Surface Normal Reconstruction from Specular Information in Light Field Data

Marcel Gutsche, Hendrik Schilling,
Maximilian Diebold, Christoph Garbe
Interdisciplinary Center for Scientific Computing
Speyerer Str 6, 69115 Heidelberg, Germany
marcel.gutsche@iwr.uni-heidelberg.de

Abstract

Specular highlights provide information about the shape of an object. Its characteristics are mostly unwanted in computer vision due to violation of the Lambertian assumption, which most algorithms require. Instead of neglecting this ubiquitous phenomenon we harvest it to extract surface normals with very high accuracy. Compared to photometric stereo our method works with multiple views and a fixed light source. We only require a low number of observation from a small part of the specular lobe to reconstruct the normal and reflection parameters.

This is achieved by jointly optimizing the normal and the light transport of surfaces points. This work is a proof of concept to demonstrate the feasibility of acquiring highly accurate surface normals from specular reflection, which can be combined with conventional methods. The model is tested and evaluated on synthetic as well as real world data acquired by a cross light field setup.

1. Introduction

Conventional approaches of 3D reconstruction often assume Lambertian surfaces. This means that the intensities observed from different viewing directions remain constant. If specular highlights occur, they are only visible for a small fraction of viewing angles. However, at these angles the object geometry cannot be fully recovered with conventional methods. Instead of disregarding these areas we want to incorporate the additional information provided by specular highlights.

In light-field recordings the surface orientation and material properties are encoded in the intensity distribution of the reflected light. The bi-directional reflectance distribution function (BRDF) describes the material specific characteristics of light absorption and distribution dependent on the ingoing and outgoing light. It can be approximated by mathematical models, with varying complexity depending on the specific use case.
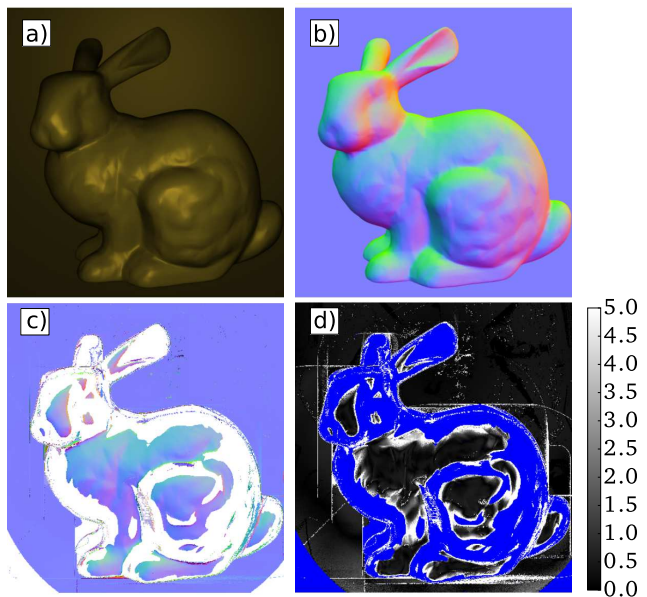

Figure 1: *a)* Original center view image. *b)* Color coded ground truth surface normals. *c)* Reconstructed surface normals. White areas were omitted, since there was not enough signal to calculate reliable estimates. *d)* Angular error in degree.

We propose a novel alternating non-linear optimization scheme to infer BRDF parameters and surface normals at the same time. Only local data is used and we make no further assumptions about neighboring regions.

We developed the an algorithm for normal and BRDF parameter extraction which exploits densely sampled light fields. We recover surface normals with high accuracy in areas where we have specular highlights. Our results are achieved by using a light field cross setup (see Figure 2), but are not limited to these. Besides spacial information, light fields also provide angular information about the scene, which is needed for the BRDF determination. In addition

they handle occlusions particularly well [6]. We explicitly handle occlusions which could otherwise corrupt our model.

2. Related Work

The first methods treating specularities in computer vision were concerned with the detection of specular highlights and the resulting exclusion of the detected areas [2, 7, 12]. To determine surface normals, it is common to use photometric stereo based methods, where objects are acquired from the same position under changing light conditions. A special case of this approach is shape from shading. Here a single image is used to determine the surface normal for known surface parameters.

Photometric stereo, first proposed by Woodham [23], uses the information that the intensity varies with the angle between the surface normal and the direction of the incoming light. Using enough measurements with different light positions, the surface normal can be constrained to a unique solution. Other work also includes specular highlights in their photometric stereo approaches [9, 14].

Early examinations of the information in specular highlights available for a moving observer with a static light source were conducted by Zisserman *et al*. [24]. They concluded that the information contained in the specular highlight by two or more images is sufficient to solve the convex/concave ambiguity, but could not further constrain the curve generated by the moving observer on the surface. Ramamoorthi and Hanrahan [17] describe the reflected light fields as a convolution of the BRDF and the lighting and the process of recovering the scene geometry as deconvolution. Jin *et al*. [11] proposed another approach which utilizes a rank based cost function in a multi-view stereo setting. Other multi-view methods exploit multiple orientations in epipolar images, or similar features [4, 22, 5, 19].

Similar to our proposed method is the work of Adato *et al*. [1] but it is restricted to mirror like surfaces. They use the apparent displacement of the surface highlight in a an optical flow frame work. Nair *et al*. [15] incorporate reflections and material properties in their stereo framework, but they do not actually handle lighting. Instead they attribute all shading effects to the diffuse color.

Oxholm *et al*. use a multi-view setup to infer geometry and reflectance properties by means of a probabilistic model[16]. In contrast to our method the need a full illumination model, whereas our method only needs the position of the strongest light source, which can be inferred by many different methods.

Jachnik *et al*. use SLAM based methods to recover the specular and diffuse components of surfaces to apply augmented reality[10]. However they assume planar surfaces and do not provide surface normal accuracies.
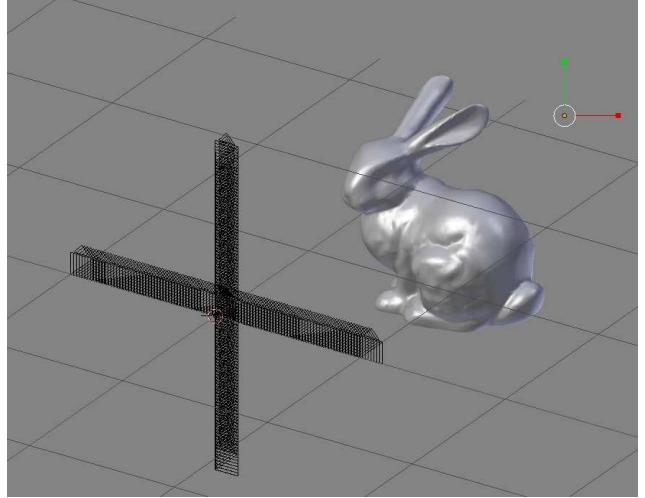
Wang *et al*. proposed a BRDF invariant theory to re-

Figure 2: Cross light field setup.

cover shape from light fields[20]. Their method generates depth and surface normals from input images, but relies on quadratic shape priors and smoothness constraints in neighboring regions.

Methods similar to ours, also use reshifting to avoid the correspondence search for better depth estimates.[18] However, their approach only works on dielectric materials, since they can not handle light color changes caused by the material.

3. Overview

To infer surface normals in 3D space one must sample viewpoints in two linearly independent directions. The intensity variation along one viewing direction constrains the surface normal in one dimension. Therefore, we use an orthogonal second direction, which is given by a cross-shaped camera array, to constrain the surface normal in two dimensions. A possible setup can be seen in Figure 2. In our approach we use the intensity distribution as observed from both acquisition directions. The resulting intensity distributions for each surface point obtained by the light-field acquisition has sufficient information to achieve high-accuracy surface normals and determine BRDF parameters.

We can picture the different intensity distributions as cuts through the 3D specular lobe. This means that we can still deliver estimates, even if only a part of the specular lobe is visible.

The processing pipeline is depicted in Figure 3. As input the method requires the calibrated camera and light source positions, as well as approximate disparity estimates. To calculate an initial disparity map we use the structure tensor method proposed by Wanner *et al*. [21].

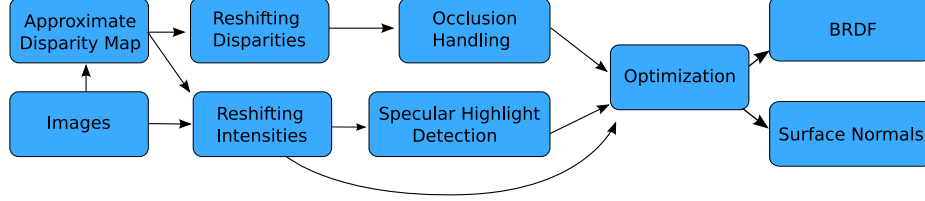Thus, we define light fields by using the Lumigraph no-

Figure 3: Process of simultaneously estimating surface normals and BRDF parameters.

tation [8, 13]. The Lumigraph is defined using two parallel planes Π and Ω. The first plane Ω denotes image coordinates $(x, y) \in \Omega$. The second plane Π contains the focal points $(s, t) \in \Pi$ of all cameras. In our special case for a cross shaped light field we sample along s and t independently, so we substitute both directions by $k \in s, t$ for simplicity.

$$L : \Omega \times \Pi \to \mathbb{R} \qquad (k, x, y) \mapsto L(k, x, y). \qquad (1)$$

To slice out an epipolar plane image (EPI), we fix the image dimension corresponding to the camera direction, *i.e.* for the horizontal direction we fix $y = y^*$ and $t = t^*$. Thus an EPI is defined as

$$S_{y^*}(k, x) := L(k, x, y^*). \qquad (2)$$

EPIs encode, besides the depth information, also the intensity distribution, along each orientation line, needed to determine material properties.

Regarding the surface points visible in the center view, we use the disparity, to map the input intensities to an image stack $L'(k, x, y)$ where $L'_{x^*, y^*}(k)$ represents the intensity distribution for a single surface point as observed from different views (see section 4). A second stack is generated accordingly using the disparity as input data. This way, occlusion maps can be computed as well, as detailed in subsection 4.1. They prevent the algorithm from mixing fore and background information. By regarding the unoccluded intensity values, we can identify surface points which exhibit specular characteristics by measuring the intensity change along the views, see subsection 4.2.

Finally, we optimize the BRDF parameters and the surface normal independently for each pixel, for the regions where we identified specular intensity variations, see subsection 5.2. Our method works on the following assumptions:

- We assume an approximate disparity map, in our implementation we utilize the structure tensor.

- The cameras, as well as the light source is calibrated in terms of location and camera intrinsics.

- The light transport is dominated by the single-bounce reflection of a point light source.

- At least two independent viewing directions are needed, *e.g.* by using a cross setup.

4. Preprocessing

To compute the intensity changes of object points in an efficient manner, we need to address all pixels related to a specific surface point. This principal is similar to the reshifting in EPI processing as introduces by Diebold and Goldluecke [6]. The specific shifting is given for the horizontal light field by

$$I'(k, x, y) = I(k, x + d \cdot (c - k), y), \qquad (3)$$

which maps all related pixels to a vertical line, which is easy to access for further processing, see Figure 4. The variable c corresponds to the index of the center view and d to the disparity of the center view at (x, y). The vertical viewing direction is handled analogously. It is important to note here, that the method is especially well suited to smooth, texture-less regions. Here, erroneous disparities have a low impact on the accuracy of the normal estimation, since even if intensity information from neighboring surface points are mistakenly used, the material properties change slowly compared to the error in the disparity. These areas are particularly difficult for conventional methods, which rely on structural information, such as strong image gradients. This way we can utilize even approximate disparities. In the case of strong material changes we expect precise disparity measurements of the structure tensor, and expect that these methods complement each other particularly well.

4.1. Occlusion Handling

To get useful intensity variations from the surface points along $I'(k)$, occlusions need to be treated carefully. To this end we handle these explicitly. Due to the dense information provided by the light field this can be done by examining the approximate disparity maps which are denoted by $d(k, x, y)$. The disparity maps are shifted in the same way as image intensities to calculate occlusion boundaries. Given

$$d'(k, x, y) = d(k, x + d \cdot (c - k), y), \qquad (4)$$

we can identify occlusions easily by looking for strong changes along $d'(k)$. To this end we apply a Gaussian derivative filter g on $d'(k)$ and threshold it depending on the noisiness of our input disparity. Our occlusion map Γ is then given by

$$a(x, y) = \frac{\partial}{\partial k}(g * d'(k, x, y)) \tag{5}$$

$$\Gamma(k, x, y) = \begin{cases} 1, & \text{if } a \geq t_o \\ 0, & \text{otherwise} \end{cases}, \tag{6}$$

where t_o is a threshold depending on the noise ratio in the disparity map. This way we can incorporate even multiple occlusions in our framework.

4.2. Specular Highlight Detection

Since we can only extract reliable information from regions with specular highlights, see subsection 5.1, we need to decide for which regions of the image we should carry out our analysis. To this end we test if the intensity variation along $I'(k)$ is above a certain threshold t_s which is calculated analogous to Equation 5. Then, we only apply our algorithm at locations where our confidence mask χ is one, see Figure 1 c), where the masked regions are white.

$$a(u, v) = \sum_k \frac{\partial}{\partial k}(g * I'(u, v, k)) \tag{7}$$

$$\chi(u, v) = \begin{cases} 1, & \text{if } a(u, v) \geq t_o \\ 0, & \text{otherwise} \end{cases}. \tag{8}$$

5. Reflectance Model

The governing equation for the light transport is the render equation. It describes the amount of light emitted from a point $\mathbf{x}$ on the surface in direction of the observer. In the following we denote the directions from the surface point to the observer with the normalized vector $\mathbf{v}$ and the direction to the light source with $\mathbf{l}$. Using this notation the rendering equation in a point $\mathbf{x}$ is given by

$$L(\mathbf{v}) = L_e(\mathbf{v}) + \int_\Omega f_r(\mathbf{v}, \mathbf{l}) \, L(\mathbf{l}) \, (\mathbf{n} \cdot \mathbf{l}) \, d\mathbf{l}, \tag{9}$$

where we integrate over all incoming light directions. The individual terms are given in the following:

- $L(\mathbf{v})$ denotes the amount of light emitted from $\mathbf{x}$ in the direction of the observer $\mathbf{v}$.

- $L_e(\mathbf{v})$ is the emission term. If $L_e(\mathbf{v}) > 0$ we have a light source at $\mathbf{x}$.

- The material interaction term $f_r(\mathbf{v}, \mathbf{l})$ is described by the BRDF with incoming direction $\mathbf{l}$ and reflecting direction $\mathbf{v}$.
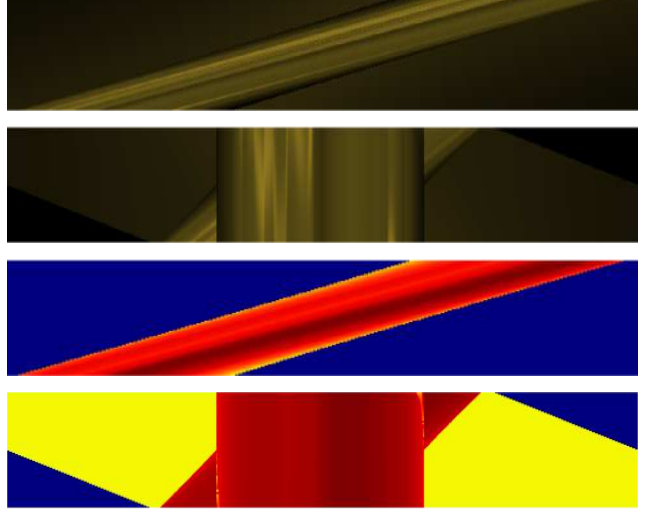

Figure 4: a) The orientation in the horizontal EPI directly relates to the disparity of the input image. b) Reshifting the EPI by the extracted disparity allows to investigate the intensity distribution by moving along the views which are here along the y-direction. c) Similarly the disparity for all views can be viewed in a "disparity EPI". d) Reshifting it leads to a straight forward computation of occlusions by looking at gradients in the y-direction.

- $L(\mathbf{l})$ describes how much light is reaching $\mathbf{x}$ from the direction $\mathbf{l}$.

- The scalar product $\mathbf{n} \cdot \mathbf{l}$ denotes the shading term, where $\mathbf{n}$ is the surface normal.

This equation describes the radiation on a unit sphere. To account for the inverse-square law we need to divide by the distance $\mathbf{r}$ the light has traveled to the object surface point. Assuming that we have only one dominant light source and no significant contributions due to multiple bounce reflection the equation simplifies to

$$L(\mathbf{v}) = \frac{L_0}{r^2} f_r(\mathbf{v}, \mathbf{l}) (\mathbf{n} \cdot \mathbf{l}), \tag{10}$$

where L_0 is the amount of light coming from a single light source.

In a calibrated cross light-field setup we know the position of the light source and our cameras which allows solving for $\mathbf{n}$ by means of a nonlinear solver. The computationally most expensive part comes from the BRDF. We test our concept on the rather simple Blinn-Phong model which is given by

$$f_r = k_d + k_s(\mathbf{n} \cdot \mathbf{h})^m, \tag{11}$$

where $\mathbf{h}$ is the half-way vector between the incident light ray and the outgoing ray to the observer. The other parameters k_d, k_s and m are non-physical and describe the

reflectance of the material. The parameter k_s and k_d denote the specular and diffuse term. The exponent m defines the sharpness of the specular peak and must be positive.

5.1. Joint Estimation of Surface Normals and BRDF Parameters

Given the visible maximum of the intensity of one surface point, as observed by a linear camera array, we can constrain the surface normal by the plane spanned by the object point and the different cameras. Given a second maximum by a linearly independent camera array we tighten that constraint to get the complete 2D surface normal. An analytical solution for this problem exists, but has the disadvantage, that both highlight maxima must be visible. To circumvent this restriction and to leverage the information contained in the light field, we optimize the surface normals of one surface point jointly with the BRDF parameters taken into account all observations. This way we can take advantage of even slight intensity variations from the tails of the specular highlights and do not require the observation of the specular peak in both viewing directions. Of course some part of the specular region must still be visible. In essence there are three different situations, see Figure 5b:

1. Along both viewing directions no intensity changes are visible. Thus, we cannot estimate the surface normal because we only see the diffuse part of our surface point. Since we have no specularities, while our method can not compute normals in this case. However, conventional methods can be applied in these cases without problems.

2. One viewing direction exhibits some intensity variation. We can say that the surface normal must be on one curve constraint by our model and can exclude all normals which would lead to an intensity variation in the second viewing direction. So basically, the wider the peak the lesser the surface normal error.

3. Both viewing directions exhibit some intensity variation. In theory, the surface normal is uniquely identifiable.

5.2. Initialization and Optimization

It is very important to choose a reasonable initial guess to ensure convergence to the global optimum. We initialize the surface normals by the half-way vector between the light source and the central camera. If the surface normal coincides with the half-way vector we have the largest possible BRDF peak, compare Figure 5b. The BRDF parameter k_d and k_s are initialized to match very roughly the given intensity distribution by calculating

$$k_s = \max(I'(k)), \tag{12}$$
$$k_d = \min(I'(k)). \tag{13}$$

To avoid ambiguous solutions we remove the common constraint of $k_s + k_d \leq 1$ and set the amount of incoming light L_0 as well as the fall of term r^{-2} equal to one. So the outgoing light will be completely determined by k_s and k_d. For each pixel we have as many measurements $\mathbf{y_i}$ as viewpoints k. Thus, we minimize

$$\underset{\mathbf{n}, k_d, k_s, m}{\arg\min} \sum_i \|\mathbf{y_i} - L_i(\mathbf{v})\|, \tag{14}$$

where i is the index variable for each view, where we have no occlusion with respect to the center view. To ease the solving procedure the surface normal is parametrized in spherical coordinates θ and ϕ. We first keep the BRDF parameter constant and vary only the surface normals. In a second step we vary the BRDF parameters and the surface normal jointly. This way we force the solver to explain a constant intensity distribution by moving the surface normal in contrast to simply reducing k_s. This uses information where we have no intensity variation and can exclude a range of surface normals which otherwise would create a visible highlight.

6. Experiments

Synthetic Evaluation   For the evaluation of our algorithm we used rendered images obtained with Blender [3]. Our cross array consists of 101 cameras in horizontal as well as in vertical direction. In Figure 7 we show how the angular error decreases with the number of views. To evaluate the error in the BRDF we use the following metric to quantify a relative error:

$$e = \sum_i \sqrt{\left(\frac{p_{i,\text{real}} - p_{i,\text{est}}}{p_{i,\text{real}}}\right)^2}, \tag{15}$$

where p_i stands for the i-th BRDF parameter.

The evaluation of the synthetic data shows that many surface normals have an angular error of less than $1°$ and almost all areas have an angular error of less than $5°$ (Figure 1). Surface normals where the intensity variation is not above a certain threshold, were not calculated.

In Figure 8 we used a constant disparity for each view to simulate unknown depth. We keep the disparity map constant at roughly the average distance of the bunny. Thus, we have a greater impact on the accuracy in the background, where the relative disparity error is larger. Despite the introduced inaccuracies the angular error is rather small (still around $1°$) for regions with strong specular highlights. Due to the constant disparity, we could not create a sensible specular mask, so we omitted the mask completely for the inaccurate case.

Evaluation on Real World Objects   To examine the feasibility of our concept for real world data we used a strongly
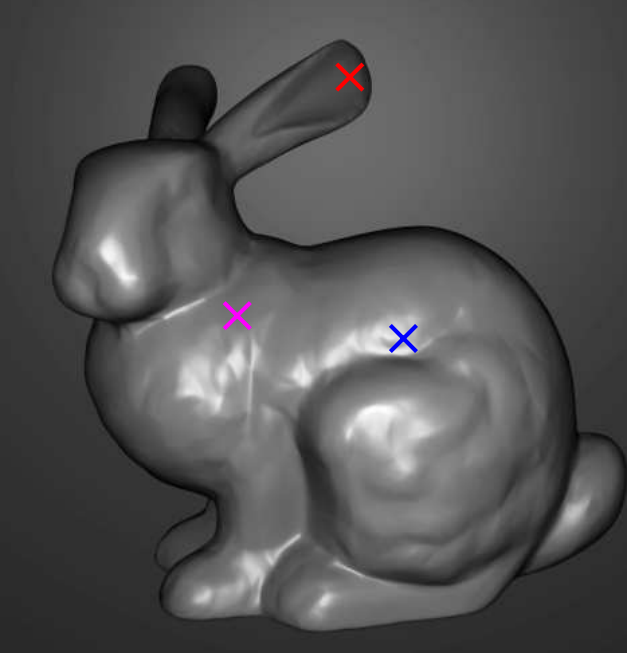
(a) Center view of the Stanford Bunny. The markers indicate the positions of the corresponding intensity distributions depicted in Figure 5b. The red marker identifies a region with Lambertion reflection, the purple one with specular components only in the vertical direction and in the last one both intensity maxima are clearly visible.
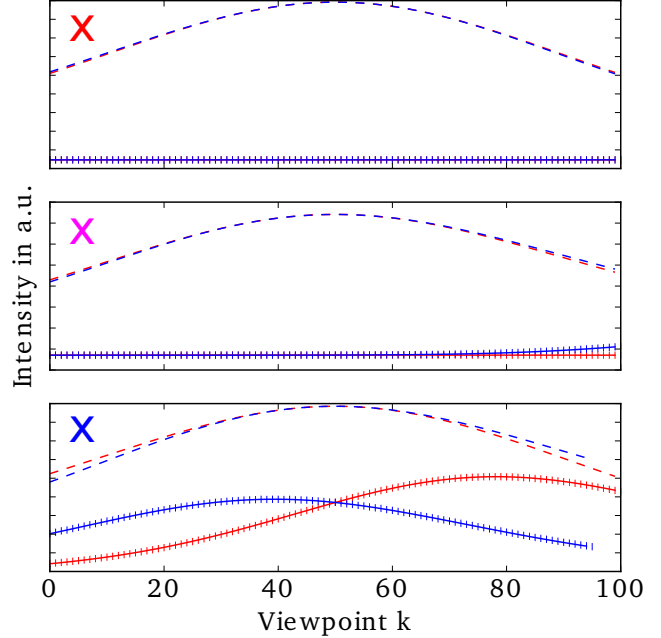
(b) Intensity distribution for different surface normals. Red corresponds to the horizontal camera array and blue to the vertical. The dashed lines indicates the model initialization, and the solid line the recovered model. From top to bottom: No intensity variation visible – the Lambertian case (marked red in Figure 5a). Variation along the vertical direction visible (purple). Both peaks visible (blue). The corresponding angular errors for the normal reconstruction at these points are $54.6°$, $4.2°$ and $0.5°$.

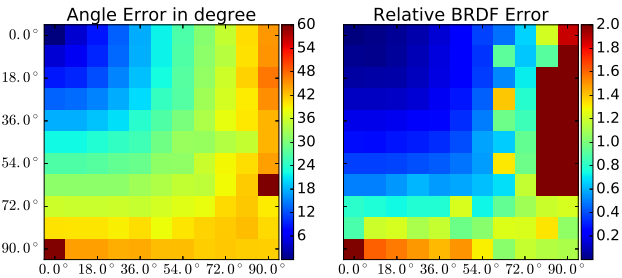Figure 5: Overview of different intensity distributions.

Figure 6: Analysis of the angular error and the relative error in the BRDF parameters, dependent on the angle error of the light direction. We rotated the light direction around the x and the y-axis up to $90°$. We took 10 different samples and took the median for each different light orientation.

specular coin (see Figure 10). To capture the data we use a gantry with two orthogonal axes to capture images in cross setup. We used 101 images for the horizontal as well as vertical axis. The surface normals reveal slight scratches. The intensity distribution together with the fit shows, that the Blinn-Phong model is not well suited to fully explain the observed intensities.

7. Conclusion

Specular highlights still pose a great challenge in computer vision and most available approaches ignore the additional information they can provide. This is mostly due to the inherent ambiguity of light source, BRDF and shading terms which influence the light transport and are difficult to separate.

Our method tackles the separation between diffuse and specular shading, while at the same time extracting surface normals to explain the observed images. We achieve extremely accurate surface normals using only a small angular range, without the need to directly observe intensity maxima.

Our optimization scheme implicitly utilizes the information of regions where no specular intensity variation is present. This is achieved by constraining the normals to those ranges which cause the specular highlight to fall outside of the observed view range.
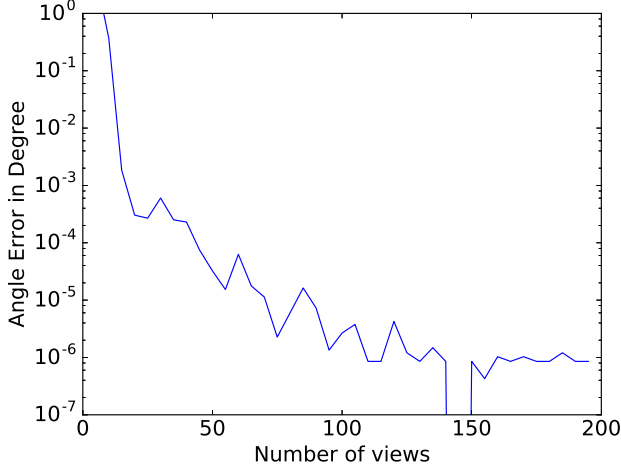
Figure 7: Analysis of the angular error, dependent on the number of views. The precisions increases remarkably up to 30 views. With more than 150 views discretization errors play a significant role and we can't increase the precision anymore.
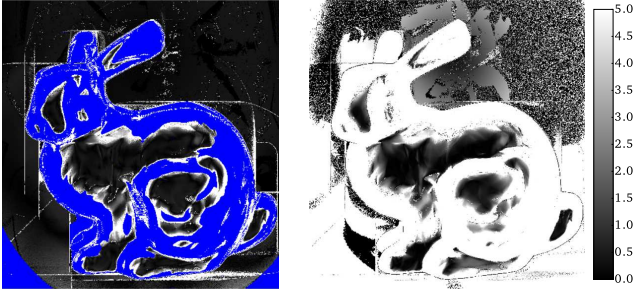

Figure 8: Comparison of angular errors for the reconstruction with correct disparity maps (left) and with a constant disparity roughly in the plane of the bunny.

8. Outlook

So far our method only makes use of local information. Introducing smoothness constraints should improve the estimates. Towards this end, it would be reasonable to integrate this BRDF based normal estimation method with a conventional depth estimation technique, within a unified framework which jointly optimizes depth, BRDF parameters and surface normals.

We make use of the Blinn-Phong model, but more complicated visual phenomena, such as the Fresnel effect, or anisotropic reflecting materials call for more complex, physically accurate models.

For simplicity we have focused on linear light fields, however this method is applicable to any at least two-dimensional distribution of viewpoints. This means that the framework can be used in a wide range of applications
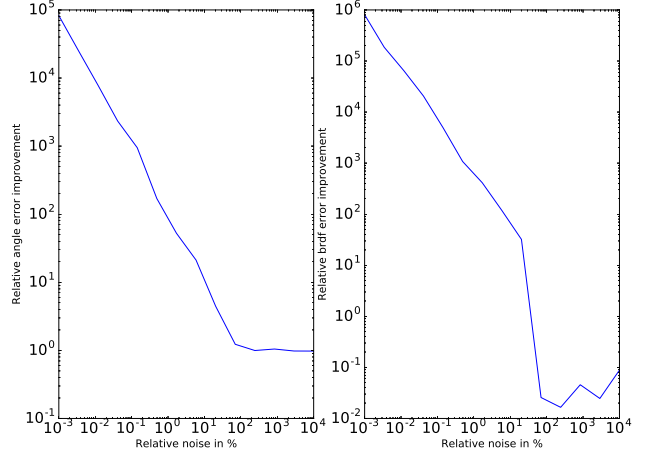

Figure 9: Comparison of angular and BRDF error improvements for the reconstruction with increasing amount of noise on the intensity values. Since the optimization may get stuck to a local minimum more easily in the presence of noise, we use the improvement yielded by the optimization with respect to the initialization as metric. The improvement is calculated by $p = \frac{v_{\mathrm{init}}}{v_{\mathrm{opt}}}$, where v stands for the angular error or the BRDF error accordingly.
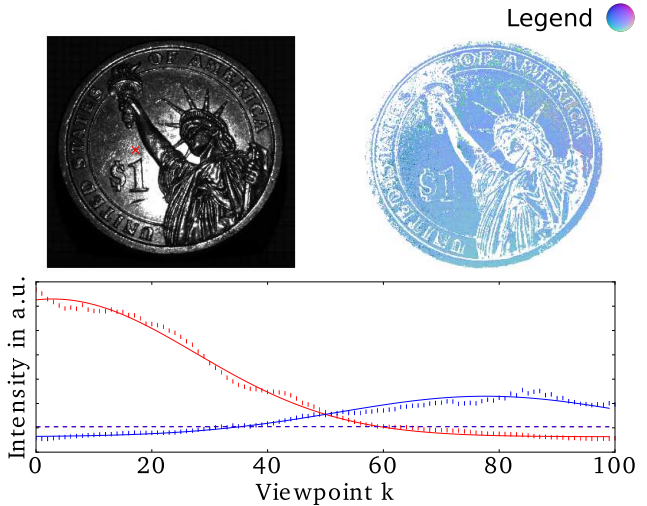

Figure 10: *a)* Original center view image. *b)* Color coded surface normals. White areas were omitted by our specular detection routine. c) Intensity curves for the horizontal and vertical view. The straight lines indicate the estimated model and the dashed line the initialization.

to improve the geometry reconstruction in the presence of specular reflections.

References

[1] Y. Adato, Y. Vasilyev, T. Zickler, and O. Ben-Shahar. Shape from specular flow. *IEEE transactions on pattern analysis and machine intelligence*, 32(11):2054–2070, 2010. 2

[2] R. Bajcsy, S. W. Lee, and A. Leonardis. Color image segmentation with detection of highlights and local illumination induced by inter-reflections. In *Pattern Recognition, 1990. Proceedings., 10th International Conference on*, volume 1, pages 785–790. IEEE, 1990. 2

[3] Blender Online Community. Blender - a 3D modelling and rendering package, 2016. 5

[4] R. C. Bolles, H. H. Baker, and D. H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, 1(1):7–55, 1987. 2

[5] A. Criminisi, S. B. Kang, R. Swaminathan, R. Szeliski, and P. Anandan. Extracting layers and analyzing their specular properties using epipolar-plane-image analysis. *Computer vision and image understanding*, 97(1):51–85, 2005. 2

[6] M. Diebold and B. Goldluecke. Epipolar plane image refocusing for improved depth estimation and occlusion handling. 2013. 2, 3

[7] R. Gershon, A. D. Jepson, and J. K. Tsotsos. The use of color in highlight identification. In *IJCAI*, pages 752–754, 1987. 2

[8] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 43–54. ACM, 1996. 3

[9] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler. A photometric approach for estimating normals and tangents. *ACM Transactions on Graphics (TOG)*, 27(5):133, 2008. 2

[10] J. Jachnik, R. A. Newcombe, and A. J. Davison. Real-time surface light-field capture for augmentation of planar specular surfaces. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, pages 91–97. IEEE, 2012. 2

[11] H. Jin, S. Soatto, and A. J. Yezzi. Multi-view stereo beyond lambert. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–171. IEEE, 2003. 2

[12] S. W. Lee and R. Bajcsy. Detection of specularity using color and multiple views. In *European Conference on Computer Vision*, pages 99–114. Springer, 1992. 2

[13] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42. ACM, 1996. 3

[14] F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato. Uncalibrated photometric stereo for unknown isotropic reflectances. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1490–1497, 2013. 2

[15] R. Nair, A. Fitzgibbon, D. Kondermann, and C. Rother. Reflection modeling for passive stereo. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 2291–2299. IEEE, 2015. 2

[16] G. Oxholm and K. Nishino. Multiview shape and reflectance from natural illumination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2155–2162, 2014. 2

[17] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 117–128. ACM, 2001. 2

[18] M. W. Tao, T.-C. Wang, J. Malik, and R. Ramamoorthi. Depth estimation for glossy surfaces with light-field cameras. In *European Conference on Computer Vision*, pages 533–547. Springer, 2014. 2

[19] Y. Tsin, S. B. Kang, and R. Szeliski. Stereo matching with reflections and translucency. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–702. IEEE, 2003. 2

[20] T.-C. Wang, M. Chandraker, A. A. Efros, and R. Ramamoorthi. Svbrdf-invariant shape and reflectance estimation from light-field cameras. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 2

[21] S. Wanner and B. Goldluecke. Globally consistent depth labeling of 4d light fields. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 41–48. IEEE, 2012. 2

[22] S. Wanner and B. Goldluecke. Reconstructing reflective and transparent surfaces from epipolar plane images. In *German Conference on Pattern Recognition*, pages 1–10. Springer, 2013. 2

[23] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139–191139, 1980. 2

[24] A. Zisserman, P. Giblin, and A. Blake. The information available to a moving observer from specularities. *Image and vision computing*, 7(1):38–42, 1989. 2