Workshop on Material Appearance Modeling (2016)
H. Rushmeier and R. Klein (Editors)

Exploring Material Recognition for Estimating Reflectance and Illumination From a Single Image

Michael Weinmann and Reinhard Klein

University of Bonn, Germany

Abstract

In this paper, we propose a novel approach for recovering illumination and reflectance from a single image. Our approach relies on the assumption that the surface geometry has already been reconstructed and a-priori knowledge in form of a database of digital material models is available. The first step of our technique consists in recognizing the respective material in the image using synthesized training data based on the given material database. Subsequently, the illumination conditions are estimated based on the recognized material and the surface geometry. Using this novel strategy we demonstrate that reflectance and illumination can be estimated reliably for several materials that are beyond simple Lambertian surface reflectance behavior because of exhibiting mesoscopic effects such as interreflections and shadows.

Categories and Subject Descriptors (according to ACM CCS): I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Intensity, color, photometry, and thresholding; Modeling and recovery of physical attributes I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture I.4.2 [Image Processing and Computer Vision]: Scene Analysis—Color; Photometry

1. Introduction

The way we perceive our environment with its objects is determined by the complex interplay of surface material characteristics, illumination conditions and the surface geometry with respect to the viewpoint. While identifying material characteristics provides important information on how we have to handle individual objects and allows reasoning about their fragility, deformability or weight, the extraction of this information is complicated by the fact that material appearance heavily depends on surface geometry, viewing conditions and the surrounding illumination conditions (see Figure 1). On the other hand, knowing illumination conditions enables a multitude of interesting applications such as e.g. inserting virtual objects in the scene. Understanding the scene, hence, requires the decomposition of the scene into its geometry, reflectance and illumination conditions.

In this paper, we focus on the reconstruction of reflectance and illumination characteristics from a single image based on an a-priori known surface geometry. While many techniques have been developed for the reconstruction of the surface geometry based on structured light scanners, multi-view stereo systems or deflectometry, the recovery of illumination and reflectance effects remains a challenging task due to its severely under-constrained nature. After the separation of the known geometry from the unknown reflectance and illumination characteristics, the latter two modalities have to be decoupled as well to get a more complete understanding of the scene. In particular, surfaces with a complex surface reflectance behavior exhibiting mesoscopic effects of light exchange such as interreflections, self-shadowing, etc. at surface structures on the scale of scratches or engravings significantly complicate this separation task.

So far, related work is either limited to Lambertian surfaces [HFB*09, BM12, BM13] or considers BRDF models for objects with homogeneous surface reflectance [LN12, LN16, ON12]. However, many of the materials encountered in daily life do not follow Lambertian surface reflectance behavior or simple BRDF models but exhibit complex spatially varying surface reflectance characteristics including mesoscopic effects such as self-shadowing, self-occlusions, interreflections, etc.. As these materials cannot be modeled using simple analytical BRDF models, more complex material models such as bidirectional texture functions (BTFs) introduced in [DNGK97] have to be used.

DOI: 10.2312/mam.20161253

Unfortunately, separating illumination and reflectance is highly challenging for objects exhibiting a complex, inhomogeneous surface reflectance behavior as some properties of the different modalities can hardly be separated and might finally be included in the wrong modality. For instance, certain colors of the surface material might be be included in the final estimate of the illumination conditions. Standard techniques [HFB*09] approach this challenging task by following an iterative approach of alternating the separate fitting of either illumination or reflectance based on the assumption that the other modality is known. However, such alternating fitting approaches are susceptible regarding local minima in the individual iterations and more recent approaches [LN12,LN16] are tailored to homogeneous objects that can be modeled using rather simple BRDF models.

In this paper, we hence introduce a novel non-iterative scheme for separating reflectance and illumination, where the estimation of material and illumination is carried out in two separate steps. Instead of an alternating or joint fitting of illumination and reflectance, we exploit the potential of standard machine learning approaches for the recognition of the surface material in a database of digitized materials. Once the material has been recognized, the surface reflectance behavior is constrained, and, taking benefit of the already recognized surface reflectance characteristics, the subsequent illumination estimation task is significantly facilitated.

Our approach relies on the availability of a database of digitized material models that covers the materials that are expected to occur in the considered scenes. A major prerequisite is given by the capability of these material models to accurately model the respective materials. Bidirectional Texture Functions (BTFs) [DNGK97] have been proven to allow such an accurate representation of surface appearance including the aforementioned mesoscopic effects for many materials. In our experiments, we use the publicly available BTF material database published in [WGK14]. However, in the future, significantly larger databases might be used.

In summary, the key contributions of this paper are:

- A novel approach for recovering reflectance and illumination characteristics from a single image based on material recognition and a subsequent illumination estimation,
- An approach that is capable of handling objects exhibiting a spatially varying surface reflectance behavior including complex effects of light exchange.

2. Related Work

The idea of separating material properties and illumination characteristics in an input image has gained attraction since several decades and can be traced back to the studies in [BT78] where the term *intrinsic image* has been introduced. First solutions to this problem have been proposed based on the Retinex theory introduced in [LM71] and its extensions such as e.g. [Bla85,Hor86,FDB92,GJAF09,STL08]. Further

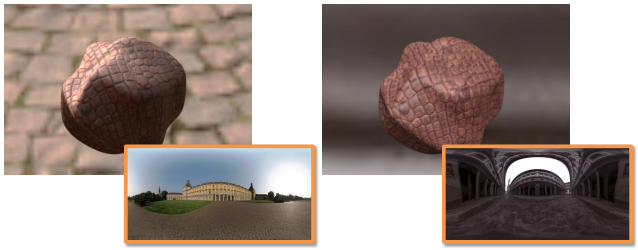

Figure 1: *Example illustration of the dependency of material appearance on surface geometry, surface reflectance, and the surrounding lighting conditions. Although surface material and geometry as well as the viewing conditions are the same in both figures, the differences in the illumination conditions induce a significantly different material appearance.*

approaches for intrinsic image decomposition for single images include [SYJL11, FPC00, GRK*11, BBS14, KPSL16]. In contrast, estimating intrinsic images from multiple views of an outdoor scene has been investigated in [LBD12].

Furthermore, a multitude of approaches exploits the potential of inverse rendering for estimating illumination and reflectance properties, i.e. during optimization of the formulated problem. For example, a more general solution to the intrinsic image problem has been analyzed in [BM12,BM13] where shape information is recovered in addition to the chromatic illumination and reflectance. In contrast, a model-based approach has been followed in [HFB*09] for photo collections. Based on an initial geometry estimation obtained via multi-view stereo, this technique solves for reflectance and illumination in an alternating manner, where the reflectance is represented using a set of basis BRDFs and illumination is represented using a set of wavelet-basis functions. Similarly, in [DS13] a frequency framework is used as well and camera response functions for the individual images of the input data formed by image collections are calculated in addition to geometry, reflectance and illumination. While all the above-mentioned approaches are based on assuming Lambertian reflectance behavior, the directional statistics bidirectional reflectance distribution function model has been used in [LN12] for the joint optimization of reflectance and illumination for objects with smooth, homogeneous surface materials from a single image. In [PCDS12], surface reflectance in terms of SVBRDFs is estimated for video sequences. However, only a few main light source directions derived from observed specular highlights are estimated in this approach. More recently, diffuse surface models have been considered in [KSH*14] for separating the input image into diffuse albedo and shading based on the Color Retinex algorithm [GJAF09]. Furthermore, the inverse rendering approach for illumination estimation for outdoor photo collections in [LM14] follows an iterative approach with separate illumination and reflectance estima-

tion steps that are alternated while keeping the other modality fixed. For this purpose, an outdoor illumination model is derived and a novel dataset is used, which contains images that are associated with their respective ground truth HDR illumination conditions. Our approach addresses the case of more complex surface reflectance behavior by recognizing the materials in the scene among materials in a database that have been measured and stored as bidirectional texture functions [DNGK97]. Based on the recognized material model, the illumination is estimated using an inverse rendering framework.

3. System Overview

An overview of the proposed approach for recovering reflectance and illumination from a single image is depicted in Figure 2. For an object with an a-priori known geometry obtained e.g. by using multi-view stereo techniques, structured light scanners, laser scanners or deflectometry, the first step in the scheme is represented by searching for the closest material in a database of synthesized images generated using different digital material models (see Section 3.2). Subsequently, the estimation of the illumination is performed based on the material model of the recognized reference in the database (see Section 3.3) and the known surface geometry. Before we focus on the corresponding technical details, we first describe our synthesis framework for generating synthesized training data in Section 3.1.

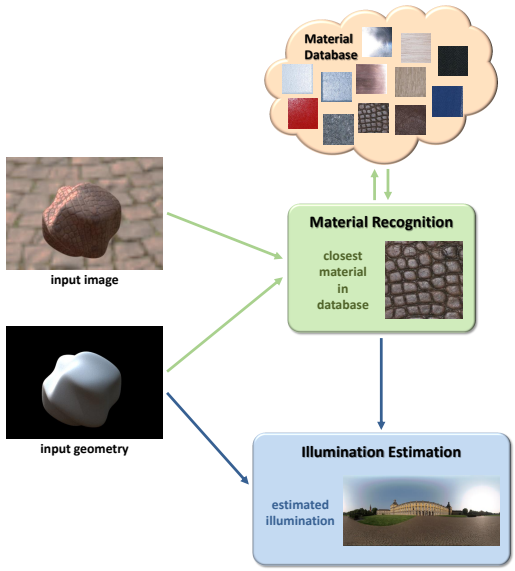

Figure 2: *The proposed approach for separating reflectance and illumination. In a first step, the surface material is recognized based on known surface geometry and a material database. Subsequently, the recognized material is used together with the surface geometry to estimate the illumination in the scene.*

3.1. Rendering Pipeline

An indispensable requirement of our approach is the availability of accurate digital material models that faithfully preserve characteristic traits of visual material appearance in the synthesized data. These material models are needed in the material recognition step for the generation of training data that depicts the appearance of the individual materials under varying conditions regarding surface geometry, illumination conditions and viewing configurations as mentioned in Section 3.2. In addition, these models are needed in the illumination estimation step, which is based on the availability of an appropriate digital material model (see Section 3.3).

For modeling spatially inhomogeneous materials which exhibit non-local effects such as interreflections and self-shadowing, as met in many daily life situations, bidirectional texture functions (BTFs) as introduced in [DNGK97] currently represent the state-of-the-art technique (see e.g. [HF11, SWRK11, HF13, SSW*14, WK15, WLGK16]). Bidirectional texture functions $\mathrm{BTF}(\mathbf{x},\omega_i,\omega_o)$ are six-dimensional functions, where the material appearance is represented for each point $\mathbf{x}$ on the surface depending on the incoming light direction ω_i and the outgoing light direction ω_o. Per surface point, the dependency of the observed reflectance is stored in terms of apparent BRDFs [WHON97] $ABRDF(\omega_i,\omega_o)$, that have the same parameters as conventional BRDFs. However, no energy conservation is enforced for ABRDFs.

Similar to the approach presented in [WGK14], we implemented a rendering pipeline based on the Mitsuba pathtracer [Wen10] where the virtual objects in the scene are illuminated via environment maps (e.g. [Deb98]). Using pathtracing for the rendering process, we obtain the exitant radiance $L_r(\mathbf{x},\omega_o)$ for each surface point $\mathbf{x}$ via the image-based relighting equation

$$L_r(\mathbf{x},\omega_o) = \int_\Omega \mathrm{BTF}(\mathbf{x},\omega_i,\omega_o)\, L_i(\omega_i)\, V(\mathbf{x},\omega_i)\, d\omega_i. \quad (1)$$

Here, the incoming and outgoing light directions are represented by ω_i and ω_o, and $L_i(\omega_i)$ denotes the radiance distribution in the environment map over the spherical domain Ω. Furthermore, $V(\mathbf{x},\omega_i)$ denotes a visibility function in form of a binary indicator function for considering if the environment map is visible from surface point $\mathbf{x}$ in the direction ω_i.

3.2. Material Recognition

Given an input image where the surface geometry is already known e.g. as a result of a preceding geometry acquisition, we first focus on estimating the surface material. Instead of fitting parametric BRDF models as in e.g. [HFB*09], our approach is based on using the BTF database presented

in [WGK14] and [WK14] to render different surface geometries with the different materials under different natural lighting conditions (here, we use environment maps for relighting the materials [Deb98]) for generating realistic training data. For this purpose, we use the pathtracer implementation in Mitsuba [Wen10].

For retrieving the material in the database that is closest to the one in the image for which the estimation of surface material and illumination has to be calculated, we use a standard retrieval method. In the feature extraction step, we densely extract color patches and SIFT descriptors per image, i.e. for the training data and the query image, which has been shown to be a valuable combination of complementary descriptor types in [WK14]. Based on the descriptors extracted for the training data, we compute a dictionary of visual words using the k-means algorithm. Following [LSAR10], we use a dictionary of $k = 150$ visual words for the color descriptors and a dictionary of $k = 250$ visual words for the SIFT descriptors. For encoding the information contained in the set of descriptors extracted for an individual image in a single vector-based representation, we make use of the VLAD representation proposed in [JDSP10] for describing the content of the masked object regions, however, other representations such as Fisher vectors [PD07] would also be adequate. Instead of simply counting the occurrences of the individual visual words assigned to the descriptors within an individual image as done in standard histogram-based bag-of-words methods, the VLAD representation is obtained by storing the sum of deviations of the descriptors $\mathbf{x}_i$ assigned to the visual word $\mathbf{c}_j$ and the respective visual word $\mathbf{c}_j$ itself, i.e.

$$\mathbf{v}_j = \sum_{\{\mathbf{x}_i | \mathrm{NN}(\mathbf{x}_i)=\mathbf{c}_j\}} \left(\mathbf{x}_i - \mathbf{c}_j\right), \tag{2}$$

where $\mathrm{NN}(\mathbf{x}_i)$ denotes the nearest neighbor of the descriptor $\mathbf{x}_i$ among the visual words that is given by $\mathbf{c}_j$. The final VLAD representation of an image is obtained by concatenating the components $\mathbf{v}_j$ for all $j = 1,\ldots,k$ according to $\mathbf{v} = [\mathbf{v}_1,\ldots,\mathbf{v}_k]$. For searching the closest reference in the database to an input image, we use a nearest-neighbor classifier based on the L_2-distance between the VLAD vectors.

We compute such VLAD descriptors for local image regions. In order to determine the material of a particular image region, the VLAD descriptor of the region is used as input to a non-linear support vector machine classifier with an RBF kernel. Both the regularization parameter and the kernel parameter are determined using grid search based on training data where materials are shown on different objects under different illumination conditions.

3.3. Illumination Estimation

Having retrieved the closest material type in the database, the next step consists in inferring the illumination conditions. Again, similar to the material recognition performed before, we take benefit of the availability of the database of highly accurate digital, data-driven material models in form of bidirectional texture functions. As illustrated in Figure 2, at this stage we have an input image with an attached depth map as well as an estimate of the surface material from the material recognition step performed before.

Similar to prior work, we aim at establishing an optimization framework for minimizing the deviation of the input image from a synthetically reproduced image of the object as seen under the illumination to be estimated. Therefore, our objective function is represented by

$$L_i^* = \arg\min_{L_i \in \mathcal{L}} \left\| I - I_{synth}(M,G,L_i) \right\|^2, \tag{3}$$

where I denotes an input image, $I_{synth}(M,G,L_i)$ represents a rendered image using the material model M, the geometry G and the lighting conditions L_i. Furthermore, we restrict the possible conditions of the incoming light L_i to the subset $\mathcal{L}$ which can be represented using a linear combination of spherical harmonics l_i similar to [HFB*09, BM12, BM13, KSES14, KSH*14]. In more detail, representing the illumination in the spherical harmonics basis with l_{ij} being the j-th harmonics and α_j being the assigned coefficient in the linear combination, we obtain

$$L_i^* = \arg\min_{L_i \in \mathcal{L}} \left\| I - \int_\Omega BTF(\mathbf{x},\omega_i,\omega_o) \cdot \right. \tag{4}$$

$$\left. \cdot\, L_i(\omega_i)\, V_i(\mathbf{x},\omega_i)\, d\omega_i \right\|^2 \tag{5}$$

$$= \arg\min_{L_i \in \mathcal{L}} \left\| I - \int_\Omega BTF(\mathbf{x},\omega_i,\omega_o) \cdot \right. \tag{6}$$

$$\left. \cdot \sum_{j=1}^{\sharp SHs} \alpha_{ij}\, l_{ij}(\omega_i)\, V_i(\mathbf{x},\omega_i) d\omega_i \right\|^2 \tag{7}$$

$$= \arg\min_{L_i \in \mathcal{L}} \left\| I - I_{synth}(M,G,\sum_{j=1}^{\sharp SHs} \alpha_j l_{ij}) \right\|^2 \tag{8}$$

$$= \arg\min_{L_i \in \mathcal{L}} \left\| I - \sum_{j=1}^{\sharp SHs} \alpha_j I_{synth}(M,G,l_{ij}) \right\|^2. \tag{9}$$

Please note that the equality of

$$I_{synth}(M,G,\sum_{j=1}^{\sharp SHs} \alpha_j l_{ij}) = \sum_{j=1}^{\sharp SHs} \alpha_j\, I_{synth}(M,G,l_{ij}) \tag{10}$$

is a result of the linear behavior of light. However, the latter is easier to implement as it only requires to render the object with the assigned material and geometry under the individual basis illuminations. As a consequence, we have to subtract a linear combination of these images obtained under the basis illuminations from the input image.

Functional (9) can further be rewritten using vector-matrix notation

$$L_i^* = \arg\min_{L_i \in \mathcal{L}} \left\| \mathbf{b} - \mathbf{A}\mathbf{x} \right\|^2, \tag{11}$$

where $\mathbf{x}$ represents a vector of the coefficients corresponding to the individual basis illuminations, $\mathbf{b}$ denotes the reshaped version of the input image in form of a vector and $\mathbf{A} = \mathbf{A}(M, G, L_i)$ contains the synthesized image values obtained under the basis illumination. This linear system can be solved efficiently in the least-squares sense. At this stage, one might argue that the uv-coordinates of the geometries used for producing the images $I(M, G, L_i)$ are different from the ones in the input image, i.e. spatially varying features of the material such as ridges in a leather do not coincide anymore introducing many outliers in the optimization. Our investigations show, that the lighting estimation still gives a rather compelling result for a multitude of materials as enough inlier data seems to be present.

Furthermore, we also consider previous investigations where it has been shown that the low-frequent effects resulting from the lighting and the high-frequent details are material-specific. For this purpose, we performed experiments using meanABRDFs, i.e. the average of all ABRDFs per acquired BTF instead of BTFs for rendering the images used in the optimization functional (9). In a certain sense, this corresponds to introducing data-driven regularization in the optimization process. In the scope of our experiments, this regularizer lead to an even more accurate recovery of reflectance and illumination. As will be demonstrated in the experimental results (see Section 4), the illumination estimates are of good quality which we expect to be a main achievement of following a two-step approach where the two involved subproblems can be solved more easily in comparison to the combined optimization.

4. Experimental Results

In our experiments, we generated synthetic data using the rendering pipeline as described in Section 3.1. In such synthesized scenarios, we have complete control on the individual factors (scene geometry, surface reflectance behavior, illumination conditions, camera parameters, ...) that contribute to material appearance. This allows to establish ground truth data for evaluating our approach.

For the first step of recognizing the surface material, we make use of a database of images where the geometry is rendered using the different surface materials we have measured in the scope of the publication [WGK14]. Throughout the experiments, we take care that the environment maps present in the input images are not used for producing the reference data used for material recognition. Based on the surface geometry and the recognized surface material, we subsequently estimate the illumination conditions. In Figure 3 and Figure 4, we show results for the separation of reflectance and illumination for different input images achieved using our technique. For each input image, we show the classified material, the estimated environment map (represented in spherical harmonics basis), the ground truth environment map (represented in spherical harmonics basis) and a rendering of the geometry with the classified material under the estimated environment map. Please note that our technique is not restricted to Lambertian surface reflectance behavior and can also deal with mesoscopic surface effects. As can be seen, the materials can be recognized quite well for the small subset of data we used so far. Furthermore, when comparing the input images with the images produced using the combination of the estimated surface material and the estimated environment map (in its spherical harmonics representation), we can observe only subtle differences which validates our approach. For evaluating the estimated illumination conditions, we also show the estimated environment map (in its spherical harmonics representation) next to the ground truth environment map. Here, we observe that the light intensities are matched quite well and that there are no significant artifacts due to the influence of the material reflectance occurring in case of a non-accurate recovery of illumination and reflectance.

Please note that the texture coordinates for the reconstructed objects have been computed using ABF++ [SLMB05] and, hence, deviate from the ones used for generating the input images. As a consequence, the material mapped onto the surface via BTF patches according to the uv-parametrization might lead to perceptual errors. For larger structures on the surface as occurring for wood or the coarse leather, slight deviations from the original input can be perceived, but still the object clearly appears to be made of the same material. In contrast, the perception of such deviations is negligible for the other examples which demonstrates the potential of our approach.

In the scope of our experiments, we focused on synthetic and, hence, controllable scenarios for validating our approach. Controllable scenarios allow us to make use of ground truth data in the scope of our experiments which represents an important aspect for inferring insights regarding the developed technique. In order to approach real-world scenarios, larger datasets of digitized materials are required which might be available in the future.

5. Conclusions

In the scope of this paper, we introduced a novel approach for recovering illumination and reflectance from single images based on an a-priori determined surface geometry and a-priori knowledge in form of a database of materials measured in standard reflectance acquisition devices. After an initial material recognition phase, our technique continues with estimating illumination conditions based on the recognized material and the surface geometry. Our results indicate that a recovery of reflectance and illumination can be performed reliably for several materials that are beyond simple Lambertian surface reflectance behavior.

In the future, the availability of large collections of digitized materials will make our approach valuable for understanding real-world scenarios.

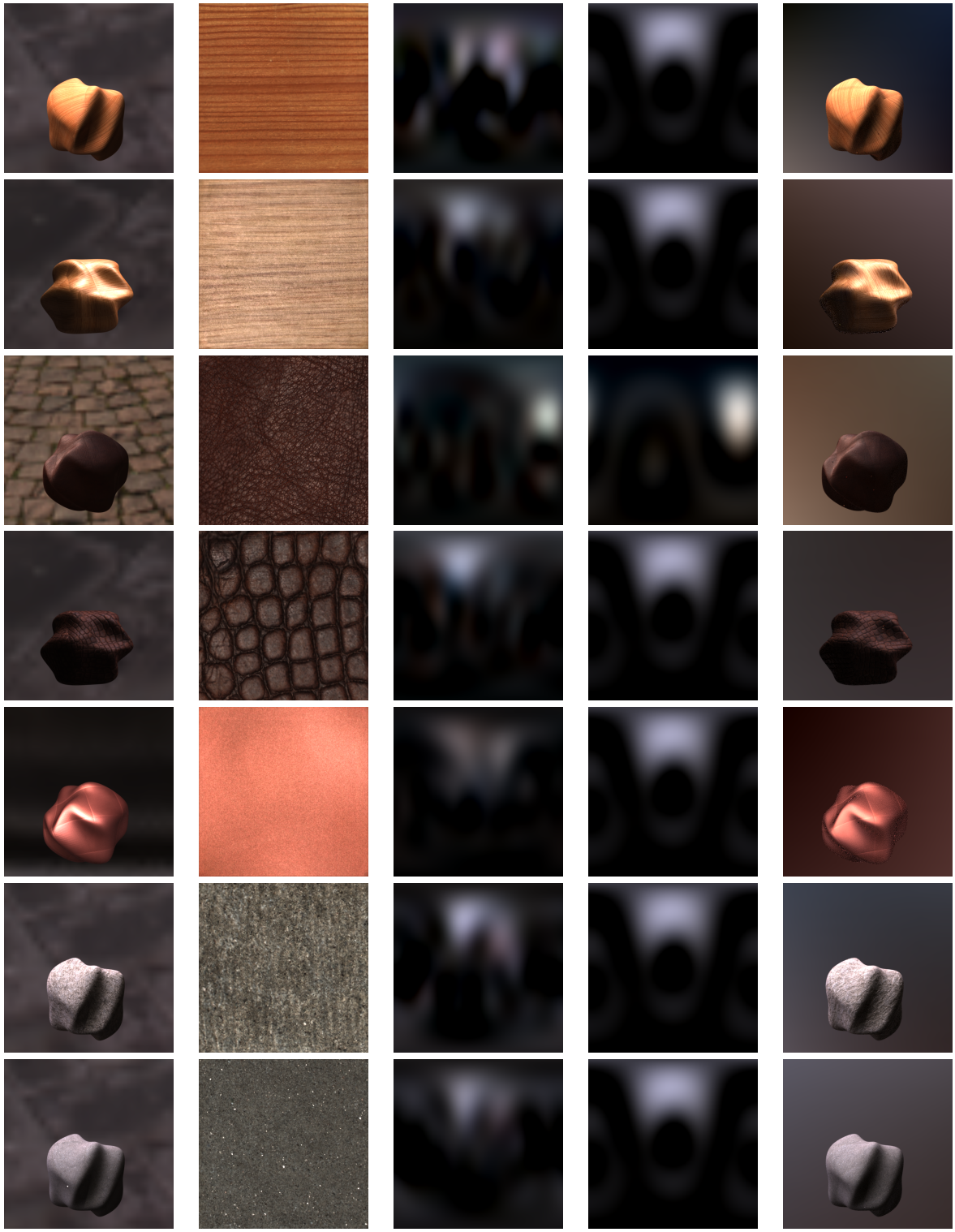

Figure 3: *Per row: Input image, classified material, estimated illumination (in spherical harmonics basis), ground truth illumination (in spherical harmonics basis) and rendering of the geometry and recognized material under relighting with the estimated environment map.*

Figure 4: *Per row: Input image, classified material, estimated illumination (in spherical harmonics basis), ground truth illumination (in spherical harmonics basis) and rendering of the geometry and recognized material under relighting with the estimated environment map.*

6. Acknowledgements

The research leading to these results was partially funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 323567 (Harvest4D); 2013-2016.

References

[BBS14] BELL S., BALA K., SNAVELY N.: Intrinsic images in the wild. *ACM Trans. Graph. 33*, 4 (2014), 159:1–159:12. 2

[Bla85] BLAKE A.: Boundary conditions for lightness computation in mondrian world. *Computer Vision, Graphics, and Image Processing 32*, 3 (1985), 314–327. 2

[BM12] BARRON J. T., MALIK J.: Shape, albedo, and illumination from a single image of an unknown object. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2012), pp. 334–341. 1, 2, 4

[BM13] BARRON J., MALIK J.: Intrinsic scene properties from a single rgb-d image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2013), pp. 17–24. 1, 2, 4

[BT78] BARROW H. G., TENENBAUM J. M.: *Recovering Intrinsic Scene Characteristics from Images*. Academic Press, 1978. 2

[Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)* (1998), pp. 189–198. 3, 4

[DNGK97] DANA K. J., NAYAR S. K., GINNEKEN B. V., KOENDERINK J. J.: Reflectance and texture of real-world surfaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (1997), pp. 151–157. 1, 2, 3

[DS13] DÍAZ M., STURM P.: Estimating Photometric Properties from Image Collections. *Journal of Mathematical Imaging and Vision 47*, 1-2 (2013), 93–107. 2

[FDB92] FUNT B. V., DREW M. S., BROCKINGTON M.: Recovering shading from color images. In *Proceedings of the European Conference on Computer Vision (ECCV)* (1992), vol. 588 of *Lecture Notes in Computer Science*, Springer, pp. 124–132. 2

[FPC00] FREEMAN W. T., PASZTOR E. C., CARMICHAEL O. T.: Learning low-level vision. *Int. J. Comput. Vision 40*, 1 (2000), 25–47. 2

[GJAF09] GROSSE R., JOHNSON M. K., ADELSON E. H., FREEMAN W. T.: Ground-truth dataset and baseline evaluations for intrinsic image algorithms. In *Proceedings of the International Conference on Computer Vision (ICCV)* (2009), pp. 2335–2342. 2

[GRK*11] GEHLER P. V., ROTHER C., KIEFEL M., ZHANG L., SCHÓLKOPF B.: Recovering intrinsic images with a global sparsity prior on reflectance. In *Proceedings of the Conference on Neural Information Processing Systems (NIPS)* (2011), pp. 765–773. 2

[HF11] HAINDL M., FILIP J.: Advanced textural representation of materials appearance. In *SIGGRAPH Asia 2011 Courses* (2011), ACM, pp. 1:1–1:84. 3

[HF13] HAINDL M., FILIP J.: *Visual Texture*. Advances in Computer Vision and Pattern Recognition. Springer-Verlag, London, 2013. 3

[HFB*09] HABER T., FUCHS C., BEKAERT P., SEIDEL H., GOESELE M., LENSCH H. P. A.: Relighting objects from image collections. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2009), pp. 627–634. 1, 2, 3, 4

[Hor86] HORN B. K. P. (Ed.): *Robot Vision*. MIT Press, Cambridge, MA, USA, 1986. 2

[JDSP10] JEGOU H., DOUZE M., SCHMID C., PÉREZ P.: Aggregating local descriptors into a compact image representation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2010), pp. 3304–3311. 4

[KPSL16] KIM S., PARK K., SOHN K., LIN S.: Unified depth prediction and intrinsic image decomposition from a single image via joint convolutional neural fields. *CoRR abs/1603.06359* (2016). URL: http://arxiv.org/abs/1603.06359. 2

[KSES14] KHOLGADE N., SIMON T., EFROS A., SHEIKH Y.: 3d object manipulation in a single photograph using stock 3d models. *ACM Transactions on Computer Graphics 33*, 4 (2014), 127:1–127:12. 4

[KSH*14] KARSCH K., SUNKAVALLI K., HADAP S., CARR N., JIN H., FONTE R., SITTIG M., FORSYTH D.: Automatic scene inference for 3d object compositing. *ACM Trans. Graph. 33*, 3 (2014), 32:1–32:15. 2, 4

[LBD12] LAFFONT P.-Y., BOUSSEAU A., DRETTAKIS G.: Rich intrinsic image decomposition of outdoor scenes from multiple views. In *ACM SIGGRAPH 2012 Talks* (New York, NY, USA, 2012), SIGGRAPH '12, ACM, pp. 51:1–51:1. 2

[LM71] LAND E. H., MCCANN J. J.: Lightness and retinex theory. *J. Opt. Soc. Am. 61*, 1 (1971), 1–11. 2

[LM14] LALONDE J.-F., MATTHEWS I.: Lighting estimation in outdoor image collections. In *Proceedings of 3DV* (2014), IEEE Computer Society, pp. 131–138. 2

[LN12] LOMBARDI S., NISHINO K.: Reflectance and natural illumination from a single image. In *Proceedings of the 12th European Conference on Computer Vision - Volume Part VI* (Berlin, Heidelberg, 2012), Springer-Verlag, pp. 582–595. 1, 2

[LN16] LOMBARDI S., NISHINO K.: Reflectance and illumination recovery in the wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 38*, 1 (2016), 129–141. 1, 2

[LSAR10] LIU C., SHARAN L., ADELSON E. H., ROSENHOLTZ R.: Exploring features in a bayesian framework for material recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2010), pp. 239–246. 4

[ON12] OXHOLM G., NISHINO K.: Shape and reflectance from natural illumination. In *Proceedings of the European Conference on Computer Vision (ECCV)* (2012), pp. 528–541. 1

[PCDS12] PALMA G., CALLIERI M., DELLEPIANE M., SCOPIGNO R.: A statistical method for svbrdf approximation from video sequences in general lighting conditions. *Comput. Graph. Forum 31*, 4 (2012), 1491–1500. 2

[PD07] PERRONNIN F., DANCE C. R.: Fisher kernels on visual vocabularies for image categorization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2007). 4

[SLMB05] SHEFFER A., LÉVY B., MOGILNITSKY M., BOGOMYAKOV A.: Abf++: Fast and robust angle based flattening. *ACM Trans. Graph. 24*, 2 (2005), 311–330. 5

[SSW*14] SCHWARTZ C., SARLETTE R., WEINMANN M., RUMP M., KLEIN R.: Design and implementation of practical bidirectional texture function measurement devices focusing on the developments at the university of bonn. *Sensors 14*, 5 (2014), 7753âĂŞ–7819. 3

[STL08] SHEN L., TAN P., LIN S.: Intrinsic image decomposition with non-local texture cues. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 0* (2008), 1–7. 2

[SWRK11] SCHWARTZ C., WEINMANN M., RUITERS R., KLEIN R.: Integrated high-quality acquisition of geometry and appearance for cultural heritage. In *Proccedings of the International Symposium on Virtual Reality, Archeology and Cultural Heritage VAST* (2011), Eurographics Association, pp. 25–32. 3

[SYJL11] SHEN J., YANG X., JIA Y., LI X.: Intrinsic images using optimization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2011), IEEE, pp. 3481–3487. 2

[Wen10] WENZEL J.: Mitsuba renderer, 2010. http://www.mitsuba-renderer.org. 3, 4

[WGK14] WEINMANN M., GALL J., KLEIN R.: Material classification based on training data synthesized using a btf database. In *Computer Vision -âĂŞ ECCV 2014*, vol. 8691 of *Lecture Notes in Computer Science*. Springer International Publishing, 2014, pp. 156–171. 2, 3, 4, 5

[WHON97] WONG T.-T., HENG P.-A., OR S.-H., NG W.-Y.: *Rendering Techniques '97: Proceedings of the Eurographics Workshop in St. Etienne, France, June 16–18, 1997*. Springer Vienna, Vienna, 1997, ch. Image-based Rendering with Controllable Illumination, pp. 13–22. 3

[WK14] WEINMANN M., KLEIN R.: Material recognition for efficient acquisition of geometry and reflectance. In *Color and Photometry in Computer Vision Workshop (ECCV 2014)*. Springer International Publishing, 2014. 4

[WK15] WEINMANN M., KLEIN R.: Advances in geometry and reflectance acquisition (course notes). In *SIGGRAPH Asia 2015 Courses* (New York, NY, USA, 2015), ACM, pp. 1:1–1:71. 3

[WLGK16] WEINMANN M., LANGGUTH F., GOESELE M., KLEIN R.: Advances in geometry and reflectance acquisition. In *Eurographics 2016 Tutorial* (2016). 3