

Using Physically Based Rendering to Benchmark Structured Light Scanners

Esdras Medeiros¹, Harish Doraiswamy², Matthew Berger³ and Claudio T. Silva²

¹Universidade Federal do Ceará, ²New York University, ³Air Force Research Laboratory

Abstract

Structured light scanning is ubiquitous in 3D acquisition. It is capable of capturing high geometric detail at a low cost under a variety of challenging scene conditions. Recent methods have demonstrated robustness in the presence of artifacts due to global illumination, such as inter-reflections and sub-surface scattering, as well as imperfections caused by projector defocus. For comparing approaches, however, the quantitative evaluation of structured lighting schemes is hindered by the challenges in obtaining ground truth data, resulting in a poor understanding for these methods across a wide range of shapes, materials, and lighting configurations. In this paper, we present a benchmark to study the performance of structured lighting algorithms in the presence of errors caused due to the above properties of the scene. In order to do this, we construct a synthetic structured lighting scanner that uses advanced physically based rendering techniques to simulate the point cloud acquisition process. We show that, under conditions similar to that of a real scanner, our synthetic scanner replicates the same artifacts found in the output of a real scanner. Using this synthetic scanner, we perform a quantitative evaluation of four different structured lighting techniques – gray-code patterns, micro-phase shifting, ensemble codes, and unstructured light scanning. The evaluation, performed on a variety of scenes, demonstrate that no one method is capable of adequately handling all sources of error – each method is appropriate for addressing distinct sources of error.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Raytracing—I.4.1 [Digitization and Image Capture]: Reflectance—Scanning

1. Introduction

Three dimensional measurement of objects (3D scanning) is an ubiquitous technology because of its great variety of applications to the industrial and scientific communities, including reverse engineering, prototyping, archaeological documentation and industrial design. There has been much research during the last two decades to create better 3D scanners, a majority of which introduces key concepts that refine and advance the state of the art. See [Bla04, LT09, Gen11] for good survey papers on this subject.

Active stereo scanners, specifically those based on digital structured-lighting (SL) projectors, are widely used because of their speed and precision. Using known calibration objects like a checkerboard, the imaging properties of the camera and the projector may be estimated. This allows a single 3D line to be drawn from each of the device's center of projection through a 3D point. The intersection of these two lines is then used to recover the depth of the

point. To overcome the correspondence problem between the camera and projector pixels, SL scanners use multiple stripes that are projected simultaneously on the object. Popular techniques for pattern projection include binary coded patterns [SCL*97, GLT96, XA09, GAVN13], and phase shift patterns [CSL08, GN12].

Compared to passive methods and laser scanners, SL scanners are often more sensitive to surface material properties. This is especially true in the presence of global illumination, where inter-reflection due to strongly reflective materials and sub-surface scattering of translucent materials, often violate assumptions made by SL scanners. Additionally, the illumination defocus of the projector also causes errors during the correspondence process. Many of the above mentioned techniques work towards solving these issues.

Despite the vast amount of work in SL scanners, there are significant shortcomings in evaluating the performance of different correspondence techniques with various shapes

and material properties. Part of the difficulty in establishing a comprehensive set of experiments for such an evaluation is the large variability of shapes and material properties. These experiments typically operate on objects for which there does not exist a computational representation of the surface from which the point cloud was measured. The different SL scanning approaches therefore resort to acquiring ground truth manually from the captured images. Moreover, in addition to the difficulty in obtaining similar objects with different materials, it is also difficult and cumbersome to manually evaluate multiple objects. An exhaustive evaluation is therefore not feasible.

1.1. Contributions

In this paper, we comprehensively evaluate SL techniques in a synthetic manner using photo-realistic rendering techniques. A synthetic setting allows us to evaluate these algorithms over a wide range of controllable conditions in a realistic manner, allowing us to perform evaluation at a scale which previous techniques fail to address. We produce realistic patterns for arbitrary SL methods by simulating various global illumination phenomena. We evaluate SL techniques with respect to three common sources of error: projector defocus, interreflections, and subsurface scattering. For each type of error, we show that the projected synthesized patterns produced via photo-realistic rendering retain the properties of a real SL scanner, where we present a scheme to validate our model on real data.

We use the simulator to compare and benchmark state of the art correspondence algorithms, namely – 1) the popular Gray-code pattern [LT09], 2) using an ensemble of codes [GAVN13], 3) the micro phase shifting technique [GN12], and 4) unstructured lighting technique [CMR11]. The algorithms are evaluated for a variety of scenarios that represent different material and shape properties. A global illumination simulator provides us with fine-grained control in evaluating different properties of interest, where we evaluate these algorithms by varying intuitive parameters for each type of experiment – see Figure 1 for an overview of our approach. More specifically, we evaluate projector defocus by varying the aperture of the projector, we evaluate interreflections by varying the level of reflectivity, and we evaluate subsurface scattering by varying the distance light is permitted to travel in the shape’s interior. Furthermore, we also consider how these methods behave when a shape is undergoing a deformation from being convex to concave, assuming a fixed BRDF. This allows us to smoothly vary the amount of interreflections present in the scene.

Our experiments highlight a number of advantages and shortcomings of existing SL techniques. In particular, we demonstrate that no one method is capable of adequately handling all sources of error – each method is appropriate for addressing distinct sources of error. We will be making

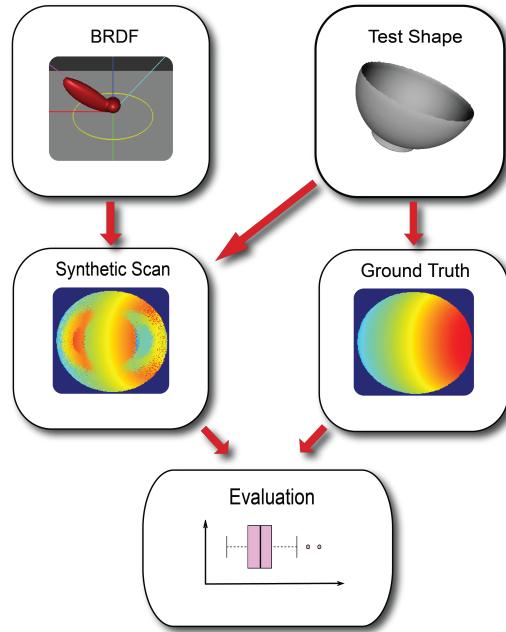


Figure 1: Overview of our evaluation scheme. A test shape along with the BRDF of its material is used by the synthetic scanner to output an image with decoded columns. This output from different correspondence algorithms are then compared with the synthesized ground truth in order to evaluate the algorithms’ performance.

the above benchmark public, so as to enable the comparisons of future algorithms with the current state of the art.

2. Related Work

There has been recent work on comparing stereo reconstruction algorithms, and creating benchmarks for them. Szeliski and Sabih [SZ00] compared the performance of multiple stereo algorithms. Scharstein and Szeliski [SS02] extended this comparison and provided an additional taxonomy of dense, two-frame stereo algorithms. In particular, they did this by assessing the different components and design decisions used in these algorithms.

Seitz et al. [SCD*06] provided a quantitative comparison of multi-view stereo reconstruction algorithms. In order to enable such a comparison of the algorithms, they introduced benchmark images registered with ground truth. They were the first to propose such a benchmark for stereo algorithms. Until this work, all comparisons were qualitative, and was performed using a physical set up, which was not only cumbersome and time-consuming, but was also not exhaustive.

More recently, Gruna and Irgenfried [GI12] developed a vision simulator to generate photo realistic images of a given scene for varying lighting conditions. This simulator was used to generating various images to be used for comparing

image processing algorithms. Berger et al. [BLN^{*}13] provided a benchmark for the evaluation and comparison of surface reconstruction algorithms. In order to do so, they simulated a triangulation-based laser scanner, and used a variety of surface meshes having different properties to obtain the point set that is used by the surface construction algorithms.

With respect to coding patterns used in structured light based scanners, the only comparison available is between the different patterns that are used. The survey presented in Battle et al. [BMS98] provided a qualitative comparison of the different coding patterns used in SL scanners. The survey by Salvi et al. [SPB04] also compares different coding patterns. However, they presented a quantitative analysis of the different patterns using a plain white surface. Additionally, they performed a qualitative comparison of the coding patterns using two surface meshes as input. Xu and Aliaga [XA09] study the performance of Gray code patterns with respect to inter-reflections, and use the analysis to design an iterative algorithm that repeatedly uses Gray code patterns to progressively obtain more accuracy. Salvi et al. [SFPL10], later again surveyed the different coding strategies used in SL scanners. In particular, they compared the quality of 3D point sets obtained using the different strategies. Gupta et al. [GAVN11] compared the effect of global illumination for different coding patterns with the aim of designing a new pattern that minimizes this effect.

All the above methods used a physical set up to perform the comparisons – a time and resource consuming process. As a result, the evaluation was restricted to very few objects and materials (usually one or two). Therefore, the effect of the materials and shape of the input object, if considered, was limited in these comparisons.

3. Background

In this section we first describe the different steps involved in the SL scanning process. The different artifacts that can be caused in this process is then discussed in Section 3.2. Finally, in Section 3.3 we briefly describe four different correspondence algorithms that we benchmark in this paper.

3.1. Scanning Process

The SL scanning pipeline can be summarized in the following steps: *image capture* → *correspondence* → *triangulation*. The *image capture* step is the physical part of the process in which light interacts with the objects in the scene until it reaches the CCD sensors. Depending on the correspondence scheme, a set of light patterns – usually some form of stripes – are projected onto the scene from a projector which is captured by the camera. The *correspondence* step processes the captured images and extracts the correspondence between each pixel of the image to a pixel of the projector. The final *triangulation* step recovers the scene

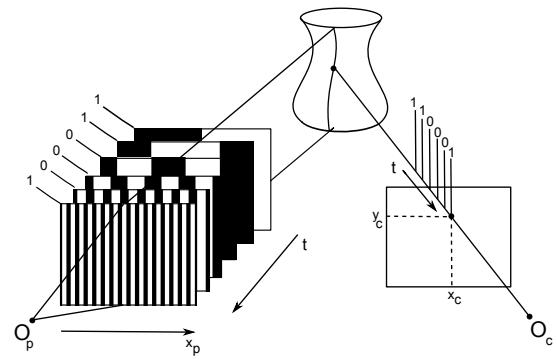


Figure 2: An illustration of the SL scanner using Gray-code patterns. Black and white interleaving striped patterns with increasing frequency is projected onto the given object and captured using a camera. The intersection of the lines passing through the optical center of the projector (O_p) and the camera (O_c), respectively, is used to recover the depth of the scene.

depth by intersecting two 3D lines emanating from the optical centres of the camera and projector respectively, that passes through each camera pixel and the associated pixel in the projector. When only vertical patterns are projected onto the object, the depth is obtained by intersecting rays with planes.

Figure 2 illustrates this process when the popular Gray code patterns (see Section 3.3) are used to perform the correspondence step. The physical step of this process, that of image capture, is prone to produce artifacts that could be caused due to both properties of the materials of the objects in the scene, as well as due to the projector. Different correspondence algorithms handle these artifacts in different ways. However, the final triangulation step is independent of the correspondence algorithm used.

3.2. Scanning Artifacts

Inaccuracies could occur in the images that capture the scene predominantly due to two causes – global illumination and illumination defocus. While the former is due to the material property of the objects in the scene, the latter is because of the aperture of the projector used. We now briefly discuss the artifacts that result because of these two properties.

3.2.1. Global illumination

The correspondence of a pixel in the captured image is obtained by analyzing the intensity of light on that pixel. However, in addition to the direct light illuminating an observed pixel, there also exists indirect (or global) light that illuminates it. This global light is due to inter-reflections and sub-surface scattering along the surface of an object. So, when the intensity of the illuminated pixel is predominantly due to indirect light rather than direct light, then the correspondence of that pixel is lost.

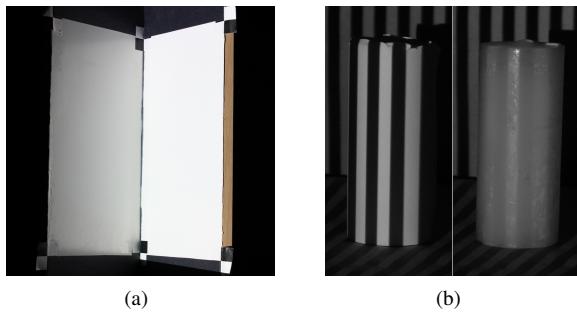


Figure 3: Artifacts due to global illumination. (a) Two adjacent planes in the form of a concave “v” shape. Even though the light is incident only on the right plane, due to strong inter-reflections the left plane is also lit. This results in incorrectly classifying pixels in the left plane. (b) The effect of subsurface scattering seen on a cylindrical candle. The candle on the left is covered with a white opaque surface. Notice the blurry effect caused by the subsurface scattering on the surface of the candle on the right.

Inter-reflections are common for objects having highly reflective surfaces (such as metals). Figure 3(a) shows an example where an unlit surface can be mistaken to be lit due to inter-reflections. Subsurface scattering is common for objects having translucent surfaces. Figure 3(b) shows an example of a candle where not only the intensity of lit regions are diminished, but those of the unlit regions are also enhanced due to the subsurface scattering of light within the candle.

3.2.2. Illumination defocus

Projectors usually have a limited depth of field due to their large apertures. Hence the patterns projected is usually focused patterns along a planar region. This results in blurred patterns being projected (and captured) on parts of the objects of the scene that are not along this plane. This could lead to incorrect classification of a pixel of the captured image. Figure 4 shows an instance of the blur captured due to illumination defocus. For a detailed analysis of the problem, see the paper by Zhang and Nayar [ZN06].

3.3. Correspondence Algorithms

We now briefly describe the four correspondence techniques that we benchmark in this paper. The detailed description of the techniques can be found in the corresponding papers.

3.3.1. Gray code patterns

In the Gray-code scheme [LT09], a set of patterns with black and white interleaving stripes are projected onto the scene. The interleaving frequency, which encodes the binary level, increases over time. An image processing operation, called

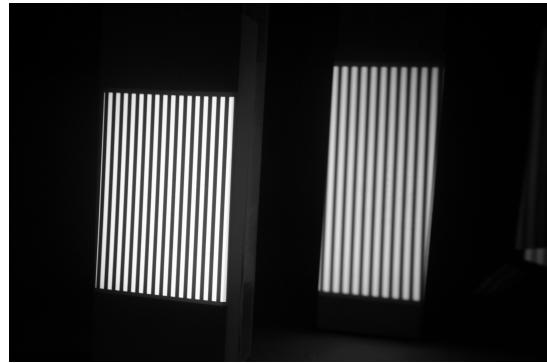


Figure 4: The illumination defocus effect of the projector. Two planes are placed one behind the other, with the focus of the projector on the front plane (left). Due to the large aperture of the projector, the projection of patterns on the back plane (right) is blurred.

Binary classification, is then performed on the captured images which transforms each pixel intensity of the image into a single bit. The correspondence between the camera and projector pixels is then performed through the concatenation of bits obtained from the above classification step.

3.3.2. Ensemble codes

The key idea [GAVN13] in this scheme is that errors made by different codes are nearly random. Therefore, if the depth values computed using two different codes are the same, then it can be assumed to be the correct value with high probability. Using this idea, a collection of four binary patterns, two catered to handle inter-reflection, and two catered to handle subsurface scattering, are projected onto the scene. Binary classification and pixel correspondences are then computed from the different images. If two of the depth values computed from the triangulation are within a small threshold, then this value is finally returned as the correct value.

3.3.3. Unstructured lighting scheme

The unstructured lighting scheme [CMR11] uses band-pass white noise patterns that are robust to errors caused due to inter-reflections. A number of random unstructured light patterns, N , are generated at a pre-selected band-pass frequency interval. These patterns are then projected one at a time. Each pixel corresponds to a N -dimensional vector in both the projected as well as the camera image. Pixel correspondences are then obtained by matching the N -dimensional vectors. The number of required projected images is much higher for this method compared to other methods.

3.3.4. Micro phase shifting

The micro phase shifting [GN12] technique projects sinusoidal patterns whose frequencies are limited to a narrow,

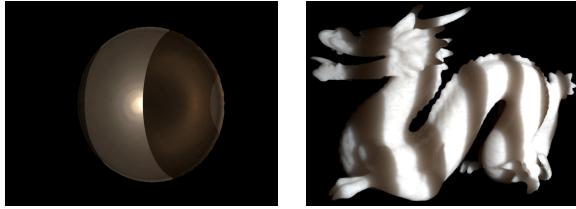


Figure 5: Simulating global illumination effects. **Left:** bowl model with inter-reflections. **Right:** dragon model with subsurface scattering.

high-frequency band. This technique was designed to handle not only global illumination errors, but also errors caused due to illumination defocus. These high frequency sinusoidal patterns produce a set of images over which the effects of both global illumination as well as illumination defocus is expected to remain constant for each point in the scene. Phase unwrapping is then performed to obtain the correspondences between the pixels.

4. The Synthetic Scanner

The synthetic scanner simulates the optical camera-projector system and the interaction of light with the objects in the scene until it reaches the CCD sensors. The simulator is implemented over PBRT, a physically based rendering software [PH] which computes images that are physically correct. An advantage of this software is that it can run on multiple cores via OpenMP [Ope], thus decreasing the rendering time.

Each range scan shares the following scene parameters, only differing in the projected patterns:

1. Camera x Projector devices:
 - a. Intrinsic parameters – field of view, aperture size, focal plane distance and resolution.
 - b. Extrinsic parameters – Position and orientation.
2. Model: Material and Geometry.

Camera and projector setup. PBRT returns a high dynamic range image which is converted to a 8-bit image in gray scale. The resolution used for the images in our simulations is 1024×768 . The camera and projector are setup along a field of view of 60° at a distance such that the scene's bounding box is contained in the view of both. The angle between camera and projector is set to 30° . We use a fixed brightness for the projector.

Simulating global illumination. The material of the model is provided using its BRDF which is used by the rendering algorithm to accurately simulate the scattering effects due to inter-reflections. The variant of the ray-tracing algorithm used to render the scan images is the Metropolis Light Transport [VG97] algorithm. Since this algorithm is unbiased, the obtained images accurately reflect the lighting as it would be

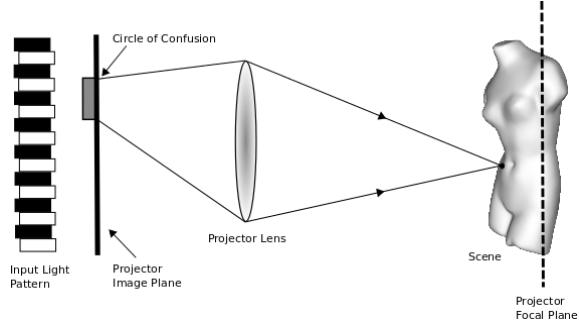


Figure 6: Illumination defocus of the projector due to circle of confusion.

in a real world scenario. PBRT also supports the rendering of the sub-surface scattering effects using BSSRDF [PH00] of the material. The algorithm implemented in PBRT employs the dipole approximation for subsurface scattering developed by Farrell et al. [FPW92] and introduced to computer graphics by Jensen et al. [JMLH01]. Figure 5 illustrates the rendering of both effects.

Validating global illumination effects. Due to the complex nature of global illumination, it is not obvious that PBRT is capable of producing images consistent with real data acquisition. We therefore have to validate the accuracy of the simulation. We accomplish this by using the results of the binary classification step of the scanning process. In particular, we compare the errors from binary classification that occur from the output of the a real scanner to that of our synthetic scanner. As shown by Nayar et al. [NKGR06], this validation is sufficient since it can also be generalized to support artifacts caused due to phase shift patterns. Due to lack of space, the details of this validation are provided in the supplementary material.

Simulating illumination defocus. Consider the schematic of a scene illuminated using the projector [ZN06], as shown in Figure 6. The circle of confusion caused due to the aperture is responsible for the blurry effect. This is used to derive the illumination defocus effect of the projector in our synthetic scanner. A surface point in the scene is lit by a subset of the light field [LH96], more specifically, the set of refracted rays emitted by the projector image. Similar to the depth of field implementation in cameras [CPC84], a random sample of points in the lens of the projector is used to simulate this effect for each pixel of the camera. The refracted light is computed from these points and the corresponding pixel intensity of the projector image is stored (see Appendix for a formal derivation). The final pixel intensity is an integration of these irradiances.

Figure 7 shows the effect of defocus when applied to a model scene. Since the simulation of the defocus effect is physically based on the optical properties of the projector, in particular its aperture, it depends neither on the scene objects nor on their material properties. Therefore its accuracy can

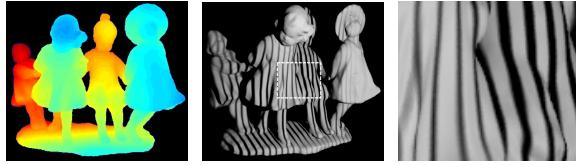


Figure 7: Simulating illumination defocus effect of the projector. (left) Depth of the scene as seen by the projector. The focal plane is placed halfway between the maximum and minimum depth; (middle) final scene as seen by the camera; (right) details of the rectangle highlighted in the middle image.

be checked by simply following the convolution property as described in Zhang and Nayar [ZN06]. Indeed, we observe that each pixel in the blurred image (Figure 7(middle)) is the result of a low pass filter of the original pattern with a kernel whose support depends on the depth of the pixel.

5. Benchmark

In this section we present a benchmark of structured lighting algorithms resulting from our experiments. In addition to this material, we will also be making public the benchmark through a web page containing all the results and the rendered images of all the datasets used in this paper.

Experimental setup. Our experimental results are broken down into three parts:

1. Projector defocus: This experiment was parameterized by the camera aperture radius. This radius is expressed in the normalized z -coordinate of the projector reference frame, where the distance from the projector image plane to the center of lens is 1. The focal plane is placed halfway in between the maximum and minimum depths.
2. Sub-surface scattering: The mean free path was used as a parameter for this experiment. The mean free path is a parameter available in PBRT that controls the average distance that the light travels in the surface. This measure is given in meters. For reference purposes, the dragon model used in the experiments has a bounding box diagonal equal to 0.267m.
3. Inter-reflection: We use various materials with different bidirectional reflectance distribution functions (BRDF). We use the BRDF database from Matusik et al. [MPBM03] which is freely available [MER06]. Due to the availability of a wide number of materials in this database, we perform a pre-processing step to identify appropriate materials to be used for benchmarking the different techniques. We do this by first using the Gray code technique at a low rate of samples per pixel to render the scenes. We then compare the scenes with the ground truth and select a subset of 20 materials that result in the highest decoding errors (see Appendix E).

In order to obtain the best possible results, the set of

patterns to be projected onto the scene in case of the Microphase and Unstructured algorithms have to be carefully chosen. In particular, if the scene has very high-frequency inter-reflections, we use higher frequency patterns. If the scene has defocus or subsurface scattering and weak inter-reflections, we use lower frequency patterns.

Evaluation measures. For each experiment, we obtain an output described by a matrix whose elements are the indices of the corresponding columns of the projector image. We use this and compute three measures to evaluate and compare the different techniques.

1. **Error:** We compute the error as the absolute difference of the above matrix with the ground truth. The ground truth is obtained by repeating the experiment on the input shape having a Lambertian reflectance model and in which the indirect component of the light is removed. Because the projected patterns and the decodification process is specific to a given technique, the ground truth is computed for each of the techniques. From the error matrix we extract the mean and variance statistics.
2. **Valid Indices:** It is possible for the different algorithms to produce outputs containing invalid column indices. These are essentially pixels for which the algorithm is unable to make a decision on the correspondence of the pixel, and therefore does not produce a column match. Such invalid correspondences are therefore not considered when computing the error above. However, having a large number of such pixels would reflect poorly on the performance of an algorithm. We therefore report the percentage of valid indices resulting from the algorithms.
3. **Outliers:** We use the Chauvenet's criterion [Cha63] to remove outliers when computing the error measure, and report the number of outliers that are removed in each experiment. This criterion can also be used as a measure to check the robustness of the different algorithms.

5.1. Defocus

In order to evaluate the impact of projector defocus, we have taken a complex shape and generated projection patterns under varying projection aperture. We selected the dancing children model because it occupies a wide area when viewed in its upright pose and it consists of surface features of varying size and shape. Both properties are useful to highlight the defocus effect of the projector. In order to emphasize the impact of defocus, scenes are rendered only with direct illumination and a Lambertian material.

We present the results of this experiment in Figure 8. We observe that variations in the shape or orientation of the object can produce different results. Nevertheless, such changes are not significant because they reflect only variations of the ratio between the area of patches that are in-focus and the total area of the object. We have also verified this claim for a simple spherical object – see the supplementary material for details.

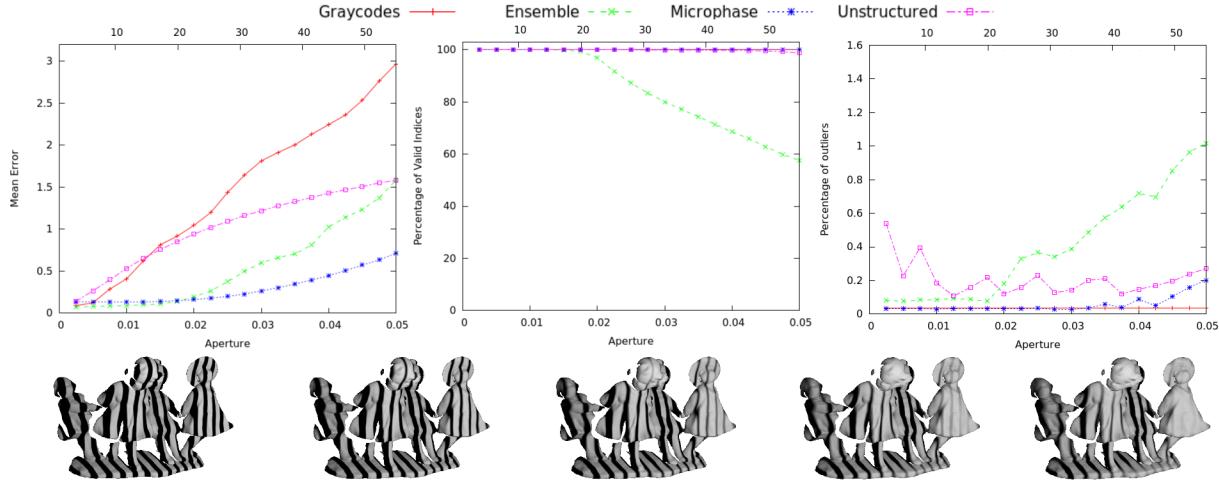


Figure 8: Defocus results parameterized by the aperture of the projector. On the bottom row, we show the artifact effect as the aperture increases. The axis on top of each plot denotes, in terms of pixel units, the size of the circle of confusion corresponding to point closest to the projector.

Note that Ensemble performs the best for small aperture in error, valid indices, and outliers, up to an aperture of approximately 0.02. Past this, however, it starts to take on a somewhat large number of outliers relative to the other methods, while Microphase supercedes it in all three evaluation measures. Although Unstructured overall produces the smallest number of outliers relative to Ensemble, its performance in mean error is overall much worse. Furthermore, even when Unstructured produces an equivalent number of outliers to these methods, it still contains higher error. This is indicative that methods which project stripes (be it binary or sinusoidal patterns) into the scene tend to combat the artifacts of defocus more effectively.

5.2. Sub-surface Scattering

To evaluate the effect of translucency in materials across the algorithms (see results in Figure 9) we employ the dragon model. This model contains varying levels of depth throughout the shape, resulting in spatially-varying translucency in the camera image. As mentioned earlier, the degree of translucency is controlled by the mean free path, which determines the distance light scatters in the shape's interior. Therefore, for each camera pixel intersecting the surface, we have the mean free path expressed in pixel units. Similar to defocus, subsurface scattering results in low-pass filtering of the incident illumination which can severely blur the projected patterns. Because of this local property, shape and pose do not change the plots significantly.

The mean error plot shows that Ensemble is superior at first, from values of 0.0002 to 0.0027 (see details in Figure 10). However, past approximately 0.004, Ensemble becomes highly unstable due to severe blurring in the patterns, making them hard to correctly decode. Here we see a signif-

icant difference in the methods, where although Ensemble, Microphase, and Unstructured all rely on high-frequency patterns at some level, we see that only Ensemble and Unstructured remain stable for highly translucent materials. In general, Microphase outperforms Unstructured as indicated by the smaller error and larger number of valid indices, though for higher translucency Microphase begins to produce a large number of outliers. Interestingly, we observe that Graycodes remains rather stable in terms of outliers at the expense a slightly higher error compared to Microphase, indicative that low-frequency patterns remain robust to high levels of translucency.

We also see that for Ensemble and Microphase, the peaks in their outliers are correlated with significant increases in mean error. This is because, at these values, Chauvenet's criterion begins to take outliers into the computation of the mean error.

5.3. Inter-reflection

In order to benchmark the algorithms with respect to inter-reflections, we conduct two experiments. The first experiment fixes the shape and considers the level of reflectivity as a parameter, in order to observe how these methods behave as the material changes. The second experiment fixes the material, and considers the concavity of a shape as a parameter, in order to measure how the methods perform under increasing amounts of inter-reflections.

For the first experiment, a shape which results in high levels of inter-reflections should be used to stress test the algorithms' performance. Therefore, we model a concave bowl, a widely used surface in the literature. We place the bowl in a pose that favors the inter-reflection of light originating from the projector.

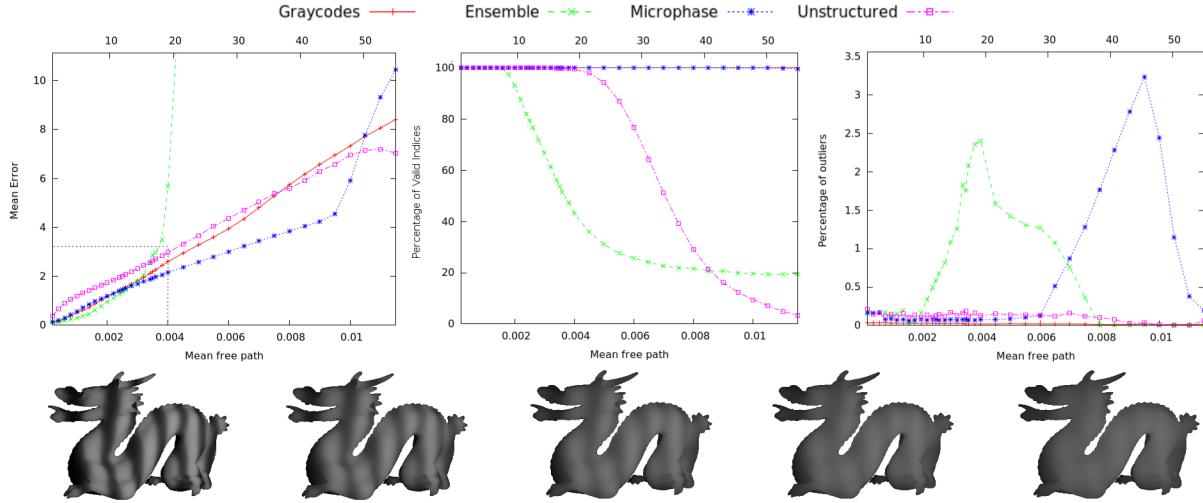


Figure 9: Subsurface scattering results parameterized by the mean free path that the light travels in the surface. On the bottom row, we show the artifact effect as the mean free path increases. The axis on top of each plot denotes, in terms of pixel units, the value of mean free path for a point placed halfway between the maximum and minimum depth of the projector.

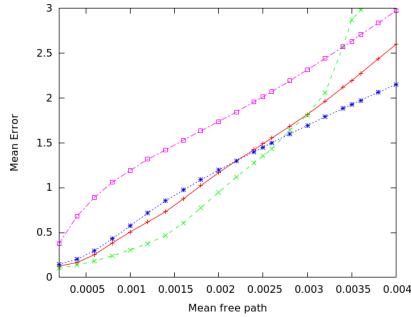


Figure 10: Zoomed inset highlighting the left plot of Figure 9.

The results are shown in Figure 11. The x-axis of the plots represents different material properties in the increasing order of the errors resulting from Graycodes. Initially, Graycodes has a small mean error, but contains lots of outliers. Then, as the mean error grows and more outliers become part of the computation, the number of outliers is reduced.

An interesting observation in the left plot is the similar performance between Ensemble, Unstructured and Microphase. Indeed, these methods completely solve the inter-reflection problem across all but three materials. The difference among these algorithms is highlighted in the middle and left plots. Microphase has the best performance in the number of valid indices followed by Ensemble and Unstructured. In the right plot, Unstructured is better whereas Ensemble followed by Microphase have much more outliers.

To evaluate inter-reflection across shapes, we fix the aluminum material (index 17 in Figure 11) and smooth the concavity of the bowl (see Figure 12). Consistent with the previous plots, we see that Unstructured performs the best

in terms of mean error, at the expense of a slightly higher number of invalid correspondences. We also note that Microphase and Ensemble perform similarly for higher levels of concavity, yet Microphase has a higher number of valid correspondences and a lower number of outliers, indicating its stability over Ensemble for increasing inter-reflections.

5.4. Discussion

The results from the different experiments provide us with various insights into the performance of the different algorithms. Below, we summarize the most relevant observations:

1. Ensemble is more appropriate for short-range effects when low parameter values are considered. At higher values Microphase tends to be more stable.
2. Despite the unfavorable results in the inter-reflection experiments, Graycodes is shown to be very resilient to short-range effects (defocus and subsurface-scattering) in producing outliers and invalid column indices. Indeed, low significant bits are represented by high frequencies which are affected progressively as the parameter value increases.
3. In general, Microphase, Ensemble and Unstructured perform well for wide range effects.

6. Conclusion

Evaluating the performance of structured lighting algorithms is difficult due to not only the physical operations involved, but also due to the large variation in the shape and material properties of the objects being scanned. In order to enable an exhaustive evaluation of the SL scanners, we first constructed a synthetic scanner using advanced physically

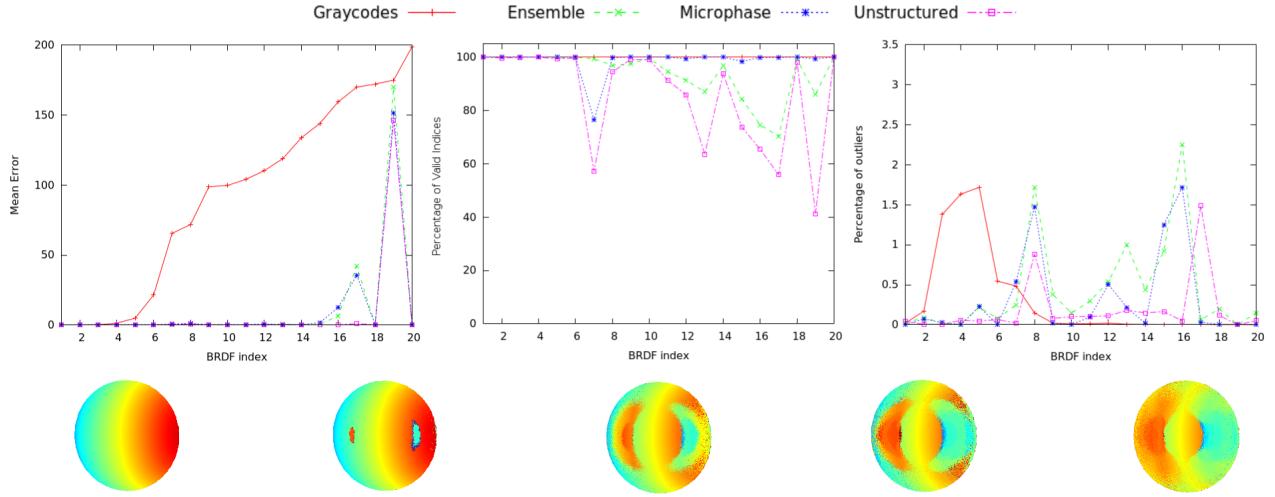


Figure 11: Inter-reflection results of a bowl with different materials. On the bottom row we show the effect of the artifact in the graycodes.

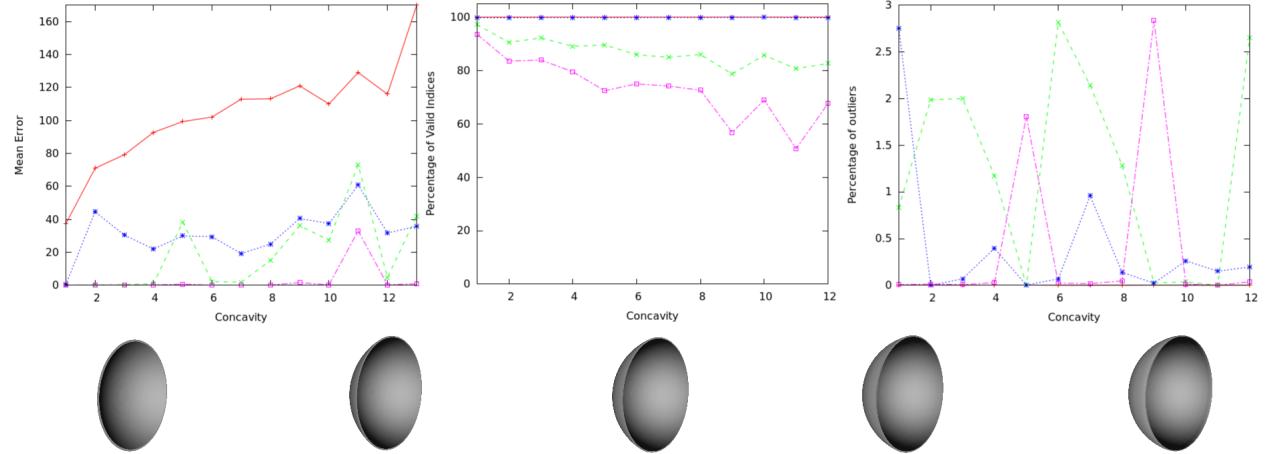


Figure 12: Inter-reflection results of a bowl with different aluminium material (reference 17 in Figure 11) and different shapes. On the bottom row we show a subset of the shapes with increasing concavity.

based rendering techniques to simulate a real scanner. Using this simulator, we create a benchmark for SL scanners using state of the art structured lighting algorithms. We did this by evaluating the performance of the different algorithms in the presence of artifacts due to global illumination, such as inter-reflections and sub-surface scattering, as well as imperfections caused by projector defocus across a variety of material and shape configurations. We expect this benchmark to enable better design of future algorithms and help compare them with the current state of the art.

While we cover three types of artifacts in our benchmark, they are by no means exhaustive. For instance, we assume that the scanner free of noise. In future, we intend to extend this benchmark by introducing different levels of noise into the scanner.

While this work focused on studying the behavior of structured lighting techniques by emphasizing on each type of artifact by isolating them, we intend to expand the benchmark in the future by evaluating the algorithms on different combinations of the artifacts. The above framework will also help us study the affect of applying post processing / progressive techniques [XA09] to existing methods. We also plan on using the benchmark to learn and automatically suggest techniques to be used for a given scene.

Acknowledgments

The authors would like to thank Jonathas Costa for his help with the PBRT rendering software. This work was supported in part by an IBM Faculty Award, the Moore-Sloan Data Science Environment at NYU, the NYU School of Engineering,

NSF award CNS-1229185, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES # 3825/11-6).

References

- [Bla04] BLAIS F.: Review of 20 years of range sensor development. *J. Electronic Imaging* 13, 1 (2004), 231–243. [1](#)
- [BLN*13] BERGER M., LEVINE J. A., NONATO L. G., TAUBIN G., SILVA C. T.: A benchmark for surface reconstruction. *ACM Trans. Graph.* 32, 2 (Apr. 2013), 20:1–20:17. [3](#)
- [BMS98] BATTLE J., MOUADDIB E., SALVI J.: Recent progress in coded structured light as a technique to solve the correspondence problem : A survey. *Pattern Recognition* 31, 7 (1998). [3](#)
- [Cha63] CHAUVENET W.: *A Manual of Spherical and Practical Astronomy: Embracing the General Problems of Spherical Astronomy, the Special Applications to Nautical Astronomy, and the Theory and Use of Fixed and Portable Astronomical Instruments, with an Appendix on the Method of Least Squares*. J. B. Lippincott & Company, 1863. [6](#)
- [CMR11] COUTURE V., MARTIN N., ROY S.: Unstructured light scanning to overcome interreflections. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (Nov 2011), pp. 1895–1902. [2, 4](#)
- [CPC84] COOK R. L., PORTER T., CARPENTER L.: Distributed ray tracing. *SIGGRAPH Comput. Graph.* 18, 3 (Jan. 1984), 137–145. [5](#)
- [CSL08] CHEN T., SEIDEL H.-P., LENSCHE H. P. A.: Modulated phase-shifting for 3d scanning. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA* (2008), IEEE Computer Society. [1](#)
- [FPW92] FARRELL T. J., PATTERSON M. S., WILSON B.: A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo. *Medical Physics* 19, 4 (July 1992), 879–888. [5](#)
- [GAVN11] GUPTA M., AGRAWAL A., VEERARAGHAVAN A., NARASIMHAN S. G.: Structured light 3d scanning in the presence of global illumination. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (Washington, DC, USA, 2011), CVPR '11*, IEEE Computer Society, pp. 713–720. [3](#)
- [GAVN13] GUPTA M., AGRAWAL A., VEERARAGHAVAN A., NARASIMHAN S.: A practical approach to 3d scanning in the presence of interreflections, subsurface scattering and defocus. *International Journal of Computer Vision* 102, 1-3 (2013), 33–55. [1, 2, 4](#)
- [Gen11] GENG J.: Structured-light 3d surface imaging: a tutorial. *Adv. Opt. Photon.* 3, 2 (Jun 2011), 128–160. [1](#)
- [GI12] GRUNA R., IRGENFRIED S.: Reflectance modeling in machine vision: Applications in image analysis and synthesis. In *Machine Vision - Applications and Systems*, Solari F., Chessa M., Sabatini S., (Eds.). InTech, Mar. 2012, pp. 227–246. [2](#)
- [GLT96] GAERTNER H., LEHLE P., TIZIANI H. J.: New highly efficient binary codes for structured light methods. 4–13. [1](#)
- [GN12] GUPTA M., NAYAR S.: Micro phase shifting. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (June 2012), pp. 813–820. [1, 2, 4](#)
- [JMLH01] JENSEN H. W., MARSCHNER S. R., LEVOY M., HANRAHAN P.: A practical model for subsurface light transport. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (2001), SIGGRAPH '01, pp. 511–518. [5](#)
- [LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1996), SIGGRAPH '96, ACM, pp. 31–42. [5](#)
- [LT09] LANMAN D., TAUBIN G.: Build your own 3d scanner: 3d photography for beginners. In *SIGGRAPH '09: ACM SIGGRAPH 2009 courses* (New York, NY, USA, 2009), ACM, pp. 1–87. [1, 2, 4](#)
- [MER06] MERL brdf database. Available at <http://www.merl.com/brdf/>, 2006. [6](#)
- [MPBM03] MATUSIK W., PFISTER H., BRAND M., McMILLAN L.: A data-driven reflectance model. *ACM Transactions on Graphics* 22, 3 (Jul 2003), 759–769. [6](#)
- [NKGR06] NAYAR S. K., KRISHNAN G., GROSSBERG M. D., RASKAR R.: Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph.* 25, 3 (July 2006), 935–944. [5](#)
- [Ope] The openmp api. [5](#)
- [PH] PHARR M., HUMPHREYS G.: A physically based rendering software. Available at <http://www.pbrt.org>. [5](#)
- [PH00] PHARR M., HANRAHAN P.: Monte carlo evaluation of non-linear scattering equations for subsurface reflection. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 2000), SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., pp. 75–84. [5](#)
- [SCD*06] SEITZ S., CURLESS B., DIEBEL J., SCHARSTEIN D., SZELISKI R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* (2006), vol. 1, pp. 519–528. [2](#)
- [SCL*97] SANSONI G., CORINI S., LAZZARI S., RODELLA R., DOCCHIO F.: Three-dimensional imaging based on gray-code light projection: characterization of the measuring algorithm and development of a measuring system for industrial applications. *Appl. Opt.* 36, 19 (Jul 1997), 4463–4472. [1](#)
- [SFPL10] SALVI J., FERNANDEZ S., PRIBANIC T., LLADO X.: A state of the art in structured light patterns for surface profilometry. *Pattern Recogn.* 43, 8 (Aug. 2010), 2666–2680. [3](#)
- [SPB04] SALVI J., PAGĀLS J., BATLLE J.: Pattern codification strategies in structured light systems. *Pattern Recognition* 37, 4 (2004), 827 – 849. [3](#)
- [SS02] SCHARSTEIN D., SZELISKI R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision* 47, 1-3 (Apr. 2002), 7–42. [2](#)
- [SZ00] SZELISKI R., ZABIH R.: An experimental comparison of stereo algorithms. In *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice* (2000), ICCV '99, Springer-Verlag, pp. 1–19. [2](#)
- [VG97] VEACH E., GUIBS L. J.: Metropolis light transport. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1997), SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co., pp. 65–76. [5](#)
- [XA09] XU Y., ALIAGA D. G.: An adaptive correspondence algorithm for modeling scenes with strong interreflections. *IEEE Transactions on Visualization and Computer Graphics* 15, 3 (2009), 465–480. [1, 3, 9](#)
- [ZN06] ZHANG L., NAYAR S.: Projection defocus analysis for scene capture and image display. *ACM Trans. Graph.* 25, 3 (July 2006), 907–915. [4, 5, 6](#)