High Quality Mesostructure Acquisition Using Specularities

Yannick Francken   Tom Cuypers   Tom Mertens   Jo Gielis*   Philippe Bekaert

Hasselt University - tUL - IBBT
Expertise Centre for Digital Media, Belgium

{firstname.lastname}@uhasselt.be

Abstract

We propose a technique for cheap and efficient acquisition of mesostructure normal maps from specularities, which only requires a simple LCD monitor and a digital camera. Coded illumination enables us to capture subtle surface details with only a handful of images. In addition, our method can deal with heterogeneous surfaces, and high albedo materials. We are able to recover highly detailed mesostructures, which was previously only possible with an expensive hardware setup.

1. Introduction

A wide variety of methods exists for scanning 3D geometry. Most methods focus on acquiring the global shape of an object, but ignore small-scale details. In this paper we present an efficient technique for scanning such small-scale surface details or *mesostructures* using cheap, off-the-shelf hardware. Our method outputs a normal map of the scanned surface. Normals can be transformed into the original 3D shape of the surface [8], or they can be combined with the output of a global shape acquisition method [23]. In computer graphics, normals can be added as texture maps to enrich 3D models with relief, and can be rendered directly using graphics hardware.

Most normal acquisition methods are based on photometric stereo [38]. These approaches usually assume a perfectly matte surface. However, many materials do not meet this assumption and are thus ill-suited for specular objects. Instead of circumventing this problem, e.g. by filtering out highlights [36], techniques have been developed specifically for specular and mirroring surfaces [14]. Some of these methods focus only on recovering global shape [27, 11, 2], or assume that the surface is an ideal reflector [32]. Other people have developed techniques that use specular highlights as cues [12], because highlights are barely affected by subsurface light transport. This makes it possible to recover

transparent, translucent and low albedo surfaces. Our work continues along this direction, and bears similarity to that of Chen *et al.* [4], since we also focus on specular mesostructure acquisition. Our method however, offers faster acquisition, and produces more detailed normal maps. High resolution specularity-based scanning of mesostructure has been demonstrated before, requiring an expensive and intricate hardware [35, 16], whereas our setup is cheap and easy to build (Fig. 1).

Since we rely only on highlights, our technique assumes that the surface is specular. We demonstrate that even for less specular (or glossy) materials, our technique still produces reasonable results.

Contributions   We introduce the following contributions: (1) improved detail and reduced acquisition time for specular mesostructures by using coded illumination; (2) robust specularity detection, which makes it possible to deal with heterogeneous materials and high albedo materials. Finally, (3) we describe a cheap acquisition system that consists of off-the-shelf components.

2. Related Work

A large body of work deals with recovering shape of real-world objects. We distinguish two trends in computer vision literature. First, stereo matching algorithms recover depth maps using triangulation, by observing a scene from two (or more) views [29]. Second, photometric stereo [38] computes surface normals from a sequence of illumination directions, while observing the scene from a single viewpoint. Our technique can be classified in the latter category.

Stereo methods usually assume that the observed materials are perfectly diffuse. Other techniques have been developed that focus on more "difficult" materials, in particular specular materials. The effect of weak specular reflection can be filtered out in order to apply techniques that assume a diffuse material. This can be realized with polarization [36, 37, 22, 34, 16], or color transformations [31, 17]. More general stereo techniques can deal with arbitrary

---

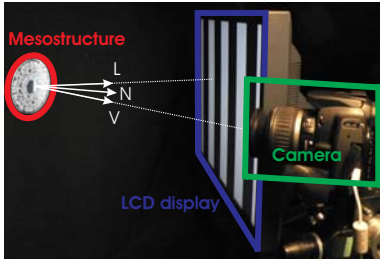*Currently at Eyetronics NV. E-mail: jo@eyetronics.com

Figure 1. Our setup consists of a digital reflex camera, LCD screen and mesostructure.

(a) 4 directions, 3 images

(b) 16 directions, 5 images

(c) 64 directions, 7 images

(d) 4096 directions, 13 images

Figure 2. Normal map reconstruction quality depends on the number of sampled light source directions, or angular sampling rate.

BRDFs [42, 10, 13, 39]. Highly specular surfaces are particularly challenging, so specialized techniques have been developed [14]. These techniques usually focus on recovering global shape [21, 11, 41, 15, 32, 2, 19, 1]. Local surface orientation can also be analyzed from specular highlights [27, 12, 4], and does not necessarily require that the surface is an ideal reflector. Our technique extends this idea by using coded illumination instead of regular light sources.

Most shape recovery methods focus on global shape, but often ignore small-scale surface detail. Techniques have been introduced specifically for recovering such details, in the form of relief (height) maps or normal maps. Rushmeier *et al.* [26] acquired normal maps from Lambertian surfaces using photometric stereo [38]. Yu and Chang [40] obtained relief maps by analyzing cast shadows. Neubeck *et al.* [24] reconstructed relief from bidirectional texture functions. Morris and Kutulakos [20] reconstructed exterior surfaces of inhomogeneous transparent scenes by capturing images from one or more viewpoints while moving a proximal light source. Wang and Dana [35] used a special hardware setup to measure normal maps, which basically estimates normal direction based on the centroid of the specular highlight. Ma *et al.* [16] computed the highlight's centroid directly by illuminating the surface with linear gradient patterns. However, their method requires an expensive spherical array of polarized light sources. Chen *et al.* [4] computed normal maps directly from specularities using just a video camera and a light source, but their technique requires many images.

Our setup is inspired by previous work that uses computer monitors as controllable illuminants. Monitors have been used for environment matting [43] and photometric stereo [5, 9]. Tarini *et al.* [32] analyzed the reflected patterns from a monitor off mirroring objects to obtain depth information. Similar to our technique, Bonfort *et al.* [2] used Gray code patterns emitted from a monitor. However, they focused on scanning the global shape of specular objects using triangulation.

3. Our Approach

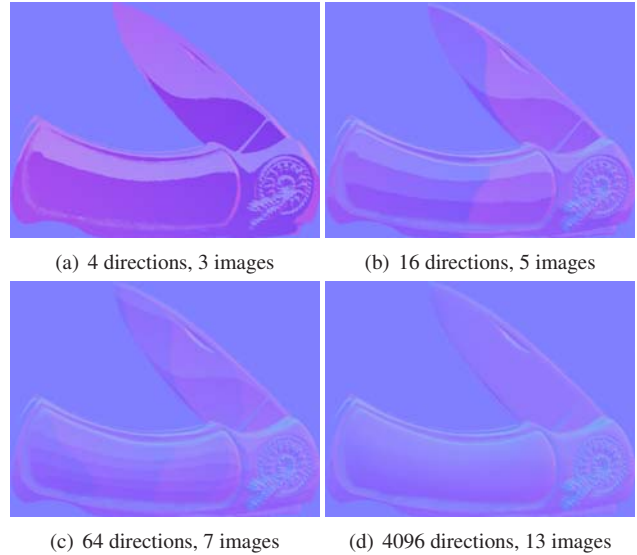In this section, we detail our method and acquisition setup.

3.1. Overview

Inferring normals from specularities is fairly straightforward. The normal associated with a specular pixel is exactly the halfway vector between the light direction and view direction [4]. The accuracy of the solution will therefore improve as more light directions are taken into account. We refer to the amount of different light directions as *angular sampling rate*. Fig. 2 shows the importance of having a high angular sampling rate. Many samples are needed to recover subtle surface details. When using a low sampling rate, certain normals will be missed (which shows up as quantization artifacts here).

We attain a high angular sampling rate by using more than one light at a time (Fig. 3). As will be explained in section 3.3, our technique samples n different light source directions, with $O(\log_2 n)$ images. In contrast, Chen *et al.*'s mesostructure-from-specularity technique requires n images for n directions [4]. In order to reach a high enough sampling rate, Chen *et al.* sample continuously using a video camera. However, video often suffers from poor image quality. Because our method requires a relatively low number of images, it becomes practical to employ a digital still camera. Therefore, better quality and specifically better resolution can be obtained, as still cameras have better optics and larger sensors, and can get away with longer exposure times.

3.2. Recovering Normal Maps from Specularities

Our goal is to find a normal map based on specularities [35, 4, 16]. To this end, we assume that the surface is an ideal reflector. If we know that illumination arrives
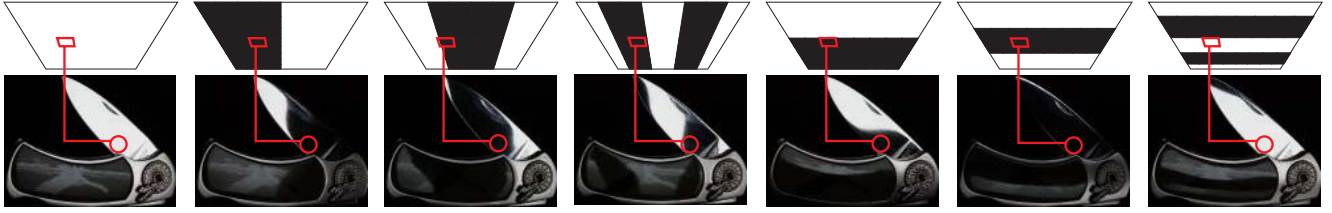
Figure 3. Top: floodlit image and Gray code patterns; bottom: corresponding images. Each location on the monitor (red square) is directly reflected (red circle). By decoding the Gray code patterns, we can recover the illumination direction, and consequently, the surface normal.

from direction $\vec{L}$ and is reflected in direction $\vec{R}$, the normal must be:

$$\vec{N} = \frac{1}{\|\vec{L} + \vec{R}\|}(\vec{L} + \vec{R}), \qquad (1)$$

Subsequently, at each specularity in the image, we use Eq. 1 to infer the corresponding normal, as done by Chen *et al*. [4]. Therefore we reasonably assume the mesostructure to be approximately planar and furthermore require a calibrated screen-camera-mesostructure setup.

Eq. 1 is only an approximation because a surface is never perfectly flat (at the microscopic level). Consequently, light is usually scattered in a small cone around the ideal reflection direction, and highlights become less clearly defined. For instance, glossy materials like plastic exhibit such behavior. We will discuss how this affects our results in section 3.5. We also ignore specular interreflection, for the sake of simplicity.

3.3. Efficient Acquisition

A high angular sampling rate is required in order to recover all possible normal orientations (see Fig. 2), which we obtain by sampling more than one light direction per image. We therefore require an array of light sources, where every source can be switched on/off individually. This array could be a computer monitor [43, 5, 9], a (hemi)spherical rig fitted with lights [27, 18, 16], or an unstructured set of controllable lights. In our implementation, we employ an LCD monitor (Fig. 1), which provides us with a high-resolution, regularly-spaced grid of light sources.

Under the assumption of ideal reflection described in the previous section, we know that a specularity can only be caused by exactly one light direction, which in turn corresponds to a location on the LCD monitor. If this location is known, we can simply compute the local normal using Eq. 1. A naive solution would be to iterate through all lights, enabling them one at a time, and observing which pixels contain highlights [4]. Unfortunately, this becomes impractical for thousands of light sources.

Instead, we identify monitor locations using Gray code patterns, which have been used before to match surface locations for stereo [28, 30]. These patterns encode discrete spatial coordinates in a bit-wise fashion, and are robust against small errors. For 2D patterns (Fig. 3), the horizontal

(a) Photo.   (b) Diffuse.   (c) Specular.   (d) Confidence.
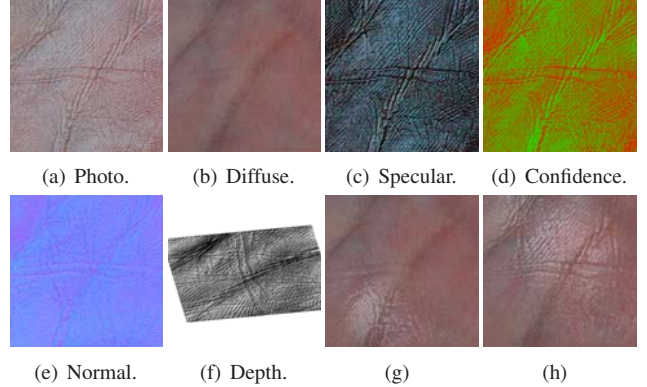
(e) Normal.   (f) Depth.   (g)   (h)

Figure 4. Acquired maps from oiled human palm skin (massage oil). (a) photo, (b) diffuse map, (c) specular map, (d) confidence map (green: high confidence, red: low confidence), (e) normal map, (f) depth map and (g,h) renderings.

and vertical coordinates are coded sequentially. Given that spatial coordinates lie in the range $[1..m] \times [1..n]$, we need only $\log_2(m) + \log_2(n) + 1$ patterns. Gray codes can also be used for an unstructured collection of light sources (even though we did not implement this).

After having recorded the surface illuminated by the patterns, we simply decode the Gray patterns by determining whether a given pixel is specular or not, using the techniques explained in the following section. See Fig. 3 for an illustration.

3.4. Specularity Detection

If the surface has a negligible albedo and is homogeneous, specularities can easily be detected using global thresholding [34, 4]. However, these assumptions do not hold in practice.

Heterogeneous objects may reflect more or less light depending on the spatial location, which makes it hard to determine a global threshold that can be applied to all pixels. We therefore normalize a pixel's intensity w.r.t. the maximally possible intensity at that pixel's location. This maximum can be found by emitting an additional floodlit pattern (all lights enabled). After normalization, we can simply apply global thresholding like before. See Fig. 9 for examples of heterogeneous surfaces ("graphics card" and "wallet").
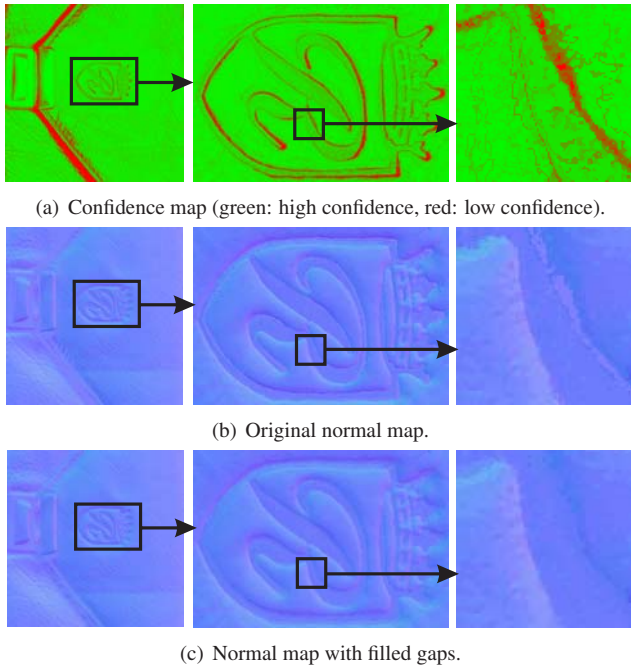
(a) Confidence map (green: high confidence, red: low confidence).

(b) Original normal map.

(c) Normal map with filled gaps.

Figure 5. Uncertain regions are filled in by interpolating normals from the nearest valid neighbors.
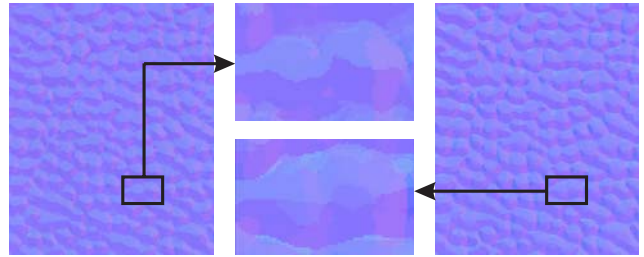

Figure 6. Comparison between results from a specular and glossy surface. The left image shows the result of capturing a glossy plastic surface. In the right image, the same surface was acquired after having applied a fine layer of oil in order to increase specularity. If the surface is more specular, finer normal variations are detected, whereas the glossy version looks more quantized.
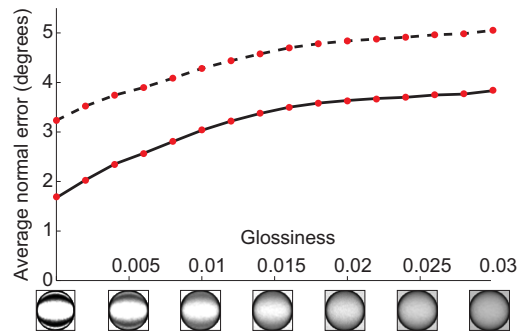

Figure 7. Error analysis for glossy materials on a synthetic example. The average angular error of our reconstruction is plotted in function of the glossiness. The solid curve shows the error measured on confident pixels only. The dashed curve shows the error when taking into account all pixels, including interpolated ones.

Even with normalization, specularity detection can be very sensitive to the chosen threshold, in particular for high albedo materials such as human skin (Fig. 4). We therefore wish to isolate the specular component in the input images using polarization [22]. The polarization state of light is approximately preserved under specular reflection, and "randomized" scattered after diffuse reflection [6]. Thus, under polarized illumination, we can get a diffuse-only image by placing a linear polarizing filter in front of the lens, oriented orthogonally to the polarization direction. Rotating the filter by 90 degrees yields an image that contains both the specular and diffuse component. Specularity isolation is achieved simply by subtracting the two images. Note that because we are using an LCD monitor, we get polarized illumination "for free". This more robust detection method comes at the cost of doubling the required number of images.

3.5. Limitations

There are two factors that negatively influence reconstruction results, namely occlusion and glossiness. This section describes how they affect our results. Note that these issues also occur with previous mesostructure-from-specularity methods [35, 4, 16].

Occlusion Self-shadowing may occur in deep grooves and pits, for instance, and results in meaningless information. We can easily detect these regions using a floodlit pattern (Fig. 3). We assign a confidence to each pixel, chosen proportional to a pixel's intensity, and fill low confidence areas by interpolating normals from the nearest valid neighbors (Fig 5).

Glossiness Our method becomes less accurate for glossier (or rougher) materials. This is visually demonstrated in Fig. 6 and verified on a ground thruth comparison in Fig. 7. For the latter figure, we generated images of a mesostructure using a photorealistic renderer (PBRT, [25]) under our Gray-coded illumination patterns. We employed a physically-based reflectance model [33], which allows for controlling the glossiness. Notice how the error increases only slightly for more gloss. Including interpolated normals of less confident pixels tend to add an overal increase to the numerical error. Fortunately, the interpolation tends to produce smooth results, and thus does not introduce visually distracting artifacts.

The reason why glossier materials result in less accurate normal maps is illustrated in Fig. 8. It shows how Gray codes are interpreted for a very specular material (narrow specular lobe) and for a more glossy object (broader lobe). As the Gray code patterns become more refined, it becomes
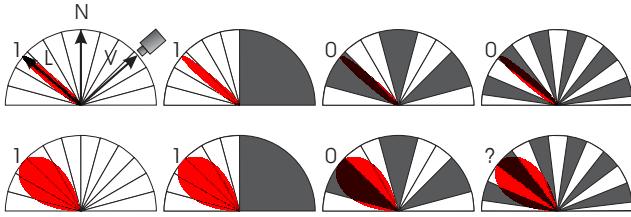
Figure 8. Illustration of precision issues associated with glossy materials. A specular BRDF directly picks up the Gray code patterns, whereas glossy BRDFs blurs the patterns. For fine patterns, this convolution makes specularity detection ambiguous, which means that no more information can be gained after a certain refinement level.

harder to distinguish between the black and white parts of the pattern, which limits the attainable angular sampling rate. We therefore stop refinement when the sum of absolute differences (SAD) between consecutive refinements drops below a threshold. Even if a material is glossy, we can still infer a reasonable normal map, albeit with less angular detail (Fig. 6,Fig. 7).

4. Results

Our results were generated with a regular 1280×1024 LCD screen, a digital SLR camera (Canon EOS 400D with default EF-S 18-55mm, f/3.5-5.6 kit lens) and a linear polarizing filter. The camera, screen and mesostructure are calibrated with respect to each other. The screen and camera are facing a similar direction and hence are calibrated using a (spherical) mirror to make the screen visible to the camera, similar to the method of Francken *et al*. [7]. The mesostructure is attached to a supporting surface with position and orientation determined using a calibration pattern [3].

Assuming the surface is in perfect focus, and depending on the mesostructure's placement, we have a spatial resolution of approximately 40 normal pixels per millimeter. The largest number of patterns we used in our setup is 15 (one floodlit, 7 horizontal, 7 vertical). The angular resolution then is approximately 1.12 normals per degree. The recoverable normal angles range from -22.5 to 22.5 degrees w.r.t. the supporting surface's normal. Using a larger screen, or multiple screens, can improve this range. Even though the normal range seems limited, it suffices for the mesostructures we scanned. We always obtained a visually plausible reconstruction. Moreover, reflected light from grazing angles is often blocked anyway due to self-occlusions.

Fig. 4 shows the different maps that we obtain from an acquired surface. In the first place, our method returns a normal map (with confidence), which can be turned into a depth map [8]. As a byproduct of the polarization-based separation discussed in Sec. 3.4, we also infer a diffuse and specular map, which were computed from the floodlit illumination pattern. These maps can easily be used as textures

material	#directions	#images	polarization
graphics card	256	18	yes
metal coaster	1024	11	no
wallet	1024	11	no
rough glass	16384	30	yes

Table 1. Number of images used in Fig. 9, in function of the number of sampled light directions. Polarization is used for high albedo surfaces, which doubles number of images (accounted for in **#images**-column).

in graphics applications, and rendered in real-time using graphics hardware. In Fig. 9 we show different normal maps and renderings, for a variety of materials: metal ("coaster", "wallet", "graphics card"), plastic ("graphics card"), semi-transparent ("rough glass"), organic ("leather") and heterogeneous ("graphics card", "wallet") materials. For each of the recovered surfaces, the angular sampling rate and number of acquired images, are given in Table 1.

5. Conclusions

We have presented an efficient acquisition method for mesostructure using specularities. Our technique employs coded illumination in order to achieve a high angular sampling rate, and robust specularity detection for heterogeneous and high albedo materials. We have demonstrated that highly detailed mesostructures can be recovered, using a cheap setup.

Currently, our method ignores the diffuse component. It would be interesting to develop a hybrid method that incorporates both components for computing a normal map.Our setup limits the range of normals that can be acquired. A larger monitor, or multiple monitors, would alleviate this problem. Another possibility would be to illuminate the mesostructure from different angles by using a turntable, and then to register and fuse the results.

Acknowledgements

References

[1] Y. Adato, Y. Vasilyev, O. Ben-Shahar, and T. Zickler. Toward a theory of shape from specular flow. In *Proceedings of ICCV*, 2007.

[2] T. Bonfort, P. Sturm, and P. Gargallo. General specular surface triangulation. In *Proceedings of ACCV*, volume 2, pages 872–881, jan 2006.

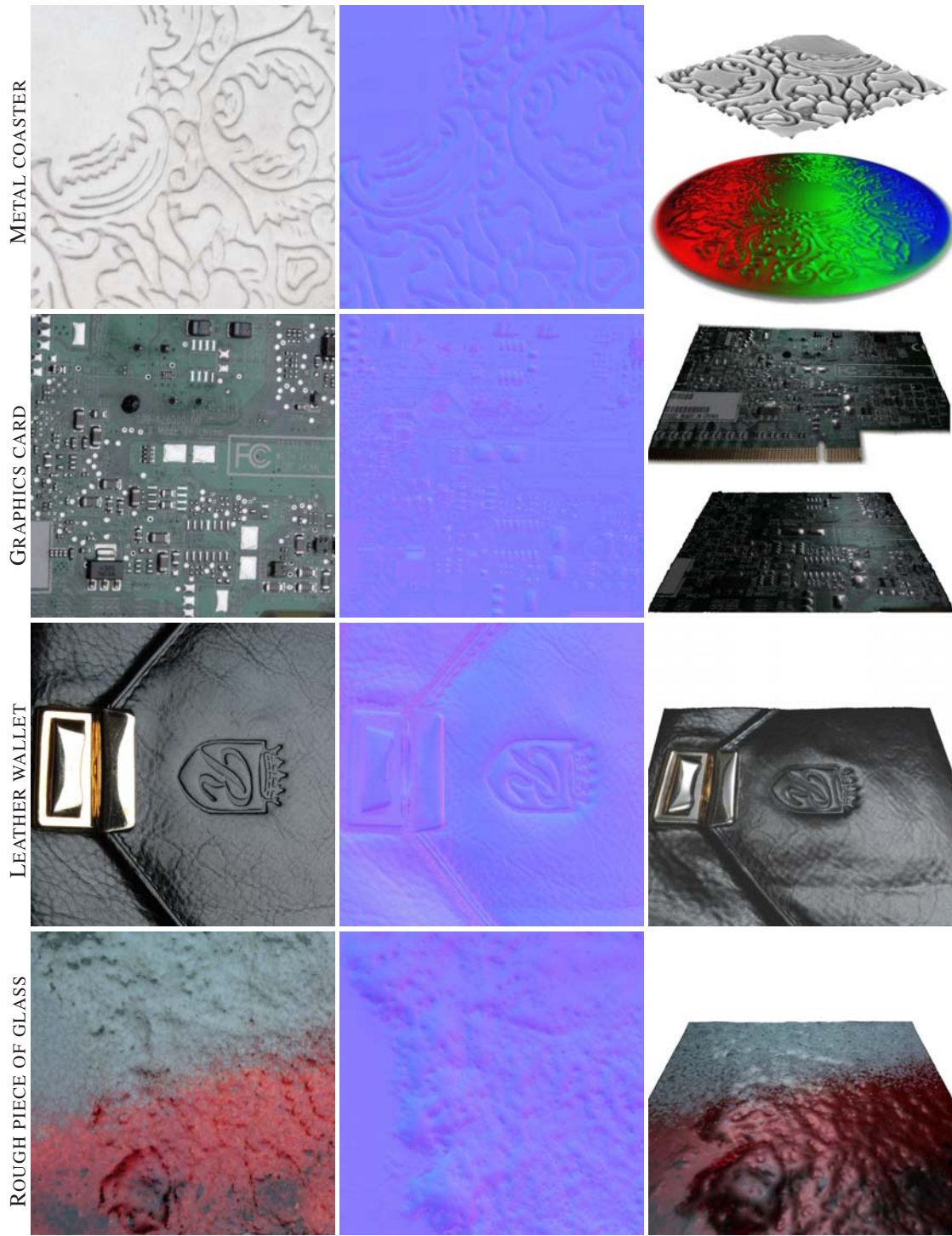[3] J.-Y. Bouguet. Camera Calibration Toolbox for MATLAB, 2006.

Figure 9. Results. Left: input images, middle: normal maps, right: renderings.

[4] T. Chen, M. Goesele, and H.-P. Seidel. Mesostructure from specularity. In *Proceedings of CVPR*, volume 2, pages 1825–1832, 2006.

[5] J. J. Clark. Photometric stereo with nearby planar distributed illuminants. In *Proceedings of CRV*, page 16, 2006.

[6] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *SIGGRAPH*, pages 145–156, 2000.

[7] Y. Francken, C. Hermans, and P. Bekaert. Screen-camera calibration using a spherical mirror. In *Proceedings of CRV*,

pages 11–20, 2007.

[8] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *PAMI*, 10(4):439–451, 1988.

[9] N. Funk and Y.-H. Yang. Using a raster display for photometric stereo. In *Proceedings of CRV*, pages 201–207, 2007.

[10] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *Proceedings of ICCV*, volume 1, pages 341–348, 2005.

[11] M. A. Halstead, B. A. Barsky, S. A. Klein, and R. B. Mandell. Reconstructing curved surfaces from specular reflection patterns using spline surface fitting of normals. In *SIGGRAPH*, pages 335–342, 1996.

[12] G. Healey and T. O. Binford. Local shape from specularity. *Computer Vision, Graphics, and Image Processing*, 42(1):62–86, 1988.

[13] A. Hertzmann. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *PAMI*, 27(8):1254–1264, 2005. Member-Steven M. Seitz.

[14] K. Ikeuchi. Determining surface orientation of specular surfaces by using the photometric stereo method. *PAMI*, 3, Nov. 1981.

[15] K. N. Kutulakos and E. Steger. A theory of refractive and specular 3d shape by light-path triangulation. In *Proceedings of ICCV*, pages 1448–1455, 2005.

[16] W.-C. Ma, T. Hawkins, P. Peers, C.-F. Chabert, M. Weiss, and P. Debevec. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proceedings of EGSR*, jun 2007.

[17] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur. Beyond lambert: Reconstructing specular surfaces using color. In *Proceedings of CVPR*, volume 2, pages 619–626, 2005.

[18] T. Malzbender, D. Gelb, and H. J. Wolters. Polynomial texture maps. In *SIGGRAPH*, pages 519–528, 2001.

[19] D. Miyazaki and K. Ikeuchi. Shape estimation of transparent objects by using inverse polarization ray tracing. *PAMI*, 29(11):2018–2030, 2007.

[20] N. J. W. Morris and K. N. Kutulakos. Reconstructing the surface of inhomogeneous transparent scenes by scatter-trace photography. In *Proceedings of ICCV*, 2007.

[21] S. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Transactions on Robotics and Automation*, 6(4):418–431, aug 1990.

[22] S. K. Nayar, X.-S. Fang, and T. Boult. Separation of reflection components using color and polarization. *IJCV*, 21(3):163–186, 1997.

[23] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi. Efficiently combining positions and normals for precise 3d geometry. In *SIGGRAPH*, volume 24, Aug. 2005.

[24] A. Neubeck, A. Zalesny, and L. V. Gool. 3d texture reconstruction from extensive btf data. In *Texture 2005 Workshop in conj. with ICCV*, pages 13–19, oct 2005.

[25] M. Pharr and G. Humphreys. *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.

[26] H. Rushmeier, G. Taubin, and A. Guéziec. *Applying Shape from Lighting Variation to Bump Map Capture*. IBM TJ Watson Research Center, 1997.

[27] A. Sanderson, L. Weiss, and S. Nayar. Structured Highlight Inspection of Specular Surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(1):44–55, Jan 1988.

[28] K. Sato and S. Inokuchi. Range-imaging system utilizing nematic liquid crystal mask. In *In proceedings of ICCV*, pages 657–661, 1987.

[29] D. Scharstein and R. Szeliski. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision*, 47(1):7–42, 2002.

[30] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proceedings of CVPR*, volume 1, page 195, 2003.

[31] K. Schlüns. Photometric stereo for non-lambertian surfaces using color information. In *Proceedings of CAIP*, pages 444–451, London, UK, 1993. Springer-Verlag.

[32] M. Tarini, H. P. A. Lensch, M. Goesele, and H.-P. Seidel. 3d acquisition of mirroring objects. *Graphical Models*, 67(4):233–259, jul 2005.

[33] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9):1105–1114, 1967.

[34] S. Umeyama and G. Godin. Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *PAMI*, 26(5):639–647, 2004.

[35] J. Wang and K. J. Dana. Relief texture from specularities. *PAMI*, 28(3):446–457, 2006.

[36] L. Wolff. Using polarization to separate reflection components. In *Proceedings of CVPR*, pages 363–369, 1989.

[37] L. B. Wolff. Material classification and separation of reflection components using polarization/radiometric information. In *Proceedings of a workshop on Image understanding workshop*, pages 232–244, San Francisco, CA, USA, 1989. Morgan Kaufmann Publishers Inc.

[38] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, jan/feb 1980.

[39] T.-P. Wu and C.-K. Tang. Dense photometric stereo using a mirror sphere and graph cut. In *Proceedings of CVPR*, volume 2, pages 140–147, 2005.

[40] Y. Yu and J. T. Chang. Shadow graphs and 3d texture reconstruction. *IJCV*, 62(1-2):35–60, 2005.

[41] J. Y. Zheng and A. Murata. Acquiring a complete 3d model from specular motion under the illumination of circular-shaped light sources. *PAMI*, 22(8):913–920, 2000.

[42] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *International Journal of Computer Vision*, 49(2-3):215–227, 2002.

[43] D. E. Zongker, D. M. Werner, B. Curless, and D. H. Salesin. Environment matting and compositing. In *SIGGRAPH*, pages 205–214, 1999.