

Light Fall-off Stereo

Miao Liao, Liang Wang, Ruigang Yang
University of Kentucky

Minglun Gong
Laurentian University, Canada

Abstract

We present *light fall-off stereo-LFS*—a new method for computing depth from scenes beyond lambertian reflectance and texture. LFS takes a number of images from a stationary camera as the illumination source moves away from the scene. Based on the inverse square law for light intensity, the ratio images are directly related to scene depth from the perspective of the light source. Using this as the invariant, we developed both local and global methods for depth recovery. Compared to previous reconstruction methods for non-lambertian scenes, LFS needs as few as two images, does not require calibrated camera or light sources, or reference objects in the scene. We demonstrated the effectiveness of LFS with a variety of real-world scenes.

1. Introduction

Recovering depth information from 2D images is one of the central problems in computer vision. Many different cues in the images have been used, such as stereoscopic disparity, shading, textures, focus and defocus. The vast majority of these methods makes a strong assumption that objects in the scene reflect light equally in all directions. Such a diffuse or Lambertian surface assumption is violated in almost all real-world objects, leading to incorrect depth estimate. Although techniques that go beyond Lambertian surfaces have been proposed, they typically require precise calibration of cameras and/or light sources, sufficient surface textures, or reference objects in the scene.

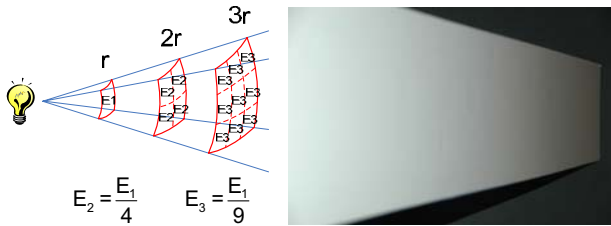


Figure 1. An illustration of the light fall-off effect. Left: light energy falls off as the square of the distance, e.g., the irradiance of a unit area at distance $2r$ is one fourth of that at distance r . Right: a real photograph showing the effect.

In this paper we exploit a different cue in the image formulation process, namely the *inverse square law* for light intensity. As illustrated in Figure 1 (left), the intensity of light observed from a source of constant intrinsic luminosity falls off as the square of the distance from the object. A real example is provided in Figure 1 (right), the scene is illuminated with a flash light, objects further away appear to be darker due to the fall-off of light intensity over distance. It can be derived that

$$I(p) = \frac{k_p}{r_p^2}, \text{ or } r_p \propto 1/\sqrt{I(p)}, \quad (1)$$

where $I(p)$ is the pixel value of a scene point p , r_p is the distance between the light source and p , and k_p is a constant related to the intensity of the light source and the reflectance and orientation of point p . In fact, in this ideal setup in which the object is flat with a uniform color, we can almost assume that k_p is identical for all surface points, therefore the pixel values directly provide a qualitative measure of depth.

However, real-world objects are usually not flat and typically have spatially varying textures. To deal with these practical issues, we can take an additional image under a different lighting configuration. The new pixel intensity then is

$$I'(p) = \frac{k_p}{r_p'^2} \quad (2)$$

If we compute the ratio of $I(p)$ and $I'(p)$, the impact of reflectance cancels out, making the ratio only related to the depth. Furthermore, if we measure the distance between the two light positions (i.e., $\Delta r = r_p' - r_p$), which is very easy to do, we can compute the scene depth to the light source.

The above formulation, which we called *light fall-off stereo* (LFS), holds for scenes with arbitrary bidirectional texture function (BTf), and therefore can be used to design a practical depth acquisition system with a single camera and a moving light source. As we will show later, LFS is easier to implement compared to existing depth acquisition methods for arbitrary BRDF/BTFs. It only needs as few as two images, making it possible to capture dynamic scenes.

The main contribution of this paper is a novel way to estimate depth information from scenes beyond lambertian re-

flectance model. We also developed a global optimization-based method that uses multiple light variations to further improve the accuracy and robustness. The effectiveness of LFS is demonstrated with a variety of real-world scenes exhibiting complex reflectance and geometries.

2. Related Work

In the literature of computer vision, a vast number of algorithms have been proposed to reconstruct the three-dimensional shape of a scene from one or multiple photographs. The set of such techniques is called Shape-from-X, where X can be Stereo, Shading, Focus/Defocus, etc. In this section we briefly discuss some classic and modern Shape-from-X methods which are mostly related to our proposed algorithm.

Shape-from-Shading. Shape-from-Shading (SFS) deals with the recovery of shape from a gradual variation of shading in the image [12, 15, 17]. Given one gray level image and known light direction, the surface shape at each pixel can be recovered by studying the image formation process [13]. Since developed, most work in SFS makes simplifying assumption, that is the recovered surface is Lambertian and with constant or known albedo [16, 21, 31, 27]. Unfortunately this assumption is violated for nearly all real world objects, leading to incorrect shape estimation. Although several extensions have been proposed [18, 2, 1, 25], satisfactory performance is still hard to achieve via SFS on real images with arbitrary surface BTFs [37]. Very recently Prados and Faugeras developed a more sophisticated imaging model that takes into account the $1/r^2$ light fall-off term [27]. This light term removes the fundamental convex/concave ambiguity. They have demonstrated some excellent results, but their formulation is still limited by Lambertian surfaces.

Photometric Stereo. Photometric Stereo methods recover the shape and albedo of an object using multiple images among which camera position is fixed, and only the lighting directions vary [33]. Although utilizing multiple lighting variations allows accurate results, classical methods follow the same image formation assumption as made in SFS [33, 3, 34]. As a result, most work cannot well handle non-Lambertian surfaces. Several methods have been addressed to overcome this limitation. Some approaches require extra constraints on the number and positions of light sources and allow only a class of diffuse non-Lambertian surfaces to be handled [19, 32]. Some require an object with BRDF similar to the unknown scene as a prior calibration knowledge [11]. In contrast the method presented herein allows depth recovery with complex BRDFs and does not require any known reference object. The light fall-off property has also been utilized in photometric stereo by [4]. This work relies on the controlled motion of a point light source that is not at infinite but relatively close to the ob-

ject. However different from our approach, their depth estimation is still based on multiple image measurements taken over time.

Stereopsis with Active Illumination. Nearly all existing methods for stereo reconstruction assume that scene reflectance is Lambertian, and make use of color constancy as a matching invariant [29]. More recently there are special metrics proposed to handle non-Lambertian cases. For example Helmholtz stereopsis allows matching of arbitrary BRDFs and uses reflectance function reciprocity as an invariant [38, 39]. By collocating point light sources with each camera it is possible to record reciprocal pairs using two different lighting conditions. However, this method requires the light sources be colocated with the optical center of each camera while ours does not require the position of camera to be even known. Some works make use of illumination variation as a correspondence aid. For instance, Spacetime stereo [6, 35] and BRDF invariant stereo [7] formulate stereo matching in the presence of illumination variation and achieves good results. While our work also builds on lighting variations, those variations come from the $1/r^2$ fall-off term of the illumination, and we only require a single un-calibrated camera.

Shape-from-Focus/Defocus. Classic Shape-from-Focus/Defocus methods collect images at multiple lens settings and define metrics that evaluate sharpness or blurring over a small spatial area surrounding the pixel [5, 20, 26, 30]. Most of these approaches follow the equalfocal assumption, i.e. the surface depth is constant within that spatial window, therefore suffers from poor performance across depth discontinuities. To alleviate this problem, Hasinoff et al. [10] and Zhang and Nayar [36] have presented new methods which allow per-pixel focus/defocus analysis to be applied and can achieve much sharper, more accurate geometric details. While nearly all Shape-from-Focus/Defocus methods have difficulty dealing with textureless regions due to focus ambiguity, illumination patterns can be projected to provide scene texture at the expenses of light source calibration [9, 36].

The idea of using the inverse-square law to recover 3D shape has been explored in two previous papers [22, 24]. However both methods proposed by these papers require multiple images with precisely controlled light source or calibrated camera. In fact [24] only presented some basic concepts without any result. Compared to their work, our approach requires a much simplified setup, i.e. one LCD projector on a linear stage together with an un-calibrated camera. And what is more, our paper demonstrates more convincing experimental results to support our LFS algorithm.

3. Methods

In this section we first present the radiometric model used in LFS and show how it can be used to provide depth estimate. We discuss the assumptions made and the associated issues in designing a practical range sensor system.

3.1. Depth Recovery for a Pivot Point

Here we define pivot point as the intersection between the line connecting the two lighting positions and the surface in the scene. We first discuss how to recover depth for the pivot point. Let us recall the image formation process. As shown in Figure 2 (left), the scene is illuminated by a single point light source L . The irradiance of the pivot point p in the scene is

$$E(p) = W \frac{\cos(\omega_L)}{r_p^2} \quad (3)$$

where W is the light radiance, r_p is the distance between point p and the light, and ω_L is the incident angle. p will reflect light toward the observation camera C . The reflected radiance value is defined as

$$L(p) = R(p, \omega_L, \omega_C) E(p), \quad (4)$$

where ω_C is the viewing direction, and $R(p, \omega_L, \omega_C)$ is the spatially varying bidirectional reflectance function (a.k.a., bidirectional texture function–BTF). It takes into consideration of surface albedo variations. Finally, the imaging system measures the irradiance value on the camera’s sensor, which is

$$I(p) = \rho L(p) \quad (5)$$

where ρ is a constant parameter determined by the imaging optics (details can be found in [14]). For the scope of this paper, we assume that the camera has a linear response, in other word, the camera is measuring relative irradiance directly. To deal with cameras with non-linear responses, standard radiometric calibration procedures (e.g. [23, 8]) should be applied to correct the pixel values.

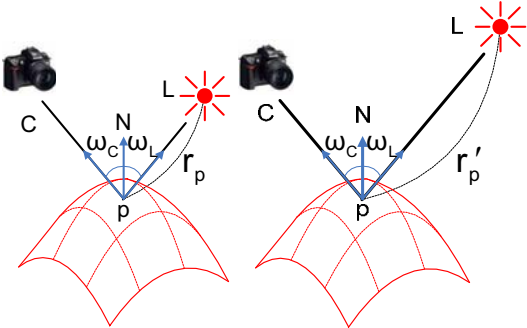


Figure 2. The pivot point illuminated by a point light source at the first (left) and the second (right) positions.

Combing equations 3, 4 and 5, we get

$$I(p) = \rho W \frac{\cos(\omega_L)}{r_p^2} R(p, \omega_L, \omega_C). \quad (6)$$

This is how equation 1 is derived.

As shown in Figure 2 (right), With both camera position and light intensity fixed, we move the point light along the direction of ω_L to a new position. In this process, every parameter in equation 6 remains the same except the distance to the light changes to r'_p . The new observed intensity $I'(p)$ measured by the camera is:

$$I'(p) = \rho W \frac{\cos(\omega_L)}{r_p'^2} R(p, \omega_L, \omega_C) \quad (7)$$

Computing the ratio of $I'(p)$ and $I(p)$ cancels out all the terms but the distance to the light source, that is

$$\frac{I'(p)}{I(p)} = \frac{r_p^2}{r_p'^2} \quad (8)$$

With $\Delta r = r'_p - r_p$ measured using a ruler, we can compute r_p as

$$r_p = \frac{\Delta r}{\sqrt{I(p)/I'(p)} - 1} \quad (9)$$

It should be noted that the depth is actually from the perspective of the light source, not from that of the camera.

In the above LFS formulation, the camera position is fixed, both the BTF term and the light intensity term are canceled out. Therefore, the depth measured by this method is invariant to lighting intensity and surface property.

3.2. Estimate a Depth Map for the Whole Scene

The above process can be extended for recovering depth for other points in the scene as well. As shown in Figure 3, the depth of all points in the scene are measured with respect to a depth reference plane, which passes through the first lighting position S and is perpendicular to the lighting direction for the pivot point p . For an arbitrary point q in the scene, here we use $I(q)$ to denote the observed intensity of the point under the lighting position S . It is worthy noting that $I(p)$ and $I(q)$ can be measured at same time as long as both points p and q are within the camera’s field of view.

In order to maintain the same incident lighting direction, when capture the intensity of point q under the second lighting position, we need to move the light along the ray qS . Here we place the light at location T' , the intersection between ray qS and the second lighting plane, and refer the new observed intensity as $I'(q)$. According to the lighting distances shown in the figure, we have:

$$\frac{I'(q)}{I(q)} = \frac{(r_q / \cos \theta)^2}{((r_q + \Delta r) / \cos \theta)^2} \quad (10)$$

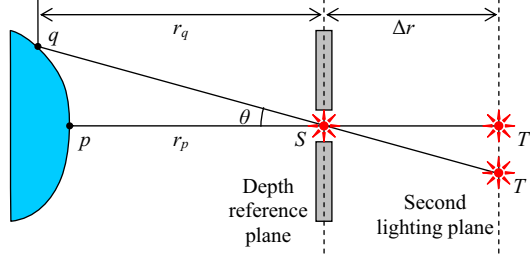


Figure 3. The setup for recovering a depth map for the whole scene.

After simplification, we get:

$$r_q = \frac{\Delta r}{\sqrt{I(q)/I'(q)} - 1} \quad (11)$$

Comparison between equation 9 and 11 suggests that, once the light is positioned at proper locations on the second lighting plane, the same equation can be used to estimate the depth of different points in the scene.

Calibrating the light position for different points in the scene can be a tedious task. Fortunately, we can simplify the problem by placing an optical occluder along the depth reference plane such that light can only go through a small hole at location S . Hence, when the light is positioned at the second lighting plane, only a small portion of the scene is illuminated and the incident lighting direction for this small portion is the same as the case when light is placed at location S . Multiple images of the scene can be captured either by moving the light along the second light plane or by building a light array at the second light plane and turning on one light at a time. Through finding the maximum intensity at each pixel location from multiple captured images, we can effectively extract the illuminated portions of the scene and merge them into a single image. The resulting image, referred as Multi-Lighting-Direction-Image, shows the appearance of the scene under different lighting directions for different portions and guarantees that the incident light direction on every surface point remains the same. As a result, it can be used to recover the depth of the entire scene using equation 11.

3.3. Practical Approximation

While the approach discussed in above subsection gives more accurate depth estimation results in theory, our experiments show that reasonable good estimations can be obtained through a much simplified procedure. By assuming the observed intensity difference of point q under lighting position T and T' is negligible, we simply capture a single image of the scene under lighting position T and use it to approximate the required Multi-Lighting-Direction-Image. In fact, this is similar as assuming that the incident lighting directions for different points in the scene are parallel.

When the object size is relatively small compared to the distance to the light, such an approximation is valid and has been used in almost all photometric stereo and shape- from-shading algorithms.

In order to evaluate the error introduced by the above approximation, an empirical experiment is conducted in which we try to reconstruct a flat piece of paper under two different settings (shown in Figure 4). In the first setting, the paper, 11 inches in width, is placed at about 20 inches from the depth reference plane. In the second setting, the paper is about 43 inches away. The evaluation result suggests that the estimated depth is noticeably distorted under the first setting where the paper spans an angle of 30 degrees in the field of view of the light. However, the error dramatically decreases when the span angle reduces to 15 degrees under the second setting.

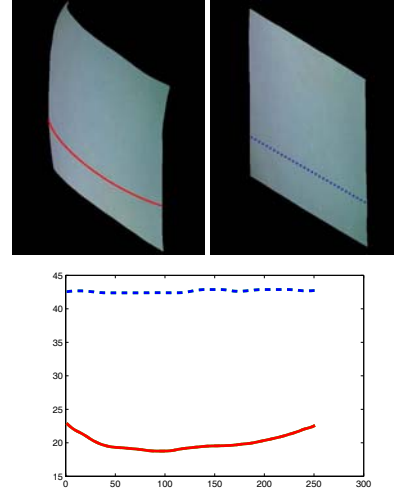


Figure 4. The effect of distortion caused by incident angle changes. Images in the top row are 3D views of the recovered depth maps when the piece of paper is positioned at the near (left) and far (right) locations. The bottom row plots the depth values extracted from a selected scanline in both depth maps. The variances of depth values in these two scanlines are 0.0376 and 1.1757, respectively.

While placing the light source far from the scene helps to minimize the distortion caused by light direction approximation, it also makes the $1/r^2$ light fall-off term less significant and hence small depth variation in the scene may become unrecoverable. Here we try to analyze the maximum depth resolution, i.e., the smallest depth change of can resolve using LFS, given a camera's limited dynamic range.

From equation 5, we know that

$$I(p) = \frac{m}{r_p^2} \quad (12)$$

where m is a constant related to camera optics, BTF and

lighting. Taking derivative on both side, we have

$$dI(p)/dr_p = -2m/r_p^3 = -2I/r. \quad (13)$$

We can see that change in I_p is only inversely proportional to r . If we say $r = 1000\text{mm}$, taking I as 127, 1 mm depth change will cause approximately 0.25 unit variation in intensity, i.e., we need to have at least 10 bits/pixel intensity resolution to resolve that. Fortunately most cameras today in fact capture 12 bits per pixel, sufficient to high accuracy capture. In addition, high-dynamic-range imaging techniques [8] can be used if necessary.

It will be interesting to study the optimal lighting distance for a given sensor dynamic range. Because of the number of factors involved in this image formulation, e.g., surface normals, albedo variations, and shininess, such an analysis is non trivial and we leave it for further work.

4. Global Method

So far we have shown that a depth map can be generated using two images of the scene illuminated by a point light source positioned at two different locations. The derived local LFS method is both efficient and simple to implement. However since the image acquisition process can be easily corrupted with noise, the depth maps obtained using local approach are sometimes noisy. In this section we reformulate the LFS problem under an energy minimization framework. The derived global optimization-based method can produce smooth and accurate depth maps since: 1) it can make use of images captured under more than two lighting configurations; and 2) it enforces the spatial consistency among adjacent pixels in the depth map.

4.1. Formulation under Energy Minimization

As illustrated in Figure 5, the scene is illuminated by a point light source located at different positions. A sequence of images of the scene, referred as $I, I^1, \dots, I^N$, are thereby acquired. Please note that the point light source is again carefully positioned so that, for a pivot point P in the scene, the incident light direction does not change. We also assume that the lighting direction changes for the remaining points on the scene are negligible.

Now if we set the depth reference plane at the first light position, according to equation 11, the following holds regardless of surface BTF and light intensity.

$$\sqrt{I_{x,y}} r_{x,y} = \sqrt{I_{x,y}^1} (r_{x,y} + \Delta r_1) = \dots = \sqrt{I_{x,y}^N} (r_{x,y} + \Delta r_N) \quad (14)$$

where $I_{x,y}$ is the intensity of pixel (x, y) in the captured image I . $r_{x,y}$ is value of pixel (x, y) in the estimated depth map, i.e., the distance between the depth reference plane and the corresponding 3D point of pixel (x, y) .

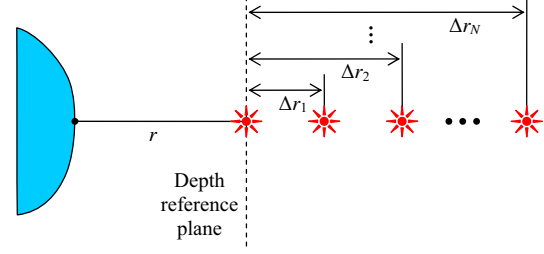


Figure 5. Images are captured under multiple lighting conditions. The light movement is carefully controlled to minimize the change of incident directions on the scene surface.

Due to the noise in the image acquisition process, when more than two images are used, for a given pixel (x, y) , we may not be able to find a value $r_{x,y}$ that satisfies the above equation. Our objective is therefore to find a depth value that minimizes the variance among different terms. Here we use $K_{0,\dots,N}$ to represent the value of the above terms:

$$\begin{aligned} K_0 &= \sqrt{I_{x,y}} r_{x,y} \\ K_i &= \sqrt{I_{x,y}^i} (r_{x,y} + \Delta r_i), \quad 1 \leq i \leq N \end{aligned} \quad (15)$$

The energy minimization objective function can then be defined accordingly using the following equation:

$$E = (1 - \lambda) \sum_{x,y} \sum_{i=0}^N (K_i - \bar{K})^2 + \lambda \sum_{x,y} (u_{x,y}^2 + v_{x,y}^2) \quad (16)$$

where u, v , defined in equation 17, are the symmetric second finite differences of the variables $r_{x,y}$. The second term is used to enforce the smoothness of the solution. λ ($0 \leq \lambda \leq 1$) represents a user defined parameter that adjusts the relative importance between the error term and the smoothness term.

$$\begin{aligned} u_{x,y} &= r_{x+1,y} - 2r_{x,y} + r_{x-1,y} \\ v_{x,y} &= r_{x,y+1} - 2r_{x,y} + r_{x,y-1} \end{aligned} \quad (17)$$

4.2. Optimization Approach

Our goal is to find a depth map that satisfies the inverse-square law at all pixel locations and in all captured images. When there is no noise, it is equivalent to finding a surface that incurs zero measurement error. However in the real world where sensor measurement is almost always corrupted with noise, we approximate the objective surface by finding a smooth surface that minimizes our cost function defined in equation 16.

Our global LFS method implements the standard Conjugate Gradient (CG) algorithm together with the line search algorithm *DBRENT* as an iterative minimization tool [28]. Given the objective function expressed in equation 16, this minimization process simply requires the construction of two functions: one that computes the objective

cost value E and the other calculates the gradient of E with respect to the vector $\vec{r}$. We also impose boundary conditions by defining the gradient to be zero at boundary pixels. Although we cannot guarantee that the solution is converged to the exact correct surface, experiment shows that our recovered depth map is indeed smooth and incurs small measurement error.

5. Experiment and Results



Figure 6. Our experimental setup. A sliding projector is in the middle and a video camera is on the right.

As demonstrated in Figure 6, in our experiments the scene is captured using a PointGrey FLEA video camera and is illuminated using a Epson powerlite 82c projector that provides very isotropic illumination. The spatial resolution of images is 1024×768 at 8 bits. To eliminate the errors introduced by the on-camera color processing algorithm, the raw images captured by the Bayer pattern sensor is used directly as input. The camera is calibrated to ensure a linear radiometric response using the techniques in [8]. A lookup table is built for each color channel since the response curves for the three color channels are slightly different.

When adjusting lighting positions, we move the projector along a guiding rail. This ensures that the light positions always lie on a straight line, the intersection of which with the scene is the pivot point. We also try to position the pivot point at the center of the scene, which helps to minimize the lighting direction changes for other points in the scene.

We demonstrate some of the results here. Figure 7 and 8 show the depth map and the 3D rendering result for two Lambertian objects. Most of the fine details are recovered. Figure 9 and Figure 10 show the depth maps of two scenes. Since our method does not deal with areas in shadow, these areas are masked out in the resulting depth map. Figure 11 illustrates the global approach presented in section 4. The scene consists of two very specular objects: a glossy book cover and a piece of silk. Note that this kind of specular surfaces is very sensitive to the light's incident angle. As

shown in the lower row, our global method demonstrates significantly improvement for this challenging scene.

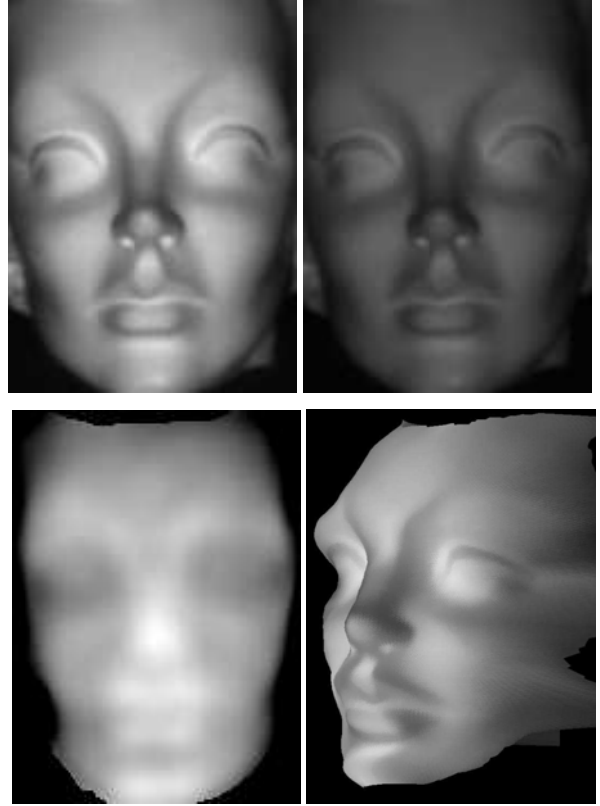


Figure 7. 1st row: images taken with light at two positions. 2nd row: left image is the depth map, and right image is a view of the recovered 3D model.

6. Conclusions and Future Work

A novel way of recovering depth information from 2D images is presented in this paper. While existing shape-from-X approaches rely on intensity matching, shading, or focusing information, our approach uses a completely different cue: the light fall-off property. The derived algorithm makes no assumption about the surfaces in the scene and can handle surfaces beyond lambertian reflectance properties and textures.

We presented two different methods for estimating depth from light fall-off. The local method requires only two input images and has a very low computational cost, but may suffer from noises. The problem can be addressed using the global method, which utilizes more than two input images for increased robustness, as well as enforces smoothness in the resulting depth map. The effectiveness of both local and global methods are demonstrated with a variety of real-world scenes.

Due to the low computational cost of the proposed approaches, it is highly possible to perform the depth recon-

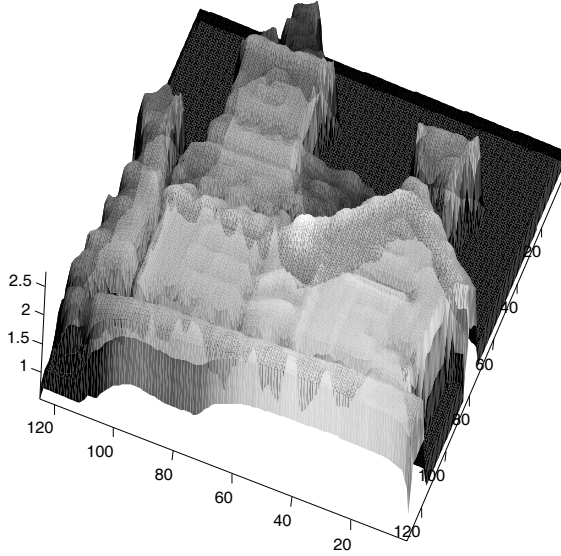


Figure 8. A toy house with very fine details. We show its depth in 3D view.

struction process in real-time. However, since a single light source is used in our current experiment and its position needs to be adjusted manually, scenes with dynamic objects cannot be captured fast enough. As an extension, we plan to use two or more synchronized projectors as light sources, which projects white light to the scene in an alternating order. The scene is captured by a high-speed video camera, which is also synchronized with the projectors such that different frames record the appearance of the scene under the lighting of different projectors. Hence, using images captured at adjacent frames as input, the proposed LFS approaches will be able to recover depth maps for dynamic

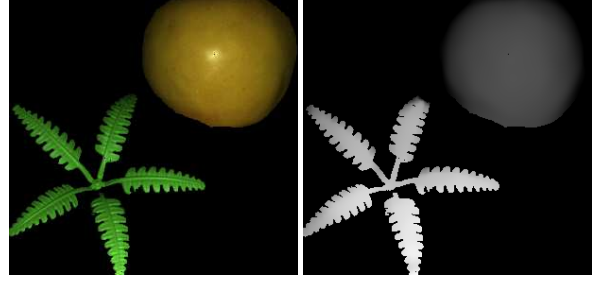


Figure 9. left image is a simple scene with plastic leaves and an apple in it. Right image is its depth map. Both the leaves and the apple are non-lambertian surface.



Figure 10. A more complex scene. There are wood, metal, plastic in this scene. Since we don't deal with shadow areas, we ignore the pixels below a certain threshold.

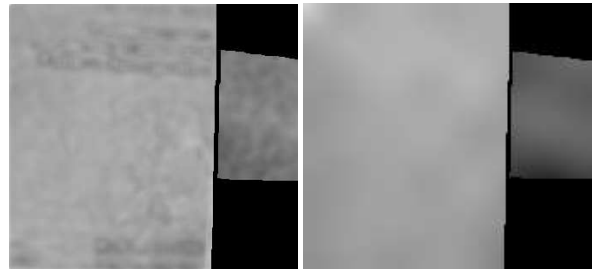


Figure 11. Comparison between local method and global method. 1st row: there are two objects in the scene at different depth. The surface of these two objects are very specular making the result quite sensitive to incident angle change. 2nd row: left image is depth map obtained by per-pixel calculation from two shaded images only. Right image is depth map processed by global method from six shaded images. (The parameter λ is set to 0.15 in our experiments)

scenes on the fly.

References

- [1] M. Asada, T. Nakamura, and Y. Shirai. Weak lambertian assumption for determining cylindrical shape and pose from shading and contour. In *Proceedings of CVPR*, pages 726–729, 1992. 2
- [2] S. Bakshi and Y. Yang. Shape from shading for non-lambertian surfaces. In *Proceedings of ICIP*, pages 130–134, 1994. 2
- [3] R. Basri and D. Jacobs. Photometric stereo with general, unknown lighting. In *Proceedings of CVPR*, pages Vol II:374–381, 2001. 2
- [4] J. Clark. Active Photometric Stereo. In *Proceedings of CVPR*, pages 29–34, 1992. 2
- [5] T. Darrell and K. Wahn. Pyramid based depth from focus. In *Proceedings of CVPR*, pages 504–509, 1988. 2
- [6] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz. Spacetime stereo: A unifying framework for depth from triangulation. *PAMI*, 27(2):296–302, 2005. 2
- [7] J. Davis, R. Yang, and L. Wang. Brdf invariant stereo using light transport constancy. In *Proceedings of ICCV*, pages 436–443, 2005. 2
- [8] P. E. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. *Proceedings of ACM Siggraph*, pages 369–378, 1997. 3, 5, 6
- [9] B. Girod and S. Scherrock. Depth from defocus of structured light. In *Proceedings of SPIE Conference on Optics, and Image Sensing for Machine Vision*, 1989. 2
- [10] S. W. Hasinoff and K. N. Kutulakos. Confocal stereo. In *Proceedings of ECCV*, pages 620–634, 2006. 2
- [11] A. Hertzmann and S. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *PAMI*, 27(8):1254–1264, 2005. 2
- [12] B. Horn. Obtaining shape from shading information. In *Proceedings of The Psychology of Computer Vision*, pages 115–155, 1975. 2
- [13] B. Horn. Understanding image intensities. *Artificial Intelligence*, 8(2):201–231, 1977. 2
- [14] B. Horn. *Robot Vision*. MIT Press, 1986. 3
- [15] B. Horn and M. Brooks. Shape and source from shading. In *International Joint Conference on Artificial Intelligence*, pages 932–936, 1985. 2
- [16] B. Horn and M. Brooks. The variational approach to shape from shading. *Computer Vision Graphics and Image Processing*, 33(2):174–208, 1986. 2
- [17] B. Horn and M. Brooks. *Shape from Shading*. MIT Press, 1989. 2
- [18] E. N. C. Jr. and R. Jain. Shape from shading for surfaces with texture and specularity. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 652–657, 1981. 2
- [19] E. N. C. Jr. and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics Image Processing*, 18(4):309–328, 1982. 2
- [20] E. Krotkov. Focusing. *IJCV*, 1(3):223–237, 1987. 2
- [21] Y. Leclerc and A. Bobick. The direct computation of height from shading. In *Proceedings of CVPR*, pages 552–558, 1991. 2
- [22] S. Magda, D. Kriegman, T. Zickler, and P. Belhumeur. Beyond lambert: reconstructing surfaces with arbitrary brdfs. In *Proceedings of ICCV*, volume 2, pages 391–398, 2001. 2
- [23] T. Mitsunaga and S. K. Nayar. Radiometric Self Calibration. In *Proceedings of CVPR*, volume 1, pages 380–387, 1999. 3
- [24] J. Mulligan and X. Brolly. Surface determination by photometric ranging. In *Proceedings of CVPRW*, volume 3, page 40, 2004. 2
- [25] A. Ortiz and G. Oliver. Shape from shading for multiple albedo images. In *Proceedings of ICPR*, pages Vol I: 786–789, 2000. 2
- [26] A. Pentland. A new sense for depth of field. *PAMI*, 9(4):523–531, 1987. 2
- [27] E. Prados and O. Faugeras. Shape from shading: a well-posed problem? In *Proceedings of CVPR*, 2005. 2
- [28] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes in C*. Cambridge Univ. Press, New York, 1988. 5
- [29] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1):7–42, May 2002. 2
- [30] M. Subbarao and G. Surya. Depth from defocus: A spatial domain approach. *IJCV*, 13(3):271–294, 1994. 2
- [31] R. Szeliski. Fast shape from shading. *Computer Vision Graphics and Image Processing: Image Underst.*, 53(2):129–153, 1991. 2
- [32] H. Tagare and R. de Figueiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *PAMI*, 13(2):133–152, 1991. 2
- [33] R. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980. 2
- [34] R. Woodham, Y. Iwahori, and R. Barman. Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength. In *University of British Columbia*, 1991. 2
- [35] L. Zhang, B. Curless, and S. Seitz. Spacetime stereo: Shape recovery for dynamic scenes. In *Proceedings of CVPR*, pages 367–374, 2003. 2
- [36] L. Zhang and S. K. Nayar. Projection defocus analysis for scene capture and image display. *Proceedings of ACM Siggraph*, 25(3):907–915, 2006. 2
- [37] R. Zhang, P. Tsai, J. Cryer, and M. Shah. Shape from shading: A survey. *PAMI*, 21(8):690–706, 1999. 2
- [38] T. Zickler, P. Belhumeur, and D. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. In *Proceedings of ECCV*, pages 869–884, 2002. 2
- [39] T. Zickler, D. K. J. Ho, J. Ponce, and P. Belhumeur. Binocular helmholtz stereopsis. In *Proceedings of ICCV*, pages 1411–1417, 2003. 2