

Applying Shape from Lighting Variation to Bump Map Capture

Holly Rushmeier, Gabriel Taubin, and André Guézic

IBM TJ Watson Research Center, Yorktown Heights, NY 10598 USA

Abstract: We describe a system for capturing bump maps from a series of images of an object from the same view point, but with varying, known, illumination. Using the illumination information we can reconstruct the surface normals for a variety of, but not all, surface finishes and geometries. The system allows an existing object to be rerendered with new lighting and surface finish without explicitly reconstructing the object geometry.

1 Introduction

Bump maps, first introduced by Blinn [4], are used in computer graphics to store complex geometry at a physical scale between explicit geometry representation and the surface microscale effects encapsulated in the bidirectional reflectance distribution function (BRDF) [2]. A bump map consists of an array of normals that are applied as perturbations to a base geometry. Bump maps are most frequently defined procedurally (e.g. see [13]). There are few reports of bump maps measured directly from existing objects. In [5] a bump map for tree bark was measured by taking an X-ray of a plaster cast of some bark, and using the resulting depth map to compute a bump map. At least on some scales, bump maps could be measured using laser scanners, reconstructing the geometry and computing the normals. In this paper we describe a simpler method for capturing the bump map from an existing object without reconstructing the object geometry using the idea of shape from lighting variations.

In computer vision there is a long history of “shape from” various lighting effects – such as shape from shading [10] and shape from specularities [9]. In the past few years the vision community has also explored the area of “appearance based matching.” In particular, various researchers have sought illumination basis images (similar in spirit to the basis images used by Nimeroff et al. [15]), that can be used to construct an image of an object under arbitrary lighting conditions [3], [14]. Recently Epstein et al. [6] used the idea of illumination basis images to develop methods for obtaining object shape from lighting variations. They addressed the general problem of obtaining geometry given a series of images of an object under unknown illumination conditions. We use a simple form of their ideas to design a system for obtaining bump maps from small sets of images captured under controlled lighting conditions.

A number of recent papers have addressed the problem of obtaining input for rendering using inexpensive equipment. Karner et al. [11] describe a method for measuring anisotropic reflectance using just a CCD camera, quartz-halogen lamp and a diffuse reference standard. Gortler et al. [7] and Levoy and Hanrahan [12] describe methods for rerendering the appearance of an object for fixed lighting conditions from a large series of video images. Ofek et al. [16] have developed a method for obtaining texture maps from non-planar, shiny, objects using a series

of video images. The bump map capture system described here complements these inexpensive techniques for obtaining complex input data for rendering.

2 Shape from lighting variations

The work by Epstein et al. for obtaining surface geometry from lighting variations begins with a simple model for the reflected radiance from a point on a surface. Assuming a Lambertian surface illuminated by a distance small source the reflected radiance L_r is given by:

$$L_r = \rho(L_o \Delta\omega / \pi) \hat{n} \cdot \hat{l} \quad (1)$$

where ρ is the Lambertian reflectance, L_o is the light source radiance, $\Delta\omega$ is the solid angle subtended by the light source, \hat{n} is the surface normal and \hat{l} is the direction to the light source (see Figure 1).

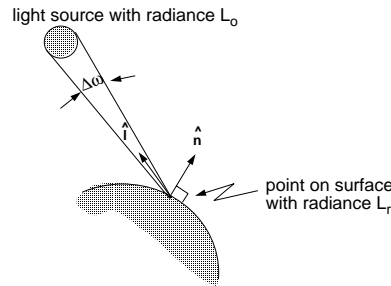


Fig. 1. Geometry of light reflectance, using the symbols defined in Eq. 1

Suppose there are three images of the object from the same view point, for three different light sources. Assume each source has the same radiance, subtends the same solid angle, and that the directions to all the sources are known. The images give us three relative values for the reflected radiance $\alpha L_{r,1}$, $\alpha L_{r,2}$, $\alpha L_{r,3}$. The values are relative (with an unknown scaling factor α , assuming approximately linear camera response) because the images are captured with a value of 0 to 255, rather than with a value for absolute radiance. We can form the following matrix equation for the surface normals \hat{n} :

$$\rho(L_o \Delta\omega / \pi) \begin{bmatrix} l_{1,1} & l_{1,2} & l_{1,3} \\ l_{2,1} & l_{2,2} & l_{2,3} \\ l_{3,1} & l_{3,2} & l_{3,3} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} \alpha L_{r,1} \\ \alpha L_{r,2} \\ \alpha L_{r,3} \end{bmatrix} \quad (2)$$

This allows us to solve for $(\rho L_o \Delta\omega / \alpha \pi) \hat{n}$. Since the magnitude of \hat{n} is one, from this result we can obtain the normal vector \hat{n} and a value $\rho_{rel} = \rho L_o \Delta\omega / \alpha \pi$ proportional to the reflectance ρ . Repeating this calculation for all visible points on the surface (i.e. for all pixels) results in estimates for the surface normal and relative reflectance ρ_{rel} . This approach is traditional photometric stereo, originally introduced by Woodham [10].

There are two well-known problems with traditional photometric stereo:

- There will be some points that will be in shadow for one or more of the images. That is, there will not always be three non-zero values of L_r to use in solving Eq. 2.
- Surfaces are not truly Lambertian and can have a specular component. That is, Eq. 1 may be a very bad model for the reflected radiance.

The computer vision literature has many approaches to deal with aspects of these problems (e.g. [17].) To design a system with simple hardware and software, we follow the approach introduced by Epstein et al.

For the problem of shadows, they identify two types of shadows – attached and cast. Points lie in attached shadows when their surface normal has a dot product of less than zero with the light source direction. Points lie in cast shadows when another part of the object blocks the light source. If we could tell when a point was in an attached rather than cast shadow, that would give us some information about the surface normal. However, since we cannot differentiate between shadow types in the image, all pixel values with low values of αL_r must be eliminated from the calculation, and more than three images will be required.

For surfaces with a specular component, they note that the highlights resulting from specular reflection have very high values of αL_r . By excluding very high pixel values, specular effects can be eliminated from the calculations. A similar approach is taken by Ofek et al. [16] to eliminate specular highlights by taking the mode of the color values recorded for a particular point on a surface to recover a texture map.

Epstein et al. go on to consider the problem of recovering geometry from the surface normals given various assumptions, and the general problem of recovering shape when the illumination in each image is not known. They show results for various cases for the problem of capturing the geometry of faces.

We now consider a different problem – the design of a system for capturing bump maps using the basic idea of employing Eq. 2 and discarding high and low values to avoid specular highlights and shadows. In addition to applying our system to relatively large scale geometries, we will also present results from our system for finer scales such as quilting and tufting on surfaces.

3 System design

Our system essentially consists of a small set of light sources, a fixed video camera and a video capture board. In setting up the system and understanding its limitations a number of factors were considered that were not discussed in [6], or were mentioned only briefly.

3.1 Design factors and limitations

Non-Lambertian surfaces Specular reflection is far from the only variation from Lambertian reflectance observed in real surfaces. Clearly, totally specular surfaces, with no diffuse component, can not be used with this approach. Furthermore, many surfaces have a “directional diffuse” component (see ref. [8]) that is not uniform. However, in many cases this component begins to have large variations in the BRDF value only for near grazing angles of incidence and

or viewing. As long as grazing angles are avoided in the measurement, reasonable estimates of the surface normals can be obtained, because the value of the reflectances is *approximately* constant. Using the strategy of discarding the highest and lowest image pixel values will assist in avoiding using data from grazing angles where reflectance deviates the most from the mean value.

For arbitrary objects, it is impossible to completely avoid near grazing viewing angles. However, for geometries that are truly suitable for representing as bump maps, near grazing angles will be avoided. As shown in Figure 2, an ideal geometry for bump map representation has small gentle variations in height relative to the size of the other two dimensions of the geometry, so that cast shadows and uneven silhouettes are not noticeable.

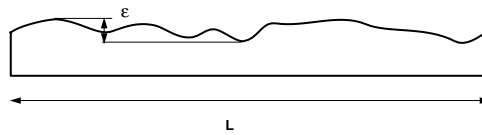


Fig. 2. An ideal surface to be represented as a bump map has small variations ϵ in one dimension relative to the scale of the object L in the other two dimensions.

The most difficult surfaces are those with a medium sized reflectance lobe in the specular direction. Narrow specular peaks are eliminated by not using high pixel values. Wide lobes such as those shown in [8] for roughened magnesium oxide at a wavelength of 0.5 microns can be sampled at approximately uniform values by avoiding grazing angles and excluding high and low values. However, surface BRDF's with medium sized lobes such as that shown in [8] for roughened aluminum at 0.5 microns will give very poor results using the lighting variation method.

The size of the highlights caused by directional reflections depends both on the width of the lobe in the BRDF and on the solid angle subtended by the light source. The area of an object that will be affected by directional reflections can be decreased by decreasing the size of the light source. We want to have the smallest light source possible that still gives adequate illumination to capture an image.

Inter-reflections Equation 1 accounts only for direct illumination, and not for indirect illumination by multiple interreflections. In designing a system we can clearly minimize indirect illumination by draping everything in the test area with black fabric. We cannot eliminate indirect illumination from the object itself. We can not expect good results for objects with deep holes, in which points on the surface can receive illumination from multiple interreflections. Even in the case of surfaces with concavities however, the effect of interreflections will be low if the surface reflectance is relatively low.

Light Source Position Epstein et al. note that the method will break down if the light source is too close. More generally, the method will break down if the

illumination of the object is not uniform. We need a source that does not itself project a pattern (i.e. patterns of light and dark that are generally produced by a flashlight bulb are not acceptable.) We also need the object to be far enough away so that $\frac{1}{r^2}$, the inverse square of the distance to the light source is nearly constant. We use as a guide the five times rule used to assess when a light source can be treated as a point [1]. That is, to capture objects of diameter D , our light sources must be a distance of at least $5D$ away.

Penumbrae For area light sources, points are not simply in or out of shadow regions. They may lie in a penumbral region in which only a portion of the light source is visible. In images, we can not differentiate between grey levels caused by surface normals turned away from the light source, and grey levels caused by a partial view of the light source. To minimize penumbral regions we again want the smallest possible light source so that the boundaries of cast shadows will be sharp.

In general, all shadows are undesirable for our measurements. Even for gently sloping nearly flat surfaces, at lighting angles near grazing there will be cast shadows of significant size. As in the case of BRDF variations, reducing cast shadows calls for avoiding lighting directions that are near perpendicular to the camera view direction.

However, for nearly flat surfaces, minimizing the angle between the light source and the view angle will increase the likelihood of having specular highlights. While it is desirable to have one light nearly co-located with the camera, the other lights should be spread out to strike a balance of being not too close to the viewing angle and not too close to perpendicular to the viewing angle.

3.2 Physical set-up

Taking into account the factors just discussed, we developed the design shown in Figure 3. We use five light positions. This gives us five image values from which we can discard the high and low values and compute the normal from the remaining three values. We place one light source next to the camera, and the other four light sources as near as possible to 45 degrees to the camera view direction as our test space would allow. Given our test space, our lights were about 35 inches from the test object, allowing us to capture objects of up to 7 inches in diameter.

For light sources, we used a small frosted 25 watt, 210 lumen, tungsten filament bulb. This is an inexpensive, commonly available bulb that is small (diameter is approximately 2.5 inches) and gives enough light to capture images.

The entire set up is draped in black felt. We used a SONY CCD XC-999 color video camera module, connected to a MRT Videopad Professional capture board on an IBM PC 750. We attached a white card painted with white diffuse paint to test the approximate consistency of incident illumination from the various light source positions. Since the center light source position is somewhat closer to the object than the other positions, the pixel values from the center lit images had to be slightly reduced in post processing to be consistent with the other images. The same aperture setting is used on the camera for each series of five images. Other than the white card and the manual setting of the aperture for a particular object, we did no other calibration or adjustments for our initial series of tests.

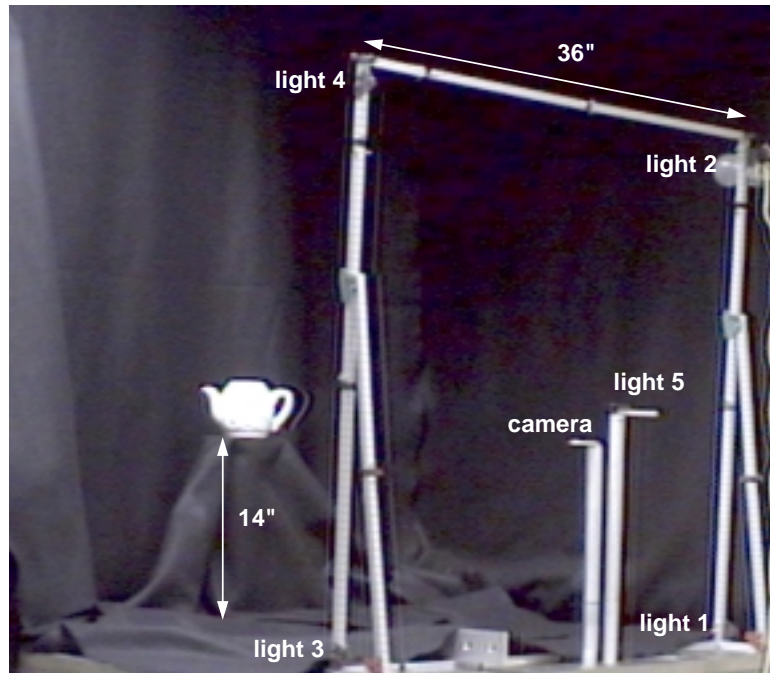


Fig. 3. Physical setup used for capturing bump maps.

3.3 Software

The software for processing the captured images has three parts – a system for image cropping, filtering, thresholding and conversion to greyscale, a program for calculating the normals and relative reflectance map, and a viewer for looking at the resulting bump maps with different light directions.

The initial images need to be cropped to include just the object of interest. A median filter is then used to remove random pop noise in the images. A threshold value is determined manually to remove the non-zero pixel values for the black felt background. Finally, the images are converted to greyscale. The cropping and threshold value are determined interactively for one of the captured images, using the IBM Visualization Data Explorer (DX) program shown in Figure 4. The other four captured images are simply processed by the same program with the same settings.

Pseudo-code for calculating the normals is shown in Figure 5. To calculate the normals, the matrix in Eq. 2 needs to be inverted. Since there are only five light source directions from which three are chosen, there are only 10 unique matrices to invert, and these are computed first. Next the program processes each pixel location. The five values are sorted, and the middle three are chosen. If they are all non-zero, the normal and relative reflectance are computed. If they are not all non-zero, a normal cannot be computed, and there will be a hole in the final results.

Finally, the results are observed in the DX program shown in Figure 6. The

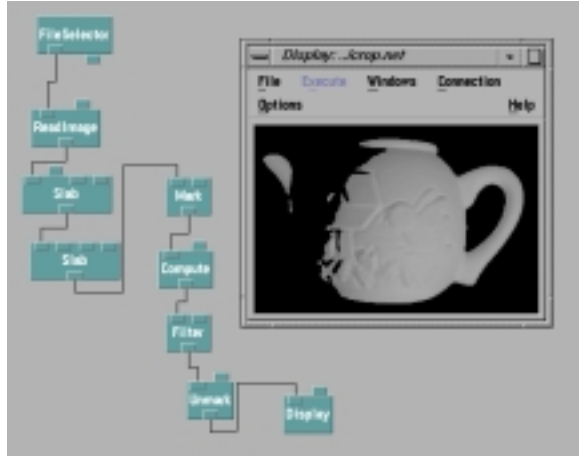


Fig. 4. Visual program used to crop and threshold images.

program takes the computed normals and produces an image from the dot product of the normals and a selected light direction. Options for closing holes where there was insufficient data can be examined. For small one or two pixel holes, applying a median filter to the bump map effectively fills in the missing values. The relative reflectance map can also be displayed. By scaling any of the captured images so that the relative RGB values are maintained but the grey level is given by the relative reflectance map, a flattened texture map of the object can be obtained. That is, pixel values $R'G'B'$ for pixel n in the flattened map are given by:

$$[R', G', B']_n = \rho_{rel,n} [R, G, B]_n / greylevel([R, G, B]_n) \quad (3)$$

4 Results

Images a-e in Figure 7 show a piece of quilted fabric captured with the five different light sources. The images are shown before scaling to account for the image taken from the center camera position being closer and so having overall higher pixel values. Figure 7f shows the bump map computed from the image illuminated from a new direction and with a white reflectance. Figure 7g shows the relative reflectance map. Some artifacts are visible in the upper and lower right hand corners of Figure 7f and g. These were caused by ill-conditioning in the light source matrices. Referring to Fig. 3, no set of three light source directions lie on a single plane. However, in practice light source combinations 1,4,5 and 2,3,5 were close enough to coplanar to produce an ill-conditioned matrix that produced poor results when these light source combinations were chosen.

The sample image shown in Figure 4 is one of 5 images of the teapot shown in Figure 3. The teapot is shown rerendered from a new lighting direction in Figure 6. Besides capturing bump maps, the method can be used to rerender larger scale objects with new surface finish and lighting.

```

Compute inverse matrices  $M[i][j][k]$  for unique combinations
of 3 light directions  $i, j, k$ ;

For each image pixel{
  Sort the 5 values  $I_1 \dots I_5$  into the list  $I'_1 \dots I'_5$ 
  and create the list  $L'_1 \dots L'_5$  of the indices
  of the associated light sources;

  If  $(I'_2, I'_3, I'_4)$  are not equal to zero{
     $[N'_1, N'_2, N'_3] = M[L'_2][L'_3][L'_4] * [I'_2, I'_3, I'_4]$ ;
     $\rho_{rel} = \text{magnitude } [N'_1, N'_2, N'_3]$ ;
     $\hat{n} = [N'_1, N'_2, N'_3] / \rho_{rel}$ 
  }
  Else{
     $\rho_{rel} = 0$ ;
     $\hat{n} = [0., 0., 0.]$ ;
  }
}

```

Fig. 5. Simple pseudo-code for processing images.

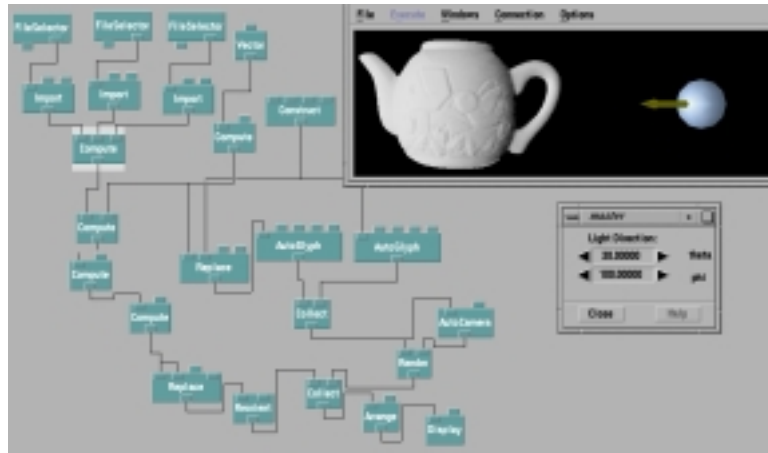


Fig. 6. Visual program used to view bump maps with varying lighting.

The left image in the color plate in the Appendix shows a captured image of a teak trivet that has white specular highlights. The image on the right shows two instances of the bump map computed for the trivet with a metallic finish applied. The patterns of light and dark are different on the two instances of the trivet, which would not be the case if a simple texture (rather than bump) map had been used.

5 Conclusions

We have demonstrated a simple system for obtaining bump maps from existing physical objects. The resulting maps can be used to re-render objects without

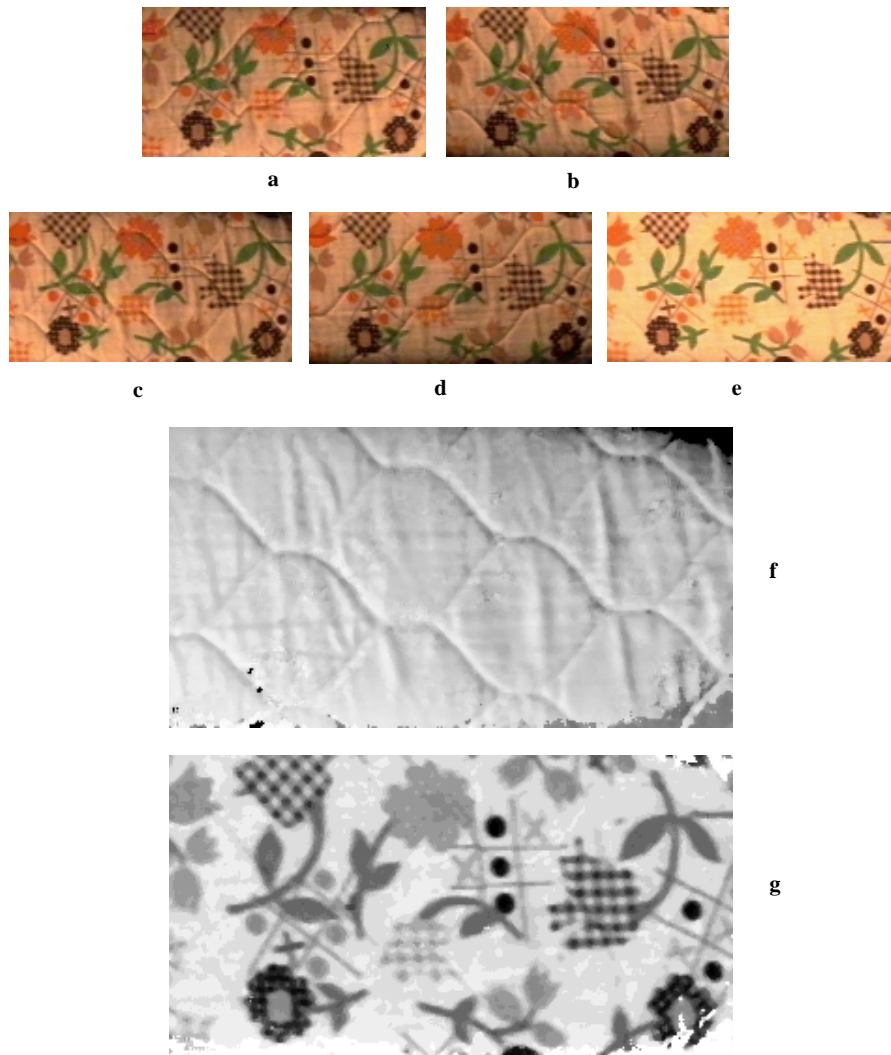


Fig. 7. A set of captured images (a-e) and the resulting bump map (f) and relative reflectance map (g.)

reconstructing the original geometry. The maps can also represent fine scale and self-shadowing geometries that would be difficult to recover from laser scans or traditional photometric stereo. Because of the assumptions required, the system will not work with objects with surfaces that do not have some range of BRDF values that are approximately uniform. The system will not work for objects with deep holes that can result in multiple intra-object interreflections.

Our system is a first effort. Clearly by using different combinations and numbers of lights there is the potential for obtaining additional rendering input data using lighting variations and video capture.

References

1. ASHDOWN, I. *Radiosity: A Programmer's Perspective*. John Wiley & Sons, Inc., 1994.
2. BECKER, B., AND MAX, N. Smooth transitions between bump rendering algorithms. *Computer Graphics (SIGGRAPH '93 Proceedings)* (1993), 183–189.
3. BELHUMEUR, P., AND KRIEGMAN, D. What is the set of images of an object under all possible lighting conditions? *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 1996* (1996).
4. BLINN, J. Models of light reflection for computer synthesized pictures. *Computer Graphics (SIGGRAPH '77 Proceedings)* 11, 2 (July 1977), 192–198.
5. BLOOMENTHAL, J. Modeling the mighty maple. *Computer Graphics (SIGGRAPH '85 Proceedings)* (July 1985), 305–312.
6. EPSTEIN, R., YUILLE, A., AND BELHUMEUR, P. Learning object representations from lighting variations. *ECCV 96 International Workshop* (April 1996), 179–199.
7. GORTLER, S., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. The lumigraph. *Proceedings of SIGGRAPH 1996* (1996), 43–54.
8. HE, X., TORRANCE, K., SILLION, F., AND GREENBERG, D. A comprehensive physical model for light reflection. *Computer Graphics (SIGGRAPH '91 Proceedings)* (1991), 175–186.
9. HEALEY, G., AND BINFORD, T. Local shape from specularity. *Computer Vision Graphics and Image Processing* (April 1988), 62–86.
10. HORN, B., AND BROOKS, M. J. *Shape from Shading*. MIT Press, 1989.
11. KARNER, K., MAYER, H., AND GERVAUTZ, M. An image based measurement system for anisotropic reflection. *Proceedings of EUROGRAPHICS 96, Computer Graphics Forum* 15, 3 (1996), 119–128.
12. LEVOY, M., AND HANRAHAN, P. Light field rendering. *Proceedings of SIGGRAPH 1996* (1996), 31–42.
13. MIYATA, K. A method of generating stone wall patterns. *Computer Graphics (SIGGRAPH '90 Proceedings)* (August 1990), 387–394.
14. NAYAR, S., AND MURASE, H. Dimensionality of illumination in appearance matching. *Proceedings of the 1996 IEEE International Conference on Robotics and Automation 2* (1996), 1326–1332.
15. NIMEROFF, J., SIMONCELLI, E., AND DORSEY, J. Efficient rerendering of naturally illuminated scenes. *Proceedings of the Fifth Eurographics Rendering Workshop* (1994), 359–373.
16. OFEK, E., SHILAT, E., RAPPOPORT, A., AND WERMAN, M. Multiresolution textures from image sequences. *IEEE Computer Graphics and Applications* (1997), 18–29.
17. TAGARE, H. D., AND DEFIGUEIREDO, R. J. P. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 2 (February 1991), 133–152.