IMAGE-BASED RELIGHTING AS THE SAMPLING AND RECONSTRUCTION OF THE PLENOPTIC ILLUMINATION FUNCTION

Tien-Tsin Wong

Department of Computer Science & Engineering
The Chinese University of Hong Kong

ABSTRACT

Image-based modeling and rendering has been demonstrated as a cost-effective and efficient approach to real-time graphics systems. In this paper, we describe an extended formulation of the plenoptic function, called the *plenoptic illumination function*, which explicitly specifies the illumination component. Techniques based on it can be extended to support *relighting* as well as view interpolation. Based on the linearity of illumination, image-based relighting can be performed with complex lighting configuration. The core of this framework is compression, and we therefore show how to exploit two types of data correlation, *intra-pixel* and *inter-pixel* correlations, in order to achieve a manageable storage size.

1. INTRODUCTION

The plenoptic function [1] was originally proposed for evaluating low-level human vision models. In the recent years, several image-based techniques [2, 3, 4, 5] that are based on this computational model have been proposed to interpolate views. Although the plenoptic function is very general, all the illumination and scene changing factors are embedded in a single aggregate time parameter. This time parameter is usually assumed fixed. Therefore, techniques based on the plenoptic function usually assume that the illumination is unchanged. Unfortunately, the capability to change illumination (*relight*) is an important function in computer graphics.

In this paper, we define an extended formulation by extracting the *illumination* component from the aggregate time parameter. We call it the *plenoptic illumination function* [6]. Thus, techniques based on this new formulation can be extended to support relighting as well as view interpolation.

Introducing an additional dimension to the plenoptic function suffers from an increase of storage requirement. To make it practical, we point out two types of data correlation, namely *intra-pixel* and *inter-pixel* data correlations. Then a compression scheme is recommended to reduce the data storage to a manageable size.

This project is supported by the Research Grants Council of the Hong Kong Special Administrative Region, under RGC Earmarked Grants (Project No. CUHK 4186/00E).

2. THE PLENOPTIC ILLUMINATION FUNCTION

Adelson and Bergen [1] proposed a seven-dimensional *plenoptic function* for evaluating the low-level human vision models. Although the original formulation of plenoptic function is very general, the illumination changing factor is assumed fixed. However, the ability to express the illumination configuration is important in computer graphics and multimedia applications. To include the illumination component, we extract an illumination component ($\vec{L}$) from the aggregate time parameter and explicitly specify it in the following new formulation, the *plenoptic illumination function* [6].

$$I = \mathrm{P_I}(\vec{L}, \vec{V}, \dot{E}, t', \lambda), \tag{1}$$

where I is the radiance; $\vec{L}$ specifies the direction of a directional light source illuminating the scene; $\vec{V}$ specifies the viewing direction originated from the viewpoint; $\dot{E}$ is the position of viewpoint; t' is the time parameter after extracting the illumination component; and λ is the wavelength.

The illumination component, $\vec{L}$, describes the lighting direction of a directional light, which emits unit radiance, illuminating the scene. Intuitively speaking, the extended formulation tells us how the environment looks like with the illumination of a directional light source. We basically treat the image synthesis as a problem of sampling and reconstruction of this plenoptic illumination function.

3. SAMPLING

Sampling the plenoptic illumination function is actually a process of taking pictures. The question is how to take these pictures. The time parameter t' is usually assumed fixed and the wavelength parameter λ is conventionally sampled and reconstructed at three positions (red, blue and green). As we are interested in constant-viewpoint images, parameter $\dot{E}$ is also fixed. The sampling of $\vec{V}$ depends on the projection manifold used. The manifold can be planar, cylindrical or spherical. Since the newly introduced dimension $\vec{L}$ is a direction, we sample it at the grid points of spherical coordinate system for simplicity (Figure 1(a)). The disadvantage is that samples are not evenly distributed on the sphere. More samples are placed near the poles of sphere. Figure 1(b) shows a 3D plot of radiance values (R channel only) associated with a pixel. The sampling rate depends on the surface property of object. Lin, Wong and Shum [7] derived a theoretical bound for this sampling rate.

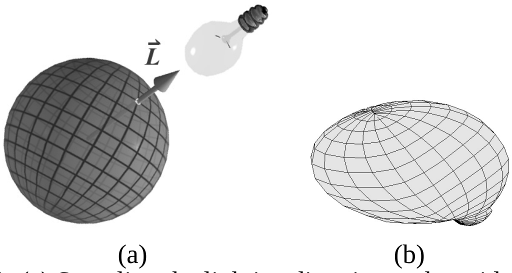


Fig. 1. (a) Sampling the lighting direction at the grid points on the sphere. (b) A sampled plenoptic illumination function.

For synthetic scenes, the samples can be easily collected by rendering images with a directional light source oriented at wish. For real scenes, spotlight positioned at a sufficiently far distance can be used to approximate a directional light source. However, precise control of the lighting direction may require the construction of a robotic arm. Instead, we developed a capture system by tracking a handheld light source using computer vision techniques [8].

The sampled data is basically a multi-dimensional table of radiance values, indexed by light vector, viewing vector (screen coordinates and viewpoints), and wavelength (color channels). Since the size of table is enormous, we describe a compression technique in Section 5.

4. RELIGHTING

Given a desired light vector which is not one of the samples, the desired image can be estimated by interpolating the samples. The interpolation on the illumination dimension is called *relighting*. In general, the result improves as we employ higher-order interpolation. But the accuracy depends on the actual scene geometry and the surface BRDF. Instead of interpolation in the spatial domain, we interpolate in the frequency domain. In our work, we use spherical harmonics as the basis functions (see Appendix for the detail equations). The major advantage is that spherical harmonic facilitates the later compression (described in Section 5).

Even though the sampled plenoptic illumination function only tells us how the environment looks like when it is illuminated by a directional light source with unit intensity, other illumination configuration can be simulated by making use of the linearity of illumination. The intensity (radiance) of a pixel can be calculated by the following formula for each pixel for each color channel (red, green and blue). This formula allows us to manipulate the direction, the color, and the number of light source.

$$\sum_{i}^{n} P_I^*(\vec{L}_i) R(\vec{L}_i), \qquad (2)$$

where n is the total number of light sources; $\vec{L}_i$ specifies the desired lighting direction of the i-th light source; $P_I^*(\vec{L}_i)$ is the interpolated plenoptic illumination function given $\vec{L}_i$ (other parameters are dropped for clarity); and R is the radiance along $\vec{L}_i$ due to the i-th light source.

Figure 2 shows the relit results for different types of light sources. They are (a) a spotlight, (b) a directional source, and (c) a slide projector source. Except for Figure 2(b), all others require the depth map in order to be relit correctly. Interested readers are referred to [6] for details. Note that if reference images exhibit shadow, the relit images also exhibit shadow (see Figures 2(a)-(c)).

One application of the plenoptic illumination model is the panorama representation [9]. Figures 2(d)-(i) show two relit panoramas, attic and city. The city panorama in Figure 2(g) is relit with a single directional light source while the attic panorama in Figure 2(d) is relit with multiple spotlights and slide projector sources. Note how the illumination in the region with occlusion (pillar & chair) is correctly accounted in Figures 2(e) and (f). The attic scene contains 50k triangles and each reference image (input sample) requires 133 seconds to render on a SGI Octane with a MIPS 10000 CPU, using the software renderer Alias|Wavefront. The city scene contains 187k triangles and requires 337 seconds for generating each reference image. On the other hand, the relighting of both image-based scenes (1024 $\times$ 256 in resolution) can be completed within a second (0.661 second) using our pure software relighting engine. This demonstrates the major advantage of image-based computer graphics – the rendering independence of scene complexity. A prototype panoramic viewer with relighting ability is available for download on our webpage (see Section 7).

5. COMPRESSION

5.1. Intra-Pixel Correlation

Without an effective compression solution, storing the plenoptic illumination function is impractical. If only the illumination parameter of the function is allowed to change (while the viewpoint, the viewing direction and the scene are all static), it is very likely that the radiance values received along the same viewing vector are strongly correlated because all geometric factors are frozen. Geometry is usually the major source of discontinuity in radiance values. Moreover, these radiance values are strongly related to the surface reflectance. Hence we first group together those radiance values related to the same pixel (or the same viewing ray) and exploit this *intra-pixel data correlation*. These values are indexed only by the light vector $\vec{L}$. In other words, it is a spherical function.

To compress them, we apply the spherical harmonic transform [10] to this spherical function. The resultant spherical harmonic coefficients are zonally sampled, quantized, and stored (Figure 3). Spherical harmonic transform has been used for compressing surface BRDF in several previous work [11]. The detailed formulae of spherical harmonic transform are described in Appendix. Spherical harmonic

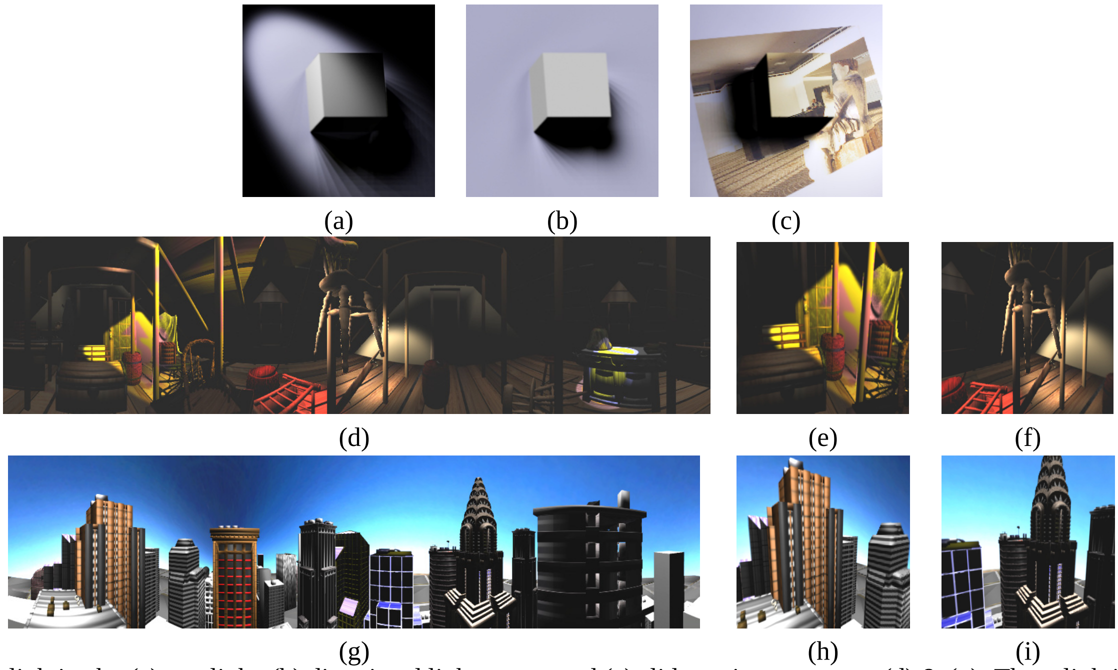


Fig. 2. Relighting by (a) spotlight, (b) directional light source and (c) slide projector source. (d) & (g): The relighting can be applied to panoramic images as well as perspective images. (e), (f), (h) & (i): The perspective snapshots.

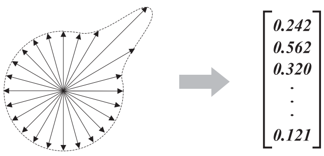


Fig. 3. Spherical harmonic transform.

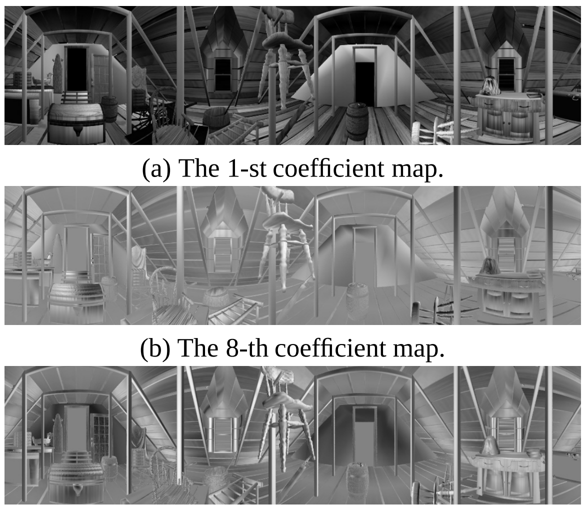
transform can be regarded as Fourier transform in the spherical domain. Just like Fourier transform, the more coefficients are used for representation, the more accurate is the reconstructed value.

The optimal (in term of image quality) number of spherical harmonic coefficients used for compression depends on the image content. Images containing specular object require more coefficients than images with only diffuse objects. Images containing shadow also require more coefficients to represent than the one without shadow. In most of our tested scenes, 25 coefficients are usually sufficient.

5.2. Inter-Pixel Correlation

Up to now, the data correlation between adjacent pixels has not yet been exploited. We pick the first coefficients from all coefficient vectors and form a map, the *SH map*. The same grouping process is applied to the second coefficients and all other coefficients. The result is a set of k SH maps if the coefficient vectors are k-dimensional. In fact, each SH map is an image. Figure 4 shows the 1-st, the 8-th and the 16-th SH maps of the red channel.

We can simply treat each SH map as an ordinary image and apply standard image compression to utilize the *inter-pixel correlation*. The SH maps are subdivided into 8×8 blocks and then DCT-encoded. It is well known that the visual artifact of a DCT-encoded image appears in the region with high contrast. This kind of artifact is not apparent in our result (see the visual comparison in Figure 5). One reason is that the DCT is applied in frequency domain (spherical harmonic domain) instead of spatial domain. For the attic panorama, it takes about 1.2 hours to encode all SH maps by a pure software DCT encoder running on SUN Ultra 5/270MHz.



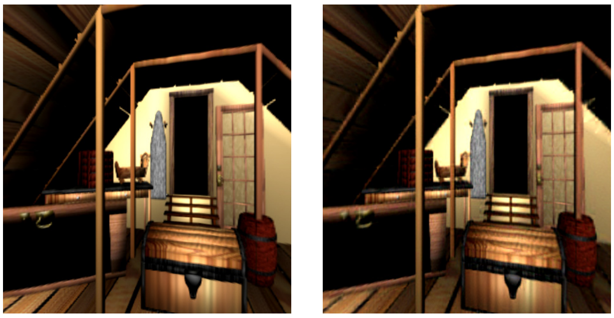
(a) The 1-st coefficient map.

(b) The 8-th coefficient map.

(c) The 16-th coefficient map.

Fig. 4. Coefficient maps.

Together with the DCT-based intra-pixel compression, a high compression ratio is achieved. For a 1024×256 panoramic image sampled under 300 illumination configu-



(a) (b)

Fig. 5. Visual comparison of images from compressed data. (a) No compression, and (b) DCT compressed.

rations, the compressed data requires only 3-4MB of storage which is obviously practical for Internet transfer.

6. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we describe an extended formulation of the plenoptic function that allows the explicit specification of illumination component. Techniques based on this new model can record and interpolate, not just viewing direction and viewpoint, but also the illumination. The trade-off is the enormous storage requirement. A compression technique is applied to exploit both intra-pixel and inter-pixel data correlations. A high compression ratio is achieved.

There is a lot of work to be done in the future. Currently, we only extract the illumination component. No other scene changing factors are investigated. If other factors are extracted, the rigidity of image-based computer graphics can be further relaxed. However, the trade-off is the further increase in data size. Another direction is to further compress the data by exploiting more sophisticated approximation methods of BRDF, such as non-linear approximation [12]. If the parameters computed by these methods also exhibit strong correlation, inter-pixel compression may also be applied to further reduce the storage.

7. WEB AVAILABILITY

Interested readers can download our interactive panoramic viewer with relighting ability at `http://www.cse.cuhk.edu.hk/~ttwong/demo/panoview/panoview.html`.

8. REFERENCES

[1] Edward H. Adelson and James R. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*, Michael S. Landy and J. Anthony Movshon, Eds., chapter 1, pp. 3–20. MIT Press, 1991.

[2] Leonard McMillan, *An Image-based Approach to Three-Dimensional Computer Graphics*, Ph.D. thesis, University of North Carolina at Chapel Hill, 1997.

[3] Shenchang Eric Chen, "QuickTime VR - an image-based approach to virtual environment navigation," in Computer Graphics *Proceedings, Annual Conference Series, SIGGRAPH'95*, August 1995, pp. 29–38.

[4] Marc Levoy and Pat Hanrahan, "Light field rendering," in Computer Graphics *Proceedings, Annual Conference Series, SIGGRAPH'96*, August 1996, pp. 31–42.

[5] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen, "The lumigraph," in Computer Graphics *Proceedings, Annual Conference Series, SIGGRAPH'96*, August 1996, pp. 43–54.

[6] Tien-Tsin Wong, Chi-Wing Fu, Pheng-Ann Heng, and Chi-Sing Leung, "The plenoptic illumination function," *IEEE Transactions on Multimedia*, vol. 4, no. 3, pp. 361–371, September 2002.

[7] Zhouchen Lin, Tien-Tsin Wong, and Heung-Yeung Shum, "Relighting with the reflected irradiance field: Representation sampling and reconstruction," *International Journal of Computer Vision*, vol. 49, no. 2-3, pp. 229–246, September-October 2002.

[8] Wai-Man Pang, "A portable capturing system for image-based relighting," M.Phil. thesis, Department of Computer Science & Engineering, The Chinese University of Hong Kong, 2002.

[9] Tien-Tsin Wong, Chi-Wing Fu, and Pheng-Ann Heng, "Interactive relighting of panoramas," *IEEE Computer Graphics & Applications*, vol. 21, no. 2, pp. 32–41, March-April 2001.

[10] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Interscience Publisher, Inc., New York, 1953.

[11] Francois X. Sillion, James R. Arvo, Stephen H. Westin, and Donald P. Greenberg, "A global illumination solution for general reflectance distributions," in *Computer Graphics (SIGGRAPH '91 Proceedings)*, July 1991, vol. 25, pp. 187–196.

[12] Eric P. F. Lafortune, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg, "Non-linear approximation of reflectance functions," in *Computer Graphics Proceedings, Annual Conference Series 1997*. August 1997, pp. 117–126, ACM SIGGRAPH.

A. APPENDIX: SPHERICAL HARMONICS

To transform a spherical function $P_I(\theta, \phi)$ to spherical harmonic domain, we use the following equation,

$$C_{l,m} = \int_0^{2\pi} \int_0^{\pi} P_I(\theta,\phi) Y_{l,m}(\theta,\phi) \sin\theta \, d\theta d\phi, \qquad (3)$$

where $P_I(\theta, \phi)$'s are the sampled radiance values,

$$Y_{l,m}(\theta,\phi) = \begin{cases} N_{l,m} Q_{l,m}(\cos\theta)\cos(m\phi) & \text{if } m > 0 \\ N_{l,0} Q_{l,0}(\cos\theta)/\sqrt{2} & \text{if } m = 0 \\ N_{l,m} Q_{l,m}(\cos\theta)\sin(|m|\phi) & \text{if } m < 0, \end{cases}$$

$$N_{l,m} = \sqrt{\frac{2l+1}{2\pi}\frac{(l-|m|)!}{(l+|m|)!}},$$

and

$$Q_{l,m}(x) = \begin{cases} (1-2m)\sqrt{1-x^2}\,Q_{m-1,m-1}(x) & \text{if } l = m \\ (2m+1)x\,Q_{m,m}(x) & \text{if } l = m+1 \\ \frac{2l-1}{l-m}x\,Q_{l-1,m}(x) - \frac{l+m-1}{l-m}Q_{l-2,m}(x) & \text{otherwise.} \end{cases}$$

and $Q_{0,0}(x) = 1$.

$C_{l,m}$'s are the spherical harmonic coefficients which are going to be zonally sampled, quantized and stored. $Q_{l,m}(x)$'s are the Legendre polynomials. From Equation 3, the integral can be computed easily if the samples are taken at the grid points on a sphere. This is another reason we choose to sample on the grid points.

To reconstruct and interpolate the radiance value, the following summation of multiplications is calculated,

$$P_I^*(\theta,\phi) = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l} C_{l,m} Y_{l,m}(\theta,\phi).$$

where $(l_{\max}+1)^2$ is the number of spherical harmonic coefficients to be used.