Vision, Modeling, and Visualization (2016)
M. Hullin, M. Stamminger, and T. Weinkauf (Eds.)

Reflection Separation in Light Fields based on Sparse Coding and Specular Flow

Antonin Sulc, Anna Alperovich, Nico Marniok and Bastian Goldluecke

University of Konstanz

Abstract

We present a method to separate a dichromatic reflection component from diffuse object colors for the set of rays in a 4D light field such that the separation is consistent across all subaperture views. The separation model is based on explaining the observed light field as a sparse linear combination of a constant-color specular term and a small finite set of albedos. Consistency across the light field is achieved by embedding the ray-wise separation into a global optimization framework. On each individual epipolar plane image (EPI), the diffuse coefficients need to be constant along lines which are the projections of the same scene point, while the specular coefficient needs to be constant along the direction of the specular flow within the epipolar volume. We handle both constraints with depth-dependent anisotropic regularizers, and demonstrate promising performance on a number of real-world light fields captured with a Lytro Illum plenoptic camera.

Categories and Subject Descriptors (according to ACM CCS): I.4.4 [Image Processing and Computer Vision]: Restoration—I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Color

1. Introduction

Many algorithms in computer vision are designed under the assumption that the appearance of a scene point does not change under different perspectives. However, this is clearly not the case for most natural scenes which are usually not perfectly Lambertian, and where surfaces can exhibit significant view-dependant specular reflectance. Thus, conventional methods for image analysis often fail on such scenes due to different motion of the specular highlights and actual physical points under change of view point.

To remedy this, current approaches often leverage that specular or glossy pixels can be decomposed under the assumption of a dichromatic reflection model [Sha85], where the observed color of a scene point is a sum of diffuse and specular components. Classical algorithms for specularity detection and removal for a single image are however strongly limited by the fact that the image does not contain information about scene geometry, and thus methods have to rely on color information.

In this work, we assume that we have captured a light field of the scene, so that we have immediate access to scene geometry. Indeed, depth reconstruction from light fields is currently a very active area of research [WG14, HRP13, TSW*15, JSG16]. For our purposes here, a light field is considered as densely sampled set of images taken from a regular grid of view points inside a common focal plane. By stacking these views and computing slices, one can directly observe the depth of pixels, as it is inversely proportional to the shift of corresponding pixels when moving the camera, i.e. disparity. The slices through stacks of views are called epipolar plane images (EPIs), see figure 1.

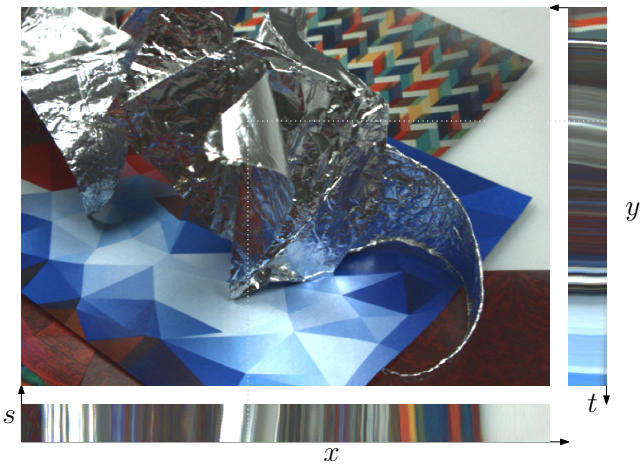

Figure 1: *Center view of a light field. The epipolar plane image along the dotted horizontal line is shown below, the one along the vertical line to the right, respectively. Coordinates (x,y) are for the image plane, coordinates (s,t) for the focal plane, so the complete light field is defined on a four dimensional ray space $\mathcal{R}$.*

However, constancy of pixel intensity is valid only for Lambertian surfaces, because specularity is view-dependent [SKS*02]. A natural way to deal with specularities is thus to separate the diffuse component from the specular while preserving constancy along the projections in the EPI only for the diffuse part. We model this in the global variational framework [GW13]. The remaining challenge is to enforce consistency of the specular part, which is the major focus of our theoretical analysis. We base our model for regularization on specular flow [AVaTZ07], which under certain practical assumptions can be computed directly from the estimated disparity.

Contributions. In this paper, we make the following main contributions to apply reflection separation to light fields:

- We extend the sparse coding approach [AO14] for reflection separation on single views to light fields by embedding it in the variational framework [GW13] for consistent light field labelling.
- We propose a novel term for the specular coefficient in epipolar volumes which is based on specular flow.
- We show how to optimize jointly for consistent diffuse and specular coefficients.
- We evaluate performance on Lytro datasets, and show improvements compared to previous work, in particular the original approach [AO14] on single views as well as the alternative light field based approach [TSW*15].

2. Related work

There is a wide range of methods for specular removal from a single image. Most of them are based on color statistics or rely on an additional knowledge about the scene. In [KSK87], they perform color space analysis by clustering the diffuse and specular component based on estimated illumination color. A similar approach was chosen in [BLL96], where they analyze color variations on the object based on hue, saturation, and brightness information. Another category of approaches are neighbour based methods, where authors additionally use spatial information [TI05, YCK06]. For an extensive survey on specular removal techniques, we refer to [ABC11]. Our own work is based on the idea in [AO14], where the authors use sparse non-negative matrix factorization to separate diffuse and specular components from a single image. They represent the color of every pixel as a sparse linear combination of a diffuse palette and one specular color based on estimated light source color. The amount of a certain color taken from the palette is controlled by a sparse set of coefficients regularized with an L^1 norm, and the palette is iteratively optimized together with the sparse coefficients.

In our paper, we leverage that having a light field available makes it much easier to separate diffuse and specular parts due to the difference in motion. Although the rich light field structure allows to obtain a lot of information about scene geometry, there is not much research in the area of light field specularity tracking and removal. An analysis of glossy and specular surfaces in light fields was done by [TSW*15], where they analysed the dichromatic model for light fields. They used the natural assumption for the diffuse component to be photo-consistent, while the specular component exhibits strong variance in subaperture views. However, in contrast to our proposed approach, they do not incorporate exact motion of specularities due to specular flow. In [LLK*02], the authors analyze diffuse-specular separation from an image sequence, which gives a structure similar to a light field. Their histogram-based approach is built upon the assumption that diffuse pixels do not vary from one view to the next, while specularities are view-dependent. By calculating the difference of color histograms between views, they localize specular regions. Another multi-view approach is proposed in [SKS*02], where they study the deviation from photometric properties under presence of specularities on the EPI level. In [PDCS13] they analyzed specularities as a part surface light field estimation. Recently, Wang et al. [WCER16] proposed a method for joint shape and reflectance estimation in light fields based on a theory for differential appearance changes for single-lobe BRDFs.

3. Reflection model and sparse light field coding

We formulate reflection separation as a sparse coding problem, similar to [AO14], but extended to the light field case. Our main contribution is to formulate and impose spatial regularizers and constraints required for a consistent separation across all subaperture views, and optimize for all of them in a unified global optimization framework. In this section, we first review the ideas in [AO14] and in particular adapt them to the light field scenario.

The basic approach to diffuse and specular separation is based on the assumption of the dichromatic reflection model introduced by [Sha85], and adopted by [TSW*15] for light fields. The key assumption is that the light field color $L(\boldsymbol{r})$ of every ray $\boldsymbol{r}$ in 4D rayspace is a sum $L(\boldsymbol{r}) = S(\boldsymbol{r}) + D(\boldsymbol{r})$ of a specular and diffuse term. Following the dichromatic model, the specular reflection is given by $S(\boldsymbol{r}) = S_c\sigma(\boldsymbol{r})$, where S_c is a globally constant RGB illumination color, and $\sigma(\boldsymbol{r})$ the specular coefficient for the ray $\boldsymbol{r}$. The assumption for the diffuse term is that the objects in the scene consist of regions of different materials, thus it is drawn from a discrete set of different RGB albedos $\{A_1,\dots,A_K\}$. Thus,

$$L(\boldsymbol{r}) = S_c\sigma(\boldsymbol{r}) + A_1\alpha_1(\boldsymbol{r}) + \cdots + A_K\alpha_K(\boldsymbol{r}), \tag{1}$$

where $\boldsymbol{\alpha}(\boldsymbol{r}) = (\alpha_1(\boldsymbol{r}),\dots,\alpha_K(\boldsymbol{r}))$ are the ray-wise diffuse coefficients. Ideally, we would like the visible material at every ray to be unique, mathematically this means the counting norm $\|\boldsymbol{\alpha}(\boldsymbol{r})\|_0$ should be equal to one. All coefficients must also be non-negative. However, as the counting norm is non-convex, we relax this constraint as well as assumption (1), and instead minimize for every ray the L^1 sparse coding energy

$$E(\boldsymbol{u}(\boldsymbol{r}), D) = \lambda_s\|\sigma(\boldsymbol{r})\|_1 + \lambda_d\|\boldsymbol{\alpha}(\boldsymbol{r})\|_1 + \|L(\boldsymbol{r}) - D\boldsymbol{u}(\boldsymbol{r})\|_2^2, \tag{2}$$

where $\lambda_s,\lambda_d > 0$ are constants controlling the amount of sparsity for the diffuse and specular coefficients, $D = \begin{bmatrix} S_c & A_1 & \cdots & A_K \end{bmatrix}$ is the global dictionary of specular and diffuse colors, the same for every ray, and $\boldsymbol{u}(\boldsymbol{r}) = \begin{bmatrix} \sigma(\boldsymbol{r}) & \boldsymbol{\alpha}(\boldsymbol{r}) \end{bmatrix}^T$ are the coding coefficients, different for every ray.

Note that while the dictionary for specular color and basis albedos is the same across the light field, equation (2) still leads to independent coefficient estimates for every ray. To make the result spatially consistent and also consistent across all subaperture views of the light field, we need to impose regularizers. The principles behind light field consistency for both diffuse and specular components will be discussed in the next section. Everything will be integrated into a global variational model with additional spatial regularization in section 5.

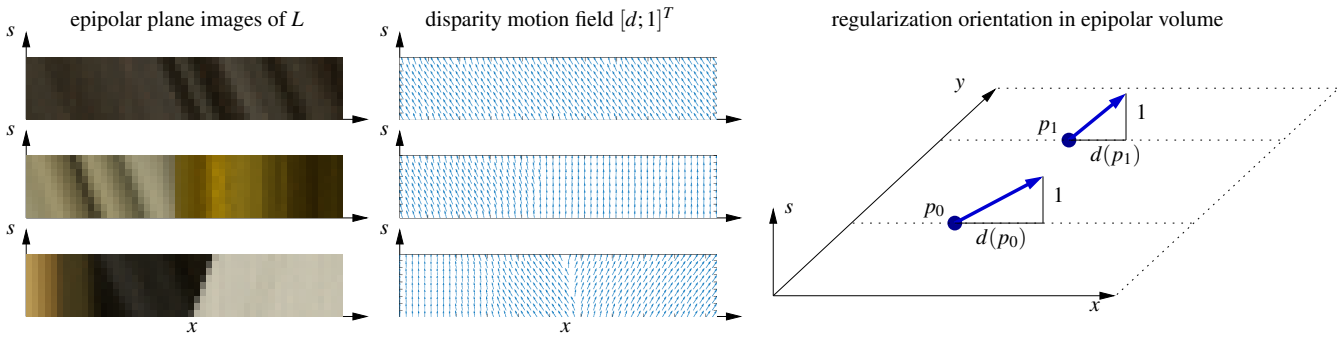

Figure 2: *Disparity-based regularization of diffuse components. For every 2D EPI, disparity induces a vector field along which any property assigned to a 3D point remains constant [GW13]. This is leveraged in Lambertian depth reconstruction, where one can search for the orientation of lines with constant color [WG14] (left and middle column). When moving the camera in s-direction, pixels shift only in x-direction, thus regularization can be treated separately for every 2D EPI [GW13]. However, one can also stack all EPIs and regularize the complete epipolar volume along a 3D vector field whose y-component is zero everywhere (right column). This is much faster as it reduces computational overhead like calling CUDA kernels, and we can improve the runtime of the original method in [GW13] by a factor of ten.*

4. Separation consistency based on depth and specular flow

As the diffuse and specular components follow different motions, we require two different reguralizers on epipolar space. We encourage consistency with the estimated motion by anisotropic smoothing in the respective direction. For this, we will first describe regularization of epipolar plane images and volumes along a vector field in general, and then turn to the specialized models for the two components.

Formally, EPIs are restrictions of a vector-valued function $\boldsymbol{u}$ to a subset of the variables. If we fix s and x, then we obtain an EPI $\boldsymbol{u}_{sx}$ in the variables (t,y), conversely, if we fix t and y, we obtain an EPI $\boldsymbol{u}_{ty}$ in the variables (s,x). In this notation, $\boldsymbol{u}_{st}$ is the subaperture view at view point coordinates (s,t). If we fix only a single view point coordinate, we obtain epipolar volumes $\boldsymbol{u}_s(t,x,y)$ and $\boldsymbol{u}_t(s,x,y)$, respectively. While in the previous work [GW13], it was only necessary to regularize on 2D EPIs, correct regularization of specular coefficients requires to work in epipolar volumes. We will explain this in detail in the remainder of the section.

In the following, let $\boldsymbol{e}$ be any epipolar plane image or volume. Ideally, we want to impose constancy of every scalar component $\boldsymbol{e}_i$ of the vector-valued function $\boldsymbol{e}$ along the vector field $\boldsymbol{v}$ (two- or three-dimensional depending on the dimension of $\boldsymbol{e}$). In practice, this cannot be achieved exactly, as we have to deal with a discrete approximation. We thus follow [GW13] and define an anisotropic regularizer which for every point $\boldsymbol{p}$ only encourages constancy in direction of the vector field $\boldsymbol{v}(\boldsymbol{p})$. It is defined as the anisotropic total variation over all components,

$$J_{\boldsymbol{v}}(\boldsymbol{e}) := \sum_i \int_{\mathrm{dom}(\boldsymbol{e})} \sqrt{\nabla \boldsymbol{e}_i^T (\boldsymbol{v}\boldsymbol{v}^T) \nabla \boldsymbol{e}_i}\, \mathrm{d}\boldsymbol{p}. \tag{3}$$

It encourages smoothing into the direction of $\boldsymbol{v}$, as it becomes larger the more $\boldsymbol{v}$ is aligned with the gradient of the components $\boldsymbol{e}_i$.

Consistency for diffuse components given a disparity map. As each scene point has a unique diffuse color independent of the view point, the diffuse coefficients $\boldsymbol{\alpha}$ must be constant along projections of the same point. Projections belonging to the same points manifest in the linear structure which is visible on the EPIs of a Lambertian scene, see figures 1 and 2. The slope of the lines is related to disparity. Let d be the disparity map defined on rayspace for the complete light field, then this disparity map induces a vector field $\begin{bmatrix}1 & d\end{bmatrix}^T$ on each EPI along which the diffuse components need to be constant. While disparity is technically a differential quantity, this is in accordance with the usual meaning that a shift of one unit in view point direction implies a shift of d of the respective image coordinate for corresponding rays. See [GW13] for more details.

In summary, the anisotropic regularizer (3) needs to be specialized to the diffuse component α_v by applying it on every EPI with smoothing in direction $\begin{bmatrix}1 & d\end{bmatrix}^T$. We thus obtain the regularizer J_d as

$$J_d(\boldsymbol{\alpha}) = \int J_{[1\ d]^T}(\boldsymbol{\alpha}_{s,x})\,\mathrm{d}(s,x) + \int J_{[1\ d]^T}(\boldsymbol{\alpha}_{t,y})\,\mathrm{d}(t,y). \tag{4}$$

In this form, the regularizer is a sum over contributions of 2D EPIs. We note, however, that when shifting the view point in a single direction, i.e. s, then the pixels shift only in x-direction, while shift in y-direction is zero. This means that we can rewrite (4) as a sum over contributions of 3D epipolar volume regularizers,

$$J_d(\boldsymbol{\alpha}) = \int J_{[1\ d\ 0]^T}(\boldsymbol{\alpha}_s)\,\mathrm{d}s + \int J_{[1\ 0\ d]^T}(\boldsymbol{\alpha}_t)\,\mathrm{d}t. \tag{5}$$

While this form does not look considerably simpler, it is now unified with the regularizer for the specular component defined below. Furthermore, it turns out that in the final implementation, it is about ten times faster to deal with a single epipolar volume regularization compared to the sum over all 2D EPIs.

Consistency for specular components given specular flow. When moving the camera, the reflection follows a motion which is different from the motion of surface points. This motion field is called the specular flow. According to the specular motion model proposed by Blake and Bülthoff [BB91], specularity displacement depends on disparity, surface curvature, and baseline of the camera. In the simple case of a specular sphere viewed along the ray that passes through the sphere's center, the specular motion vector is proportional to the camera speed and radius of the sphere, and inverse proportional to the square of the depth.

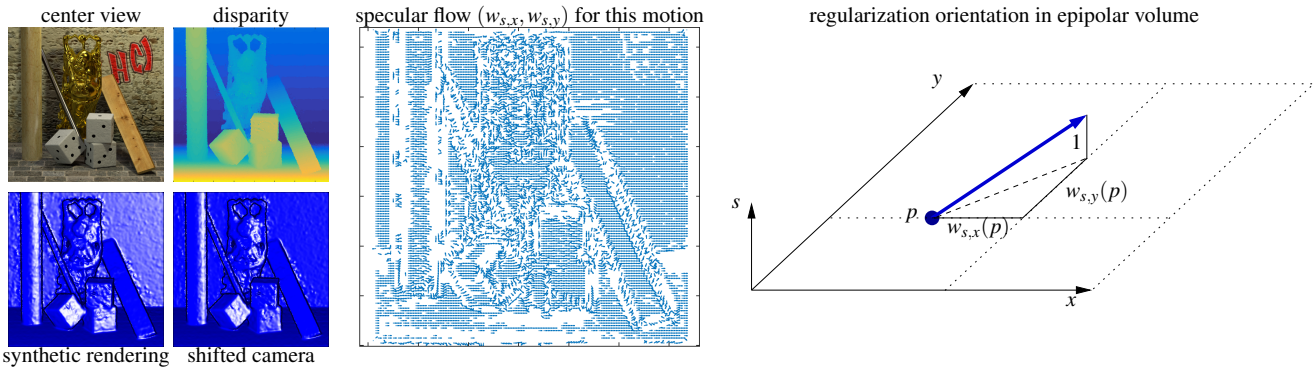

Figure 3: *Specular flow based regularization of specular components. Left column: synthetic renderings of the* buddha2 *light field [WMG13] using the estimated depth map as a geometry proxy, with the camera shifting towards the right. Middle column: specular flow for differential motion in positive s-direction, illustrating the differential shift of reflection patterns. Specular flow is a 2D vector field $(w_{s,x}, w_{s,y})$ in the image domain (left column), which depends in a complex way on curvature and surface normals. Thus, specular components have to be constant along the integral curves of the vector field $(1, w_{s,x}, w_{s,y})^T$ (right column), which can be encouraged with an anisotropic regularizer on the epipolar volume in (s,x,y)-space. Similarly, one defines a second specular regularizer on (t,x,y) space for movement in t-direction.*

A generic model for specular motion along an arbitrary surface was studied by Swaminathan et al. [SKS*02] and Adato et al. [AVaTZ07]. Swaminathan et al. detect specular motion based on deviations in EPI stripes, while Adato et al. recover 3D geometry from the specular flow motion. We do it exactly the other way around, and infer the theoretical specular motion for the given disparity map by specializing the approach detailed in [AVaTZ07].

Let f be the scene surface parametrized by (x,y) coordinates on a view, which can be inferred from a calibrated light field camera given the disparity estimate d. We assume a distant light source, and first compute the reflection vector on the surface in spherical angles according to [AVaTZ07], equation (5),

$$\alpha = \operatorname{atan}\left(\frac{2\|\nabla f\|}{1-\|\nabla f\|^2}\right), \quad \beta = \operatorname{atan}\left(\frac{f_y}{f_x}\right). \tag{6}$$

All derivatives are computed with a finite difference approximation using filters with optimized rotation invariance [WS02]. For camera motion in the direction of $\boldsymbol{q}$ parallel to the image plane, the environment motion field $\boldsymbol{\omega}$ can be approximated as

$$\boldsymbol{\omega} = \begin{bmatrix} d\alpha/dt \\ d\beta/dt \end{bmatrix} = \begin{bmatrix} \nabla\alpha^T \\ \nabla\beta^T \end{bmatrix} d\boldsymbol{q}. \tag{7}$$

From this, we recover the specular flow $\boldsymbol{w}$ as $\boldsymbol{w} = J^{-1}\boldsymbol{\omega}$ with the Jacobian

$$J = \frac{1}{\|\nabla f\|(1+\|\nabla f\|^2)} \begin{bmatrix} f_x f_{xx} + f_y f_{xy} & f_x f_{xy} + f_y f_{yy} \\ f_x f_{xy} - f_y f_{xx} & f_x f_{yy} - f_y f_{xy} \end{bmatrix}. \tag{8}$$

See e.g. [AVaTZ07], equations (7) and (8). An example for estimated specular flow is shown in figure 3.

In this way, when shifting the view point in direction $\boldsymbol{q} = [s\ 0]^T$, we obtain a specular flow $\boldsymbol{w}_s = (w_{s,x}, w_{s,y})$. The meaning of this flow field is that within the epipolar volume for constant t, the specular coefficient should stay constant along the direction $\begin{bmatrix} 1 & w_{s,x} & w_{s,y} \end{bmatrix}$. We thus need to regularize σ_t accordingly. Similarly, we obtain for a shift in direction $\boldsymbol{q} = [0\ t]^T$ a second specular flow $\boldsymbol{w}_t = (w_{t,x}, w_{t,y})$.

Having directions of specular flow computed, the regularizer J_s for the specular component σ can be written as anisotropic regularization (3) of the epipolar volumes in the directions given by the specular flow, i.e.

$$J_s(\sigma) = \int J_{[1\ w_{s,x}\ w_{s,y}]^T}(\sigma_t)\,\mathrm{d}t + \int J_{[1\ w_{t,x}\ w_{t,y}]^T}(\sigma_s)\,\mathrm{d}s. \tag{9}$$

In the following section, we will embed the regularizers encouraging consistency of the specular and diffuse components in a variational optimization framework for the complete light field.

5. Final variational energy and optimization

When obtaining the separation for the complete light field, the unknown is the function $\boldsymbol{u} : \mathcal{R} \to \mathbb{R}^{1+K}$ of coefficients on ray space. We optimize over all of the energies defined previously, in particular the ray-wise separation model based on sparse coding (2) over all of rayspace, and epipolar plane consistency regularizer for the diffuse component (5) and specular component (9). In addition, we use the image-driven total generalized variation ITGV_2 [RGPB12] as a spatial regularizer on each individual subaperture view. In effect, this gives the final variational model

$$\begin{aligned} \underset{\boldsymbol{u}=[\sigma\ \boldsymbol{\alpha}]}{\operatorname{argmin}} \Big\{ & \int_{\mathcal{R}} E(\boldsymbol{u}(\boldsymbol{r}), D)\,\mathrm{d}\boldsymbol{r} + \mu_d J_d(\boldsymbol{\alpha}) + \mu_s J_s(\sigma) + \dots \\ & \dots + \rho_d \int \mathrm{ITGV}_2(\boldsymbol{\alpha}_{st})\,\mathrm{d}(s,t) + \rho_s \int \mathrm{ITGV}_2(\sigma_{st})\,\mathrm{d}(s,t) \Big\}. \end{aligned} \tag{10}$$

Note that the overall energy is convex in $\boldsymbol{u}$ for a fixed dictionary D, which we pre-compute before solving (10). The constants μ_d, ρ_d and μ_s, ρ_s are the smoothness weights for diffuse and specular regularization, respectively.

Initializing the dictionary. To estimate illumination color S_c corresponding to the first column of the dictionary D, we use the approch proposed by Yang et al. [YGL15]. The key idea presented in their work is that most of natural images contain grey pixels under a white light source. The authors identify those pixels with an illuminant invariant measure (IIM), which is computed for local image patches. As a necessary condition for a pixel to be grey, they

Step 1: initialization

- Initialize dictionary D according to section 5.
- Allocate vector-valued functions $\boldsymbol{u} = [\sigma, \boldsymbol{\alpha}]$ on the subset $\mathcal{C}$ of views, with $\sigma(\boldsymbol{r}) \in \mathbb{R}$, $\boldsymbol{\alpha}(\boldsymbol{r}) \in \mathbb{R}^K$.
- Pre-compute the disparity map using [WG14], and the specular flow fields for horizontal and vertical camera shifts, see section 4.
- For every ray $\boldsymbol{r}$, pre-compute regularization orientation for disparity and specular flow for the regularizers (5) and (9) according to section 4.
- Set step size $\tau = 1/(K+1)$.
- Set diffuse and specular smoothness weights $\mu_d, \rho_d, \mu_s, \rho_s$ and diffuse and specular sparsity weights λ_d, λ_s.

Step 2: iterate until convergence

- gradient descent for data term in (2),

$$\boldsymbol{u}(\boldsymbol{r}) \leftarrow \boldsymbol{u}(\boldsymbol{r}) - 2\tau D^T (D\boldsymbol{u}(\boldsymbol{r}) - L(\boldsymbol{r})).$$

- subgradient descent for sparsity norm in (2), i.e. a prox operator for the L^1-norm (shrinkage),

$$\boldsymbol{\alpha}(\boldsymbol{r})_i \leftarrow (\boldsymbol{\alpha}(\boldsymbol{r})_i - \tau\lambda_d)_+ \operatorname{sgn}(\boldsymbol{\alpha}(\boldsymbol{r})_i),\ i = 1, \dots, K,$$
$$\sigma(\boldsymbol{r}) \leftarrow (\sigma(\boldsymbol{r}) - \tau\lambda_s)_+ \operatorname{sgn}(\sigma(\boldsymbol{r})).$$

- subgradient descent for diffuse regularizer (5) and specular regularizer (9) using [PC11],

$$\boldsymbol{\alpha} \leftarrow \operatorname{prox}_{\tau\mu_d J_d}(\boldsymbol{\alpha}) \text{ and } \sigma \leftarrow \operatorname{prox}_{\tau\mu_s J_s}(\sigma).$$

- subgradient descent for the spatial regularizers as implemented in [RGPB12], i.e. for every subaperture view $(s,t) \in \mathcal{C}$,

$$\boldsymbol{\alpha}_{st} \leftarrow \operatorname{prox}_{\tau\rho_d \mathrm{ITGV}_2}(\boldsymbol{\alpha}_{st}) \text{ and } \sigma_{st} \leftarrow \operatorname{prox}_{\tau\rho_s \mathrm{ITGV}_2}(\sigma_{st}).$$

Note that the proximation operators for the diffuse and spatial regularizers separate into one each for the vertical and horizontal slit of the crosshair, respectively.

Figure 4: *Overview of the complete algorithm to solve the separation problem (10) with consistent regularization. To improve efficiency, we follow [GW13] and only run it on a crosshair-shaped set $\mathcal{C}$ of views around the center view (s_c, t_c), i.e. $\mathcal{C}$ consists of all views (s,t) for which $s = s_c$ or $t = t_c$.*

propose that the non-zero IIM is the same for all RGB channels. Having grey pixels detected, they compute illumination color as the mean RGB value of those pixels. The remainder of the dictionary is initialized as an overcomplete basis of $\mathbb{R}^3$, where we interpolate the RGB unit vectors to span K columns.

Optimizing for the coefficients. When optimizing for $\boldsymbol{u}$ with fixed D, the energy (10) is convex. In principle, it fits the overall optimization framework for consistent light field assignment [GW13]. However, there are more regularizers to be taken into account and there are two orientations to consider for the EPI regularization. The technique is still the same and inspired by FISTA [BT09]: iterate over a gradient descent step for the differentiable part of the energy, as well as an implicit subgradient descent step for every convex term. As a reminder, if J is convex, then an implicit subgradient descent step with step size $\tau > 0$ starting from $\boldsymbol{u}$ is performed by computing the prox operator [Roc97]

$$\operatorname{prox}_{\tau J}(\boldsymbol{u}) = \operatorname*{argmin}_{\boldsymbol{w}} \left\{ \frac{\|\boldsymbol{w} - \boldsymbol{u}\|_2^2}{2\tau} + J(\boldsymbol{w}) \right\}. \tag{11}$$

Thus, a subgradient descent step for total generalized variation or anisotropic total variation amounts to solving an instance of L^2-denoising with the respective regularizer, while a subgradient descent step for the L^1 sparsity norm is computed with the shrinkage operator T_μ, defined component-wise as [BT09]

$$\operatorname{prox}_{\tau\lambda\|\cdot\|_1}(\boldsymbol{u})_i = T_{\tau\lambda}(\boldsymbol{u})_i := (|\boldsymbol{u}_i| - \tau\lambda)_+ \operatorname{sgn}(\boldsymbol{u}_i). \tag{12}$$

We solve all denoising subproblems with the primal-dual algorithm [PC11] with diagonal pre-conditioning. For a comprehensive overview, all steps of the algorithm are summarized in figure 4.

6. Experimental evaluation

To demonstrate the capability of our method on real-world scenes, We captured multiple light fields under different natural or synthetic lighting with the Lytro Illum light field camera.

We compare our method to two different approaches. First, the original single view approach proposed in [AO14], which we extended to the light field setting in order to show that having the light field structure helps with getting better results. Although suggested parameter settings are given in their paper, we adapted them to the individual light fields to obtain results as good as possible. As input to their method, we give the center view of the light field. The second method we compare to is [TSW*15], a state-of-the-art approach for specular separation tailored specifically to light fields. We kept default parameters as it was recommended by the authors and evaluated results of their implementation.

Our method requires a precomputed disparity map, which is calculated on the input light field with the first order structure tensor [WG14] and regularized with image-driven TGV [RGPB12]. The Lytro Illum was calibrated using [DPW13].

Results can be observed in figures 5 and 6. Parameters are determined empirically and individually for each light field. Note that all results for the specular component have been normalized to make them more visible, so diffuse and specular part do not sum up to the input. We will briefly comment on each individual dataset. On the dataset *Bowl*, both previous methods failed on specular highlight removal because illumination has a very similar color to the that of the surface, while our method could successfully differentiate between the two motions using specular flow. However, we can see some bluish artifacts in the diffuse component, which are due to channel saturation in the input images, making the exact additive model partially invalid. Improving this will be investigated in upcoming work. On the *Swan* dataset, we obtain very good results in comparison to [TSW*15], as their method cannot successfully handle regions with larger specularities. However, the original method provides quantitatively similar results. On both, *Dragon* and *Owl*, our method gets a better diffuse component than [AO14], as well as a better specular component in comparison to [TSW*15]. Very promising results are obtained on *Leaf*, where despite a very noisy pre-computed disparity map we obtained a very reliable diffuse

component. Strongly filtering the disparity map before computing the specular flow helped here. Finally, on the *Rose* dataset, our results are significantly better than [TSW*15]. Compared to [AO14], we have a much less noisy diffuse component due to our proposed regularization, but a very similar result for the specular component.

In summary, in most cases we obtained a similar or better specular component, with usually substantial improvements for the diffuse one. Compared to [TSW*15], our method is more accurate in regions with larger specularities and more reliable than the original method [AO14] due to our contributions of diffuse and specular flow reguralization based on the light field structure.

7. Conclusion

We propose a method for specular detection and removal in light field domain. Our approach leverages the geometry of light fields in order to obtain more robust results. In particular, we embed the previous approach [AO14], which is based on the idea that the observed color is a sparse linear combination of a few diffuse basis colors and a single specular color, in a global optimization framework for light field labelling [GW13]. Here, we regularize the diffuse component based on disparity, and propose a novel regularization scheme for the specular component based on specular flow [AVaTZ07]. In many experiments on real-world light fields captured with the Lytro Illum, we demonstrate that we outperform both the original method [AO14] as well as the state-of-the-art approach [TSW*15] for specular removal in light fields.

Acknowledgements

This work was supported by the ERC Starting Grant "Light Field Imaging and Analysis" (LIA 336978, FP7-2014) and the SFB Transregio 161 "Quantiative Methods for Visual Computing".

References

[ABC11] ARTUSI A., BANTERLE F., CHETVERIKOV D.: A survey of specularity removal methods. *Computer Graphics Forum 30*, 8 (2011), 2208–2230. 2

[AO14] AKASHI Y., OKATANI T.: Separation of reflection components by sparse non-negative matrix factorization. In *Asian Conf. on Computer Vision* (2014). 2, 5, 6, 7, 8

[AVaTZ07] ADATO Y., VASILYEV Y., ANS T. ZICKLER O. B.-S.: Toward a theory of shape from specular flow. In *Proc. International Conference on Computer Vision* (2007). 2, 4, 6

[BB91] BLAKE A., BULTHOFF H.: Shape from specularities: computation and psychophysics. *Phil. Trans. R. Soc. Lond. B 331*, 1 (1991), 237–252. 3

[BLL96] BAJCSY R., LEE S. W., LEONARDIS A.: Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation. *International Journal of Computer Vision 17*, 3 (1996), 241–272. 2

[BT09] BECK A., TEBOULLE M.: Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM J. Imaging Sciences 2* (2009), 183–202. 5

[DPW13] DANSEREAU D., PIZARRO O., WILLIAMS S.: Decoding, calibration and rectification for lenselet-based plenoptic cameras. In *Proc. International Conference on Computer Vision and Pattern Recognition* (2013), pp. 1027–1034. 5

[GW13] GOLDLUECKE B., WANNER S.: The variational structure of disparity and regularization of 4D light fields. In *Proc. International Conference on Computer Vision and Pattern Recognition* (2013). 2, 3, 5, 6

[HRP13] HEBER S., RANFTL R., POCK T.: Variational shape from light field. In *Int. Conf. on Energy Minimization Methods for Computer Vision and Pattern Recognition* (2013), pp. 66–79. 1

[JSG16] JOHANNSEN O., SULC A., GOLDLUECKE B.: What sparse light field coding reveals about scene structure. In *Proc. International Conference on Computer Vision and Pattern Recognition* (2016). 1

[KSK87] KLINKER G. J., SHAFER S. A., KANADE T.: Using a color reflection model to separate highlights from object color. In *Proc. International Conference on Computer Vision* (1987), vol. 87, pp. 145–150. 2

[LLK*02] LIN S., LI Y., KANG S. B., TONG X., SHUM H.: Diffuse-specular separation and depth recovery from image sequences. In *Proc. European Conference on Computer Vision* (2002). 2

[PC11] POCK T., CHAMBOLLE A.: Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *International Conference on Computer Vision (ICCV 2011)* (2011). 5

[PDCS13] PALMA G., DESOGUS N., CIGNONI P., SCOPIGNO R.: Surface light field from video acquired in uncontrolled settings. In *Digital Heritage International Congress (DigitalHeritage), 2013* (2013), vol. 1, IEEE, pp. 31–38. 2

[RGPB12] RANFTL R., GEHRIG S., POCK T., BISCHOF H.: Pushing the limits of stereo using variational stereo estimation. In *IEEE Intelligent Vehicles Symposium* (2012). 4, 5

[Roc97] ROCKAFELLAR R. T.: *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks. 5

[Sha85] SHAFER S.: Using color to separate reflection components. *Color Research & Application 10*, 4 (1985), 210–218. 1, 2

[SKS*02] SWAMINATHAN R., KANG S. B., SZELISKI R., CRIMINISI A., NAYAR S. K.: On the motion and appearance of specularities in image sequences. In *Proc. European Conference on Computer Vision* (May 2002), vol. I, pp. 508–523. 2, 4

[TI05] TAN R. T., IKEUCHI K.: Separating reflection components of textured surfaces using a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27*, 2 (2005), 178–193. 2

[TSW*15] TAO M., SU J. C., WANG T. C., MALIK J., RAMAMOORTHI R.: Depth estimation and specular removal for glossy surfaces using point and line consistency with light-field cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2015). 1, 2, 5, 6, 7, 8

[WCER16] WANG T. C., CHANDRAKER M., EFROS A., RAMAMOORTHI R.: SVBRDF-invariant shape and reflectance estimation from light-field cameras. In *Proc. International Conference on Computer Vision and Pattern Recognition* (2016). 2

[WG14] WANNER S., GOLDLUECKE B.: Variational light field analysis for disparity estimation and super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence 36*, 3 (2014), 606–619. 1, 3, 5

[WMG13] WANNER S., MEISTER S., GOLDLUECKE B.: Datasets and benchmarks for densely sampled 4D light fields. In *Vision, Modelling and Visualization (VMV)* (2013). 4

[WS02] WEICKERT J., SCHARR H.: A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *Journal of Visual Communication and Image Representation 13*, 1/2 (2002), 103–118. 4

[YCK06] YOON K., CHOI Y., KWEON I. S.: Fast separation of reflection components using a specularity-invariant image representation. In *Int. Conf. on Image Processing* (2006), pp. 973–976. 2

[YGL15] YANG K., GAO S., LI Y.: Efficient illuminant estimation for color constancy using grey pixels. In *Proc. International Conference on Computer Vision and Pattern Recognition* (2015). 4

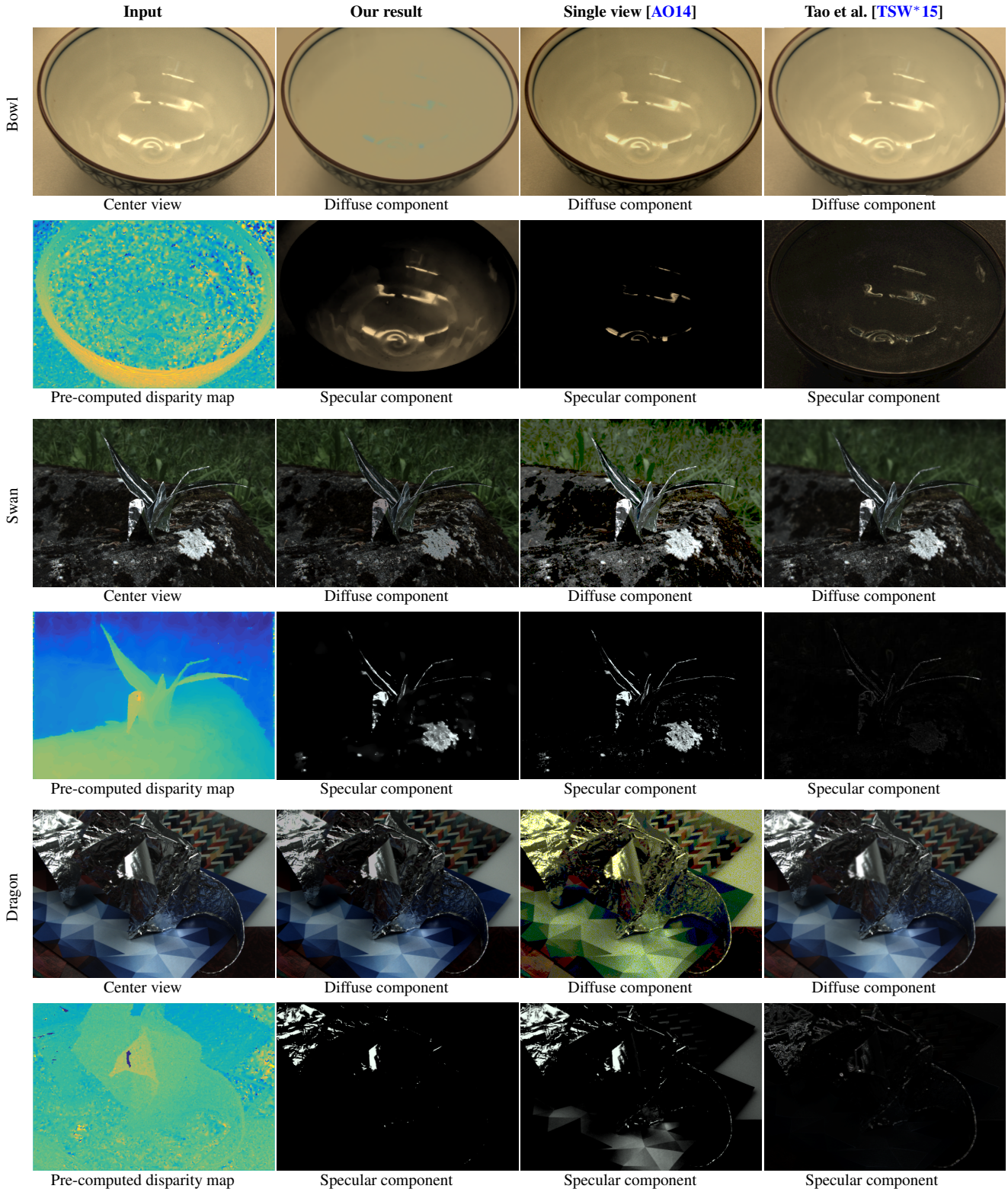

Figure 5: *First set of results comparing our method to the previous single view method [AO14] as well as a light-field based approach [TSW*15]. All results for specular terms are normalized for better visibility, so specular and diffuse terms do not sum up to the input. Note that frequently, previous methods recover a specular term which is qualitatively acceptable, but much too dark, so that the diffuse component stays the same. See section 6 for a detailed discussion.*

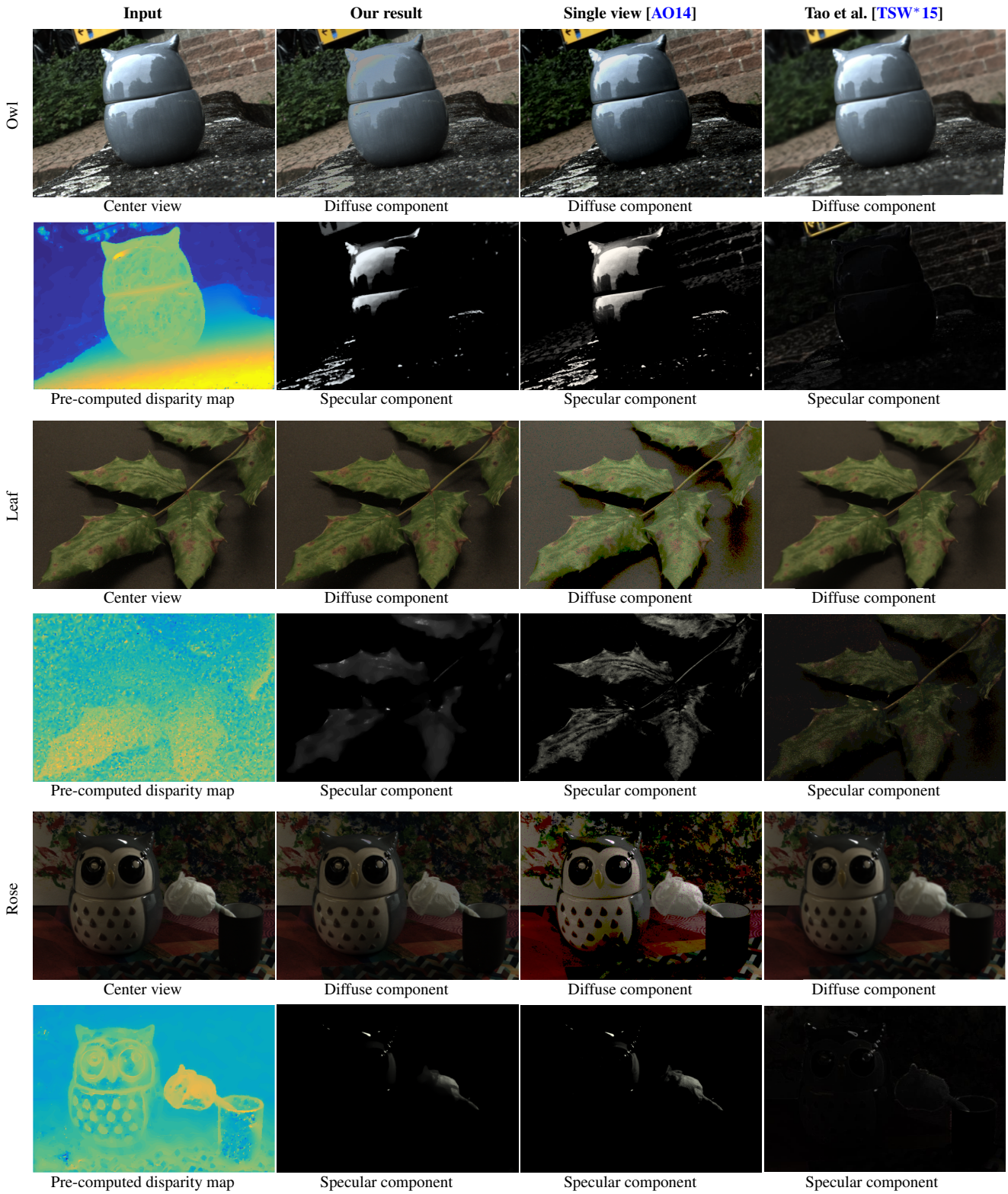

Figure 6: *Second set of results comparing our method to the previous single view method [AO14] as well as a light-field based approach [TSW*15]. All results for specular terms are normalized for better visibility, so specular and diffuse terms do not sum up to the input. Note that frequently, previous methods recover a specular term which is qualitatively acceptable, but much too dark, so that the diffuse compoent stays the same. See section 6 for a detailed discussion.*