Base Materials for Photometric Stereo

David Tingdahl[1], Christoph Godau[2], and Luc Van Gool[1]

[1] Katholieke Universiteit Leuven, ESAT/IBBT/VISICS, Belgium
[2] Technische Universität Darmstadt, IDD, Germany

Abstract. Image-based capture of material appearance has been extensively studied, but the quality of the results and generality of the applied methods leave a lot of room for improvement. Most existing methods rely on parametric models of reflectance and require complex hardware systems or accurate geometric models that are not always available or practical. Rather than independently estimating reflectance properties for each surface point, it is common to express the reflectance as a combination of base materials inherent to each particular object or scene.

We propose a method for efficient and automatic extraction of base materials in a photometric stereo system. After jointly estimating per-pixel reflectances and refined surface normals using these materials, we can render photo-realistic images of complex objects under novel lighting conditions in real time.

1 Introduction

After decades of development, the acquisition of object geometry using specialized scanners and image-based methods is now a relatively mature technology. The accurate capture of surface appearance and reflectance on the other hand, remains a major challenge for all but the simplest materials. Image-based capture of material appearance has been extensively studied [1], but the quality of the results and generality of the applied methods leave a lot of room for improvement. In particular, most existing methods rely on parametric models of reflectance and require complex hardware systems or accurate geometric models that are not always available.

Leaving aside the complex effects of subsurface scattering, transparency and inter-reflections, the appearance of a homogeneous material can be accurately described by its bidirectional reflectance distribution function (BRDF). This four-dimensional function defines the percentage of light reflected for every possible combination of incoming and outgoing light directions.

The acquisition of an independent BRDF for each surface location is not practical. A high enough sample density would require thousands of images for a single object. Rather than estimating a full BRDF for each surface point, we can express the reflectance as a combination of BRDFs from base materials inherent to each particular object. The variation of materials and surface geometry across the object allows for a reasonably dense sampling of these base materials from a relatively small number of images. This concept is illustrated in Fig. 1, where a strawberry is reconstructed using the method presented in this paper.

A. Fusiello et al. (Eds.): ECCV 2012 Ws/Demos, Part II, LNCS 7584, pp. 350–359, 2012.

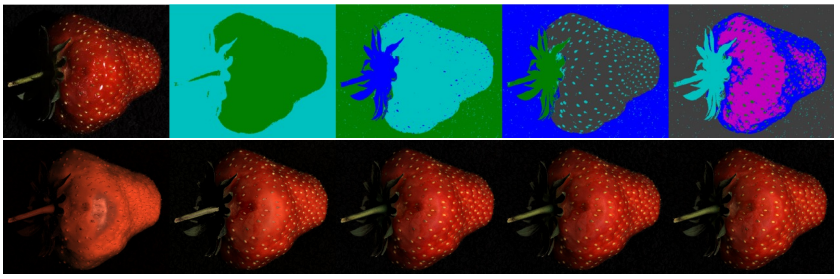

Fig. 1. Material clusters and reconstructed images with the number of base materials varying from 1 (left) to 5 (right). The ground truth image is in the top-left.

After reviewing related works in Section 2, we propose a method for efficient and automatic extraction of base materials from images in Section 3, which are then used to enhance results obtained from a traditional photometric stereo system in Section 4. We show that our conceptually simple, data-driven reflectance descriptor outperforms model-based approaches for material segmentation in terms of robustness to errors in the estimated geometry (Section 5.2). Using the captured reflectance and geometry information, we can produce photo-realistic renderings under arbitrary lighting conditions in real time (Section 5.3).

2 Background and Related Works

Photometric stereo infers information about a surface from images taken under varying light directions. The traditional approach [2] assumes Lambertian reflectance because of its simplicity and linearity. Surfaces violating this assumption can still be reconstructed by employing robustification to detect and discard outliers such as shadows and specular highlights [3,4]. More recent photometric stereo algorithms attempt to go beyond the Lambertian model by estimating a full bidirectional reflectance distribution function (BRDF) [5,6].

A BRDF describes the reflectance characteristics of a surface for varying directions of incoming (v_{in}) and outgoing (v_{out}) light and can capture complex phenomena including off-specular highlights and iridescence (goniochromatism).

BRDFs are commonly parametrized in spherical coordinates over the four-dimensional domain $(\theta_{in}, \phi_{in}, \theta_{out}, \phi_{out})$, where the surface normal is at the north pole ($\theta_{normal} = 0$) and the surface tangent (at $\phi = 0$) defines the orientation of an anisotropic surface.

We use the reparameterization proposed by Rusinkiewicz [7] based on half and difference angles: $(\theta_h, \phi_h, \theta_d, \phi_d)$. The halfway vector (θ_h, ϕ_h) is the vector bisecting v_{in} and v_{out}. The difference angles (θ_d, ϕ_d) are the spherical coordinates of v_{in} after rotating the halfway vector to the north pole.

Exploiting the symmetries and isotropy in commonly encountered materials Romeiro et al. [8,9] propose discarding the azimuthal components. This results in a bivariate BRDF representation using only (θ_h, θ_d), as illustrated in Fig. 2.

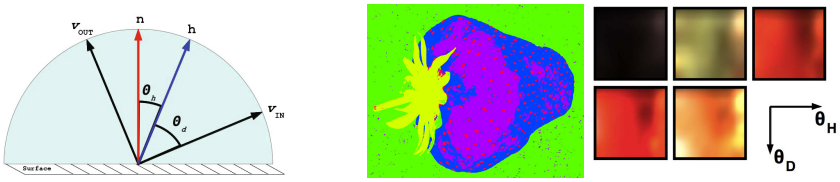

Fig. 2. Left: The bivariate BRDF model is parameterized by θ_h and θ_d with the normal pointing towards the north pole. Right: Base materials extracted from a strawberry, represented with the bivariate model.

They show that this simplification only results in a small decrease in rendered image quality for a wide range of materials.

Sophisticated hardware systems like the multi-camera, multi-projector system constructed by Weinmann et al. [10] achieve impressive results, but they are expensive, and their size makes them unusable for on-site acquisition. Fortunately, most real-world objects are composed of a limited number of base materials. The variation of normals and materials across the surface makes it possible to obtain a large number of samples for each base material from a relatively small number of images. As shown by Lensch et al. [11], combinations of these materials can describe the reflectance at each surface point. Additional constraints can impose spatial smoothness or limit the number of base materials for each combination [5,12,6,13].

In order to obtain reflectance data for these base materials, some methods make use of reference objects [5] or custom-made material charts [14] that have to be included in the scene, but this is not always possible or practical. Goldman et al. [6] cluster the albedos obtained from photometric stereo to obtain an initial set of base materials, but this is not directly applicable for specular surfaces [15], as their albedo is poorly defined. Other approaches [6,11,15] try to fit analytic reflectance models such as Ward's and Lafortune or hemispherical harmonics basis functions to each surface element and apply a clustering algorithm to their parameters. These models can be highly non-linear, and the resulting complexity and sensitivity to noise can be problematic, especially when geometry and lighting are not precisely known. Alldrin et al. [13] make use of the bivariate BRDF representation by globally optimizing for base materials, weights and normals. By employing a non-parametric reflectance descriptor for robust material clustering, we extend on their work by separating the tasks of material classification and optimization into two. This allows us to reconstruct high-resolution images of complex samples within an hour on a desktop PC.

3 Extracting Base Materials

The input to our algorithm is a set of M photometrically calibrated images $\mathcal{I}_1, ...\mathcal{I}_M$ acquired with a static camera and varying but known lighting directions. We assume an orthographic projection model and point-like light sources at infinity. These assumptions hold for objects which are small in comparison

to the distance to both the light source and the camera. We also assume that the target surface can be described with the bivariate BRDF model described in Section 2.

We find initial surface normals $\mathcal{N} = \{n_1, n_2, ..., n_N\}$ using Lambertian photometric stereo. Our implementation is robust to non-Lambertian characteristics and produces a usable, albeit noisy normal map even for specular and other intricate surface effects. In theory, the method can recover the normal if a surface point exhibits Lambertian characteristics for at least three of the images, which is the case for most materials [5].

Before we can extract full BRDFs for the base materials, we need to cluster the surface points into groups of similar reflectance. This requires a compact and robust representation of the local reflectance characteristics at each point, as well as a suitable distance measure.

3.1 Describing Local Reflectance

For every pixel, we have a series of RGB measurements $\mathcal{P} = \{p_1, ..., p_M\}$, one for each of the M light directions. The incoming light intensity varies with geometry and is compensated for by dividing each measurement with $cos\,(\alpha)$, where α is the angle between incoming light and surface normal.

We now sample individual bivariate BRDF descriptors for each surface element (pixel) from the set of RGB values in $\mathcal{P}$. First, the M light directions and the camera direction (v_{in} resp. v_{out} in Fig. 2) are rotated such that the normal is pointing towards the north pole. This gives us a pair of angular coordinates (θ_h, θ_d) for each measurement in $\mathcal{P}$. We sample $\mathcal{P}$ on a grid of $u \times v$ bins over the angular space by computing the average measurement for each bin. This is done separately for each color channel. Finally, we transform each sampled bin into the $L^*a^*b^*$ color space to render distance measurements more perceptually meaningful. One material descriptor d thus consists of a vector of $3 \times u \times v$ elements. The whole BRDF domain is rarely sampled at any one pixel and d is thus usually a sparse vector.

3.2 Clustering into Base Materials

We use k-means to cluster the descriptors into k clusters $\{c_1, c_2, ..., c_k\}$ of similar reflectance. For the moment, assume that k is known. Starting from a random labeling, we compute each cluster center as the mean of all its descriptors, discarding any missing values. For element j in cluster c, we have:

$$c_j = \frac{1}{\sum_{d \in c} g(d_j)} \sum_{d \in c} g(d_j) d_j, \quad j = [1, ..., 3uv]\,. \tag{1}$$

where $g(d_j)$ is an indicator function, returning 1 if element j exists in descriptor d and 0 otherwise. The dissimilarity between descriptor d and cluster c is computed as the mean squared euclidean distance between their overlapping elements:

$$dist(d,c) = \frac{1}{\sum_j [g(d_j)g(c_j)]} \sum_{j=1}^{3uv} g(d_j)g(c_j)(d_j - c_j)^2. \quad (2)$$

As d is sparse, Equation 2 is not suitable for directly computing the dissimilarity between two descriptors. However, since k-means only requires the distance between d and the denser c, this measure is applicable as long as the angular span of c is large enough to include d. To reduce the sparseness of d, we use a relatively small dimensionality of the sample grid ($u = v = 3$ in all our experiments). As our results show, this low angular resolution still allows for an efficient clustering of the reflectance space.

Since the output of k-means depends on the initial labeling, we run the clustering 25 times and select the labeling with the lowest error, measured as the accumulated distance between the descriptors and their corresponding cluster centers:

$$E = \sum_d dist\,(d, c_d)\,. \quad (3)$$

Note that the clusters themselves are not used as base material BRDFs due to their low angular resolution. Instead, we use the resulting labeling to sample high-resolution BRDFs, as explained in Section 4.1.

3.3 How Many Base Materials?

The number of base materials obviously varies depending on the object. Fig. 1 shows the reconstruction of a strawberry for various number of base materials, rendered using the output from our photometric stereo pipeline (as described in Section 4). Intuitively, this scene seems to consist of four materials: background cloth, red body, green leaf and yellow seed capsules. However, the visual quality for three and even two base materials is surprisingly realistic, while adding a fifth material does not appear to significantly improve the result. To determine the number of base materials, we use down-sampled (300×300 pixels) versions of the original images. Starting with a single material ($k = 1$), we increase k until convergence, i.e. until $E_{k+1}/E_k > 0.9$ using the error measure from Equation 3. This algorithm detects five base materials for the strawberry.

4 Photometric Stereo

4.1 BRDF Sampling

The low resolution BRDF descriptors are efficient for classification, but do not provide enough angular resolution for realistic rendering. To this end, we sample a high resolution BRDF from each detected cluster. Again, we use the bivariate representation over (θ_h, θ_d), but this time with a larger grid size (50×50 in our experiments). This results in a set of base material BRDFs: $\mathbf{B} = \{B_1, B_2, ..., B_k\}$. We interpolate empty bins and smooth the BRDF with a Gaussian kernel. Examples of extracted base materials can be seen in Fig. 2.

4.2 Computing Weight Map and Normals

We express each surface point as a linear combination of the base materials by computing a weight vector $w = [w_1, w_2, ..., w_k]^T$ for each pixel. The RGB values for base material j are given as a function $\mathbf{b}_j(n, l)$ of surface normal n (depending on the pixel) and light direction l (depending on the image). We can then render a pixel as a function of surface normal, light direction and material weights:

$$\hat{p}(n, l, w) = \sum_{j=1}^{k} \mathbf{b}_j(n, l) w_j \, cos(\alpha). \tag{4}$$

For each pixel, we then define a cost function C based on the ΔE^*_{ab} color difference between predicted ($\hat{p}$) and measured (p) values in each of the M images:

$$C(n, w) = \sum_{m=1}^{M} \Delta E^*_{ab}(\hat{p}(n, w, l_m), p) \tag{5}$$

where l_m is the (known) light direction in image m.

For each pixel, Equation 5 is minimized with respect to n and w using Levenberg-Marquard. The normals are initialized from the Lambertian reconstruction and the weights are initialized according to the k-means labeling: unity for the labeled material and zero for all others. We obtain the surface normals and material weight vectors that result in the lowest color difference across all images. Since the calculations are independent for each pixel, the computations are simple to parallelize.

5 Results

5.1 Acquisition System

We captured all images using a custom-built dome-shaped acquisition device. The device consists of 260 LEDs placed on a hemisphere, and a downward-looking camera at the top. Note that our method is not tied specifically to this device, for example reflective spheres could be used to obtain the light directions. For each sequence, we recorded 260 images. All images were used for the Lambertian initialization, while it proved sufficient to only use every third image for the remainder of the algorithm (87 images), significantly reducing computation time.

5.2 Robustness to Noise

In order to test the robustness of our algorithm, we use a simple object with only two materials: a matte piece of paper with a grid-like pattern of transparent glue shown in Fig. 3. The perceived color of the painted grid is very close to the original paper surface, but with severe specular highlights. Material clustering that takes into account the full BRDF should be able to separate these materials.

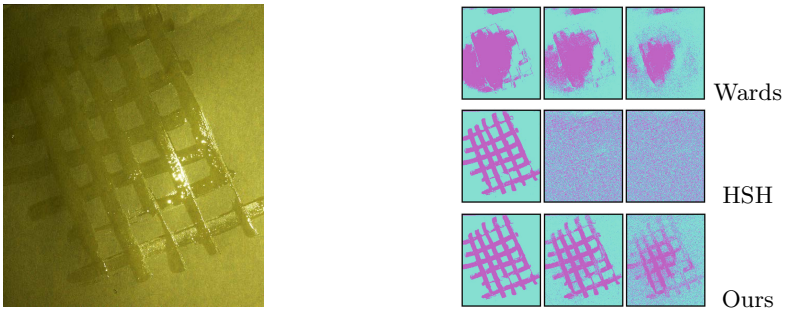

Fig. 3. Noise robustness of material clustering with three different reflectance models. (a) One of the input images, shown with increased contrast. (b) Clustering results with increasing normal noise, from left to right: no added noise, $\sigma = 0.05$ and $\sigma = 0.1$. Ward's model (top) does not result in a usable clustering, while HSH (middle) fails for even small amounts of noise. Our parameter-free description (bottom) exhibits graceful degradation for increasing levels of noise.

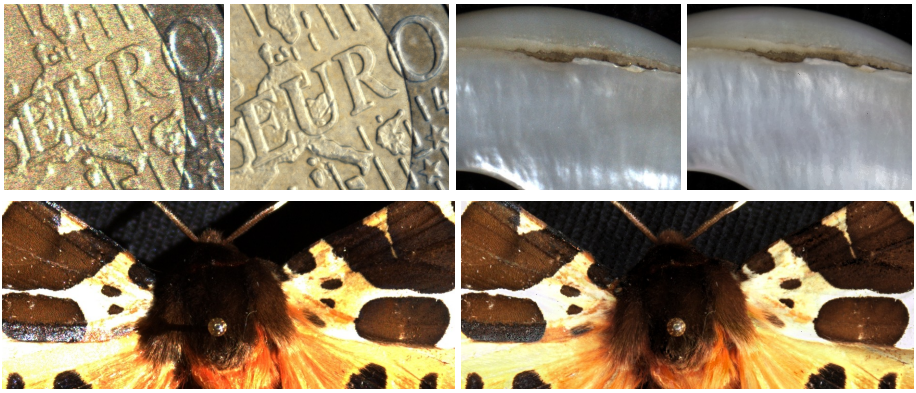

Fig. 4. Views of reconstruction results. The original image (left) was removed from the input set, its light direction was then used in the reconstructions (right).

We compare our non-parametric model to two others that have previously been used for material classification: Ward's anisotropic model [16], and hemispherical harmonics (HSH) [17,15]. For both of these models, we independently fit a BRDF to each color channel. We use the three-parameter version of Ward's model as in [16], computed using Levenberg-Marquard with analytic Jacobian. The HSH model is of second degree, resulting in 9 coefficients per color channel, computed using linear least squares.

Fig. 3 shows the resulting labeling from k-means with the number of materials fixed to $k = 2$. HSH and our model perform similarly without added noise, while Ward's model apparently not well suited for this clustering approach. When adding Gaussian noise to the normals, our method shows a graceful degradation

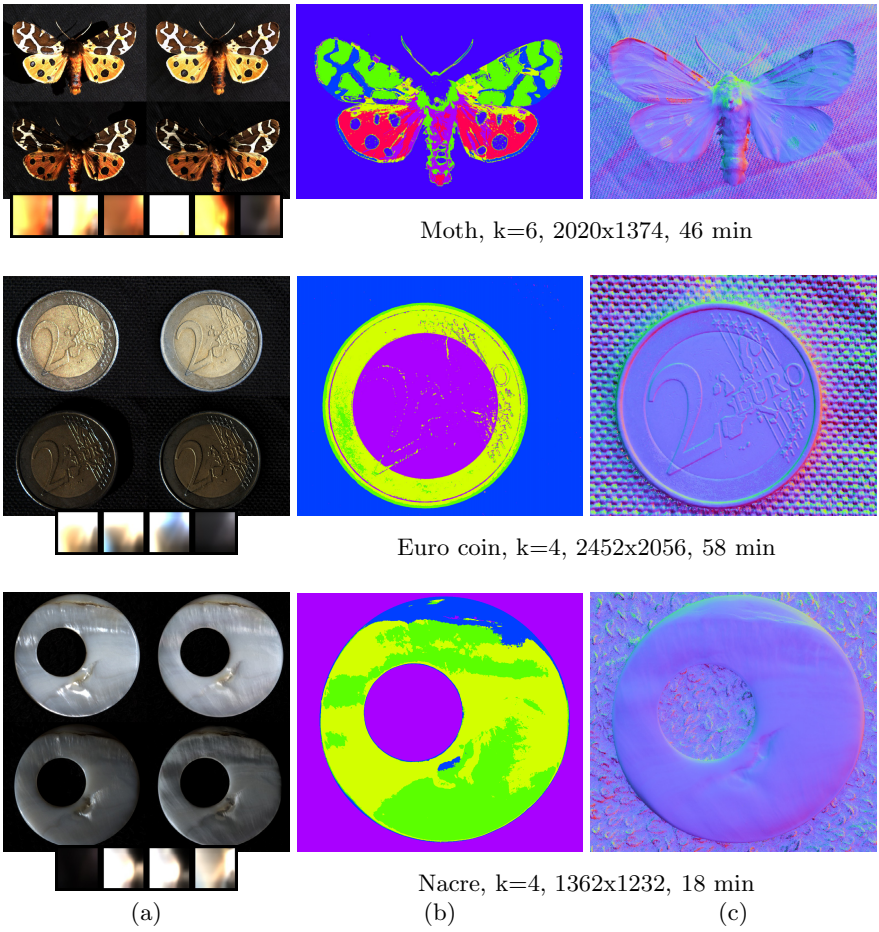

Fig. 5. Photometric stereo results showing: (a) Original photographs (left) and rendered results (right) for two novel light directions, along with the set of extracted base materials. (b) Final material labels. (c) Surface normals. The rendered results are all visually convincing. The piece of nacre is translucent and highly specular, showing the robustness of the proposed method to violations of the underlying reflectance model.

with increasing levels of noise, while HSH performance decreases dramatically even for low levels of noise. This illustrates the difficulty of fitting non-linear, analytic BRDF models to noisy measurements.

5.3 Material Clustering and Reconstruction Results

Using our enhanced photometric stereo system, we acquire and reconstruct three objects: A Euro coin, a Garden Tiger moth, and a small ornamental piece of nacre (mother of pearl). All experiments were executed and timed on a desktop PC with a Quad Core Intel i7 CPU. Fig. 5 shows the rendered output,

material segmentation, surface normals and computation time for each object. The simplicity of the bivariate BRDF model allows for fast rendering: our single-core CPU based renderer easily achieves real-time performance for all our examples. The rendered images are obtained for a novel light direction that was removed from the input set of images. Our model convincingly reproduces the iridescence of the moth wings (Fig. 5a, top row) where the hue changes from yellow to red when the incoming light angle increases. This behavior can also be seen in the captured base materials. In the images of the Euro coin, mutual interference of reflected wavefronts create a noisy speckle pattern in the image, visible in in the close-up of Fig. 4, while the rendered image is virtually noise-free. The correct locations of shadows show that the estimated geometry is true to the original, even for small details such as the individual countries. Nacre is a particularly challenging material due to its combination of translucency and strong specularities. Nevertheless, the rendered output is visually acceptable, which demonstrates the robustness of our algorithm to violations of the bivariate BRDF model. Overall, the reconstructions are of high quality, except for highly specular details which are difficult to capture with low dynamic range input images.

6 Concluding Remarks

We have shown that the acquisition of scene-specific base materials using a non-parametric BRDF descriptor can reliably be used to enhance a traditional photometric stereo system, providing high-quality rendered results, material segmentations and normal maps for several challenging objects. The results are surprisingly robust to noise and violations of the underlying assumptions about reflectance that cause difficulties for approaches based on non-linear analytic reflectance models.

A dedicated acquisition device such as our dome allows for rapid capture of 260 images. However, such a large number is likely to be redundant; it is well worth investigating how to adaptively reduce the number while maintaining high photometric and geometric fidelity. Note also that our fixed-camera setup limits the domain of the sampled BRDFs which may prohibit novel-viewpoint rendering.

Going forward, there are numerous possible generalizations: multi-view environments, unknown lighting, multi-spectral and HDR imaging, uncalibrated cameras, and even more general reflectance models including anisotropic materials. New developments in BRDF acquisition and representation may provide possibilities beyond those of existing approaches.

Acknowledgments. We gratefully acknowledge support from the EC 7th Framework Program under grant agreement n. 231809 (IP project 3D-COFORM) and the KULeuven Research Council grant Center for Archaeological Studies.

References

1. Weyrich, T., Lawrence, J., Lensch, H., Rusinkiewicz, S., Zickler, T.: Principles of Appearance Acquisition and Representation. Foundations and Trends in Computer Graphics and Vision 4(2), 75–191 (2007)
2. Woodham, R.J.: A Photometric Method for Determining Surface Orientation from Multiple Images. Optical Engineering 19(1), 139–144 (1980)
3. Chandraker, M., Agarwal, S., Kriegman, D.: ShadowCuts: Photometric Stereo with Shadows. In: CVPR, pp. 1–8 (2007)
4. Verbiest, F., Van Gool, L.: Photometric stereo with coherent outlier handling and confidence estimation. In: CVPR, pp. 1–8 (2008)
5. Hertzmann, A., Seitz, S.M.: Example-based photometric stereo: shape reconstruction with general, varying BRDFs. PAMI 27(8), 1254–1264 (2005)
6. Goldman, D.B., Curless, B., Hertzmann, A., Seitz, S.M.: Shape and Spatially-Varying BRDFs from Photometric Stereo. PAMI 32(6), 1060–1071 (2010)
7. Rusinkiewicz, S.: A new change of variables for efficient BRDF representation. In: Rendering Techniques, pp. 11–22 (1998)
8. Romeiro, F., Vasilyev, Y., Zickler, T.: Passive Reflectometry. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 859–872. Springer, Heidelberg (2008)
9. Romeiro, F., Zickler, T.: Blind Reflectometry. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 45–58. Springer, Heidelberg (2010)
10. Weinmann, M., Schwartz, C.: A multi-camera, multi-projector super-resolution framework for structured light. In: 3DIMPVT, pp. 397–404 (2011)
11. Lensch, H.P.A., Goesele, M., Kautz, J., Heidrich, W., Seidel, H.-P.: Image-based reconstruction of spatially varying materials. In: Eurographics Workshop on Rendering Techniques, pp. 103–114 (2001)
12. Haber, T., Fuchs, C., Bekaer, P., Seidel, H., Goesele, M., Lensch, H.: Relighting objects from image collections. In: CVPR, pp. 627–634 (2009)
13. Alldrin, N., Zickler, T.: Photometric stereo with non-parametric and spatially-varying reflectance. In: CVPR (2008)
14. Ren, P., Wang, J., Snyder, J., Tong, X., Guo, B.: Pocket reflectometry. ACM Transactions on Graphics 30(4), 45:1–45:10 (2011)
15. Wang, O., Gunawardane, P., Scher, S., Davis, J.: Material classification using BRDF slices. In: CVPR, pp. 2805–2811 (2009)
16. Chung, H.: Efficient photometric stereo on glossy surfaces with wide specular lobes. In: CVPR (2008)
17. Gautron, P., Krivanek, J., Pattanaik, S.: A novel hemispherical basis for accurate and efficient rendering. In: Eurographics Symposium on Rendering (2004)