from a single rgb-d image. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013. 2

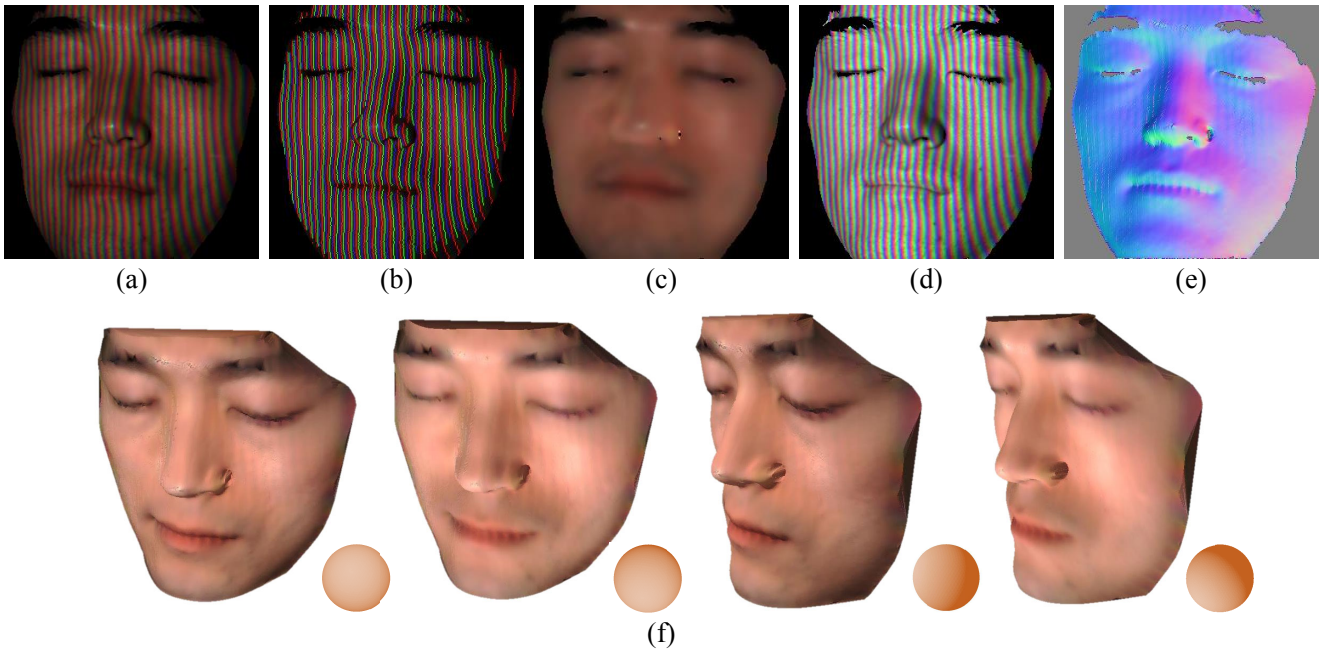


Figure 8. Relighting human face as an application of our algorithm. (a) A pattern image (input). (b) Visualization of local intensity maxima. (c) Estimated albedo map. (d) Estimated shading image with projector pattern. (e) Estimated surface normal. (f) Relighting examples with various light directions at different viewpoints. These are rendered by the Lambertian shading model. Small spheres on the bottom-right side of the faces are rendered with the same lighting conditions.

[2] J. T. Barron and J. Malik. Shape, illumination, and reflectance from shading. In *Tech Report*, 2013. 1, 4

[3] D. Caspi, N. Kiryati, and J. Shamir. Range imaging with adaptive color structured light. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 20(5):470–480, 1998. 1, 2, 5

[4] G. Choe, J. Park, Y.-W. Tai, and I. S. Kweon. Exploiting shading cues in kinect IR images for geometry refinement. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014. 1

[5] D. B. Goldman and J.-H. Chen. Vignette and exposure calibration and compensation. In *Proceedings of International Conference on Computer Vision (ICCV)*, 2005. 5

[6] M. Gupta and S. Nayar. Micro Phase Shifting. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, Jun 2012. 1, 5

[7] Y. Han, J.-Y. Lee, and I. S. Kweon. High quality shape from a single rgb-d image under uncalibrated natural illumination. In *Proceedings of International Conference on Computer Vision (ICCV)*, 2013. 1

[8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004. 5

[9] P. S. Huang, Q. Hu, F. Jin, and F.-P. Chiang. Color-encoded digital fringe projection technique for high-speed three-dimensional surface contouring. *Optical Engineering*, 38(6):1065–1071, 1999. 1, 3

[10] C. Je, S. W. Lee, and R.-H. Park. Color-phase analysis for sinusoidal structured light in rapid range imaging. In *Proceedings of Asian Conference on Computer Vision (ACCV)*, 2004. 1, 3, 7

[11] J.-Y. Lee, Y. Matsushita, B. Shi, I. S. Kweon, and K. Ikeuchi. Radiometric calibration by rank minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 35(1):144–156, 2013. 5

[12] D. Moreno and G. Taubin. Simple, accurate, and robust projector-camera calibration. In *3DIMPVT - Joint 3DIM/3DPVT Conference*, 2012. 5

[13] J. Park, S. N. Sinha, Y. Matsushita, Y.-W. Tai, and I. S. Kweon. Multiview photometric stereo using planar mesh parameterization. In *Proceedings of International Conference on Computer Vision (ICCV)*, 2013. 1

[14] J. Park, S. N. Sinha, Y. Matsushita, Y.-W. Tai, and I. S. Kweon. Calibrating a non-isotropic near point light source using a plane. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014. 7

[15] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface prolometry. *Pattern Recognition (PR)*, 43(8):2666–2680, 2010. 1

[16] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003. 1, 5

[17] M. F. Tappen, W. T. Freeman, and E. H. Adelson. Recovering intrinsic images from a single image. In *Annual Conference on Neural Information Processing Systems (NIPS)*, 2002. 1

[18] Y. Zhou, D. Zhao, Y. Yu, J. Yuan, and S. Du. Adaptive color calibration based on one-shot structured light system. *Sensors*, 12(8):10947–10963, 2012. 7