

# Gloss and Normal Map Acquisition of Mesostructures Using Gray Codes

Yannick Francken, Tom Cuypers, Tom Mertens, and Philippe Bekaert

Hasselt University - tUL - IBBT  
Expertise Centre for Digital Media  
Wetenschapspark 2, 3590 Diepenbeek, Belgium  
{firstname.lastname}@uhasselt.be

**Abstract.** We propose a technique for gloss and normal map acquisition of fine-scale specular surface details, or mesostructure. Our main goal is to provide an efficient, easily applicable, but sufficiently accurate method to acquire mesostructures. We therefore employ a setup consisting of inexpensive and accessible components, including a regular computer screen and a digital still camera. We extend the Gray code based normal map acquisition approach of Francken et al. [1] which utilizes a similar setup. The quality of the original method is retained and without requiring any extra input data we are able to extract per pixel glossiness information. In the paper we show the theoretical background of the method as well as results on real-world specular mesostructures.

## 1 Introduction

During the last few decades, computers have become increasingly important for performing a wide variety of tasks. One of these tasks is generating images of virtual scenes. Nowadays convincing rendering techniques are applied in many applications such as computer games. Even photo-realistic images can be generated offline, for example to be used in movies. Therefore fast and/or accurate rendering techniques have been developed, approximating or accurately simulating the light transport within the virtual world.

However, even if light interaction could be simulated in a physically correct manner, scene data still has to be provided in the form of a 3D model. If the input scene data does not contain small-scale surface details such as scratches, imperfections, etc, the scene will probably be judged as unrealistic. Hence, manually modeling the world in such a level of detail can be a tedious task, suggesting automatic 3D scanning methods.

Throughout the years, many techniques have been proposed to digitize the world around us. These techniques typically capture either the (a) light in the scene, (b) the geometry, or (c) the reflectance properties, or any combination. In this paper we will mainly focus on capturing the reflectance properties, although we extend a fine-scale geometry acquisition system.

Even though several techniques already exist for scanning reflectance properties as well as fine-scale geometry, users tend to stick to their manual approach. One of the reasons for this is the complexity of currently available methods. Many approaches

require special purpose setups containing exotic hardware components, time consuming calibration procedures, difficult implementations and scanning procedures etc.

In this paper the goal is to make small-scale appearance acquisition available to the public, bridging the gap between current research and practical usage. This is achieved by presenting an *efficient, easy to implement* approach, employing solely *off-the-shelf hardware components* consisting of a regular still camera and a computer screen that functions as a planar illuminant.

## 2 Related Work

In this related work section we will distinguish between fine-scale geometry acquisition and reflectance acquisition.

### 2.1 Fine-scale Geometry

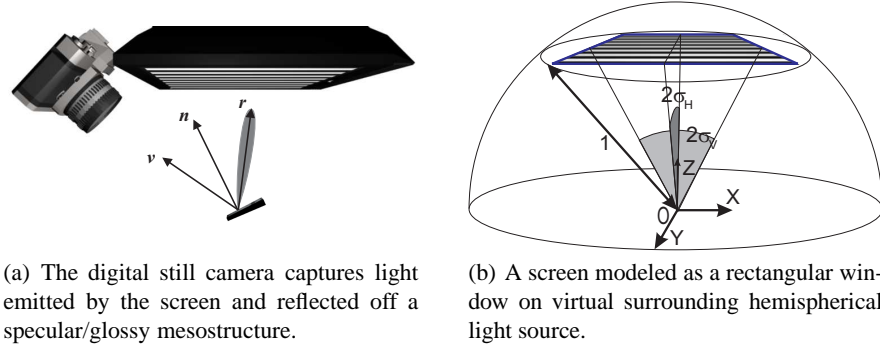
Several techniques have been introduced specifically for recovering small-scale surface details, in the form of relief (height) maps or normal maps, assuming various types of materials [2–5]. The majority of these methods requires a specialized hardware setup [6–10], have long acquisition/processing times [11, 12] or are not able to scan specular surfaces [2, 3]. In our work we will use a slightly adapted version of the Gray code based approach of Francken et al. [1], employing a screen-camera setup as acquisition setup. Because of the use of a planar illuminant and Gray codes, fine-scale specular surface geometry can accurately be measured using only up to 40 input images.

### 2.2 Reflectance

Acquiring spatially varying reflectance usually requires a complicated hardware setup, which measures the Bidirectional Reflectance Distribution Function (BRDF) [13] at each spatial location. This is a four dimensional function describing the surface’s response given the exitant (light) and incident (observation) direction. Our method is much simpler and cheaper. Even though we assume a simplified BRDF model, our technique is able to reproduce the mesostructure’s appearance faithfully.

Numerous representations exist for storing either modeled or captured BRDFs [14–17]. As storing individual data samples of densely sampled BRDFs is memory inefficient, often approximating models are fitted through the large data collection. This is either achieved by fitting an analytical model [18–22], or projecting the data to polynomial [23, 6], spherical harmonics [24–26] or wavelet bases [27, 28]. For the sake of simplicity as well as compatibility with known tools, in our work we will employ a simple analytical Phong model [29] where the glossiness is represented by a single exponent parameter.

Previous methods tend to focus mainly on improving BRDF quality, and less on acquisition speed and practical usability for a large class of users. Often very specialized setups or long and tedious procedures are required. As we focus on increasing the wide applicability rather than improving the quality of recent BRDF methods, an approximate glossiness acquisition suffices for our purposes.



**Fig. 1.** Setup.

The most closely related approach was presented by Ghosh et al. [30]. They estimate roughness as well as anisotropy from second order spherical gradient illumination. In their approach they make use of a specialized hardware setup. In our work we take an alternative approach as we want to avoid non-trivial hardware setups. We achieve this by using a screen-camera setup consisting of off-the-shelf and omnipresent hardware components.

In our work, we start from an existing Gray code based mesostructure acquisition system [1] and show that glossiness information can easily be extracted from the already available data necessary for shape reconstruction. The original method only has to be slightly modified by replacing the polarization based specular-diffuse separation with the use of pattern complements instead. No extra data is required, and besides LCD screens, also non-polarization based illuminants such as CRT screens can now be employed.

### 3 Setup

The proposed setup consists of a digital still camera that serves as light sensor and a computer screen that serves as planar illuminant (Fig. 1). Current digital still cameras are relatively inexpensive and are able to accurately measure light reflections. LCD as well as CRT computer screens are also inexpensive and omnipresent making it an ideal controllable light source.

In order to turn a screen-camera setup into a mesostructure acquisition setup, a geometric calibration step is required to relate 2D screen pixels to 3D location with respect to the camera. As the screen is not directly visible to the camera, a spherical mirror is employed for the geometrical calibration [31]. To find the internal camera parameters and the mesostructure's supporting plane, we use a standard calibration toolbox which makes use of a checkerboard pattern [32]. Radiometric calibration, which relates emitted and captured light intensities, is not essential as we are using binary (Gray code) illumination patterns.

## 4 Acquiring Surface Normals and Gloss

Acquiring local surface orientation and glossiness is achieved by placing the target object in front of a CRT or LCD monitor which acts as a light source, and recording the corresponding images using a camera. As in the normal map acquisition technique [1], we display stepwise refining vertical and horizontal Gray code patterns. We also display each pattern’s complement in order to robustly separate diffuse from specular reflection. The specular reflections then efficiently encode discrete spatial screen coordinates in a bit-wise fashion. In a geometrically calibrated setup, this allows for estimating the ideal reflection direction for each pixel. This enables us to estimate the surface normal  $\mathbf{n}$  by taking the halfway vector between the reflection vector  $\mathbf{r}$  and viewing vector  $\mathbf{v}$ , as depicted in Fig. 1 (a). In this section, we will extend this system by performing an additional glossiness analysis step.

### 4.1 Overview

In order to extract glossiness information from the recorded mesostructure taken under Gray code illumination, we require some additional illumination patterns. More specifically, complements of the original Gray code patterns are introduced. Fortunately, these render the use of polarization based separation redundant so the number of required patterns does not increase. This is due to the fact that specular highlights are considered much stronger than diffuse reflections [33, 11] and hence a binary decision (white or black reflection) can robustly be made by comparing the pixels illuminated by the pattern and the pattern’s complement.

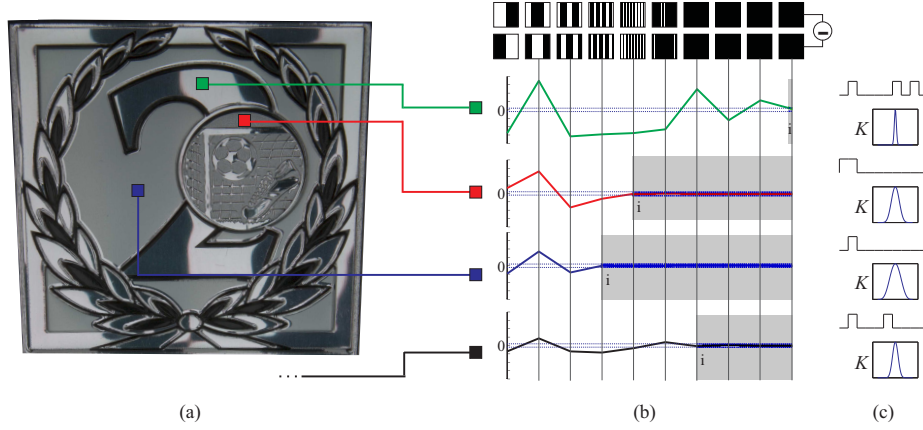
As indicated by the grey area in Fig. 2 (b), after a certain number of pattern refinements, no extra information will be gained as the intensity differences between reflected patterns and their complements will converge to zero. We analyze this convergence process to obtain glossiness information. Without requiring additional input images, we are now able to obtain a per pixel shininess coefficient as well as a surface normal.

The more pattern refinements that can be discerned, the more specular the material will be, and vice versa. This is the case because glossy reflections blur the reflected incoming light pattern. More precisely, the reflected pattern is convolved with a BRDF kernel around the ideal reflection direction [34]. The number of refinements thus is proportional to the shininess of the material. The size (or narrowness) and shape of the kernel is defined by the specular lobe of the BRDF. For the sake of simplicity as well as compatibility with known tools, we assume a Phong reflection model. This symmetric lobe is then described by a single exponent value  $n$  which is stored in the gloss map.

### 4.2 Theory

We will now formalize the concept proposed in the previous section. Therefore a model will be built that describes the captured radiance  $L$  of an imaged surface point, observed from a direction  $\mathbf{v}$ , illuminated by a given light pattern  $P$ . The equation is given by:

$$L(\mathbf{v}) = \int_{\Omega} P(\boldsymbol{\omega}) [R_d(\boldsymbol{\omega}, \mathbf{n}) + R_s(\mathbf{r}, \boldsymbol{\omega}, \mathbf{n})] d\boldsymbol{\omega} \quad (1)$$



**Fig. 2.** Acquisition pipeline. (a) mesostructure, (b) intensity differences depending on the pattern refinement level, (c) detected normal codes and Phong kernels.

The following assumptions are made before applying this equation for determining the gloss level:

**Specular + diffuse:** The imaged surface is assumed to be a combination of a specular component  $R_s(\mathbf{r}, \boldsymbol{\omega}, \mathbf{n})$  and the diffuse component  $R_d(\boldsymbol{\omega}, \mathbf{n})$ , where  $\boldsymbol{\omega}$  is the incoming light direction,  $\mathbf{n}$  the surface normal and  $\mathbf{r}$  the specular reflection vector depending on the observation direction  $\mathbf{v}$ .

**Distant hemispherical illumination:** The mesostructure is assumed to be a point in front of the center of the screen, illuminated by a rectangular part of the hemisphere  $\Omega = [\frac{\pi}{2} - \sigma_v, \frac{\pi}{2} + \sigma_v] \times [\frac{\pi}{2} - \sigma_h, \frac{\pi}{2} + \sigma_h]$  (Fig.1 (b))

**Inter-reflections and occlusions:** Both inter-reflections as occlusions are ignored for reasons of simplicity.

Under uniform illumination  $u$ , where  $P_u(\boldsymbol{\omega}) = 1$  for each incoming light direction  $\boldsymbol{\omega}$ , the equation can be simplified.

$$L_u(\mathbf{v}) = \int_{\Omega} R_d(\boldsymbol{\omega}, \mathbf{n}) d\boldsymbol{\omega} + \int_{\Omega} R_s(\mathbf{r}, \boldsymbol{\omega}, \mathbf{n}) d\boldsymbol{\omega} \quad (2)$$

$$= L_d + L_s(\mathbf{v}) \quad (3)$$

As we use Gray code patterns, we will define the patterns  $P_i$  in terms of the pattern refinement level  $i$ . For each incoming light direction  $\boldsymbol{\omega} \in \Omega$ , the pattern  $P_i(\boldsymbol{\omega})$  is either 0 or 1. The precise pattern definition for vertical patterns  $P_i^v$  and horizontal patterns  $P_i^h$  are given in equation (4) and (5), where  $(\theta, \phi) \in \Omega$ . Notice that the Gray code patterns are basically modeled as a phase shifted ( $\frac{1}{4}$  of the period) square wave in the vertical or horizontal interval  $[\frac{\pi}{2} - \sigma, \frac{\pi}{2} + \sigma]$ . Each pattern refinement from  $i$  to  $i+1$  the frequency of the wave doubles as  $\frac{i^{(i+1)-2}}{i^{i-2}} = 2$ .

$$P_i^v(\omega) = \frac{1}{2} \Psi \left( \frac{2^{i-2}(\theta - \frac{\pi}{2} + \sigma_v)}{2\sigma_v} + \frac{1}{4} \right) + \frac{1}{2} \quad (4)$$

$$P_i^h(\omega) = \frac{1}{2} \Psi \left( \frac{2^{i-2}(\phi - \frac{\pi}{2} + \sigma_h)}{2\sigma_h} + \frac{1}{4} \right) + \frac{1}{2} \quad (5)$$

The integer function  $\Psi$  is defined as

$$\Psi(x) = \begin{cases} +1 & \text{if } x - \lfloor x \rfloor \in [0, 0.5) \\ -1 & \text{if } x - \lfloor x \rfloor \in [0.5, 1) \end{cases} \quad (6)$$

Also the complements of the patterns need to be defined. They are referred to as  $P_i^{c,v}$  and  $P_i^{c,h}$ .

$$P_i^{c,v}(\omega) = 1 - P_i^v(\omega) \quad (7)$$

$$P_i^{c,h}(\omega) = 1 - P_i^h(\omega) \quad (8)$$

The captured radiance can now be modeled applying the previous definitions. The remainder part of this section will focus on the use of horizontal Gray code patterns only. However, an analogous derivation can be done for vertical patterns.

$$L_i(\mathbf{v}) = \frac{1}{2} \left[ L_d + L_s(\mathbf{v}) + \int_{\Omega} \Psi \left( \frac{2^{i-2}(\phi - \frac{\pi}{2} + \sigma_h)}{2\sigma_h} + \frac{1}{4} \right) R_d(\omega, \mathbf{n}) d\omega \right. \\ \left. + \int_{\Omega} \Psi \left( \frac{2^{i-2}(\phi - \frac{\pi}{2} + \sigma_h)}{2\sigma_h} + \frac{1}{4} \right) R_s(\mathbf{r}, \omega, \mathbf{n}) d\omega \right] \quad (9)$$

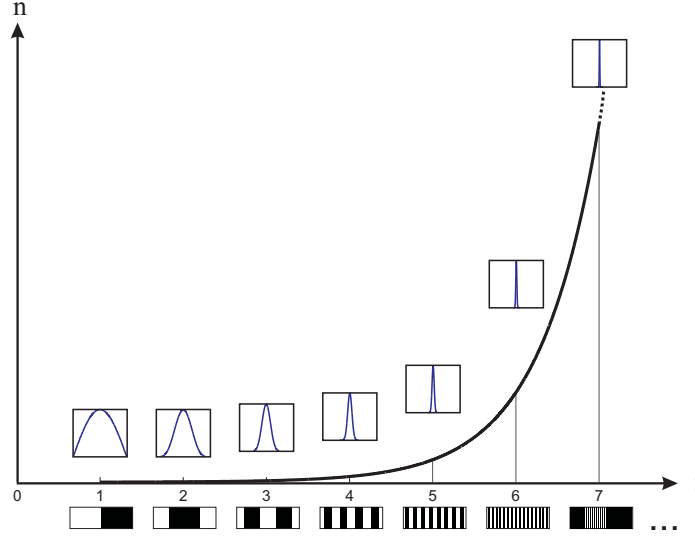
If the frequency of the pattern  $i$  is sufficiently large, the Lambertian term is approximately zero, as shown by Lamond et al. [35]. The underlying reason for this is that the Lambertian reflection can be seen as an applied low frequency convolution filter blurring away the high frequency pattern. Hence the following form can be obtained:

$$L_i(\mathbf{v}) = \frac{1}{2} \left[ L_d + L_s(\mathbf{v}) + \int_{\Omega} \Psi \left( \frac{2^{i-2}(\phi - \frac{\pi}{2} + \sigma_h)}{2\sigma_h} + \frac{1}{4} \right) R_s(\mathbf{r}, \omega, \mathbf{n}) d\omega \right] \quad (10)$$

When the pattern frequency is high compared to the size of the specular lobe, the same holds for the specular term, meaning that it also converges to zero. Hence, the wider the specular lobe, the faster this term converges to zero. As the same reasoning applies for the pattern complement  $P_i^c$ , the difference between the radiance of a scene illuminated by  $P_i$  and  $P_i^c$  converges to zero after a certain pattern refinement level  $i$ :

$$|L_i(\mathbf{v}) - L_i^c(\mathbf{v})| = \frac{1}{2} |L_d + L_s(\mathbf{v}) - L_d^c - L_s^c(\mathbf{v})| = 0 \quad (11)$$

Concretely, the smallest pattern number  $i$  has to found, such that for *all* the subsequent patterns  $j \geq i$  the intensity differences  $|L_j(\mathbf{v}) - L_j^c(\mathbf{v})|$  drop below a given



**Fig. 3.** Relation between pattern refinement level  $i$  and the gloss level  $n$ .

threshold (Fig. 2 (b)). When  $i$  is found, it is converted into a corresponding Phong kernel  $K(\omega) = \cos^n(\omega)$  (Fig. 2 (c)). Therefore we propose a simple heuristic which takes into account the following constraint: The surface area  $\int_0^{\pi/2} K(\omega) d\omega$  under the kernel  $K$  has to halve if the assigned  $i$  increases one level. We have empirically established that this relation can be well-approximated by a simple exponential function:

$$n = 4^{(i-1)} \quad (12)$$

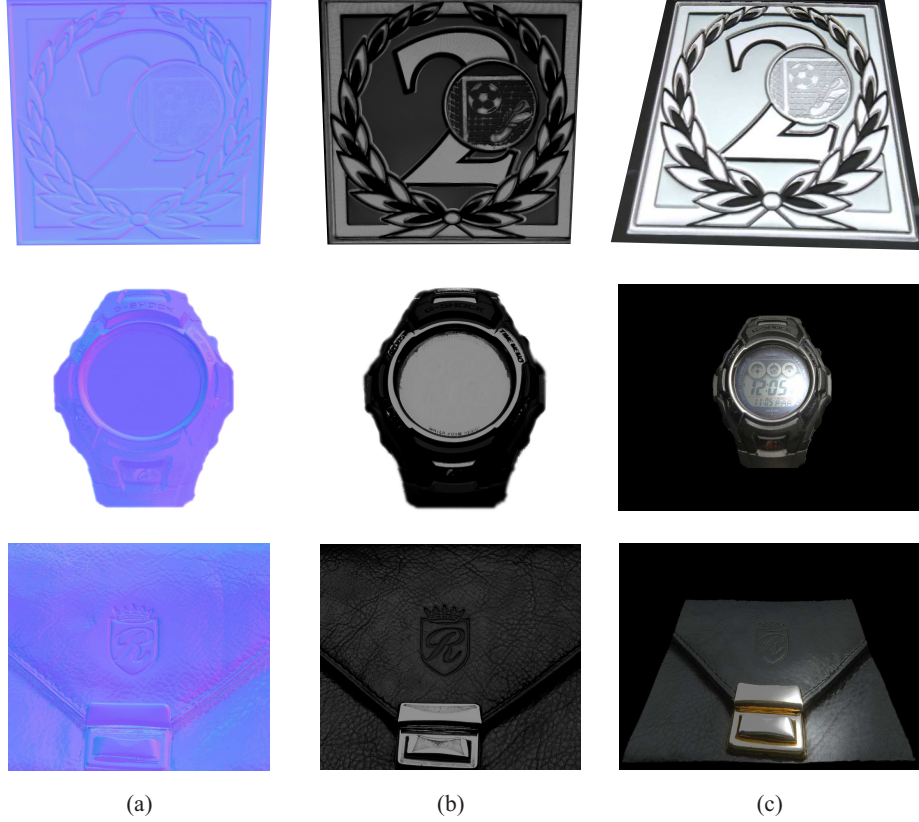
This relation is illustrated in Fig. 3. Note we assign no  $n$  value when  $i = 0$  as the material is then meant to be perfectly diffuse.

Notice that this kernel fitting is approximate because of the limited number of input images we require and the reflection model we employ. However, as we focus more on the efficiency and easy applicability than on pure accuracy, it yields sufficiently precise results, as can be seen in the next section.

## 5 Results and Discussion

We have created a proof of concept implementation of the described procedure. The setup we employed consists of a 19 inch LCD monitor and a Canon EOS 400D camera. Experiments were done on different specular materials including plastics, leather, metals, glass and polished marble. For all our results 40 input images were recorded, 10 for each direction plus their complements.

Results on real-world examples are illustrated in Fig. 4. Column (a) shows the acquired normal maps stored in the red, green and blue color channel. Column (b) contains the gloss maps. The gloss values range from black to white. Black values indicate



**Fig. 4.** Results. (a) normal maps obtained from detected codes, (b) gloss maps containing Phong exponents, (c) renderings.

diffuse reflections, white values represent highly specular reflections, and intermediate grey values represent glossy reflections. The results show for example that different metal coatings yield different gloss values (top row), the glass of the watch is more specular than the plastics (middle row), and scratches on the wallet’s hasp make it locally less specular (bottom row). However, also notice in this image that self-shadowed regions in the pores and grooves of the leather are mistakenly classified as non-specular since the occlusion assumptions did not hold. Column (c) shows virtual renderings of the scanned surfaces under point light illumination, taking into account the displayed normal and gloss maps as well as regular texture maps.

Besides inherent problems of the relief acquisition, such as light occlusions, the main limitation is the small number of available per pixel samples due to efficient binary encoding. Only  $p$  possible gloss levels can be distinguished using our technique, where  $p$  is the number of patterns used (in our case 10). In addition, as the kernel width is directly dependent on the exponentially decreasing pattern’s stripe width, only a few



possible kernels can be assigned to diffuse materials. A more kernel size distribution may be desirable in this case.

## 6 Conclusions

In this paper we showed how a straightforward extension of a Gray code based normal scanning can provide us with a very simple BRDF approximation in the form of a single Phong exponent. However, taking into account these approximate gloss maps in addition to traditional texture and normal maps tends to considerably improve rerenderings of heterogeneous materials.

## 7 Future Work

Improvements are possible regarding the convolution kernel approximations. Currently we are looking into recovering more general BRDFs by adding extra and more optimal patterns to allow for a more precise kernel fitting. Furthermore we believe this work can function as basis for an integrated normal map acquisition system, where the type of pattern/method depends on which mesostructure regions are processed.

## Acknowledgements

Part of the research at EDM is funded by the ERDF (European Regional Development Fund) and the Flemish government. Furthermore we would like to thank our colleagues for their help and inspiration.

## References

1. Francken, Y., Cuypers, T., Mertens, T., Gielis, J., Bekaert, P.: High quality mesostructure acquisition using specularities. In: Proceedings of Conference on Computer Vision and Pattern Recognition, IEEE (2008) 1–7
2. Rushmeier, H., Taubin, G., Guéziec, A.: Applying Shape from Lighting Variation to Bump Map Capture. IBM TJ Watson Research Center (1997)
3. Hernández, C., Vogiatzis, G., Brostow, G.J., Stenger, B., Cipolla, R.: Non-rigid photometric stereo with colored lights. In: Proceedings of International Conference on Computer Vision. (2007)
4. Morris, N.J.W., Kutulakos, K.N.: Reconstructing the surface of inhomogeneous transparent scenes by scatter-trace photography. In: Proceedings of International Conference on Computer Vision. (2007)
5. Chung, H., Jia, J.: Efficient photometric stereo on glossy surfaces with wide specular lobes. In: Proceedings of Conference on Computer Vision and Pattern Recognition, IEEE (2008)
6. Malzbender, T., Gelb, D., Wolters, H.J.: Polynomial texture maps. In: SIGGRAPH. (2001) 519–528
7. Han, J.Y., Perlin, K.: Measuring bidirectional texture reflectance with a kaleidoscope. In: SIGGRAPH, ACM (2003) 741–748

8. Neubeck, A., Zalesny, A., Gool, L.V.: 3d texture reconstruction from extensive btf data. In: Texture 2005 Workshop in conjunction with ICCV 2005. (2005) 13–19
9. Wang, J., Dana, K.J.: Relief texture from specularities. Transactions on Pattern Analysis and Machine Intelligence **28** (2006) 446–457
10. Ma, W.C., Hawkins, T., Peers, P., Chabert, C.F., Weiss, M., Debevec, P.: Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In: Proceedings of Eurographics Symposium on Rendering. (2007)
11. Chen, T., Goesele, M., Seidel, H.P.: Mesostructure from specularity. In: Proceedings of Conference on Computer Vision and Pattern Recognition. Volume 2. (2006) 1825–1832
12. Holroyd, M., Lawrence, J., Humphreys, G., Zickler, T.: A photometric approach for estimating normals and tangents. SIGGRAPH Asia **27** (2008) 133
13. Nicodemus, F., Richmond, J., Hsia, J., Ginsberg, I., Limperis, T.: Geometrical considerations and nomenclature for reflectance. NBS monograph **160** (1977) 201–231
14. He, X.D., Torrance, K.E., Sillion, F.X., Greenberg, D.P.: A comprehensive physical model for light reflection. In: SIGGRAPH, New York, NY, USA, ACM (1991) 175–186
15. Ashikhmin, M., Shirley, P.: An anisotropic phong brdf model. Journal of Graphics Tools **5** (2000) 25–32
16. Ashikhmin, M., Premože, S., Shirley, P.: A microfacet-based brdf generator. In: SIGGRAPH, New York, NY, USA, ACM Press (2000) 65–74
17. He, X.D., Heynen, P.O., Phillips, R.L., Torrance, K.E., Salesin, D.H., Greenberg, D.P.: A fast and accurate light reflection model. SIGGRAPH **26** (1992) 253–254
18. Ward, G.J.: Measuring and modeling anisotropic reflection. SIGGRAPH **26** (1992) 265–272
19. Lafortune, E.P.F., Foo, S.C., Torrance, K.E., Greenberg, D.P.: Non-linear approximation of reflectance functions. In: SIGGRAPH, New York, NY, USA, ACM Press/Addison-Wesley Publishing Co. (1997) 117–126
20. Lensch, H.P.A., Goesele, M., Kautz, J., Heidrich, W., Seidel, H.P.: Image-based reconstruction of spatially varying materials. In: Proceedings of Eurographics Symposium on Rendering, London, UK, Springer-Verlag (2001) 103–114
21. Gardner, A., Tchou, C., Hawkins, T., Debevec, P.: Linear light source reflectometry. In: SIGGRAPH, ACM (2003) 749–758
22. Ngan, A., Durand, F., Matusik, W.: Experimental analysis of brdf models. In: Proceedings of Eurographics Symposium on Rendering, Eurographics Association (2005) 117–226
23. Koenderink, J.J., Doorn, A.J.v., Stavridi, M.: Bidirectional reflection distribution function expressed in terms of surface scattering modes. In: Proceedings of European Conference on Computer Vision, London, UK, Springer-Verlag (1996) 28–39
24. Westin, S.H., Arvo, J.R., Torrance, K.E.: Predicting reflectance functions from complex surfaces. In: SIGGRAPH, New York, NY, USA, ACM (1992) 255–264
25. Ramamoorthi, R., Hanrahan, P.: Frequency space environment map rendering. In: SIGGRAPH, New York, NY, USA, ACM (2002) 517–526
26. Basri, R., Jacobs, D.W.: Lambertian reflectance and linear subspaces. Transactions on Pattern Analysis and Machine Intelligence **25** (2003) 218–233
27. Lalonde, P., Fournier, A.: A wavelet representation of reflectance functions. Transactions on Visualization and Computer Graphics **3** (1997) 329–336
28. Ng, R., Ramamoorthi, R., Hanrahan, P.: All-frequency shadows using non-linear wavelet lighting approximation. In: SIGGRAPH, New York, NY, USA, ACM (2003) 376–381
29. Blinn, J.F.: Models of light reflection for computer synthesized pictures. In: SIGGRAPH, New York, NY, USA, ACM (1977) 192–198
30. Ghosh, A., Chen, T., Peers, P., Wilson, C.A., Debevec, P.: Estimating specular roughness and anisotropy from second order spherical gradient illumination. In: Proceedings of Eurographics Symposium on Rendering. (2009)

31. Francken, Y., Hermans, C., Bekaert, P.: Screen-camera calibration using gray codes. In: Proceedings of Canadian Conference on Computer and Robot Vision, IEEE Computer Society (2009)
32. Bouguet, J.Y.: Camera calibration toolbox for matlab (2006)
33. Umeyama, S., Godin, G.: Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *Transactions on Pattern Analysis and Machine Intelligence* **26** (2004) 639–647
34. Ramamoorthi, R., Hanrahan, P.: A signal-processing framework for inverse rendering. In: SIGGRAPH, New York, NY, USA, ACM (2001) 117–128
35. Lamond, B., Peers, P., Ghosh, A., Debevec, P.: Image-based separation of diffuse and specular reflections using environmental structured illumination. In: Proceedings of International Conference on Computational Photography. (2009)