ICICS-PCM 2003
15-18 December 2003
Singapore

Surface Reflectance Classifying under Natural Illumination

Laiyun Qing[1, 2]; Wen Gao[1, 2]; Shiguang Shan[2]

[1]Graduate School, Chinese Academy of Sciences
[2] ICT-ISVISION Joint R&D Laboratory for Face Recognition , , Chinese Academy of Sciences

Abstract

Though a point light source is more suitable to measure the BRDF of the surface, the natural illuminations in the real-world are not point light source and very complex. Fortunately, the complex natural illuminations exhibit some statistical regularity [3]. These statistical properties of the natural illuminations lead to predictable image statistics for a surface with given reflectance properties. We develop an algorithm for classifiying a surface according to its reflectance from a single photograph under unknown illumination by learning relationships between surface reflectance and certain features computed from the observed image. The statistics of the natural illuminations and the relationships learning are performed in frequency domain because the reflection equation is a rotational convolution and it is convenient to analyze it in space-frequency domain.

Key words: illumination statistics, reflection equation, BRDF estimation

1. Introduction

In the signal-processing framework proposed by Ramamoorthi and Hanrahan [8], the reflected light field (part of it is the image we see) is convolution of the illumination and BRDF (Bi-directional Reflectance Distribution Function) of the surface. The illumination is regarded as the input signal, the BRDF of surface is taken as the filter and the reflected light field is the output signal. According to this, a single point light source is suitable to measure the BRDF of the surface, because the input signal, a point light source is a Delta function and then the output is the filter, the BRDF of surface. This is the principle of many BRDF measurement systems, such as Marschner's [6].

However, the natural illuminations in the real world are not point light source but highly complex, consisting of reflected light from every direction as well as distributed and localized primary light sources. Therefore estimating the BRDF of the surface under the natural illuminations makes more sense.

It has been proven that the natural images exhibit statistical regularity, especially in the frequency and wavelet domains. The statistics of the natural image have been applied successfully in many applications such as image compression, image denoising and object recognition [2, 5, 7]. Inspired by this, we exploited the statistics of the natural illuminations. Fortunately, the natural illuminations in the real world do share some similar statistical properties, though they are high dynamic range, panoramic images. Dror has proved that this [3] and used it in surface reflectance estimation from a single photograph under unknown natural illumination [4]. Weiss [10] has decomposed a set of images of the same scene under different illumination into intrinsic "illumination" and "reflectance" images by assuming statistics on the illumination images.

The problem we study in this paper is as same as Dror's et al. [4]. They solved the problem by learning the reflectance of a surface and the statistics of its image depending on the spatial structure of the real world illuminations. Our method differs from Dror's in that we educed the features used for classification by exploiting the statistical properties of natural illumination in frequency based on the reflection equation. It has proved in [8] that BRDF estimation under unknown illumination was ill posed even with the knowledge of the total reflected light field because of the associativity of convolution. With statistical regularity of the natural illuminations (for example, the natural illuminations have high frequency signal as well as low frequency signal), estimation surface reflectance properties from a single photograph under natural illumination is practicable.

The remainder of this paper is organized as follows. The statistics of the natural illumination in frequency domain is described in the next section. Then we introduce the proposed algorithm in section 3. Section 4 shows the experimental results. Section 5 concludes our paper and discusses the future work.

2. Natural Illumination Statistics in Power Spectra

Many researchers have observed that 2D power spectra of the natural images typically fall off as $1/f^{2+\eta}$, where f represents the modulus of the frequency and η is a small constant that varies from scene to scene.

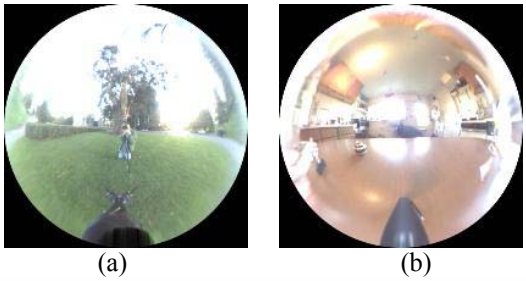
Spherical harmonic on 3D surface is equivalent Fourier transform on 2D planar. Spherical harmonics form a countable orthonormal basis for square integrable functions on the sphere. Associated with each basis function is an order l, a nonnegative integer analogous to frequency. The $2l+1$ spherical harmonics of order l span a space that is closed under rotation. If the regularity observed in the natural images statistics literature carries over to spherical illumination maps, the average power of the spherical harmonics at order l will fall off as $1/l^{2+\eta}$.

We computed spherical harmonic coefficients for every illumination map and obtained average power at each order l as the mean of squares of the coefficients at that order. The statistics was done on the nine illumination maps in Debevec's Light Probe Image Gallery (http://www.debevec.org/Probes/) and 200 nodes in MIT city scan project acquired by Teller et al. (http://city.lcs.mit.edu/data). The illumination maps in Debevec's Light Probe Image Gallery include four indoor settings and five outdoor settings. The illumination maps in MIT city scan project are the environments in MIT campus.

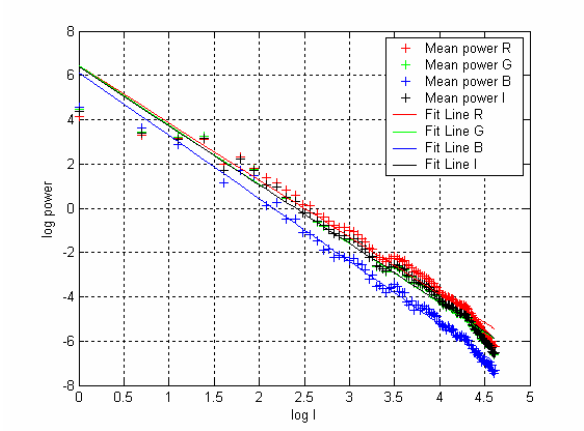
Figure 1 shows the relationship between average power and harmonic order for two illumination maps, one outdoor illumination and one indoor illumination, when pixel value is proportional to luminance. The fit line is $k/l^{2+\eta}$. We can see the power of R, G, and B channels of the illumination maps, outdoor or indoor, are almost the same, that is, the illuminations in the real world is approximate white. All the two images have power spectra that lie close to the fit line (the slope is $-(2+\eta)$) on log-log axes. But the match between the fit line and the actual power spectra in (d) is not as good as that in (c). This is because the illumination map in (b) contains intense, localized light sources that dominate the power spectrum and have smooth power spectra that remain flat at low frequencies before falling off at higher frequencies. Dror [3] has proved that if the brightest pixel values in these illumination maps are clipped, the match will be better.

The great majority of other illumination maps in both databases exhibit similar behavior. The differences between the power spectra of the illumination maps in MIT campus are smaller because they are all urban outdoor environments.

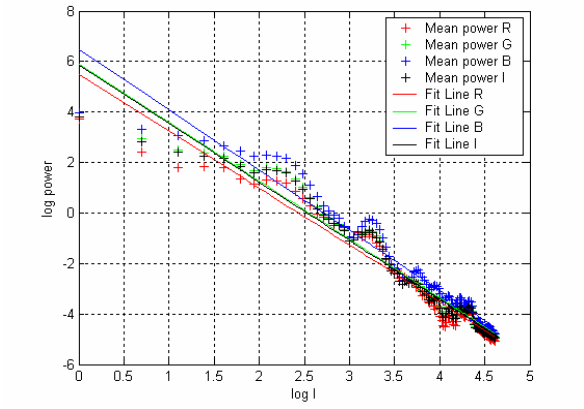
Two extreme illuminations are uniform light source and a single point light source. In the case of uniform source, only the power of first order is not zero; as for the single point light source, the average power of every order is same ($1/4\pi$). But we observe that these two cases are rare in the real-world.



(a) (b)



(c)



(d)

Figure 1: Spherical harmonic power spectra and the fit line of illumination maps. (a): an outdoor illumination map; (b): an indoor illumination map; (c) and (d) are the power spectra and the fitting line of (a) and (b). η is 0.65 for (c) and 0.31 for (d).

3. The Proposed Method

We first introduce the reflectance model and its representation in frequency domain in section 3.1 and then we educe the features to classify surface reflectance in section 3.2.

3.1 Reflectance model

We select Ward's parameterized reflectance model from computer graphics to gain insight into the relationship between illumination statistics, surface reflectance and

surface image statistics. The Ward reflectance model [9] is:

$$\rho(\theta_i', \phi_i', \theta_o', \phi_o') = \frac{\rho_d}{\pi} + \frac{\rho_s}{\sqrt{\cos\theta_o' \cos\theta_i'}} \frac{\exp(-\tan^2\delta/\sigma^2)}{4\pi\sigma^2} \cdot \quad (1)$$

Where δ is the angle between the surface normal and a vector bisecting the incident and reflected directions. ρ_d is the fraction of incident energy reflected by the diffuse component and ρ_s is the fraction of energy reflected by the specular component. σ is surface roughness measured as the standard deviation of surface slope. Higher σ implies a more blurred specular component. σ is usually small for real materials, i.e., $\sigma \le 0.2$.

For diffuse component, $\hat{\rho}_l$ are the spherical harmonics coefficients of the BRDF function $\cos\theta'$. The analytic formulae of $\hat{\rho}_l$ are given by [1, 8]:

$$\hat{\rho}_0 = \sqrt{\pi/4}\,,\ \hat{\rho}_1 = \sqrt{\pi/3}$$
$$\hat{\rho}_l = 0, l > 1, odd\,,$$
$$\hat{\rho}_l = 2\pi\sqrt{\frac{2l+1}{4\pi}} \times \frac{(-1)^{l/2-1}}{(l+1)(l-1)} \times \left[\frac{l!}{2^l[(l/2)!]^2}\right] \quad (2)$$
$$l > 1, even$$

It can be shown that $\hat{\rho}_l$ vanishes for odd values of $l > 1$, and even terms fall off very rapidly as l^{-2} .

As for specular component, the spherical harmonics coefficients of the BRDF with $\rho_s = 1$ can be approximated as [8]:

$$\Lambda_l \hat{\rho}_l \approx \exp[-(\sigma l)^2]. \quad (3)$$

The coefficients fall off as a Gaussian with width of order $\sqrt{1/\sigma}$. The specular component of BRDF behaves in the frequency domain like a Gaussian filter, with the filter width controlled by the roughness. When $\sigma = 0$, the coefficients of every order are equal to Λ_l .

3.2 Relationship between image statistics and surface reflectance

In the signal-processing framework [8], the illumination is regarded as input signal, the BRDF of surface is taken as filter and the image is the output signal. Surface reflectance classifying is analogical to estimate filter given the output signal. If the input signal has some statistical properties, the problem is solvable to some extent.

As the reflection equation is rotation rotational convolution, it is natural to analyze it in frequency-space. We computed the average power Spectra of the reflected light field B_l . One problem is that spherical harmonics transformation is global spatial support but only one image is available. As we don't make any bias for view direction, it is reasonable to map the image to the back part of the reflected light field.

We deduce the features to classify surface reflectance according to its parameters based on the signal-processing framework as follows.

■ **Features used to distinguish Lambertian from the others**

We define two features to distinguish Lambertian (pure diffuse) surfaces from the others. v_1 is the percent of the average power of the first three orders ($l = 0,1,2$). v_2 is the average ratio between the average power of consecutive odd order and even odd for $l \ge 3$. It is predictable that the value of v_1 of Lambertian surface are very large (usually larger that 99%) and the value of v_2 of pure diffuse surface are close to zero. This is because that $\hat{\rho}_l$ of Lambertian surface vanishes for odd values of $l > 1$, and even terms fall off very rapidly as l^{-2} . As for the other surfaces, the values of v_1 will not so large and the values of v_2 are usually larger than 1 because the average power of the lower orders (odd order) is larger than that of the high orders (even order).

To illustrate this, the feature space of v_1 and v_2 for the eight surfaces is given in Figure 2. The images of the eight spheres and the values of their BRDF parameters are given in Figure 3 in section 4. The sky plastic spheres are close to Lambertian surface because the specuar component is much smaller compared with diffuse component ($0.025/0.78 = 0.03$).

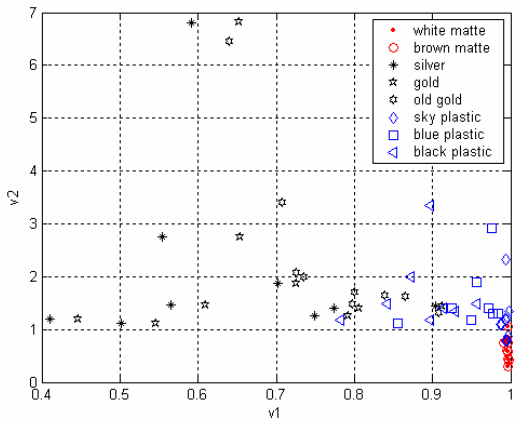


Figure 2. The feature space of v_1 - v_2 .

■ **Features used to distinguish different roughness**

The roughness of surface controls the blur of specular component and it is no meaning for Lambertian surface.

Because $\hat{\rho}_l$ of diffuse component vanishes for odd values of $l > 1$, the powers in odd orders of $l > 1$ include only specular component. We define four features based on bandwidth filters for odd orders of $5 \le l < 15$, $15 \le l < 25$, $25 \le l < 35$, $35 \le l < 45$ respectively. The definitions the four features are similar except that the range of l are different. As an example, the feature v_3 is defined as:

$$v_3 = \frac{1}{5}\sum_l \left(B_l / B_3\right),\ 15 \le l < 25, odd\ . \qquad (4)$$

where B_l is the average power of order l of the reflected light field.

■ **Miscellaneous Features**

To completely distinguish the surfaces reflectance, we need to know the information such as ρ_s / ρ_{dI}, ρ_{dR} / ρ_{dI}, ρ_{dG} / ρ_{dI}, and ρ_{dB} / ρ_{dI}. Therefore we define some other features in this subsection.

The feature v_1 and v_2 has some information of ρ_s / ρ_{dI}. The smaller ρ_s / ρ_{dI} is, the larger the value of v_1 is, and the smaller the values of v_2 is. Because the smaller ρ_s / ρ_{dI} is, it looks more like a Lambertian surface. Another feature to distinguish ρ_s / ρ_{dI} is defined as:

$$v_7 = \arg\min_L \left(\sum_{l=0}^{L} B_l \ge 98\% \right) . \qquad (5)$$

The larger the value of v_7 is, the larger ρ_s / ρ_{dI} is. The feature v_7 also has some information about the roughness σ. The larger the value of v_7, the smaller σ is, i.e., the surface is more smoother.

As for the color of the surface, that is ρ_{dR} / ρ_{dI}, ρ_{dG} / ρ_{dI} and ρ_{dB} / ρ_{dI}, we simple use the average power in the zero order of the respective color channel. This is because the color of the natural illumination is approximate white.

4. Experimental Results

To test the accuracy of the classifying, we use synthetic images and photographs. Synthetic images are rendered with Ward's Radiance package, which can be downloaded freely from http://radsite.lbl.gov/radiance/. The materials include the eight kinds of most popular materials in the real-world: matte plastics, smooth plastics, rough metals, smooth metals. The nine illumination maps we used are the Debevec's Light Probe Image Gallery. Figure 3 shows the rendered images of the eight spheres under the illumination maps in Figure 2 (b). The photographs we used are same with Dror's [4] (http://ai.mit.edu/people/rondror/sphere_photos/).

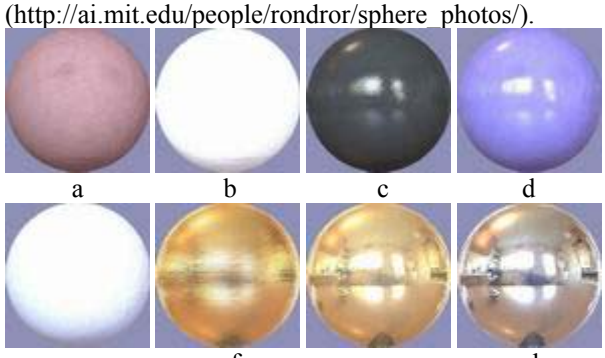


Figure 3: The synthesized images of eight spheres. The parameters of Ward model of the eight materials are follows ($\rho_{dR}, \rho_{dG}, \rho_{dB}, \rho_s, \sigma$): (a)brown matte: 0.2, 0.1, 0.1, 0, 0; (b) white matte: 0.7, 0.7, 0.7, 0, 0; (c) black plastic: 0.02, 0.02, 0.02, 0.025, 0.05; (d) blue plastic: 0.13, 0.1, 0.3, 0.05, 0.03; (e) sky plastic: 0.42, 0.42, 0.56, 0.025, 0; (f) old gold: 1, 0.7, 0.3, 1, 0.1; (g) gold: 0.8, 0.6, 0.3, 0.9, 0; (h) silver: 0.7, 0.7, 0.7, 1, 0. These parameters were chosen such that these materials include the main popular material and several spheres look similar.

There is ambiguity between the overall strength of illumination and the overall lightness of the surface. For example, a white matte sphere under dim illumination and a gray matte sphere under bright illumination will produce identical images. Human's eyes can adapt to the lightness of the environment quickly and the exposure time of camera can be adjusted to have best quality image. Therefore it is reasonable to eliminate this ambiguity by normalizing the images for mean strength of the illumination, as measured by the mean brightness of the image of the white sphere.

Because the number of available sample images is limited, we preformed a leave-one-out cross-validation. We left out the images under one illumination and using the others for training in turn. .

The classifier is very simple Nearest Neighbor. Table 1 lists the performance of synthetic images and photographs. It shows that the performance of the synthetic images is almost perfect while the performance of the real photographs is not so good. There are several reasons. The main reason is that the map for the back part of the sphere. The view direction of the synthesized images is forward and it is reasonable to map. While the view direction is 45 degree angle from the horizontal. The difference of the illumination between the upper part and the low part are very large while the difference between the frontal part and the back part is little in statistics. This leads large difference between the upper part and the low part of the reflected light field. The other reason is that the noises in photographs are larger than those of synthetic images, including the low dynamic range and the assumed parallel projection of photographs. It is expectable that the

performance of our system will be better if the entire reflected light field is known because no noise is introduced by mapping.

We have not yet performed experiments on the much larger database, the illumination maps in MIT city scan project yet. As listed in [4] by Dror, the performance is better than that of Devevec illuminations. This is because the illumination maps in the database are urban environments and they are much more similar with each other. We are sure that the performance will be better in our systems.

Table 1: The Results of Classifying

Databases	**Our Results**	**The Best Results of Dror by NN [4]**
Synthetic images	94.4%	78.8%
Photographs	53%	67.7%

All the four misclassifications in synthesized images are caused by gold and old gold. This is because they are so similar that the difference between the two materials is smaller than the difference between different illuminations.

5. Conclusion

The natural illuminations in the real-world exhibit statistical regularity which may facilitate reflectance estimation under unknown illumination, an otherwise ill-posed problem. This paper demonstrates the feasibility of reflectance classification from a single image in unknown real-world scenes based on the statistical prosperities in frequency of natural illumination. Although the performance our classification algorithm is impressive, it leaves a number of open research questions.

First, we have not solved the problem in an explicitly Bayesian manner. While we have analyzed the relationship between illumination statistics, image statistics, and reflectance, we do not translate the results directly to an optimal technique for reflectance classification. We would like to put our classification method on a more rigorous theoretical foundation.

Second, we believe that significant performance gains may be attained through the use of a more general feature set. We are focusing on space-frequency analysis and the features used in this paper are all based on power spectra of the image. Spherical harmonics transformation is time consuming due to its global spatial support. The view point and view direction the photographs of Dror's [4] make that it is not suitable to map the image to the back of the sphere therefore the performance descends compared with the synthesized images. It is more advisable to use wavelet coefficients for this purpose because wavelet localizes in both space and frequency domain. Some other features base on the pixels intensity used in Dror's [4] could be also incorporated into our classification scheme.

Acknowledgement

This research is partly sponsored by Natural Science Foundation of China (No.69789301), National Hi-Tech Program of China (No.2001AA114190 and No. 2002AA118010), ISVISION Technologies Co., Ltd and 100 Talents Foundation of Chinese Academy of Sciences.

References

[1] R.Basri and D.Jacobs. "Lambertian Reflectance and Linear Subspaces". Proc. Eighth IEEE International Conference Computer Vision, pp. 383-390, 2001.

[2] R. W. Buccigrossi and E. P. Simoncelli, "Image compression via joint statistical characterization in the wavelet domain," IEEE Transactions on Image Processing, Vol.8, pp.1688 - 1701, August 1999.

[3] R. O. Dror, T. K. Leung, E. H. Adelson, and A. S. Willsky, "Statistics of real-world illumination". Proc. IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, December 2001.

[4] R.O. Dror, E.H. Adelson, and A.S. Willsky, " Surface Reflectance Estimation and Natural Illumination," Proc. Second International Workshop on Statistical and Computational Theories of Vision at ICCV, Vancouver, CA, July 2001.

[5] J. Huang, Statistics of Natural Images and Models. PhD thesis, Brown University, May 2000.

[6] S. R. Marschner, S. H. Westin, E. P. F. Lafortune, and K. E. Torrance. "Image-based Bidirectional Reflectance Distribution Function Measurement," Applied Optics, vol. 39, no. 16 (2000).

[7] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Adaptive Wiener denoising using a Gaussian scale mixture model in the wavelet domain," Proc. 8th International Conference on Image Processing, Thessaloniki, Greece, October, 2001

[8] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering," Proc. ACM SIGGRAPH 2001, pp.497-500, Los Angeles, California, August 2001.

[9] G. J. Ward, "Measuring and modeling anisotropic reflection," Computer Graphics (SIGGRAPH), pp.265-272, 1992.

[10] Y. Weiss, "Deriving intrinsic images from image sequences," Proc. IEEE Conference on Computer Vision, volume II, pages 68–75, Vancouver, Canada, July 2001.