

## Depth from Sliding Projections

Peer-reviewed author version

HERMANS, Chris; FRANCKEN, Yannick; CUYPERS, Tom & BEKAERT, Philippe (2009) Depth from Sliding Projections. In: CVPR: 2009 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, VOLS 1-4. p. 1865-1872..

Handle: <http://hdl.handle.net/1942/10357>

# Depth from Sliding Projections

Chris Hermans    Yannick Francken    Tom Cuypers    Philippe Bekaert  
Hasselt University - tUL - IBBT  
Expertise Centre for Digital Media, Belgium  
{firstname.lastname}@uhasselt.be

## Abstract

*In this paper we present a novel method for 3D structure acquisition, based on structured light. Unlike classical structured light methods, in which a static projector illuminates a scene with time-varying illumination patterns, our technique makes use of a moving projector emitting a static striped illumination pattern. This projector is translated at a constant velocity, in the direction of the projector's horizontal axis. Illuminating the object in this manner allows us to perform a per pixel analysis, in which we decompose the recorded illumination sequence into a corresponding set of frequency components. The dominant frequency in this set can be directly converted into a corresponding depth value. This per pixel analysis allows us to preserve sharp edges in the depth image. Unlike classical structured light methods, the quality of our results is not limited by projector or camera resolution, but is solely dependent on the temporal sampling density of the captured image sequence. Additional benefits include a significant robustness against common problems encountered with structured light methods, such as occlusions, specular reflections, subsurface scattering, interreflections, and to a certain extent projector defocus.*

## 1. Introduction

The digitization of physical entities has become increasingly important in the last few decades. More and more often 3D scans of physical objects are required for use in computer games, CAD/CAM systems, 3D-TV and many other applications. Throughout the years, many different camera based shape acquisition techniques have been proposed, all having their advantages and disadvantages.

A very well investigated group of algorithms employs multiple cameras in order to obtain depth from matching corresponding features [22, 24]. Due to typical correspondence problems that occur in the presence of sharp edges, specular highlights, uniformly colored regions, etc., new techniques needed to be developed in order to increase the robustness of this process.

In indoor scenes, where the illumination can be controlled, a commonly used method to overcome the correspondence problem is structured light. One of the cameras is replaced by a projector, creating easily detectable features [20, 6, 4, 23]. These methods perform very well for purely Lambertian materials, but in practice the majority of materials have more complex reflectance properties, e.g. the specular reflections on a sheet of plastic or the subsurface scattering in organic materials such as fruit or human skin. This complex reflectance behavior leads to mismatches, resulting in erroneous depth values. In our work we only assume the presence of a reasonable amount of Lambertian reflection, neglecting specular reflections and diffusion due to scattering, as these phenomena do not affect the principal per pixel frequency of the shifted stripe projections [8, 9]. Correct depth values are obtained because they can be related to frequencies on a local per pixel basis. In a similar fashion, our method is robust against interreflections because we use high frequency patterns that are smoothed out by higher order reflections [19].

Another issue with current approaches is the limited depth-of-field of off-the-shelf projectors, narrowing down the available scanning volume for traditional scanners [17]. As the detection of the principal frequency of our projected stripe pattern is rather insensitive to blur from projection defocus, the scanable volume is extended. Note that the only requirement for our projector is the ability to project a high-frequency stripe pattern. This opens up the possibility to use cheap custom-made hardware, e.g. a box containing an LED behind a slide. This way a number of disadvantages of standard light projectors such as light spill, shallow depth of defocus, clunky mass and high price can be reduced [21].

To summarize, the main contributions of our method are:

1. Insensitivity to non-Lambertian reflections
2. Insensitivity to interreflections
3. Reduced sensitivity to occlusions
4. Sharp edge preservation
5. Extended scanable volume

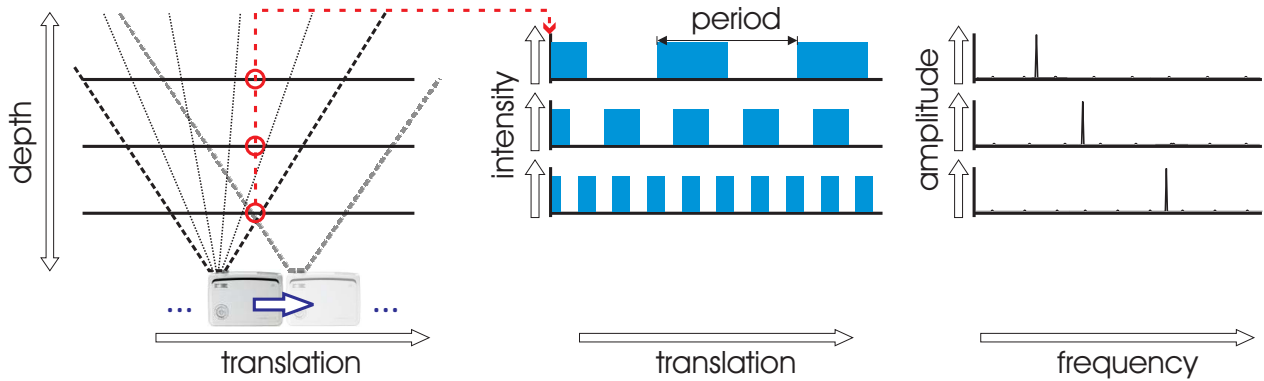


Figure 1. Method overview: a projector emitting a stripe pattern is translated according to its horizontal axis (left). As a result, scene points (indicated by red circles) at different depths  $\{d, 2d, 3d\}$  will be illuminated with stripe patterns with different periods  $\{p, 2p, 3p\}$  (center). Converting the observed changes in intensity over time into an equivalent frequency domain representation (right) will allow us to find the dominant frequency, which can be directly converted into a corresponding depth value.

## 2. Related Work

In this section we provide a brief overview of the state-of-the-art of camera based shape acquisition methods. First, we distinguish between two categories of acquisition methods: passive and active. Passive techniques do not alter the scene that is being captured, while active methods influence the scene by controlling the illumination. Our method belongs to the latter category.

We start our overview by discussing the passive methods. These methods typically match features observed by two or more cameras [22, 24], a moving camera with a known camera path [5], or a moving camera with an unknown camera path [25], estimating the depth information for each set of matched feature points. There exist many different implementations of these (multi-view) stereo algorithms, typically suffering from a number of problems caused by feature matching difficulties. Typical difficulties for this class of algorithms include uniform regions, depth discontinuities, specular highlights and other more complex BRDFs. Interesting surveys on this material are provided by Scharstein *et al.* [22] and Seitz *et al.* [24].

The active methods also make use of one or more cameras, but will also control the lighting in the scene to gain extra information from the captured reflections. Seminal work in this field was done by Horn [13] and Woodham [27] who reconstructed local surface orientation for diffuse surfaces illuminated by point light sources. More in general, these photometric stereo approaches try to reconstruct surface orientations from the recorded images, requiring additional illumination and reflectance information [16, 12, 11, 15]. Reconstructing the surface by integrating the estimated local surface orientation often proves to be a hard problem. In contrast, our method avoids this step, as we directly estimate the depth information.

In another slightly different approach for multi-camera setups, texture information is projected onto the scene, augmenting the original scene with new distinct features that can be used for robust feature matching, even in previously textureless regions [28]. We can take this approach a step further by using the controlled illumination in the scene to spatially label the volume containing the scene, once again using projected light instead of point/area light sources. These techniques, referred to as structured light approaches [23], are well understood for Lambertian surfaces. Hence, many spatial labeling codes already exist [6, 4], e.g. binary codes [20], Gray codes [23], phase shifting patterns [14, 30, 8, 9], and ratios [7]. Throughout the years, many improvements and extensions have been proposed: scanning of dynamic scenes, by taking into account both spatial and temporal information [10], scanning translucent materials [8], suppressing specular highlights by cross-polarization [26] or removing interreflections [19]. Although our method also employs a camera and a projector for controlled illumination, our technique fundamentally differs from the techniques mentioned above, as our depth estimation is based purely on frequency analysis. As such, we do not suffer from difficulties associated with spatial labeling or feature matching.

Other structured light methods exist where the depth is measured with respect to the projector instead of the camera. As shown by Zhang [29] and Moreno-Noguer [18], depth information can be retrieved by observing the amount of defocus introduced by the projector, focused in front of the scene. These techniques perform very well on Lambertian and slightly specular materials. However, since translucent materials tend to blur sharp projections, introducing additional defocus which cannot be differentiated from projector defocus, these methods cannot deal with these types of materials.

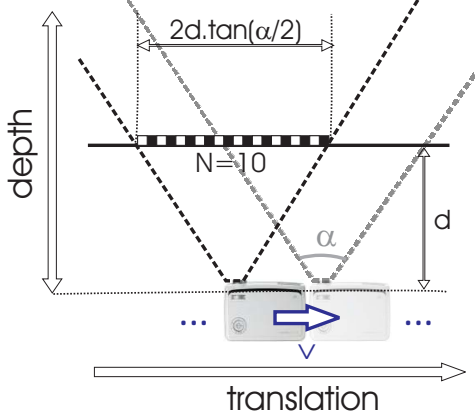


Figure 2. A linear relationship exists between the depth  $d$  (in meters) of the scene points and the period  $p$  (in frames,  $p = \frac{1}{f}$ ) of their observed wave patterns. If a projector with a fov of  $\alpha$  radians, translating at  $v$  meters per second, emits  $N$  stripes (waves), and the camera captures these at a rate of  $r$  frames per second, the distance traveled (in meters) by a wave in a single period at depth  $d$  is equal to  $\frac{v \cdot p}{r}$ . Using basic trigonometry, we can also show this to be equal to  $\frac{2d \tan(\alpha/2)}{N}$ , which leads to a uniquely defined linear relationship between depth and observed frequency.

Liao *et al.* [17] exploit the quadratic light attenuation from a scene translated with respect to the light source. Our method also measures depth with respect to a projector instead of camera. Due to its insensitivity to specular outliers and projector defocus, our method can accurately scan translucent and specular material, in addition to Lambertian surfaces.

In our work, we use a translating light source projecting black/white stripe patterns. Depth can then robustly be recovered for individual pixels from frequency analysis. Since the depth sampling is mainly dependent on controllable quantities, such as the projector’s translation speed, the spatial resolution of the stripe pattern and the camera’s capturing speed, a dense depth sampling can be obtained. In addition, the method preserves sharp edges and performs well in the presence of occlusions. Although Lambertian surfaces are assumed, in practice specular highlights and subsurface scattering have an insignificant impact on our frequency analysis.

### 3. Our Approach

Our method produces a depth estimate for each individual pixel, based solely on the intensities observed for that particular pixel. As no additional spatial information is used, our algorithm lends itself well to parallelization.

## 3.1. Theory

### 3.1.1 Basic Concept

An overview of our approach is shown in Figure 1. In the left image we see a projector translating at a constant velocity, in the direction of the projector’s horizontal axis. If we project a stripe pattern onto the scene, the illumination arriving at a random scene point will be dependent on the distance between the point and the principal plane of the projector. During the rest of this paper, when we refer to the depth of an observed scene point, we are talking about this distance. It is important to note that the depth of a scene point is defined in terms of the projector, and not in terms of the camera.

If we take a closer look at the illumination arriving at the different scene points in Figure 1, it becomes clear that there exists a linear relationship between the depth of the points and the period of their observed wave patterns. Converting the observed changes in intensity over time into an equivalent frequency domain representation will allow us to find the dominant frequency  $f$ , which can be directly converted into a corresponding depth value  $d$ :

$$d = \frac{s}{f} \quad (1)$$

The scaling factor  $s$  is uniquely defined by the parameters of the projector (translating at  $v$  meters per second, emitting a pattern of  $N$  stripes, using a field-of-view of  $\alpha$  radians), and the camera (recording at  $r$  frames per second):

$$s = \frac{N \cdot v}{2 \tan(\alpha/2) \cdot r} \quad (2)$$

Equations (1) and (2) are illustrated and explained in more detail in Figure 2.

### 3.1.2 Scene Properties

In this section we will describe the influence of the different scene properties on the wave pattern of the temporal intensity function  $I(t)$ . These observed intensities are the result of a moving stripe pattern behind a moving point light source (the intensities  $P(t)$  emitted by the projector), multiplied with the visible part of the reflectance function  $R(t)$ :

$$I(t) = P(t) \cdot R(t) \quad (3)$$

If we manage to estimate the underlying reflectance function  $R(t)$ , multiplying  $I(t)$  with the function  $\frac{1}{R(t)}$  removes its influence on the signal. What remains is an approximately sinusoidal pattern  $\tilde{P}(t)$  from which we can accurately determine the depth. This is illustrated with several examples in Figure 3.

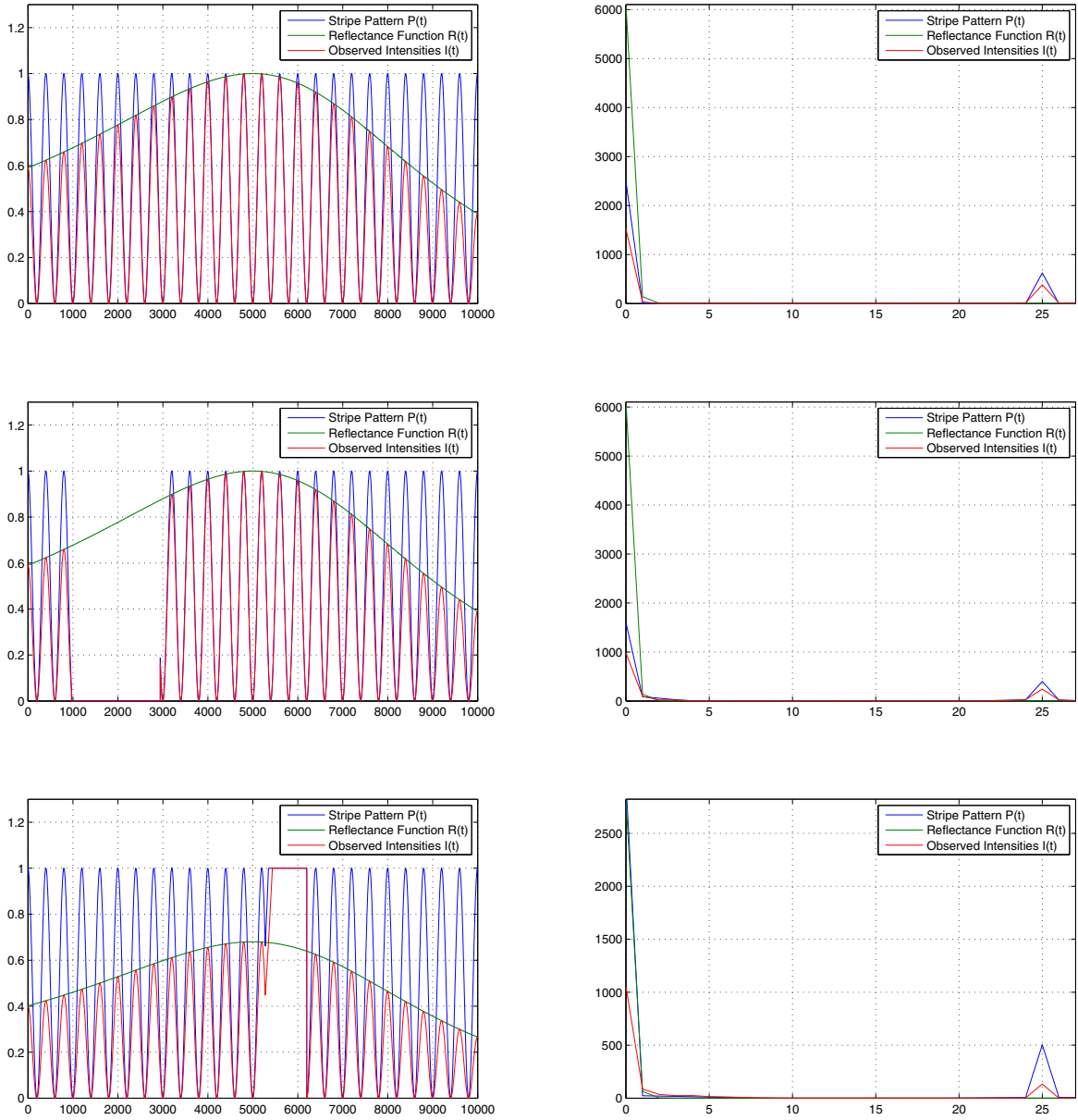


Figure 3. The observed intensities  $I(t)$  are the product of a wave pattern  $P(t)$  emitted by the projector and the visible part of the reflectance function  $R(t)$ . Several material and scene properties can affect the observed intensities. We illustrate three theoretical examples: (a) a diffuse material without occlusions, (b) a diffuse material, with occlusions, (c) a specular highlight. In the left column, we see the three temporal functions  $I(t)$ ,  $P(t)$  and  $R(t)$ . In the right column, we show the corresponding Fourier power spectra.

**Influence of the Reflectance Function** We shall now illustrate that in practical examples, it is not always necessary to actually estimate  $R(t)$  in order to recover a good depth estimate. Due to the convolution theorem, which states that a multiplication ( $\cdot$ ) in the time-domain is equal to a convolution ( $\otimes$ ) in the frequency domain, equation (3) is equal to:

$$\mathfrak{F}[I(t)] = \mathfrak{F}[P(t)] \otimes \mathfrak{F}[R(t)] \quad (4)$$

where  $\mathfrak{F}$  indicates the Fourier transformed of the original function. In this form, assuming  $P(t)$  is a perfect sinusoidal wave pattern, we can see that  $\mathfrak{F}[P(t)]$  is the sum of an impulse function with  $\delta$  equal to zero, and an impulse function with  $\delta$  equal to the dominant frequency  $f$  of the stripe pattern. As it is this  $f$  we are interested in, we neglect the lower end of the frequency spectrum in order to avoid false positives.

For most materials, the reflectance function  $\mathfrak{F}[R(t)]$  consist solely of a few low frequency components. Because of this,  $f$  will both be the dominant frequency in  $\mathfrak{F}[P(t)]$  and  $\mathfrak{F}[I(t)]$ . Only when  $\mathfrak{F}[R(t)]$  takes on more exotic shapes, e.g. the reflectance function of translucent materials,  $\mathfrak{F}[R(t)]$  might amplify low-amplitude noise to a such degree that  $f$  is no longer the dominant frequency.

**Scene Properties** Because of projector occlusions, not every point in the scene is illuminated with the stripe pattern for the entire duration of the recording. This 'incomplete' stripe pattern  $P(t)$  will still produce our familiar  $\mathfrak{F}[P(t)]$ , with the addition of a few minor low frequency components due to the missing stripes (Figure 3b, blue curve). In a similar matter, specular highlights can be interpreted as occlusions with a maximal intensity value instead of a minimal one (Figure 3c, blue curve). Once again, a few low frequency components are introduced to  $\mathfrak{F}[P(t)]$ , but as in the previous case these highlights on their own are insufficient to influence the detection of the dominant frequency  $f$ . Only when  $\mathfrak{F}[R(t)]$  shows irregular behaviour (and we do not compensate for it by a prior estimation of  $R(t)$ ), this low-amplitude noise can result in a deterioration of the results.

**Interreflections and Stray Light** Eq. 3 assumes an ideal scenario, in which each point in the scene is independent of the rest of the scene. In reality however, additional terms should be added to this equation, describing the intensities related to interreflections and subsurface scattering. However, these terms are relatively small compared to the main term described in eq. 3, and we therefor ignored them in our explanation.

Interreflections (second order reflections produced by another stripe pattern with another dominant frequency) are a source of noise in the frequency domain. However, the reflected intensities from second order reflections are significantly lower than their directly illuminated counterparts. The reason for this is twofold: (a) the distance traveled by light of second order reflections is greater than the distance traveled by light from direct illumination, and (b) light reflected by a surface is only a fraction of the light that arrived at the surface. Also, this noise is not accumulated in a single frequency, but spread out across a large set of frequencies. This is due to the fact that the incoming light angle is different for each timestep.

In a similar way, intensities arriving by subsurface scattering should not be able influence the detection of  $f$ , as the amount of noise is marginal with respect to the frequencies produced by direct illumination, and spread out among all frequencies.

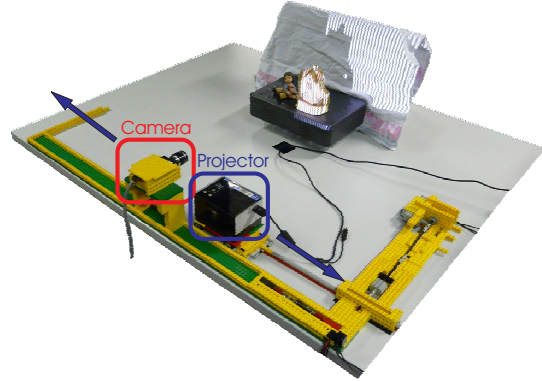


Figure 4. A picture of our setup: a projector is translated in the direction of its horizontal axis, while a static camera observes the scene.

### 3.1.3 Projector depth vs. camera depth

We have defined the depth  $\mu_i$  of an observed scene point  $i$  as its distance to the principal plane of the translating projector. This is a very different depth than the conventional camera pixel's depth  $\lambda_i$ . How are these two related, and when can we derive one from the other?

The first thing to note is that unless any initial information is known about the principal plane of the sliding projector, we cannot derive the set  $\{\lambda_i\}$  from  $\{\mu_i\}$ , because the number of unknowns will always outnumber the number of equations by 4 plane parameters. On the other hand, if these plane coordinates  $(a, b, c, d)$  with respect to the camera are known, the camera pixel depths  $\lambda_i$  are uniquely defined by the point-plane distance equation:

$$\lambda_i = \frac{\mu_i \cdot \sqrt{a^2 + b^2 + c^2} - d}{ax_i + by_i + c} \quad (5)$$

where  $(x_i, y_i)$  represents the pixel's normalized camera coordinates, assuming the camera's intrinsic parameters are known.

In practice, this means that in order to obtain camera depth values instead of projector depth values, we have to (a) align the camera's principal plane with that of the sliding projector, or (b) align the *orthonormal* camera's principal plane with a plane parallel to the principal plane of the sliding projector.

## 3.2. Practical Considerations

In the discussion above, we have always implicitly assumed a continuous function  $I(t)$ . In practice however, we are limited to a discrete sampling of this function, as well as a discrete sampling of its frequency domain (obtained using the Discrete Fourier Transform). As such we have to take several considerations into account if we want to maximize our depth resolution.

**Spatial Resolution of our Stripe Pattern** The spatial resolution of our pattern needs to be maximized, under the following constraints:

- Projector defocus has to stay within reasonable limits for the depth range of the scanned scene.
- Projector translation speed has to be low enough, in order to stay above the Nyquist sampling frequency: we do not want to miss a stripe in our projected pattern, at any depth.

**Translation Speed & Acquisition Speed** The speed of the projector translation needs to be minimized and constant, as it is directly responsible for the sampling frequency of the observed scene. Halving the speed of the projector results in doubling the number of captured frames over the same scanning volume, doubling the sampling frequency. Increasing the capturing speed of our camera results in a similar improvement.

**Projection Defocus** Adding projector defocus to equation (3) introduces a smoothing kernel, depending on the points depth, to the wave pattern  $P(t)$ . As a result, the amplitude of the dominant frequency  $f$  in  $\mathfrak{F}[P(t)]$  becomes less pronounced as the level of blur increases. At a certain point, we will no longer be able to distinguish between the amplitude of the dominant frequency and random noise. Thus, projection defocus poses a physical limit on the range of depths we can measure using the described technique.

## 4. Discussion

### 4.1. Setup

An image of our setup can be found in Figure 4. The bulk of our setup is made using off-the-shelf Lego Technic [2]. A light-weight pocket projector (Samsung Pocket Imager SP-P310ME Projector) is placed on a platform. This platform is translated in the direction of the projector’s horizontal axis, by applying rotational movement on the rail attached to the platform. Activating the motor on the right results in a slow translational movement in the direction of choice. A single camera (Point Grey Research Grasshopper [3]) is encapsulated into our Lego [2] framework, observing the scene. Our results were generated with variable camera resolutions at 60 frames per second. We used stripe patterns with variable stripe sizes, varying from 2 to 8 pixels, depending on the required depth range. The native resolution of the used projector was  $800 \times 600$ .

### 4.2. Results

Results from our method are shown in Figure 5. Each row contains information of the same scene. In the first column, we show a color image of the maximal intensity observed at each pixel. The second column shows a depth

map produced by our system. The third column provides a magnified zoom, focusing on certain details provided in the second column. And finally, in the fourth column we rerender the original scene in a variety of ways.

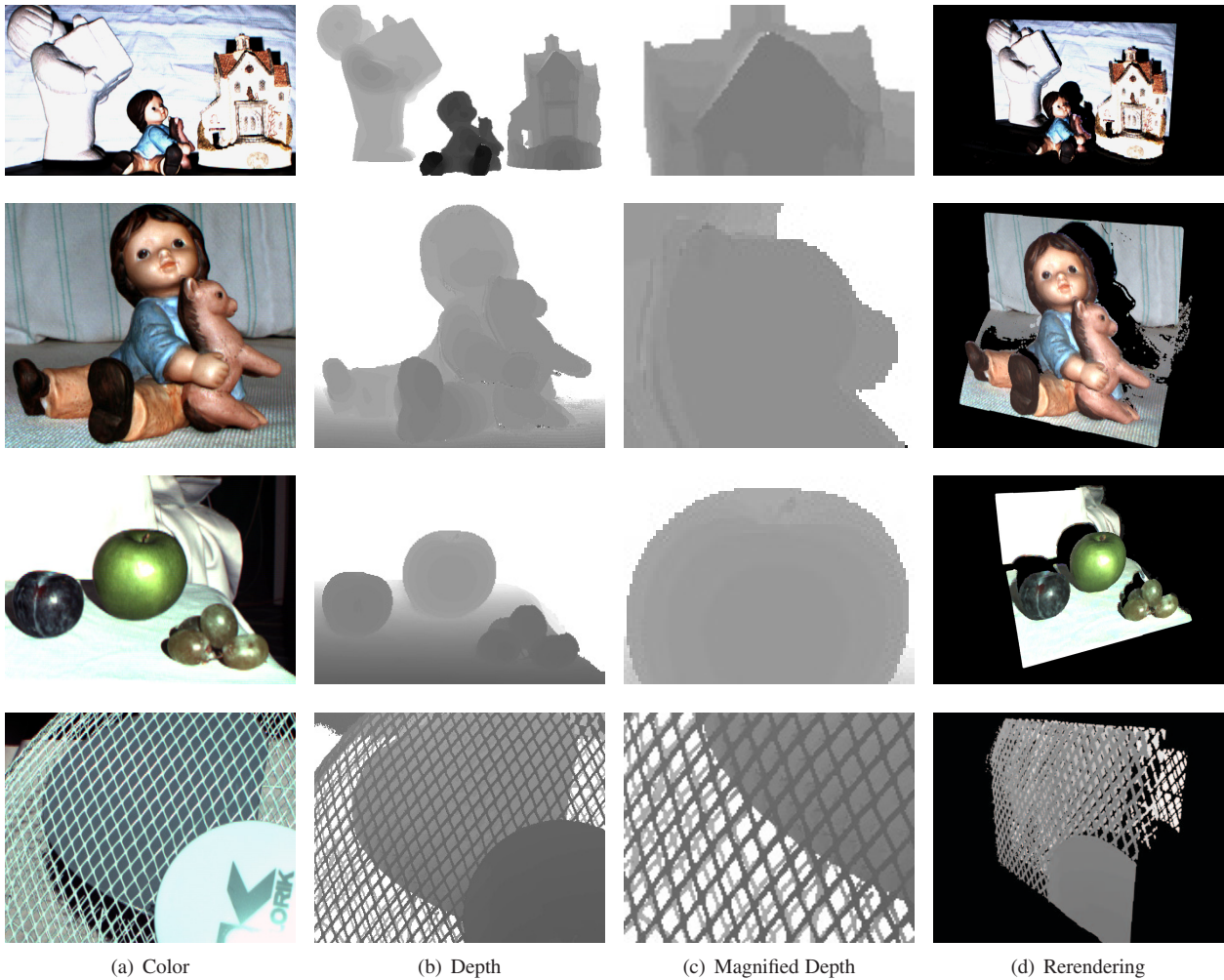
The first two rows show a dense scanning of approximately Lambertian surfaces. In the first example we have used a wide stripe pattern and a high resolution, in order to capture a large range of depths. The second example shows a close-up scan of the middle figurine, in which we applied a finer stripe pattern. Because of the higher spatial resolution, our depth planes are now concentrated in a smaller depth range. The third row shows a selection of fruits, objects with strong specular and translucent properties. Finally, we end with the fourth row, showing a fan segment with a very fine grid structure. This example illustrates our methods ability to preserve sharp edges, and cope with specular materials. For a comparison of our technique against a ground truth, we refer to Figure 6, in which we compare our obtained depth values against the values returned by our synthetic dataset. This data was generated with Blender [1] using a simple diffuse shading model, and consisted of 2000 frames. We placed the projector above the projector translation trajectory, and aligned its principal plane with that of the projector.

Processing the results shown in our paper took about 5 to 10 minutes on average. The bulk of this processing time was allocated to hard drive access, and not due to computational complexity.

## 5. Conclusions and Future Work

In this paper we have presented a novel dense 3D shape acquisition method, based on the per pixel frequency analysis of a scene illuminated with translating stripe patterns. Because material properties have a limited impact on the recovery of the dominant frequency, our algorithm works well for a wide variety of scenes, including those with materials where classical structured light methods commonly have problems.

In the near future, we will investigate several extensions to our approach. First, we will look into the use of a low-cost alternative for the projector in our setup, employing a ‘pinhole projector’: a box with a LED inside, and a grid in front of the box. Alternatively, we will explore the use of a laser stripe pattern in order to avoid problems with the limited depth range due to defocus. Also, the high intensities allow for faster capturing rates, which in turn increase our sampling rate. We will also be looking into different configurations for our system, such as extensions to 2D translations or a turntable configuration. Finally, we will investigate the use of more cameras or sliding projectors, in order to avoid the specific camera-projector configurations required to convert projector depths into the more common camera depths.



(a) Color

(b) Depth

(c) Magnified Depth

(d) Rerendering

Figure 5. An example of results produced by our method. We only display depth values that were present within our predefined depth range. If no depth could be assigned to a pixel, it was assigned the value of the depth plane at the largest distance from the projector.

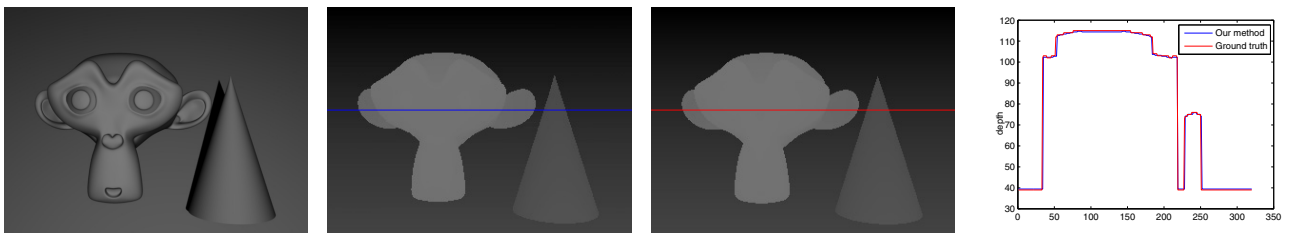


Figure 6. Comparing our results on a synthetic dataset with ground truth: in the left image we have our result, the center image shows the ground truth, and the right image shows a plot of a scanline at the same vertical coordinates in both images.

## Acknowledgements

The authors acknowledge financial support by the ERDF (European Regional Development Fund), the European Commission (FP7 IP "2020 3D media") and the Flemish government. Furthermore, we would like to thank our colleagues and reviewers for their useful comments and suggestions.

## References

- [1] Blender. <http://www.blender.org/>.
- [2] Lego. <http://www.lego.com/>.
- [3] Point grey research inc. <http://www.ptgrey.com/>.
- [4] J. Batlle, E. Mouaddib, and J. Salvi. Recent progress in coded structured light as a technique to solve the correspondence problem: a survey. *Pattern Recognition*, 31(7):963–982, jul 1998.



- [5] R. Bolles, H. Baker, and D. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *IJCV*, 1(1):7–55, 1987.
- [6] K. L. Boyer and A. C. Kak. Color-encoded structured light for rapid active ranging. *PAMI*, 9(1):14–28, 1987.
- [7] B. Carrihill and R. Hummel. Experiments with the intensity ratio data sensor. 32(3):337–358, December 1985.
- [8] T. Chen, H. P. A. Lensch, C. Fuchs, and H.-P. Seidel. Polarization and phase-shifting for 3d scanning of translucent objects. In *Proceedings of CVPR*, 2007.
- [9] T. Chen, H. Seidel, and H. Lensch. Modulated phase-shifting for 3D scanning. *Proceedings of CVPR*, pages 1–8, June 2008.
- [10] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz. Spacetime Stereo: A Unifying Framework for Depth from Triangulation. *PAMI*, pages 296–302, 2005.
- [11] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *Proceedings of ICCV*, volume 1, pages 341–348, 2005.
- [12] A. Hertzmann. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *PAMI*, 27(8):1254–1264, 2005. Member-Steven M. Seitz.
- [13] B. Horn. Obtaining shape from shading information. In *The Psychology of Computer Vision*, pages 115–155, McGraw-Hill, New York, USA, 1975.
- [14] P. Huang, C. Zhang, and F. Chiang. High-speed 3-D shape measurement based on digital fringe projection. *Optical Engineering*, 42:163, 2002.
- [15] I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich. State of the art in transparent and specular object reconstruction. In *STAR Proceedings of Eurographics*, pages 87–108, 2008.
- [16] K. Ikeuchi. Determining surface orientation of specular surfaces by using the photometric stereo method. *PAMI*, 3(6):661–669, nov 1981.
- [17] M. Liao, L. Wang, R. Yang, and M. Gong. Light fall-off stereo. In *Proceedings of CVPR*, pages 1–8, 2007.
- [18] F. Moreno-Noguer, P. Belhumeur, and S. Nayar. Active Refocusing of Images and Videos. *Proceedings of SIGGRAPH*, Aug 2007.
- [19] S. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast Separation of Direct and Global Components of a Scene using High Frequency Illumination. *Proceedings in SIGGRAPH*, Jul 2006.
- [20] J. Posdamer and M. Altschuler. Surface measurement by space-encoded projected beam systems. *Computer Graphics and Image Processing*, 18(1):1–17, 1982.
- [21] R. Raskar, H. Nii, B. deDecker, Y. Hashimoto, J. Summet, D. Moore, Y. Zhao, J. Westhues, P. Dietz, J. Barnwell, S. Nayar, M. Inami, P. Bekaert, M. Noland, V. Branzoi, and E. Bruns. Prakash: lighting aware motion capture using photosensing markers and multiplexed illuminators. *Proceedings of SIGGRAPH*, 26(3):36, 2007.
- [22] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1-3):7–42, 2002.
- [23] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proceedings of CVPR*, volume 1, page 195, 2003.
- [24] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proceedings of CVPR*, pages 519–528, Washington, DC, USA, 2006. IEEE Computer Society.
- [25] P. F. Sturm and B. Triggs. A factorization based algorithm for multi-image projective structure and motion. In *Proceedings of the 4th European Conference on Computer Vision-Volume II (ECCV '96)*, pages 709–720, London, UK, 1996. Springer-Verlag.
- [26] L. Wolff. Using polarization to separate reflection components. In *Proceedings of CVPR*, pages 363–369, 1989.
- [27] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, jan/feb 1980.
- [28] M. Young, E. Beeson, J. Davis, S. Rusinkiewicz, and R. Ramamoorthi. Viewpoint-coded structured light. In *Proceedings of CVPR*, jun 2007.
- [29] L. Zhang and S. K. Nayar. Projection Defocus Analysis for Scene Capture and Image Display. *Proceedings of SIGGRAPH*, Jul 2006.
- [30] S. Zhang and S. Yau. High-resolution, real-time 3D absolute coordinate measurement based on a phase-shifting method. *Optics Express*, 14(7):2644–2649, 2006.