

Relighting Objects from Image Collections

Tom Haber
Hasselt University
tUL - IBBT, EDM, Belgium

Christian Fuchs
Ulm University

Michael Goesele
TU Darmstadt

Philippe Bekaert
Hasselt University
tUL - IBBT, EDM, Belgium

Hendrik P. A. Lensch
Ulm University

Hans-Peter Seidel
MPI Informatik

Abstract

We present an approach for recovering the reflectance of a static scene with known geometry from a collection of images taken under distant, unknown illumination. In contrast to previous work, we allow the illumination to vary between the images, which greatly increases the applicability of the approach. Using an all-frequency relighting framework based on wavelets, we are able to simultaneously estimate the per-image incident illumination and the per-surface point reflectance. The wavelet framework allows for incorporating various reflection models. We demonstrate the quality of our results for synthetic test cases as well as for several datasets captured under laboratory conditions. Combined with multi-view stereo reconstruction, we are even able to recover the geometry and reflectance of a scene solely using images collected from the Internet.

1. Introduction

Estimating scene reflectance using image-based techniques is a well-studied topic in the graphics and vision literature. Common approaches are based on controlled, active illumination [7] or special setups to record the illumination during scene capture [31, 15]. In specific cases it is also possible to estimate illumination from shadows [26] or from other properties of the scene. An alternative line of work assumes that multiple images of the scene are captured under identical illumination, which is then estimated together with the surface reflectance [21, 30]. In contrast, our goal is to estimate the reflectance of each scene point from a set of photographs captured under unknown, constant or varying distant illumination.

This has two striking implications: First, we do not require special lighting or capture equipment to shoot our images. Instead, we can just casually acquire a set of photographs with distant lighting. Second, it is not even necessary to capture the images ourselves. Instead, we are able to make use of existing photographs, even images taken from

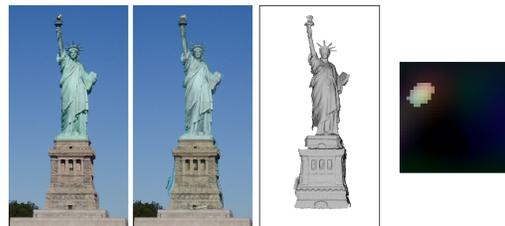


Figure 1. Overview of our reconstruction pipeline. From left to right: an example image taken from Flickr [9], estimated geometry, recovered environment map for this particular input view, and rendered result using the geometry, environment and estimated reflectance properties.

online community photo collections (CPC). Combined with a robust multi-view stereo approach for CPCs [12], we can estimate the scene geometry from these images as well. To our knowledge, this is the first system that recovers such a complete scene model from images acquired under such general conditions.

Our system consists of two main components: First, we use an *all frequency relighting framework* based on a wavelet representation of the local visibility and the current estimate of illumination and scene reflectance to render the scene efficiently using the triple-wavelet product integral [18]. Factoring illumination and scene reflectance requires *solving a bilinear system of equations*. We therefore secondly employ an iterative optimization to estimate illumination given scene reflectance and vice versa. To summarize, our contributions are as follows:

- We simultaneously estimate the reflectance and illumination of a scene captured under varying distant illumination. This simplifies the capture process and makes our system applicable to a larger range of scenes and existing datasets.
- We demonstrate first reconstructions solely based on images captured from an Internet photo sharing site. Scene geometry, reflectance, and distant illumination are all estimated from these images.

The achievable quality of our system is bound by fundamental limitations such as the bandwidth of the reflectance and the frequency content of the illumination [24, 25]. We nevertheless achieve good quality results for the datasets presented in this paper.

2. Related Work

2.1. Reflectance from Known Illumination

A common strategy for relighting is to sample the incident illumination densely using a controlled acquisition setup [7]. Other reconstruction approaches use a sparse set of images captured under known, point light illumination and combine them with additional assumptions. Assuming *spatial coherence*, Marschner et al. [17] estimate a single BRDF from a set of images of an object while Zickler et al. [32] reconstruct a relightable model from a single image. Lensch et al. [15] capture multiple images of a heterogeneous object. They cluster the surface into regions with similar appearance and determine a BRDF for each cluster which serve as *basis BRDFs*. The BRDF of each surface point is then determined as linear combination of these basis BRDFs. Weistroffer et al. [29] extend this approach and determine the reflectance of a surface point as a weighted sum of materials, each consisting of several basis BRDFs.

Due to the lack of control, reflectance recovery for outdoor scenes generally operates under complex illumination. Yu and Malik [31] recovered the photometric properties of architectural scenes from photographs using an explicit outdoor illumination model. Debevec et al. [8] first captured the scene geometry and reflectance samples. They then used an inverse rendering approach to create a relightable model of the Pantheon from images with known, distant illumination. Most recently, Romeiro et al. [25] introduced a passive reflectometry approach that estimates a bi-variate representation of an isotropic BRDF from a single image of a curved surface captured under known distant illumination. All of these systems require either known or controlled incident illumination and often also a detailed, captured or manually created scene model and are thus unable to reconstruct a scene’s reflectance from existing imagery.

2.2. Illumination Estimation

Distant illumination in a scene can be directly measured using a light probe [6] or estimated from the scene. Sato et al. [26] estimate the illumination distribution of a scene with known geometry from shadows cast by an occluder. They define an adaptive refinement criterion based on the intensity distribution of direct light sources. Later, Okabe et al. [22] also investigated illumination estimation from cast shadows. For the case of lighting expressed in a Haar wavelet basis, they start with a coarse representation and iteratively add basis functions with increasing resolution.

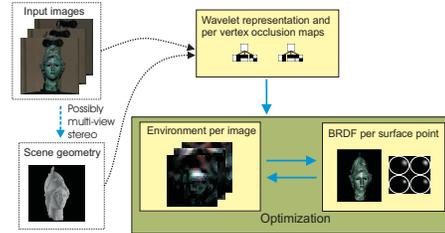


Figure 2. System overview.

2.3. Estimating Reflectance and Illumination

Ramamoorthi and Hanrahan [24] give an algorithm to factor the reflected light field of a scene captured under unknown illumination into the reflectance and illumination components. They also provide a theoretical framework that predicts under which conditions this can be achieved and how the recoverable detail of the reflectance functions is determined by the frequency content of the incident illumination. Their analysis applies directly to our system. Based on a sparse set of images, Nishino et al. [21, 20] first separate the diffuse texture from the specular reflectance of an object captured under fixed, unknown lighting. Given the object’s geometry, they then compute an initial estimate for the illumination and determine both the illumination and the parameters of a simplified Torrance-Sparrow model. Both of these methods assume known scene geometry and constant specular reflectance over the surface. Georgiades [10] uses generalized photometric stereo to estimate geometry, positions of a single point light source, and the parameters of a single reflectance model for a sparse set of images with fixed viewpoint.

Yu et al. [30] recover both lighting and surface reflectance from a small set of images captured under constant distant illumination. Given the scene geometry, they compute the diffuse albedo and determine the parameters of a Gaussian filtered mirror BRDF model for each surface point. Illumination is modelled in a spherical harmonics framework. Our system is closely related to this work but has two important advantages. First, we allow varying illumination between the input views. This simplifies acquisition and makes our method applicable to existing datasets captured under arbitrary distant lighting conditions. Second, we operate in an all-frequency framework which simplifies modelling of complex illumination and allows the use of arbitrary BRDFs.

3. Overview and Assumptions

Our system takes as input a set of photographs of an object or a scene captured under unknown and possibly varying distant illumination (see Figure 2). We assume that these images are already in a photometrically linear space. Overexposed pixels cannot be trusted and receive a weight of zero in the reconstruction process.

A triangle mesh representing the scene geometry is either given (e.g. manually modeled or captured in a separate acquisition process), or reconstructed using a multi-view stereo approach [12]. The mesh is manually cleaned and cropped from spurious or inaccurate geometry. We assume that the mesh models the complete relevant scene geometry including blockers that cast shadows on the object.

The system operates within an all-frequency relighting framework [18, 30] based on Haar wavelets. In a preprocessing step, we compute an occlusion map per mesh vertex, i.e. determine which part of the hemisphere of illumination directions is occluded, and transform the occlusion map into the wavelet basis. We then set up an efficient rendering pipeline based on triple-wavelet products, which allows us to easily update the current reflectance and illumination estimates. Local surface reflectance is expressed as a weighted sum of basis materials. We assume that inter-reflections in the scene can be neglected. The relighting task leads to a set of bilinear equations, which is solved iteratively yielding a reflectance estimate per surface point and an illumination estimate per input image. Provided that the illumination used for image capture and the BRDF contains sufficient high frequencies [24, 25], we are able to reconstruct a detailed model of the scene, reflection properties, and of the illumination.

4. Rendering Pipeline

A central part in our estimation process is the simulation of light propagation in the scene based on the reflectance equation, which is evaluated for every rendered pixel:

$$L(\vec{x}, \omega_o) = \int_{\Omega} \rho(\vec{x}, \omega_o, \omega_i) V(\vec{x}, \omega_i) \tilde{L}(\vec{x}, \omega_i) d\omega_i, \quad (1)$$

where ρ is the BRDF including the cosine term for the incident illumination from direction ω_i at position \vec{x} and the outgoing direction ω_o . \tilde{L} is the incident illumination and V represents the occlusion at \vec{x} . Since we iteratively optimize for ρ and \tilde{L} , we need an efficient evaluation scheme.

4.1. All Frequency Relighting

Efficient techniques for solving this rendering problem under the distant illumination assumption have been investigated in the field of precomputed radiance transfer [27], where the incident illumination and the reflectance functions are represented using spherical basis functions such as spherical harmonics (SH) or wavelet bases. In these bases, the otherwise costly integral reduces to a dot product. In order to not limit the resolution capabilities, we follow the all-frequency relighting approach by Ng et al. [18].

The local visibility, the local environment map, and the slice of the BRDF that corresponds to the viewing direction ω_o are expressed in the Haar wavelet basis Ψ defined over the hemisphere using the hemi-octrahedral parameter-

ization introduced by Praun and Hoppe [23]:

$$V(\omega) = \sum_l V_l \Psi_l(\omega) \quad (2)$$

$$\tilde{L}(\omega) = \sum_m \tilde{L}_m \Psi_m(\omega) \quad (3)$$

$$\rho(\vec{x}, \omega) = \sum_k \rho_k(\vec{x}) \Psi_k(\omega) \quad (4)$$

In practice, we compute local visibility only per mesh vertex and interpolate it for any point in a triangle. Note that representing the BRDF in the wavelet basis (Equation 2) allows us to use arbitrary BRDFs (including isotropic and anisotropic, measured and analytical BRDFs) in the rendering pipeline.

Once all three components are represented in the same Haar wavelet basis the integral over the hemisphere is evaluated as a triple-wavelet product integral [18]

$$L(\vec{x}, \omega_o) = \sum_k \sum_l \sum_m C_{klm} \rho_k V_l \tilde{L}_m \quad (5)$$

with tripling coefficients $C_{klm} = \int_{\Omega} \Psi_k \Psi_l \Psi_m d\omega$. Exploiting the hierarchical nature and sparseness of the wavelet representation drastically reduces the computation effort.

4.2. Lighting Rotation

One problem of this approach is that for every pixel each slice of the BRDF needs to be represented in the wavelet basis Ψ . While Ng et al. [18] used precomputed 6D BRDF data sets, parameterized by ω_i , ω_o , and the surface normal \vec{n} , we employ the Efficient Wavelet Rotation algorithm [28] to transform the environment map \tilde{L} in the octahedral parameterization into the local coordinate frame of each pixel and the corresponding hemi-octahedral representation. This step reduces the dimensionality of the BRDF per pixel to 4D resulting in reduced memory overhead at the cost of additional computation. In a preprocessing step, we discretize the sphere into 32^2 normal directions $\{\vec{v}\}$ and compute a rotation matrix $\mathbf{R}_{\vec{v}}$ for each of them. $\mathbf{R}_{\vec{v}}$ encodes the influence of every wavelet basis coefficient in the global coordinate frame onto the wavelet coefficients in the rotated space. A rotated environment map $\tilde{L}(\vec{n})$ for an in-between normal is constructed by interpolation from the rotated maps of the four nearest normals with bilinear interpolation weights $\phi_{\vec{v}}(\vec{n})$:

$$\tilde{L}(\vec{n}) = \sum_{\vec{v} \in \text{Neighborhood}(\vec{n})} \phi_{\vec{v}}(\vec{n}) \cdot \mathbf{R}_{\vec{v}} \tilde{L}. \quad (6)$$

4.3. Bilinear Model

Given the object geometry and a set of images $\{I_i\}$ captured under distant illumination, we can extract the set of measurements $\{y_p\}$ that show the response of the scene to an environment map $\tilde{L}_{I(y_p)}$ recorded in image $I(y_p) \in \{I_i\}$ measured at surface point $\vec{x}(y_p)$ on the object's surface S .

The visibility term V_l in Equation 5 is constant for a given surface point \vec{x} as it encodes the portion of the hemisphere that is blocked by the object geometry. It can there-

fore conceptually be combined with the tripling coefficients to a new set of coefficients

$$T_{\vec{x},klm} = \sum_l C_{klm} V_l. \quad (7)$$

Using matrix notation, Equation 5 simplifies then for a particular y_p to a bilinear system in L and ρ

$$L_p = \rho_{\vec{x}(y_p)}^T \mathbf{T}_{\vec{x}(y_p)} \tilde{L}_{I(y_p)}. \quad (8)$$

5. Optimization

We can now define our inverse rendering task. From a set of images $\{I_i\}$ we seek to determine the reflectance functions $\rho(\vec{x})$ and the incident illumination \tilde{L}_I for each I_i . This is equivalent to minimizing the following objective function in a least squares sense with respect to \tilde{L} and ρ :

$$\mathcal{O} = \sum_{p=1}^N \alpha_p (y_p - \rho_{\vec{x}(y_p)}^T \mathbf{T}_{\vec{x}(y_p)} \tilde{L}_{I(y_p)})^2 \quad (9)$$

with $\tilde{L} = \{\tilde{L}_0, \dots, \tilde{L}_{Q-1}\}$ and $\rho = \{\rho_{\vec{x}|\vec{x} \in S}\}$. y_p is a measurement of surface point \vec{x}_p seen in an input image. We additionally introduce confidence values α_p per observation to reduce the effect of less reliable data, such as measurements at grazing angles and overexposed pixels.

Since our system depends linearly on both ρ and \tilde{L} we have to solve a bilinear system of equations. This can be performed iteratively by solving linear problems once for \tilde{L} and once for ρ [5, 1]. Note that solving for both simultaneously would yield clear advantages [3] but is impractical due to the size of our problem. The solution can only be unique up to a global scale factor. In addition, we constrain both \tilde{L} and ρ to be strictly positive in order to ensure physically plausible results. Each subproblem can now be efficiently solved with arbitrary linear equality and inequality constraints by posing it as a convex Quadratic Programming problem. We apply an implementation of the primal-dual interior point algorithm [11] for this task.

5.1. Optimization of Environment Maps

Given an initial guess of the reflectance ρ we first solve for the environment \tilde{L} minimizing the linear system

$$Y_I = \mathbf{M} \tilde{L}_I \quad (10)$$

where we combined the constant terms into a single matrix $\mathbf{M}_p = \rho_{\vec{x}(y_p)}^T \mathbf{T}_{\vec{x}(y_p)}$. This inverse lighting problem is in general ill-posed [16, 24]. The recoverable information of \tilde{L} depends on the bandwidth of the reflectance function ρ and on the geometry of the object, as shadows carry additional information about the illumination [26]. In general, it will not be possible to recover all details of the illumination.

To obtain a sensible solution despite the ill-posedness of the problem, we apply a small amount (e.g. 0.1% of the maximum value in the system) of regularization by imposing smoothness on the environment maps. This improves the stability by constraining the otherwise underconstrained

high-frequency parameters. Currently, we enforce smoothness by minimizing the laplacian of the environment map.

We improve robustness to measurement noise, inaccurate geometry, interreflection and other unmodeled effects by applying an iteratively reweighted least squares approach [2] which effectively reduces the contribution of pixels with high reconstruction error. Reweighting is started from scratch in each iteration of estimating the lighting, guaranteeing that no data point is lost due to aggressive pruning.

5.2. BRDF Optimization within a Linear Basis

In order to obtain a linear system for the reflectance $\rho(\vec{x})$, we follow the approach of Weistroffer et al. [29] and represent the reflectance at each surface point as a linear combination of materials. Each material in turn consists of a linear combination of basis BRDFs

$$\rho(\vec{x}) = \sum_k (w_k(\vec{x}) \sum_b \lambda_{kb} \rho_b). \quad (11)$$

The optimization for each λ_{kb} uses all available samples that are part of material k (i.e. $w_k > 0$). This two-level approach improves the stability for sparse and noisy data by restricting the estimated reflectances to a physically plausible subspace [29, 15]. The recoverable BRDF information is limited by the bandwidth of the illumination and the geometry of the object, similar to the environments. However each independent lighting condition adds additional constraints to the reconstruction and leads to a more precise estimate of the BRDFs. Furthermore, the more lighting conditions are available, the more stable the estimation will be against errors in the recovered illumination.

We again solve this bilinear problem by alternating between the estimation of material weights λ_{kb} and blending weights $w_k(\vec{x})$ per surface point

$$\mathcal{O} = \sum_{\{p|\vec{x}(y_p)=\vec{x}\}} \alpha_p \left(y_p - \rho(\vec{x})^T \mathbf{T}_{\vec{x}(y_p)} \tilde{L}_{I(y_p)} \right)^2 \quad (12)$$

for each surface point \vec{x} . We again reduce the effect of less reliable data using the confidence value α_p . All coefficients $w_k(\vec{x})$ and λ_{kb} are constrained to be strictly positive. To improve numerical stability while minimizing the number of used basis BRDFs, we modify this system by additionally minimizing $\beta \sum_k \lambda_{kb}^2$, where β is a small number controlling this constraint.

To decrease computation times, we define this system on mesh vertices and then later bilinearly interpolate the blending weights for every surface point. Using a sufficiently dense mesh, the additional smoothing introduced is barely noticeable.

One of the advantages of our framework is that we can use arbitrary BRDFs (including measured or analytical BRDFs) as basis provided they can be expressed in the wavelet basis. For the examples shown in this paper, we performed k-means clustering on the fitted Cook-Torrance pa-

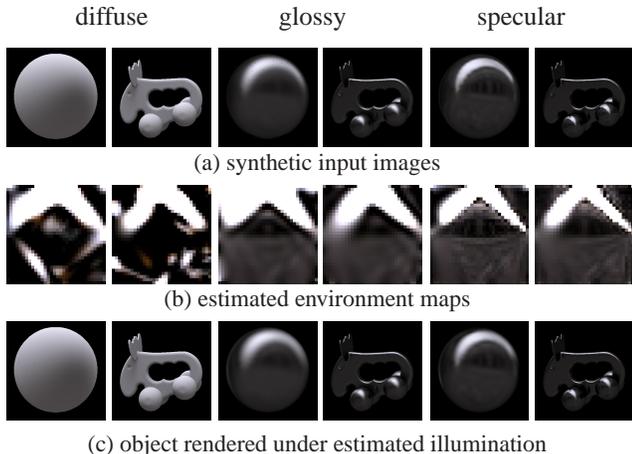


Figure 3. Estimated environment maps for synthetic diffuse, glossy, and specular test cases with a fixed BRDF. The amount of detail recoverable in the environment maps depends on both the bandwidth of the BRDF as well as the scene’s geometry.

parameters for the MERL BRDF database [19] and selected the specular lobes from representatives of the ten largest clusters. The actual basis $\{\rho_b\}$ for a reconstructed scene then consists of a subset of these lobes plus a Lambertian BRDF to model the diffuse reflectance.

To compute an initial estimate of the material and blending weights, we transform the averaged per-vertex samples into HSV colorspace and cluster the vertices using the parametrization $(\cos(2\pi \cdot H), \sin(2\pi \cdot H), S)$. This transformation reduces the influence of shadows and specular highlights. We use k-means clustering to get the separation into regions. For each region, a different material is constructed by computing a diffuse color based on the median of the per-vertex averages, the other BRDF weights are set to zero. The material blending weights are obtained from the distance to the centroids.

6. Results

We present results for three different types of input data: synthetic datasets, datasets captured under controlled conditions, and a dataset solely based on image collections from Internet photo sharing sites (see also Table 1).

6.1. Synthetic Data

We render two objects (a sphere and an elk, see Figure 3a) with varying surface properties (diffuse, glossy with Lafortune exponent $N = 50$, and specular with $N = 500$ [14]) using the *Uffizi* environment map (Figure 4a). We then discard the lighting information and reconstruct it from the rendered image, known object geometry, and BRDF (Section 5.1). The resulting environment maps (Figure 3b) clearly show that the lighting is recovered faithfully. Note how remarkably close the sky region reconstructed from the

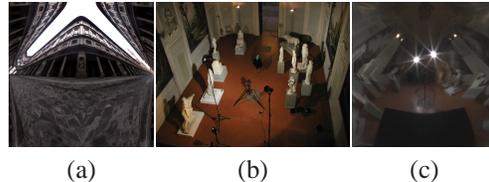


Figure 4. a) The *Uffizi* environment map [6] in the octahedral representation as used in our experiments. b) The acquisition setup to capture the *Minerva* dataset. The two main light sources are facing towards the head. c) The ground truth environment map for the *Minerva* dataset.

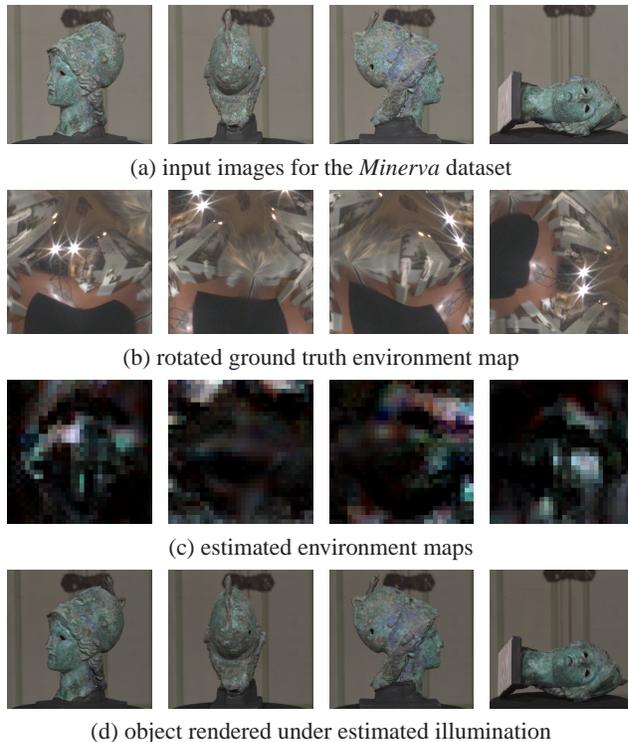


Figure 5. Overview of the *Minerva* dataset showing four of the eight input images used in the estimation.

diffuse sphere is to the original. As predicted by Ramamoorthi and Hanrahan [24], the amount of detail in the lighting increases with the specularity and thus with the bandwidth of the BRDF. In all cases, illuminating the model with the recovered lighting yields results hardly distinguishable from the input (Figure 3c).

6.2. Captured Data

Figure 5 gives an overview over the *Minerva* dataset. The dataset consists of a detailed geometry model captured with a 3D scanning system and a set of high-dynamic range (HDR) images captured under distant illumination in a church-like environment. Lighting and camera were fixed while the object was rotated in front of the camera (see Figure 4 b) and c)). This dataset is therefore an ideal test case to compare the environment maps reconstructed for differ-

ent input images.

We used a collection of 13 images captured under general lighting (Figure 5a) shows 4 of these images) and reconstructed the per-view illumination and the per-pixel BRDF using 6 basis materials each composed of 7 BRDFs. Figure 5d clearly demonstrates that the resulting model rendered with the estimated illumination matches the input images very well. Even small highlights on the temple and chin are faithfully recovered. The estimated environment maps Figure 5c locate the main light sources correctly and otherwise approximate roughly the ground truth illumination since the head features limited specularly. Nevertheless, the materials are well estimated which is also shown in the albedo map (Figure 9) which reveals hardly any geometric feature or shadow, as expected but show some color artifacts which cause a color shift when the model is relit.

As second example, we selected 30 images from the *van Gogh* dataset [4] (see Figure 6a). This dataset was captured under static illumination using a moving camera. Although we reconstruct an individual environment map per view, Figure 6b and the relighting of a novel view with the average of all environment maps (Figure 7) clearly show how consistent our reconstructions are even without assuming a constant environment as in Yu et al. [30]. Note that the estimated environment maps are not identical but consistent for the different views. Missing geometric detail in the input mesh prevents us from reconstructing a high resolution environment map and also reduces the quality of the material estimation especially in the hair region. As we reconstruct materials per vertex, the signature on the bust is highly undersampled. Yu et al. [30] additionally refine the normals to clear up some of these geometric issues.

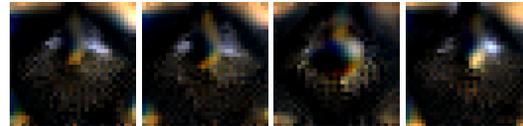
The relit models shown illuminated by the *Uffizi* environment and under point light illumination (Figure 6d) are of high quality. They show clearly that we were able to reconstruct more high frequency details than Yu et al. [30]. However, our color estimates are slightly desaturated. Here we observe the fact that the inverse rendering problem can only be solved up to scale. The red channel is slightly underestimated in the BRDF while it is slightly overestimated in the environment.

6.3. Internet Data

We downloaded image datasets of the Statue of Liberty and Venus de Milo from the Internet photo sharing site Flickr [9] and reconstructed the scene geometry using a multi-view stereo approach [12]. The resulting meshes were manually cleaned and simplified to reduce its complexity. Images found on the Internet are generally neither HDR nor photometrically calibrated. We therefore assume that they are encoded according to the sRGB standard and convert them into photometrically linear space by inverse gamma mapping, but one can alternatively use a technique such as



(a) input images for the *van Gogh* dataset



(b) estimated environment maps



(c) object rendered under estimated illumination



(d) object rendered under the *Uffizi* environment map and under point light illumination.

Figure 6. Overview of the *van Gogh* dataset showing 4 out of the 30 input images.

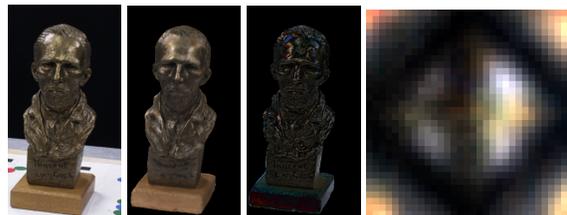


Figure 7. An image (left) from the *van Gogh* dataset that was not used in the reconstruction process is regenerated using the average environment map (right). The difference image shows that the largest error occurs in the specular highlights and in places where geometric detail is missing.

Kuthirummal et al. [13] to perform the photometric calibra-

model	#vertices	#mat	#brdfs	#imgs	time	RMSE
Minerva	100K	6	7	13	7h	.18
Liberty	100K	2	4	6	3h	.20
van Gogh	32K	3	4	30	14h	.11
Venus	40K	2	4	13	14h	.22

Table 1. Overview of the different models. “#mat”, “#brdfs”, “#imgs” and “RMSE” refer to the number of materials used, number of basis brdfs incorporated, number of input images and average RMS error, respectively.

tion.

For the Statue of Liberty, only six images were used in the lighting and reflectance estimation, they are shown in Figures 8 and 1. The estimated environment maps clearly locate the sun direction correctly. Note that the light source might wrap around borders due to the parametrization of the environment maps. Its narrow size can only be achieved by including the visibility information and the use of an all frequency framework. Other regions in the environment map are however reconstructed with low frequency due to their smaller intensity. Again, the color separation between materials and environment is not perfect. This could also be due to the different white balancing of images found on the Internet. Nevertheless, the estimated materials are consistent and of high quality as evident in Figures 8 c) and d).

Figure 10 shows our reconstructions of the Venus Milo, on a set of 13 images. The environment maps shows the presence of two main lightsources: daylight coming in through the window and a tungsten interior light.

7. Discussion and Conclusion

In the presented inverse rendering framework we estimate both the reflection properties of an object and the incident illumination for each individual input view. While this problem is ill-posed in nature, we have demonstrated that it is in principle possible to obtain a meaningful separation. The reconstructed environment maps are the best estimates given the input data and the bandwidth of the BRDFs. By restricting the space of BRDFs to linear combinations of basis BRDFs and by slightly enforcing smoothness in the environment we constrain the space to plausible solutions even in cases where the BRDF limits the reconstruction. The quality of the estimated reflection properties typically increases with the number of images as each image adds novel constraints on the BRDF. The quality of the environment maps will only benefit from more images indirectly through the more precise BRDFs. Furthermore, the quality of the results is to some extent influenced by the precision of the input geometry. Another strength of our approach is that no assumptions are placed on the image set. We flexibly can incorporate image collections gathered from various sources.

A problem that remains is the ambiguity between illu-



(a) input images for the *Liberty* dataset



(b) estimated environment maps



(c) object rendered under estimated illumination



(d) object rendered under point light illumination and under the *Uffizi* environment map.

Figure 8. Overview over the *Liberty* dataset.



(a) (b) (c) (d)

Figure 9. (a) shows the recovered albedo map for the *Minerva*, while (b)-(d) show the relit model using the *Uffizi* environment.



Figure 10. Results for the *Venus* dataset. The images (*left*) are input images while the images in the *middle* show our rendering using the recovered lighting (*right*) and BRDFs.

mination and surface color. Currently, this can only be resolved by having sufficiently different input environments. A couple of extensions of the current framework would be interesting. So far, our system does not take interreflections into account. An additional geometry optimization

step could help with the reconstruction, since more precise geometry and normals will improve the estimation. This is supported by the synthetic test cases where we see a close to perfect reconstruction. Furthermore, it would be interesting to incorporate other ways to stabilize the estimation. For this, techniques such as reflectance sharing [32] could be incorporated.

Acknowledgments

We would like to thank all the flickr users for contributing their photographs. The first author was funded by "tUL impulsfinanciering". Part of the research at EDM is funded by the IBBT. This work has been partially funded by the DFG Emmy Noether fellowships (Le 1341/1-1, GO 1752/3-1) and the Max Planck Center for Visual Computing and Communication (BMBF-FKZ01IMC01).

References

- [1] E.-W. Bai and Y. Liu. Least squares solutions of bilinear equations. *Systems & Control Letters*, 55(6):466–472, June 2006. 4
- [2] Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, 1996. 4
- [3] A. M. Buchanan and A. W. Fitzgibbon. Damped newton algorithms for matrix factorization with missing data. In *CVPR*, pp. 316–322, 2005. 4
- [4] W.-C. Chen, J.-Y. Bouguet, M. H. Chu, and R. Grzeszczuk. Light field mapping: efficient representation and hardware rendering of surface light fields. In *SIGGRAPH*, 2002. 6
- [5] S. Cohen and C. Tomasi. Systems of bilinear equations. Tech. Rep. CS-TR-97-1588, Stanford University, 1997. 4
- [6] P. Debevec. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH*, pp. 189–198, 1998. 2, 5
- [7] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *SIGGRAPH*, pp. 145–156, 2000. 1, 2
- [8] P. Debevec, C. Tchou, A. Gardner, T. Hawkins, C. Poullis, J. Stumpfel, A. Jones, N. Yun, P. Einarsson, T. Lundgren, M. Fajardo, and P. Martinez. Estimating surface reflectance properties of a complex scene under captured natural illumination. Tech. Rep. ICT-TR-06.2004, USC, 2004. 2
- [9] Flickr. <http://www.flickr.com>, 2008. 1, 6
- [10] A. S. Georghiades. Recovering 3-d shape and reflectance from a small number of photographs. In *EGSR*, pp. 230–240, 2003. 2
- [11] E. M. Gertz and S. J. Wright. Object-oriented software for quadratic programming. *ACM Trans. Math. Softw.*, 29(1):58–81, 2003. 4
- [12] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz. Multi-view stereo for community photo collections. In *ICCV*, 2007. 1, 3, 6
- [13] S. Kuthirummal, A. Agarwala, D. B. Goldman, and S. K. Nayar. Priors for Large Photo Collections and What They Reveal about Cameras. In *ECCV*, pp. 74–87, 2008. 6
- [14] E. P. F. Lafortune, S.-C. Foo, K. E. Torrance, and D. P. Greenberg. Non-linear approximation of reflectance functions. In *SIGGRAPH*, pp. 117–126, 1997. 5
- [15] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM TOG*, 22(2):234–257, 2003. 1, 2, 4
- [16] S. Marschner. *Inverse rendering for computer graphics*. PhD thesis, Cornell University, 1998. 4
- [17] S. Marschner, S. Westin, E. Lafortune, K. Torrance, and D. Greenberg. Image-based BRDF measurement including human skin. In *EGWR*, pp. 139–152, 2000. 2
- [18] R. Ng, R. Ramamoorthi, and P. Hanrahan. Triple product wavelet integrals for all-frequency relighting. In *SIGGRAPH*, pp. 477–487, 2004. 1, 3
- [19] A. Ngan, F. Durand, and W. Matusik. Experimental analysis of BRDF models. In *EGSR*, pp. 117–126, 2005. 5
- [20] K. Nishino, K. Ikeuchi, and Z. Zhang. Re-rendering from a sparse set of images. Tech. Rep. DU-CS-05-12, Drexel University, 2005. 2
- [21] K. Nishino, Z. Zhang, and K. Ikeuchi. Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis. In *ICCV*, pp. 599–606, 2001. 1, 2
- [22] T. Okabe, I. Sato, and Y. Sato. Spherical harmonics vs. haar wavelets: Basis for recovering illumination from cast shadows. In *CVPR*, pp. 50–57, 2004. 2
- [23] E. Praun and H. Hoppe. Spherical parametrization and remeshing. *ACM TOG*, 22(3):340–349, 2003. 3
- [24] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. In *SIGGRAPH*, pp. 117–128, 2001. 2, 3, 4, 5
- [25] F. Romeiro, Y. Vasilyev, and T. Zickler. Passive reflectometry. In *ECCV*, 2008. 2, 3
- [26] I. Sato, Y. Sato, and K. Ikeuchi. Illumination from shadows. *Trans. PAMI*, 25(3):290–300, 2003. 1, 2, 4
- [27] P.-P. Sloan, J. Kautz, and J. Snyder. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In *SIGGRAPH*, pp. 527–536, 2002. 3
- [28] R. Wang, R. Ng, D. Luebke, and G. Humphreys. Efficient wavelet rotation for environment map rendering. In *EGSR*, pp. 173–182, 2006. 3
- [29] R. P. Weistroffer, K. R. Walcott, G. Humphreys, and J. Lawrence. Efficient Basis Decomposition for Scattered Reflectance Data. In *EGSR*, pp. 207–218, 2007. 2, 4
- [30] T. Yu, H. Wang, N. Ahuja, and W.-C. Chen. Sparse luminograph relighting by illumination and reflectance estimation from multi-view images. In *EGSR*, pp. 41–50, 2006. 1, 2, 3, 6
- [31] Y. Yu and J. Malik. Recovering photometric properties of architectural scenes from photographs. In *SIGGRAPH*, pp. 207–217, 1998. 1, 2
- [32] T. Zickler, R. Ramamoorthi, S. Enrique, and P. Belhumeur. Reflectance sharing: predicting appearance from a sparse set of images of a known shape. *Trans. PAMI*, 28(8):1287–1302, 2006. 2, 8