

# IMAGE-ASSISTED GEOMETRY SIMPLIFICATION FOR THE PLENOPTIC SAMPLING

Chang-Jian Zhu, and Li Yu

School of Electron. and Inform. Comm., Huazhong Univ. of Sci. & Tech., Wuhan, China  
[{changjianzhu,hustlyu}@hust.edu.cn](mailto:{changjianzhu,hustlyu}@hust.edu.cn)

## ABSTRACT

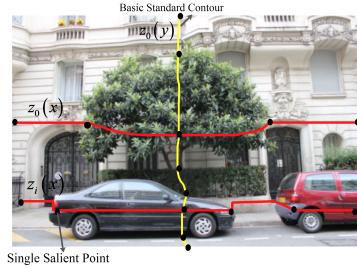
In this paper, a new method is proposed to generate image-assisted geometry simplification for estimating minimum sampling rate of image-based rendering (IBR). If some geometry information (e.g., the shape of object surface and proxy geometry) on a scene is known, we can decompose the scene geometry into a collection of simpler structures on a block-by-block basis. Our framework predicts the characteristics of simpler structure such as irregular object. Predictions on the frequency content can then be used to control sampling rates for IBR. The new method allows the sampling of the IBR to be analyzed and estimated for the non-uniform sampling. Furthermore, the minimum sampling rate of the IBR necessary for alias-free rendering will be reduced as the number of simpler structures increases.

**Index Terms**— Image-based rendering, plenoptic sampling, simpler structure, geometry simplification.

## 1. INTRODUCTION

Image-based rendering (IBR) is one of the most basic and fascinating techniques in 3D computer graphics [1], [2]. In general, the IBR uses many images but does not require any geometrical information to directly create novel views [3], [4]. However, by Chai *et al.* [5], for IBR, as more geometrical information becomes available, fewer images are necessary. To this case, an important decision is the needed geometrical information determination, that best balance between the number of images necessary and reconstructed view quality[5].

The answers to the above question are related to the sampling problem of the IBR which is mentioned for the first time by Gortler *et al.* [6] and Levoy *et al.*[7]. It is also analyzed in detail by Shum *et al.* [8], Chai *et al.* [5], Zhang *et al.* [9], Do *et al.* [10], and Gilliam *et al.* [11]. The above methods are based on the spectral analysis point of view to determine the minimum sampling rate of the IBR. These efforts have produced an extensive array of interesting results that shed light on various aspects of the problem. In particular, a special case of little amount of depth information utilized is shown in [5] to appropriately reduce the sampling rate necessary. They used a method named layered depth [12]-[14] to approximate a complex scene. Further results for the method of layered depth from Zhang *et al.* [9] are investigated to analyze the



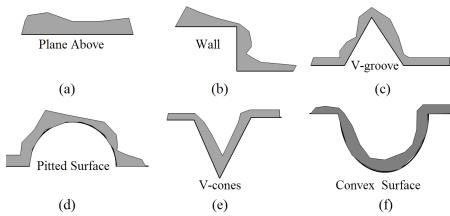
**Fig. 1.** Conceptual illustration of image-assisted geometry simplification to the IBR based on different surface shapes.

occluded complex scenes. Reference [15] also provides a fast automatic layer-based method to reduce the number of images necessary for alias-free rendering in IBR. The theory has been adapted to approximate the geometry of a scene. In [16], Chaurasia *et al.* presented a new approach which uses the introducing silhouette aware variational warping to compensate for incorrect or incomplete geometric information in IBR. Recently, Shidanshidi *et al.* [17] extends the concept of effective sampling density (ESD) [18]. Measuring the variation of the ESD on an irregular surface of a scene would allow us to reproduce the geometry information accurately. Similarly, a signal-processing framework for light transport [19] has been presented. From a frequency analysis of light transport, the radiance of light ray and how it is altered by phenomena such as shading, occlusion, and transport can be studied.

The success of these methods depends on accurate 3D proxies. In general, the scene geometry is unknown; little amount of geometry simplification is challenging topic from dense range scans for a complex scene [16]. Inspired by the results in [20], our approach does not depend on dense accurate geometric reconstruction; instead we compensate for sparse 3D information by a set of simpler structures as depicted in Fig. 1. In particular, we formulate the spectral support of each simpler structure that preserve single salient points. Our contribution to the existing theory is simplification of a complex scene to simple structures. Additionally, this method can be applied to study the non-uniformly sampling in IBR.

## 2. THE STRUCTURE OF THE SCENE SURFACE

As mentioned earlier, the scene geometry for irregular shape is unknown and its quantification is very complicated, such as



**Fig. 2.** Six common scene surface geometry situations. We show that these all have similar structures.

the leaves of a tree, and mathematical analysis seems hopeless [20]. However, we believe many common scene geometry situations have simpler structures, some of which are illustrated in Fig. 2. If some information (e.g., the shape of the surface and proxy geometry) on the simpler structures is known, for the original complex geometrical model, it is simplified as a set of simpler structures. Each structure will have a much tighter bound than the whole scene. How tight the bound is will depend on the accuracy of the geometry. Thus, given the appropriate number of image samples an average user can afford, appropriate geometric information is utilized while still guaranteeing the same rendering quality.

### 2.1. Simpler structure selection

Structure selection can be manually authored in each input image or computed automatically. For the computed automatically, shape similarity and shape retrieval are very important step [16]. Now, we assume that the shape of object is given in advance, and it is represented using  $S$ . And then there are  $K$  single salient points according to shape.  $K$  also relates to the spacing  $\Delta t$  between the cameras and the pixel spacing  $u$  and the depths. Without loss of generality,  $K$  is determined by

$$K \approx f\Delta t A h S, \quad (1)$$

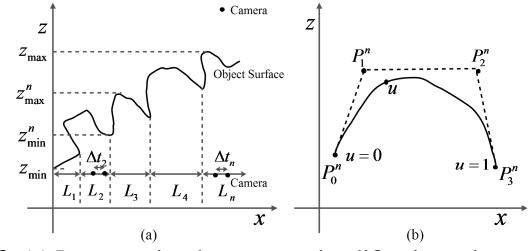
$h = \left[ \frac{1}{z_{\min}} - \frac{1}{z_{\max}} \right]$  and  $A \leq 0.5/u$  is the highest image bandwidth in reference to [15], and  $f$  is the focal length.

How to the number  $N$  of simpler structures is determined? Taking into account for the scene distribution, our approach requires pre-annotated single salient points (Fig. 1). By dividing the scene into a set of simpler structures, this allows us to determine the minimum number of simpler structures,  $N \leq K - 1$ .

### 2.2. Parameterization

To intuitively show this method used to simplify the whole scene geometry, we present an example, which is illustrated in Fig. 3(a). In this example, the whole scene geometry consists of a plane above, a wall, two V-grooves and a pitted surface. The whole scene geometry is approximated by these structures. Note that a scene is described using 3D coordinates in a space as Fig. 1. Here, to simplify, we only present the parameterization in the 2D space. Thus, the whole scene geometry can be represented as

$$z(x) = \{z_n(x) : x \in [x_n, x_{n+1}], n = 1, \dots, N, 0 \leq x \leq L\}, \quad (2)$$



**Fig. 3.** (a) Image-assisted geometry simplification using a set of simpler structures on a block-by-block basis. (b) A simpler structure is described using Bezier curves.

where  $z_n(x)$  denotes the geometry function of the  $n^{\text{th}}$  structure, and  $L$  is the length of scene. Further, the expression of the geometry with the  $n^{\text{th}}$  structure is described as

$$z_n(x) = b_n Q_n(x), \text{ for } x \in [x_n, x_{n+1}], \quad (3)$$

where  $b_n$  is a scaling factor of the  $n^{\text{th}}$  simpler structure, and  $Q_n(\cdot)$  is the surface geometry function of the  $n^{\text{th}}$  simpler structure between  $z$  and  $x$ . As shown in Fig. 3(a), the geometry of the  $n^{\text{th}}$  simpler structure is mainly described by its minimum depth  $z_{\min}^n$  and the maximum depth  $z_{\max}^n$  and its length  $L_n$  in  $x$ -axis.

Surface geometry parameterization is a extremely complicated work, and well knows methods such as polygonal mesh, linear fitting, and so on. Here, a surface curve is approximated by Bezier curves [21] which using four points and a primary function to represent. The precision of the fitting depends on the four points as shown in Fig. 3(b). In (2),  $Q_n(x)$  can be described using Bezier, and then the surface geometry mainly depends on four points,  $P_0^n(x_0^n, z_0^n), P_1^n(x_1^n, z_1^n), P_2^n(x_2^n, z_2^n)$  and  $P_3^n(x_3^n, z_3^n)$ :

$$Q_n(x) \approx \sum_{k=0}^3 P_k^n B_k^n(x), \quad (4)$$

where  $B_k^n(x)$  is a geometry function for the  $k^{\text{th}}$  element of the  $n^{\text{th}}$  simpler structure, and it mainly depends on the surface geometry of the  $n^{\text{th}}$  simpler structure. Using (4), we have  $z_{\min}^n = \min(z_0^n, z_1^n, z_2^n, z_3^n)$ ,  $z_{\max}^n = \max(z_0^n, z_1^n, z_2^n, z_3^n)$ .

$$L_n = |x_0^n - x_3^n|. \quad (5)$$

Then, as a very coarse approximation, combining (2) and (4), the whole scene geometry consists of all the simpler structures. Here, we only use the control points  $P_k^n$  to approximately represent the scene geometry information.

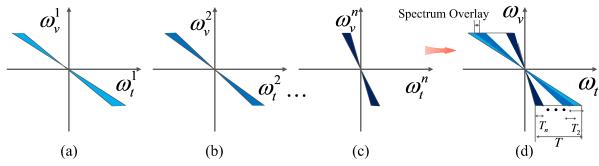
## 3. 2D ANALYSIS OF EACH STRUCTURE

### 3.1. Spectral analysis for each simpler structure

Now, assuming the camera line  $t$  coincides with the  $x$  coordinate system, the authors link a light ray arriving at  $(t, v)$  to its point of origin on the surface at  $(x, z_n(x))$  using the following geometric relationship

$$t = x - z_n(x) \tan(\theta) = x - \frac{z_n(x)v}{f}, \quad (7)$$

where  $\theta$  is the viewing angle [5]-[11]. It is named as epipolar plane image (EPI) [22] which is represented by  $p_n(t, v)$  in



**Fig. 4.** Diagrams (a)-(c) showing the spectral supports for the  $n^{th}$  simpler structure; (d) The spectral support of the whole scene which consists of  $N$  simpler structures.

the 2D space. Thus, the spectrum of the  $n^{th}$  simpler structure is expressed as

$$P_n(\omega_t^n, \omega_v^n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_n(t, v) e^{-j(\omega_t^n t + \omega_v^n v)} dt dv. \quad (8)$$

Substituting (7) to (8), the spectrum of plenoptic function [5] with the properties of the simpler structure can be analyzed. However, these related works have been studied by Chai and Gilliam *et al.* [5]-[11]. Here, we will not go into the details of its derivation. Based on the results in [5]-[11], the spectrum of the  $n^{th}$  simpler structure is represented as

$$\omega_v^n f \approx \omega_t^n b_n \sum_{k=0}^3 P_k^n B_k^n(x). \quad (9)$$

### 3.2. Spectral support of each simpler structure

In (9), the spectral support of each simpler structure is closely related to the geometry shape. As mentioned earlier, it can be described by four points in Fig. 3(b), also, the texture information in the scene surface, occlusion, non-Lambertian, and so on. The spectrum with these information is represented by  $B_k^n(x)$  as in (9). In addition, these factors influence on the spectral support of the plenoptic function have been studied by Chai and Gilliam *et al.* [5]-[11]. And then, we extend their results to study the non-uniform sampling of the plenoptic function based on the method of simpler structures. Figs. 4(a)-(c) show the spectral support for each simpler structure. Additionally, according to (2) and (9), the spectral support for the whole scene is given as

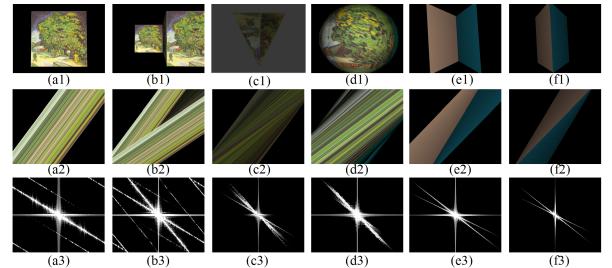
$$\omega_v f \approx \omega_t \left\{ b_n \sum_{k=0}^3 P_k^n B_k^n(x) : n = 1, \dots, N \right\}. \quad (10)$$

Finally, the spectral support of the whole scene consists of all the spectral supports of simpler structures as Fig. 4(d).

## 4. SAMPLING THEOREM OF THE IBR BASED ON THE SIMPLER STRUCTURES

### 4.1. Non-uniform sampling theory

In [23], Zhang *et al.* demonstrated that the non-uniform sampling approaches outperform the traditional uniform methods. Now, based on the method of simpler structures, we present a non-uniform sampling method for plenoptic sampling. As in Fig. 3(a), the geometry surface consists of  $N$  simpler structures. This diagram also shows that  $\Delta t_n$  is the spacing between the cameras based on the non-uniform sampling for the  $n^{th}$  simpler structure. According to (5) and (9), the maximum camera spacing of the  $n^{th}$  simpler structure is given as [5]



**Fig. 5.** Diagrams (a1)-(f1) show the different simpler structures. Diagrams (a2)-(f2) of the EPI-volumes for the corresponding simpler structures. Diagrams (a3)-(f3) show the spectrums for the corresponding simpler structures.

$$\Delta t_n = \frac{2\pi z_{\min}^n z_{\max}^n}{\Omega_v^n f(z_{\max}^n - z_{\min}^n)}, \quad (11)$$

where  $\Omega_v^n$  is the maximum frequency in the  $\omega_v^n$ -axis.

### 4.2. Uniform sampling theory

Using the method of the simpler structures, we also present a uniform sampling theory for plenoptic sampling. For the uniformity sampling, we use the average depth of all the simpler structures. Then, by (10), similarly to (11), the maximum camera spacing is expressed as

$$\Delta t_{\max} = \frac{\pi \sum_{n=1}^N (z_{\min}^n + z_{\max}^n)}{N \Omega_v f (z_{\max} - z_{\min})}, \quad (12)$$

where  $\Omega_v$  is the maximum frequency in the  $\omega_v$ -axis.

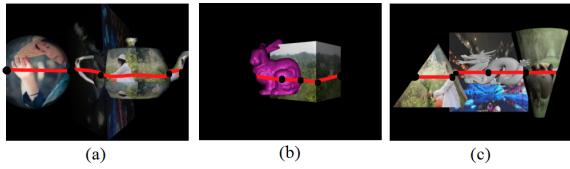
## 5. EXPERIMENTAL RESULTS

### 5.1. Validation of the spectrums for different structures

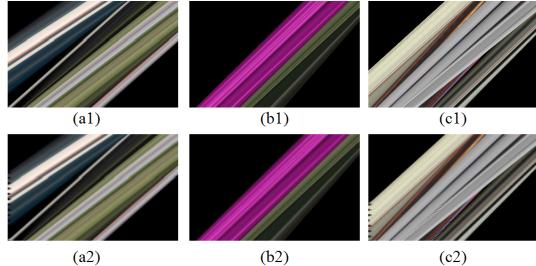
To testify our proposed method can be applied to perform the geometry simplification, the spectral supports for different simpler structures are measured on a line in a space. Fig. 5 shows that although the minimum depth and the maximum depth are the same for different simpler structures, but the spectral supports are different. For example, the spectral supports in Figs. 5(b3)-(d3) are different. The reason can be seen by the geometry shape and texture information are different between the three structures. This demonstrates that spectral supports of the plenoptic function also depend on the geometry shape.

### 5.2. View synthesis

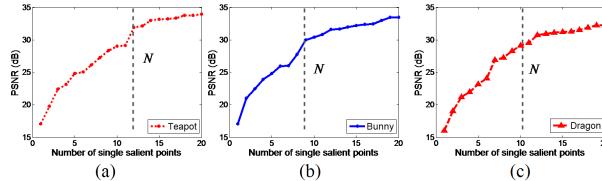
To estimate our proposed method for the geometry simplification, synthesis results are carried out by a set of captured images. The images are captured by a set of cameras which are non-uniformly placed in a line  $l = \{(x, 0, 0)\}$ . Here,  $x \in [-100.0, 100.0]$  cm. The spacing between the cameras for each simpler structure is calculated by (11). Each view is rendered using the method of image interpolation [24]-[27]. Three scenes shown in Fig. 6 are used to perform the synthesis. For the synthesis, first we need the depth of the  $K$  single salient points for all of the input images and the view to be synthesised as Fig. 6. The black points indicate the single salient points. Now, 200 views are rendered for each scene.



**Fig. 6.** (a) Teapot scene. (b) Bunny scene. (c) Dragon scene.



**Fig. 7.** Diagrams (a1)-(c3) illustrate the ground truths for teapot, bunny and dragon scenes. Diagrams (a2)-(c2) show the reconstruction EPI volumes for the corresponding scenes.



**Fig. 8.** PSNR with respect to number of single salient points.

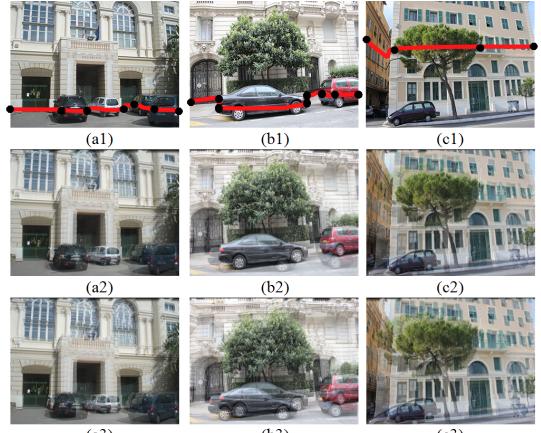
Fig. 7 shows the rendered EPI volumes of 200 views. It can be seen that given the lack of accurate geometry for foreground objects, the rendered results have no ghosting artifacts.

To evaluate the influence of single salient points on the sampling rate, we also use the three scenes to render the 200 views. The number of captured images is the same, and then using different number of single salient points. We validate this analysis and the effectiveness of our method in Fig. 8 which shows the variation of PSNRs with the number of single salient points over all the evaluation data sets. This demonstrates that the PSNR increases as the number of single salient points increases. But then the improvement of PSNR becomes less significant with increasing number.

### 5.3. Comparison results

To enable comparison of accuracy with alternative solutions, the method of silhouette-aware warping (SAW) [16] and its datasets are applied to perform the experimental evaluation. The data sets are named as aquarium-20, street-10 and tree-18 in Figs. 9(a1)-(c1). For each scene, one view is rendered by 12 captured images, and the rendered method is using image interpolation [24]-[27]. In addition, the geometry simplification is performed using our proposed method and the SAW.

The results which are rendered by our proposed method are shown in Figs. 9(a2)-(c2). Figs. 9(a3)-(c3) show the corresponding rendering results for using SAW. It can be observed in either Figs. 9(a2)-(c2) or Figs. 9(a3)-(c3) that seri-



**Fig. 9.** (a1)-(c1) Ground truths; (a1) Aquarium-20; (b1) Street-10; (c1) Tree-18 [16]. (a2)-(c2) The rendered virtual views using our proposed method. (a3)-(c3) The corresponding rendered virtual views using SAW.

**Table 1.** The PSNRs using our proposed method and the SAW for different scenes, (dB).

Methods	Aquarium-20	Street-10	Tree-18
SAW	25.87	26.76	27.54
Our method	26.90	27.68	28.13

ous aliasing in the rendered views of the these scenes exists, due to the number of captured images too small (i.e., under-sampled). Apparently, the aliasing in Figs. 9(a3)-(c3) is more serious than in Figs. 9(a2)-(c2). Furthermore, the PSNRs in Figs. 9(a2)-(c2) and (a3)-(c3) are presented in Table I. Form the above analysis, we can consider that the method of simpler structures can be applied to simplify the geometry of a scene for estimating minimum sampling rate of the IBR. And the performance of the geometry simplification is precede SAW.

## 6. CONCLUSION

In this paper, a new method is proposed to decompose a original complex geometrical model into a collection of simpler structures. The purpose is to reduce the number of images necessary for alias-free rendering in IBR. The influence of the geometry variation for each structure on the spectrum has been studied. Based on the support of the spectrum, the spacing between the cameras needed to reconstruct the continuous light field up to a certain camera plane frequency can be determined. Furthermore, this geometry simplification method can be applied to analyze the non-uniform sampling in IBR.

## 7. ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 61231010, in part by the 863 High-Tech Research and Development Program under Grant 2015AA015901 and Grant 2015AA015903, and in part by the Research Fund for the Doctoral Program, 20120142110015.

## 8. REFERENCES

- [1] H.-Y. Shum, S. Kang, and S.-C. Chan, “Survey of image-based representations and compression techniques,” *IEEE Trans. Circ. and Syst. for Video Tech.*, vol. 13, no. 11, pp. 1020-1037, Nov. 2003.
- [2] C. Zhang, and T. Chen, “A survey on image-based rendering-representation, sampling and compression,” *EURASIP Signal Process., Image Commun.*, vol. 19, no. 1, pp. 1-28, 2004.
- [3] H.-Y. Shum, S. C. Chan, and S. B. Kang, “Image-Based rendering,” New York, NY, USA: Springer-Verlag, 2007.
- [4] C. Zhang, and T. Chen, “Light field sampling,” *Synthesis lectures on image, video, and multimedia processing*, 6, 2006.
- [5] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, “Plenoptic sampling,” in *Proc. SIGGRAPH*, New York, NY, USA, 2000, pp. 307-318.
- [6] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen, “The lumigraph,” in *Proc. SIGGRAPH*, 1996, pp. 43-54.
- [7] M. Levoy, and P. Hanrahan, “Light field rendering,” in: *Proc. SIGGRAPH*, New Orleans, LA, USA, 1996, pp. 31-40.
- [8] H.-Y. Shum, and L.-W. He, “Rendering with concentric mosaics,” in *Proc. Computer Graphics (SIGGRAPH’99)*, Aug. 1999, pp. 299-306.
- [9] C. Zhang, and T. Chen, “Spectral analysis for sampling image-based rendering data,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 11, pp. 1038-1050, Nov. 2003.
- [10] M. N. Do, D. Marchand-Maillet, and M. Vetterli, “On the bandwidth of the plenoptic function,” *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 708-717, Feb. 2012.
- [11] C. Gilliam, P. Dragotti, and M. Brookes, “On the spectrum of the plenoptic function,” *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 502-516, Feb. 2014.
- [12] J. Shade, S. Gortler, L.-W. He, and R. Szeliski, “Layered depth images,” in *Proc. SIGGRAPH*, 1998, pp. 231C242.
- [13] H.-Y. Shum, J. Sun, S. Yamazaki, Y. Li, and C.-K. Tang, “Pop-up light field: An interactive image-based modeling and rendering system,” *ACM Trans. Graph.*, vol. 23, no. 2, pp. 143-162, 2004.
- [14] J. Berent and P. L. Dragotti, “Plenoptic manifolds,” *IEEE Signal Process. Mag.*, vol. 24, no. 6, pp. 34-44, Nov. 2007.
- [15] J. Pearson, M. Brookes, and P. L. Dragotti, “Plenoptic layer-based modeling for image based rendering,” *IEEE Trans. on Image Processing*, vol. 22, no. 9, pp. 3405-3419, 2013.
- [16] G. Chaurasia, O. Sorkine-Hornung, and G. Drettakis, “Silhouette-aware warping for image-based rendering,” in *Proc. of Computer Graphics Forum*, vol. 30, no. 4, pp. 1223-1232, 2011.
- [17] H. Shidanshidi, F. Safaei, and W. Li, “Estimation of signal distortion using effective sampling density for light field-based free viewpoint video,” *IEEE Trans. on Multimedia*, DOI: 10.1109/TMM.2015.2447274, 2015.
- [18] H. Shidanshidi, F. Safaei, and W. Li, “Objective evaluation of light field rendering methods using effective sampling density,” in *Proc. 2011 IEEE 13th International Workshop on Multimedia Signal Processing (MMSP)*, Hangzhou, Oct. 2011, pp. 1-6.
- [19] F. Durand, N. Holzschuch, C. Soler, E. Chan, and F. X. Sillion, “A frequency analysis of light transport,” *ACM Trans. on Graphics (TOG)*, vol. 24, no. 3, pp. 1115-1126, 2005.
- [20] R. Ramamoorthi, M. Koudelka, and P. Belhumeur, “A fourier theory for cast shadows,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 288-295, 2005.
- [21] B. Beizer, “Black-box testing: techniques for functional testing of software and systems,” *John Wiley and Sons, Inc.*, 1995.
- [22] R. Bolles, H. Baker, and D. Marimont, “Epipolar-plane image analysis: An approach to determining structure from motion,” *Int. J. Comput. Vis.*, vol. 1, no. 1, pp. 7-55, 1987.
- [23] C. Zhang, and T. Chen, “Non-uniform sampling of image-based rendering data with the position-interval-error (PIE) function,” in *Proc. IEEE Visual Communications and Image Processing*, Jun. 2003, pp. 1347-1358.
- [24] C. Buehler, M. Bosse, L. McMillan, S. J. Gortler, and M. F. Cohen, “Unstructured lumigraph rendering,” in *Proc. SIGGRAPH*, 2001, pp. 425-432.
- [25] S. E. Chen, and L. Williams, “View interpolation for image synthesis,” in *Proc. SIGGRAPH*, pp. 279-288, 1993.
- [26] M. Lhuillier, and L. Quan, “Image interpolation by joint view triangulation,” in *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, vol. 2, 1999.
- [27] D. Mahajan, F. C. Huang, W. Matusik, R. Ramamoorthi, and P. Belhumeur, “Moving gradients: a path-based method for plausible image interpolation,” *ACM Trans. on Graphics (TOG)*, vol. 28, no. 3, 2009.