https://doi.org/10.2352/issn.2168-3204.2017.1.0.23
© 2017; Society for Imaging Science and Technology

Image-Based Relighting using Environment Maps

Michael Tetzlaff and Gary Meyer; Computer Science and Engineering Department, University of Minnesota; Minneapolis, Minnesota

Abstract

An image-based relighting algorithm has been extended so that it can accommodate environment based lighting. Camera mounted flash photographs, employed in the original relighting algorithm, are also used to achieve the environment map based relighting results. In addition to preserving the simple equipment and setup utilized in the original relighting approach, the new method allows professional studio lighting effects, simulation of museum gallery illumination, and outdoor lighting at particular times of the day and year.

Introduction

Image-based rendering, used in combination with photogrammetry, is a powerful way to document cultural heritage artifacts. Photographs of the object are first taken and the original camera positions are determined from the pictures using photogrammetric techniques. Careful blending [1] of the photographs using image-based rendering techniques allows one to observe the object from a position where no pictures were initially taken. In comparison to other computer graphic techniques, the result is faithful to the original pictures because it is derived directly from them.

When the pictures are taken using a camera mounted flash the object can also be relit. Simple point lighting can be achieved by combining the original photographs according to their proximity to the ideal specular direction, incorporating their position relative to a virtual light source instead of just their location with respect to a desired viewing position [2]. This approach has practical advantages when compared to other relighting methods because the equipment necessary is simple and the photographic skill required is not high. New lighting setups and photographic poses can also be created at a later date using the original flash pictures.

In this paper we demonstrate how to extend the flash-based image relighting technique so that it supports full environment map based lighting. This approach has the advantage of providing the lighting facilities of a professional photographic studio. It also makes it possible to visualize how the object will look in a variety of environmental situations. This allows a museum professional to consider how the artifact will look in a particular gallery circumstance, and it permits an archeologist to consider lighting determined by the hour of the day or even the time of the year.

Related Work

Environment maps were first introduced by Blinn and Newell as “reflection maps” for rendering perfect mirror reflections of an environment on a surface [3]. Miller and Hoffman later described how to simulate reflections of an environment by other, more diffuse kinds of surfaces [4]. Debevec recently showed how to perform environment-based lighting using a high-dynamic-range environment texture [5]. There are various image-based acquisition techniques that produce a parameterized material which can be lit by an environment using traditional techniques such as these (for instance, Debevec et al. [6]).

Recently, Schwartz et al. showed that a bidirectional texture function (BTF) acquired from photographs can be environment-mapped [7]. However, to acquire a BTF means photographing the object from many combinations of viewing and lighting directions, which is not practical without the use of expensive gantries. Ludwig and Meyer developed a method for applying high dynamic range environment maps to reflectance transformation images [8]; however, in such a representation, the viewpoint is fixed, so while this method is effective for 2D objects, it does not provide general 3D viewing. What is missing is a truly image-based technique that preserves the photographic quality of the images of the object, which supports free-viewpoint rendering simultaneously with relighting, using a photographic configuration that is practical for cultural heritage.

Our previous work discussed different ways of capturing image-based reflectance, but focused on *backscattering*, a technique in which the light source is attached to the camera, which can be as simple as a commercially-produced camera with a built-in flash [2]. It was noted that several approaches have used backscattering configurations where the camera and light source are fixed and the object turns to expose multiple viewpoints. These methods ultimately fit the backscattering images to a parameterized model, so they could be rendered using existing environment map based relighting techniques, as was demonstrated, for instance, by Riviere et al. [9]. However, all of these prior works have either imposed limitations on the reflectance model [10] [11] [12] [13] or the geometry [14] [9].

Image-Based Rendering with Relighting

The initial steps necessary to apply the image-based relighting technique are straightforward. A set of flash pictures are taken, and the photographs are processed using a commercial photogrammetry tool such as Agisoft Photoscan to obtain the camera locations and the object mesh. If necessary, a diffuse texture map for each picture can be obtained relatively easily (see our recent paper [2] for details), and the specular residual for each image is then determined by subtracting the diffuse contribution for that photograph from the original flash picture.

Projective texture mapping is then used to map the specular residuals onto the geometric mesh of the object, blending them as they are being projected. In spirit, it is similar to the light field rendering algorithm proposed by Buehler et al. [15] and improved by us in a recent publication (Berrier et al. [1]). However, to achieve relighting, the images are weighted based on the similarity of the image’s halfway vector to the virtual halfway vector, not the similarity between the viewing vectors as is the case for light field

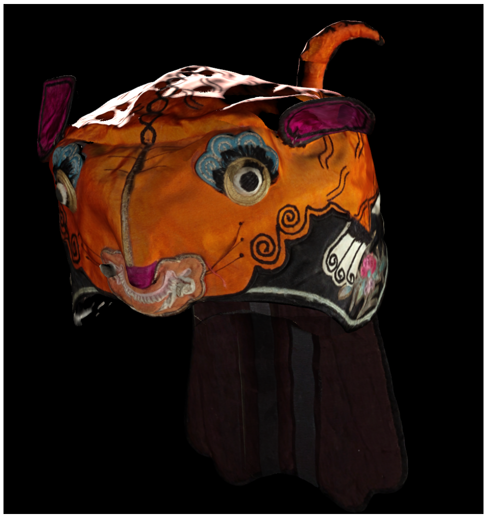

Figure 1. *A synthetic image of a child's tiger hat (67 views). A video of the hat with relighting is available here: VIDEO LINK*

rendering.

This is done because a large category of reflectance models are based on microfacet distributions, which means that the most important variable in the BRDF is the halfway vector that bisects the viewing direction and the lighting direction. Given the viewing direction $\mathbf{v}$ and the illumination direction $\mathbf{l}$, the halfway vector is computed as:

$$\mathbf{h} = \frac{\mathbf{l}+\mathbf{v}}{||\mathbf{l}+\mathbf{v}||} \tag{1}$$

A heuristic is then evaluated to find the blending weight for each photograph, using $\mathbf{h}$, the virtual halfway vector between the desired viewpoint and light source, and $\mathbf{h}_i$, the actual halfway vector between the real viewpoint and light source in the i^{th} original photograph [2]:

$$w(i) = \frac{1}{1-(\mathbf{h}_i \cdot \mathbf{h})^{\alpha}} - 1 \tag{2}$$

Two examples rendered using the image-based relighting algorithm are shown in Figure 1. These images were rendered using $\alpha = 16$ in the heuristic weighting function. Each image was lit by three distinct virtual point light sources shining at the object from different directions. Neither example in Figure 1 has a Lambertian diffuse component, so the results were obtained by simply blending the original flash photographs. In both cases the results are superior to the simple diffuse texture mapping that is commonly used in cultural heritage work.

Extending the Relighting Capabilities

There are a couple of reasons why it is important to extend our novel image-based rendering system so that it supports environment map based relighting. The addition of this feature raises the lighting proficiency of the system to the same level as its representational capability. It also means that it is possible to duplicate professional lighting setups that are used in photography studios and on movie sets. This extension of the system's lighting abilities is also critical in order to fully appreciate the surface reflectance properties of the materials from which the object is constructed. This ability to better appreciate the color appearance of an object is also important when fitting a spatially varying reflectance function to the object's surface.

In general, all that needs to be done to support such light sources is a change to the weighting function. Currently, for point light sources, the closer the halfway vector in the sample photograph is to the halfway vector of the virtual light source, the higher the weight. For an environment map, the weights are instead based on a lookup into the environment map.

First, for each sample photograph, the halfway direction $\mathbf{h}$ is computed. Next, given the virtual viewing direction $\mathbf{v}$, the light vector $\mathbf{l}$ that corresponds to the halfway direction of the photograph is computed by reflecting the view vector across the half vector. Then, this light vector is used to perform a texture lookup into the environment map to determine how much light is cast on the object from that direction. Finally, the photograph is projected onto the object and modulated by the environment map color.

Another way of looking at this technique is that it is essentially an evaluation of the rendering equation [16] where the sample directions for the integral are determined by the photographs, so that the photographs can then be referenced to find the BRDF at each direction. There does remain a challenge of correctly normalizing the sum of these samples, for which a good reference is the total reflectance when illuminated by a uniform hemisphere. Under an unstructured sampling procedure like ours, without any correction, the total reflectance will be off by an arbitrary scaling factor.

We address this scaling problem as follows. As in our previous work, we are basing the image-based rendering technique on the Cook-Torrance model [17] for specular reflection:

$$R_S(\mathbf{l},\mathbf{v}) = \frac{F(\lambda,\mathbf{v})D(\mathbf{h})G(\mathbf{l},\mathbf{v})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})} \tag{3}$$

This equation expresses specular reflectance in terms of the Fresnel reflectivity $F(\lambda,\mathbf{v})$, a microfacet distribution $D(\mathbf{h})$ which describes the probability of a microfacet normal facing in a particular direction, and a geometric attenuation factor $G(\mathbf{l},\mathbf{v})$ that accounts for masking and shadowing. For simplicity in this work, we make an approximation that the Fresnel reflectivity $F(\lambda,\mathbf{v})$ and geometric attenuation $G(\mathbf{l},\mathbf{v})$ are constant:

$$R_S(\mathbf{l},\mathbf{v}) = \frac{F_0 D(\mathbf{h})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})} \tag{4}$$

In order to normalize the sum of environment map samples, each of which is associated with a halfway direction $\mathbf{h}_i$, we rearrange the preceding equation to estimate the value of the microfacet distribution $D(\mathbf{h}_i)$ for each sample:

$$D(\mathbf{h}_i) = \frac{R_S(\mathbf{l}_i,\mathbf{v}_i)4(\mathbf{n}\cdot\mathbf{l}_i)(\mathbf{n}\cdot\mathbf{v}_i)}{F_0} \tag{5}$$

R_S can be obtained from the images as in our prior work [2]. The normal vector $\mathbf{n}$ can be derived from the scene geometry. The vectors $\mathbf{l}_i, \mathbf{v}_i$ are the light and view directions under which the original photo was taken. Because flash photography is used, both $\mathbf{l}_i$ and $\mathbf{v}_i$ are known from the camera locations returned by the photogrammetry software.

The remaining variable is F_0, which should be the specular reflectivity of the surface. Although our prior work estimated

© 2017 SOCIETY FOR IMAGING SCIENCE AND TECHNOLOGY

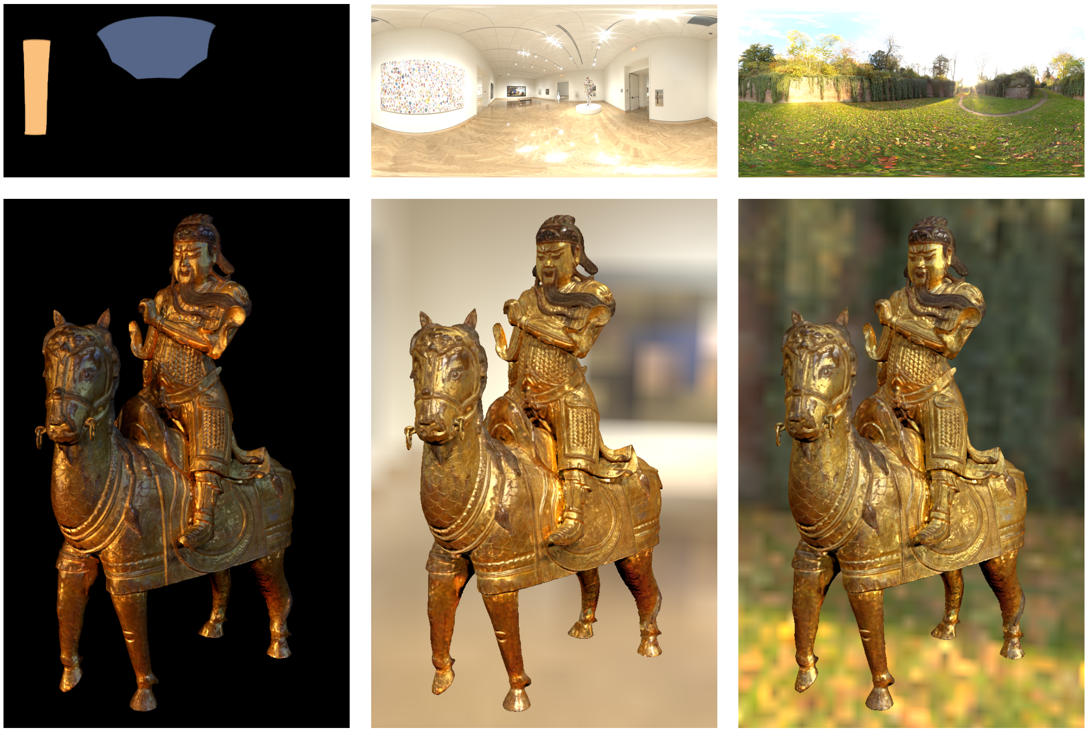

Figure 2. *Image-based environment-mapping on a bronze statue of Kuan Yu. The environments used were, from left to right, a studio with color lights, a gallery at the Minneapolis Institute of Arts, and an outdoor scene.*

the reflectivity under the assumption that it was uniform over the surface, for our new technique it is necessary to have a more accurate spatially-varying reflectivity. We obtain this by using the Levenberg-Marquardt algorithm [18] [19] to fit the parameters of the Cook-Torrance model [17], including the specular reflectivity. (We used the GGX distribution [20] as the model for the microfacet distribution $D(\mathbf{h})$.)

Once we are able to obtain an estimate of the microfacet distribution at each sample, we maintain a sum of microfacet distribution values along with the sum of environment-map-weighted photographs. If the sampling were unbiased and properly weighted, the sum of microfacet distribution values would have summed to one. In reality, this will not be the case, but the scaling factor can be corrected by dividing the weighted sum of photographs by the sum of microfacet distribution values.

Finally, for best results, the environment map needs to be blurred or downsampled to approximately the same angular resolution with which the photographs were taken to prevent aliasing. A good heuristic is that the environment map should have no more pixels than there are photographs in the dataset. A convenient way to achieve this is to specify a mipmap level when performing the texture lookup in the pixel shader. If the original size of the environment map (as a single panoramic image) is $W \times H$ and there are n photographs of the object, then the appropriate mipmap level L can be computed using the following formula:

$$L = \left\lceil \frac{1}{2} \log_2 (WH/n) \right\rceil \tag{6}$$

Results

Figure 2 shows a bronze statue of Kuan Yu from the Minneapolis Institute of Arts lit under a variety of different illumination conditions, using 56 flash photographs. Each environment has a noticeable effect on the appearance of the object, while preserving the photographic quality of the original images.

Figure 3 shows a synthetic example in order to validate this work. The object, a statue of Horus from the British Museum, was downloaded from SketchFab[1]. The object's reflectance was set to be metallic and then the necessary flash photographs (100 views) were synthesized using the ray-tracing software LuxRender[2]. LuxRender was also used to create a reliable rendering of what the object should look like under various environments. Figure 3 compares the ray-tracing results with the result of real-time, image-based rendering. The similarity of the results is evidence for the validity of our image-based environment-mapping technique.

Future Work

Our technique lacks several lighting effects, specifically the Fresnel effect at grazing angles and environment shadows. These are more difficult to implement in real-time for environments than for point lights, so they have been omitted in this work for simplicity. However, implementing these features would improve the quality of the rendering.

This work only considers environment maps assumed to be infinitely far away from an object. This is a reasonable approx-

---

[1] `https://sketchfab.com/models/f16eb0a359aa4354af848744c6916c87`

[2] `http://www.luxrender.net`

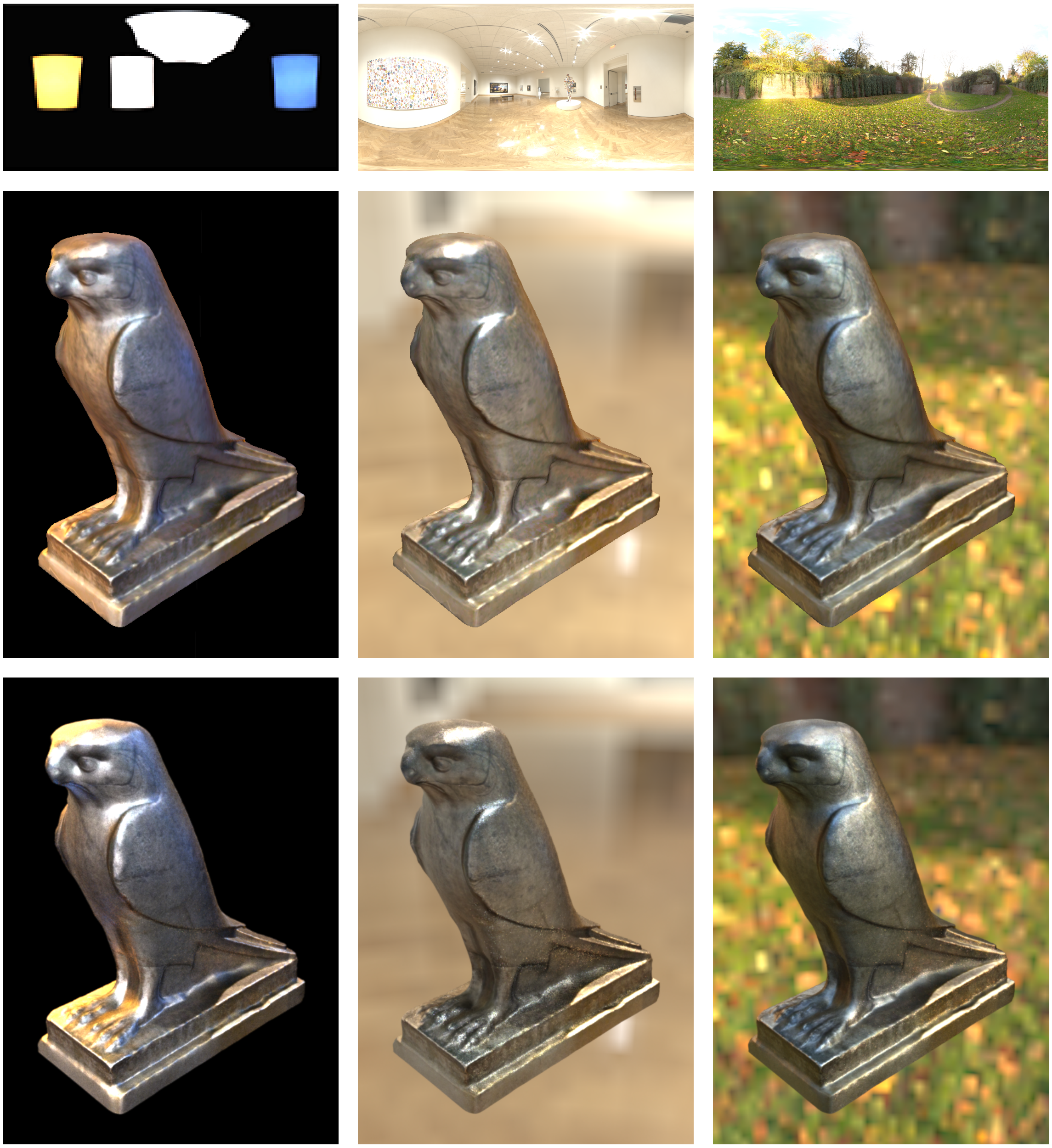

Figure 3. *Middle row: Image-based environment-mapping on a synthetic example (from left to right: studio with colored lights, gallery, outdoor scene). Bottom row: The same scenes ray-traced for comparison.*

imation for most lighting configurations made up of area light sources illuminating a single object. There could, however, be situations where an object needs to be lit explicitly using true area light sources. A simple way to extend the approach presented here would be with another simple change to the weighting function: compute the light vector in the same way as for an environment map lookup, and if it falls within the virtual area light source the corresponding photograph would be given a weight of 1. All other photographs are given a weight less than 1, based on their distance to the boundary of the area light source, using a formula similar to the one used with point light sources, to prevent sudden changes in reflectance when a photograph moves outside of the applicable range for the light source (analagous to the aliasing problem discussed earlier with environment maps). The implementation of an approach such as this would depend on the representation of the area light source.

Finally, we acknowledge that even with some considerations for aliasing such as downsampling an environment map, there is still the possibility of aliasing issues if all of the object has not been sampled well. Texture synthesis, such as Perlin noise, could be one way to fill in the missing information and present an acceptable rendering. Another option would be to capture the object twice, the first time using backscattering, and the second time using a more general illumination configuration that could be used to provide missing information about the specular reflectance.

Conclusions

Environment map based lighting has many advantages for cultural heritage imaging applications. Its use allows almost total flexibility in how artifacts are lit, and it provides museum photographers, curators, and exhibit designers with a powerful tool that can be used to consider an object in the museum's collection under a wide variety of illumination conditions. This paper has demonstrated how a particular type of image-based relighting can be extended to include environment map based lighting without changing the method's underlying simplicity. The results suggest that camera mounted flash photography can be used once to document an object and that new images under novel lighting schemes can be obtained at any time using the original photographic set and the environment mapped based relighting techniques described in this paper.

Acknowledgments

This work was partially funded by DARPA (contract HR0011-16-C-0024). Special thanks to the Minneapolis Institute of Arts and Charles Walbridge, who photographed the Kuan Yu example shown in this paper, and Seth Berrier, who photographed the gallery environment. Thanks to the British Museum for making a 3D model of the Horus statue publicly available on Sketchfab.

References

[1] S. Berrier, M. Tetzlaff, M. Ludwig, and G. Meyer, "Improved appearance rendering for photogrammetrically acquired 3d models," in *Digital Heritage 2015*, vol. 1. IEEE, 2015, pp. 255–262.

[2] M. Tetzlaff and G. Meyer, "Using flash photography and image-based rendering to document cultural heritage artifacts," in *Eurographics Workshop on Graphics and Cultural Heritage*. The Eurographics Association, 2016, pp. 137–146.

[3] J. F. Blinn and M. E. Newell, "Texture and reflection in computer generated images," *Communications of the ACM*, vol. 19, no. 10, pp. 542–547, 1976.

[4] G. Miller and C. R. Hoffman, "Illumination and reflection maps: Simulated objects in simulated and real environments," in *SIGGRAPH 84 Advanced Computer Graphics Animation Seminar Notes*, vol. 190. ACM, 1984.

[5] P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *SIGGRAPH 2008 Classes*. ACM, 2008, pp. 32:1–32:10.

[6] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 145–156.

[7] C. Schwartz, M. Weinmann, R. Ruiters, and R. Klein, "Integrated high-quality acquisition of geometry and appearance for cultural heritage." in *Proceedings of the 12th International Conference on Virtual Reality, Archaeology and Cultural Heritage*. Eurographics Association, 2011, pp. 25–32.

[8] M. Ludwig and G. Meyer, "Environment map based lighting for reflectance transformation images," in *Digital Heritage 2015*, vol. 1. IEEE, 2015, pp. 385–388.

[9] J. Riviere, P. Peers, and A. Ghosh, "Mobile surface reflectometry," in *SIGGRAPH 2014 Posters*. ACM, 2014, p. 68.

[10] Y. Sato, M. D. Wheeler, and K. Ikeuchi, "Object shape and reflectance modeling from observation," in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. ACM Press/Addison-Wesley Publishing Co., 1997, pp. 379–387.

[11] S. R. Marschner, "Inverse rendering for computer graphics," Ph.D. dissertation, Cornell University, 1998.

[12] M. Dellepiane, M. Callieri, M. Corsini, P. Cignoni, and R. Scopigno, "Improved color acquisition and mapping on 3d models via flash-based photography," *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 2, no. 4, pp. 9:1–9:20, 2010.

[13] H. Wu and K. Zhou, "Appfusion: Interactive appearance acquisition using a kinect sensor," *Computer Graphics Forum*, vol. 34, no. 6, pp. 289–298, 2015.

[14] J. A. Paterson, D. Claus, and A. W. Fitzgibbon, "Brdf and geometry capture from extended inhomogeneous samples using flash photography," *Computer Graphics Forum*, vol. 24, no. 3, pp. 383–391, 2005.

[15] C. Buehler, M. Bosse, L. McMillan, S. Gortler, and M. Cohen, "Unstructured lumigraph rendering," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 2001, pp. 425–432.

[16] J. T. Kajiya, "The rendering equation," in *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 1986, pp. 143–150.

[17] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Transactions on Graphics (TOG)*, vol. 1, no. 1, pp. 7–24, 1982.

[18] K. Levenberg, "A method for the solution of certain nonlinear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[19] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[20] B. Walter, S. R. Marschner, H. Li, and K. E. Torrance, "Microfacet models for refraction through rough surfaces," in *Proceedings of the 18th Eurographics Conference on Rendering Techniques*. Eurographics Association, 2007, pp. 195–206.

Author Biography

Michael Tetzlaff is working on his PhD degree at the University of Minnesota Department of Computer Science and Engineering. He is a graduate of Bethel University and he has published papers on several different image-based rendering techniques.

Gary Meyer is an Associate Professor in the Department of Computer Science and Engineering at the University of Minnesota. He received his BS degree from the University of Michigan, his MS degree from Stanford University, and his PhD degree from Cornell University. Meyer specializes in the simulation of color appearance using computer graphic reflectance modeling techniques.