STAG: Smart Tools & Apps for Graphics (2015)
Silvia Biasotti and Marco Tarini and Andrea Giachetti (Editors)

Recovering Intrinsic Images by Minimizing Image Complexity

N. Stefani, A. Fusiello

Università di Udine - Dipartimento di Ingegneria Elettrica Gestionale e Meccanica
Via Delle Scienze, 208 - 33100 Udine
nome.cognome@uniud.it

Abstract

This paper tackles the problem of decomposing a single image into two intrinsic images – a shading image (the illumination at each point) and a reflectance image (the colour at each point). Assuming a linear response of the camera, the acquired image $I(x,y)$ is modelled as the product of the shading $S(x,y)$ and the reflectance $R(x,y)$ (collectively called intrinsic images*): the goal is to recover S and R from $I(x,y)$. The proposed method stems from the observation that R is "simpler" than I, in some sense related to its information content. This allows to formulate the problem as the minimization over all the possible S of a cost function describing the complexity of a tentative reflectance image given a shading image S. Given a 3D model of the scene, the orientation of the camera, and an illumination model, S can be parameterized with the position of light sources on a hemisphere. Preliminary experiments in a simulated environment validate the substance of the method, although many details will be subject of further improvement.*

1. Introduction

Images of a three-dimensional (3D) scene depends on many physical characteristics of the surfaces, such as lighting, local geometry, depth, and reflectance. All these information are confounded by the imaging process into an array of integer values, the sensed image, which reveals the physical parameters only indirectly. An *intrinsic image* is an image that represents one of these physical properties, which are *intrinsic* to the surfaces in the scene.

The ability to decompose an image into its intrinsic components is a major step toward scene understanding, because algorithms often rely exclusively on one of the intrinsic characteristics of the scene. For example, shape-from-shading techniques require image data with no changes in colour (or albedo). In the context of 3D modelling, the reflectance image is used as a texture map: being devoid of illumination effects the model can be re-illuminated without artifacts.

Assuming a linear response of the camera (or, equivalently, after conversion from sRBG to linear RGB), the input image $I(x,y)$ is modelled as the product of the *shading* image $S(x,y)$ (the illumination at each point) and the *reflectance* image $R(x,y)$ (the colour at each point):

$$I(x,y) = S(x,y) \cdot R(x,y). \tag{1}$$

This model is the same behind the "light mapping" technique [Abr] in Computer Graphics, but we address here the inverse problem of decomposing $I(x,y)$ into two intrinsic images (see Fig. 1).

The main intuition behind our approach is that R is simpler than any other image depicting the same scene: if we compress R, for example, we should expect that it takes less bits than I, because R is typically flatter than I. Several measures of image complexity are available, which will be surveyed in Sec. 2. Let us assume for the moment that a suitable measure $\mathcal{C}$ is available. This allows to formulate the problem as the minimization over all the possible S of the cost function $\mathcal{C}(I/S)$. Unfortunately this problem is ill-posed and practically infeasible unless one is able to find a low-dimensional parameterization for S.

Nevertheless, if a 3D model of the scene is known together with the orientation (i.e., position and angular attitude) of the camera, given an illumination model, S can be parameterized with the position of the light sources on a hemisphere, making the problem tractable.

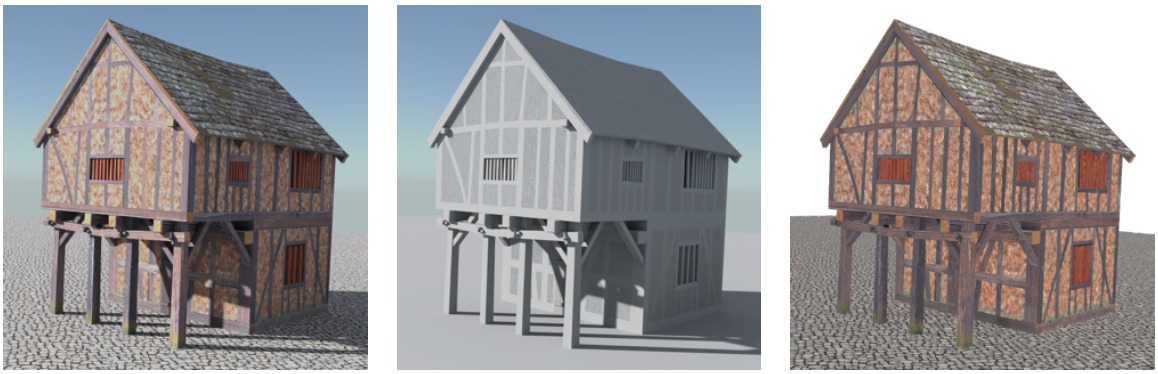

Figure 1: *From left to right: a synthetic image I, its shading component S and reflectance R. Please note how colour and texture information are contained in R, while S encompasses only shading variations, which modulates R.*

The motivation of this work is to produce correct (i.e., reflectance) textures for 3D models recovered by structure-from-motion (e.g., [TGFF15]) and multi-view stereo (e.g. [TFG*13]). Therefore the assumptions can be fulfilled.

As the focus of this paper is exploratory, we report only preliminary results on synthetic images.

1.1. Related works

The concept of intrinsic images was first introduced by Barrow and Tenenbaum [BT78]. With the term *intrinsic*, they refer to those characteristics that belong to a surface itself, such as local geometry, depth, reflectance, and incident illumination. They propose a mid-level decomposition in which every image can be thought as being the product of a reflectance component and an illumination component. Even if this decomposition does not make all the image features explicit, the authors state that this is still extremely useful for solving many computer vision problems.

An important single-image-based work is the one developed in [SA93]. The authors defined a new domain made of painted polyhedra, with no object occlusions, no cast shadows and a single distant light source. Authors aimed to solve the problem by analyzing the shape of every junction present in the scene and then with a successive global analysis on edge patterns.

A different approach was followed in [FHD02b], where the intrinsic components were recovered using an invariant grey-scale image from a previous work [FH01], where authors showed that an image invariant to illumination could be retrieved under particular illumination circumstances. The invariant image is used to find the location of shadow edges through appropriate thresholding. Building on the same invariant image [FHD02a] develop an automatic method to remove shadows by incorporating shadow edges location in the original retinex algorithm. In [FF07] the information contained in the grey-scale invariant image is used to compare the derivatives in the original image and classify them as being caused by reflectance or illumination.

[TFA05] took a completely different approach and managed to recover intrinsic images from a single image using color information and a classifier trained to recognize grey-scale patterns and discriminate between reflectance derivatives and illumination ones. Even if this algorithm does not put any restriction on the domain, it is still weak in a way that the classifier needs knowledge about the structure of the surface and about how it appears when illuminated. An evolution of this work was presented in [TAF06] where the authors tried to recover intrinsic images using non-linear regression on small patches of the image in order to avoid the problem of high dimensionality. Tappen's original idea was also used in [STL08] where a new non-local cue was introduced to better deal with ambiguous local inferences: the key idea is that distinct points with the same texture configuration generally have the same reflectance value. [SLX09] introduced the new concept of *colour invariant edge* to refine the classification process to classify reflectance and shading edges. [SYJL11] made the assumption that neighbouring pixels, in a local window, that have similar intensity values, should have the same reflectance. The intrinsic components are retrieved by optimizing an energy function with the aid of user scribbles, a similar approach was also used in [BPD09].

Another important algorithm is the one proposed in [FDL04]. The main idea was to recover intrinsic images from a single image by minimizing the entropy of a resulting invariant image. This algorithm performs surprisingly good even if the camera is unknown, eliminating the process of camera calibration, that was a necessary step in previous similar works.

All these mentioned works followed an approach that is called discriminative since these algorithms focus on discriminating between reflectance edges and illumination

edges; on the other hand the generative approach tries to create possible surfaces and reflectance patterns that explain the image and then use a model to choose the one the best fits. [SL03] presents an Independent Component Analysis related probabilistic model where illumination and reflectance are referred in a log space by a generalized autoregressive process and a Hidden Markov Random field. Other generative approaches are those proposed in [SY11, RKZ*11] where the key idea is that global reflectance can be modeled as a sparse set of reflectances since neighbour pixels have the same reflectance if their chromacities are similar. Lately the works by Barron and Malik showed outstanding results [BM11,BM12a,BM12b,BM15]: they proposed a global unified method, called *SIRFS*, to recover shape, illumination and reflectance from shading at the same time managing to outperform all other algorithms tested using MIT dataset [GJAF09]. Other references include [BBS14, GMLMG12].

Considering approaches treating multiple images, the most important work was proposed by Weiss, who assumed a sequence of images representing the same scene with different illumination [Wei01]. The reflectance is expected to be constant over time, while only the illumination component is changing in each image. All the images are filtered horizontally and vertically and then the median over k samples is taken before the information are combined back together to obtain the final result. Multiple images were also used to remove shadows from surveillance images in [MNIS03]. An evolution of this approach was proposed in 2004: in [MLKS04] the author recover intrinsic images by studying the effect of biased illumination on derivatives distribution and defining constraints on frames, pixels and making the assumption of smoothness. The first method designed explicitly for video was presented in [YGL*14].

In [LBP*12] the authors exploit several images of the same scene under different viewpoints and lighting conditions. They use multi-view stereo to automatically reconstruct 3D points and normals from which we derive relationships between reflectance values at different locations across multiple views. This work is the closest to our, as it builds on the availability of multiple images, with the difference that they assume also different lighting conditions (as it is usual in community photo collection), whereas our method works also with constant lighting conditions.

With the exception of [LBP*12] and our approach, previous work addressed the intrinsic decomposition problem assuming a single image or multiple images with static camera and different illumination. The availability of multiple images taken from different viewpoints or, equivalently, of a 3D model of the scene has been never exploited.

2. Image complexity

As stated in the introduction, our plan is to recover the intrinsic images by minimizing $C(I/S)$, the complexity of the reflectance image. Therefore a suitable measure of image complexity should be adopted.

In [YW13] an overview on this problem is reported and the correlation between the compression of an image with different encoders and three spatial operators is analysed.

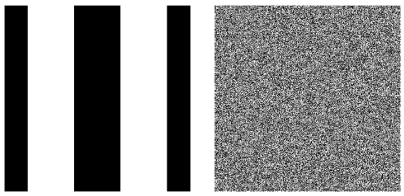

Figure 2: *Two images with the same entropy, but different complexity.*

The concept of complexity that suits our need cannot be reduced to Shannon's information theory only. For example, the two images in Fig. 2 have the same entropy but in some sense the right image is much more complex than left one because of the spatial arrangement of pixels, that is ignored by entropy. Compression-based complexity measures seem to be the most appropriate, in particular the so called compression ratio:

$$CR = \frac{s(I)}{s(C(I))} \tag{2}$$

where $s(I)$ is the file size of the uncompressed image I, and $s(C(I))$ is the file size of the output of compressor C. Unfortunately these measures are quite expensive to compute, but in [YW13] it is shown that $1/CR$ correlates very well with some spatial operators, in particular with

$$SI_{mean} = \frac{1}{N}\sum\sqrt{s_h^2 + s_v^2} \tag{3}$$

where s_h and s_v are the results of the filtering of original image I with horizontal and vertical Sobel kernels respectively, and N is the number of pixels.

Hereinafter, we will use SI_{mean} as the complexity measure C.

3. Method

We assume to know the 3D model of the scene, the orientation of the camera, and we are given an image of the scene I (taken from that camera) that needs to be decomposed in its intrinsic components, according to Eq. (1). We are only interested in recovering the *relative* reflectance, for the estimated reflectance and shading images are each allowed to be any scalar multiple of the true ones.

The idea behind the method is the following: if one were able to generate all the possible $S(x,y)$, he could recognize the right reflectance

$$R(x,y) = \frac{I(x,y)}{S(x,y)} \tag{4}$$

as the one with the lowest complexity $\mathcal{C}$.

The problem remains of how to bound the tremendous complexity of minimizing it over all the shading images, which is also an ill-posed problem.

The knowledge of the 3D model of the scene and some other assumptions (to be further specified) allows us to cut this huge search space down to a few dimensions, that corresponds to the unknown position of the light sources.

For the sake of simplicity, we assume a single distant light source, that can be parameterized as a point (ψ,ϕ) in a hemisphere. Given a value of (ψ,ϕ) a tentative shading image S is rendered using a grey, diffusive 3D model. From this, a tentative R is obtained, whose complexity $\mathcal{C}$ is evaluated. In the end, the complexity of $R(x,y) = \frac{I(x,y)}{L(x,y)}$ depends only on (ψ,ϕ), so we can write $\mathcal{C}(\psi,\phi)$ and solve:

$$(\psi_o,\phi_o) = \arg\min_{(\psi,\phi)} \mathcal{C}(\psi,\phi) \tag{5}$$

to find the parameters (ψ_o,ϕ_o) that gives the simplest reflectance image.

This conjecture is supported by the following observation. Consider the simple diffusive object depicted in Fig. 3. This figure represents two possible decomposition of the same image, the first is correct, the second is wrong. We can tell it because the reflectance image in the bottom row contains shading artifacts, and the algorithm can recognize it as well because it has a higher complexity than the correct one.

As a matter of fact, if we look at the function $\mathcal{C}(\psi,\phi)$ plotted over its angular domain (Fig. 4) we see that the minimum is exactly in the correct position of the light: $(\psi_o,\phi_o) = (45^\circ, 180^\circ)$.

In the previous example the light was a point source at a known distance, however the same behaviour is obtained if we add some complexity to the scene (Fig. 5), by introducing a sun-like illumination, a wooden texture and enabling ambient occlusions.

It should be clear that the model (Eq. 1) does not cater for specular behaviour, which would require an extra additive term $C(x,y)$:

$$I(x,y) = S(x,y)\cdot R(x,y) + C(x,y). \tag{6}$$

Our method assumes diffusive behaviour, hence wrong results are obtained if a specular component is added to the materials, as shown in Fig. 6. Nevertheless, the method is able to recover the light direction correctly, hence the specular component can be estimated in a second stage. This will be subject to further research.

We should also expect failures when realistic global illumination models are applied, since in that case, the shading component should be coloured according to secondary reflections, which is not accounted for in our simple model.

In summary, the algorithm proposed here relies only on the correct determination of the light source position knowing the 3D model of the scene and knowing the position of the camera. The intrinsic reflectance component can be determined, provided that the illumination model and the material properties describe sufficiently well the image formation process.

If illumination is constant over multiple images, the light position can be estimated by minimizing an aggregate cost function that has one term for each image; otherwise the light position in estimated independently for each image.

4. Experiments

In order to generate our 3D scene to test the algorithm we used *Blender*, an open-source 3D modelling software, which also includes a built-in global rendering engine called *Cycles*. The reason for this choice is to be mainly found in the fact that *Blender* has a built-in *Python* interpreter, that allows the user to utilize both the *Blender* API and all the other available 3rd party *Python* modules at the same time. In particular these are the main *Python* modules used in our algorithm:

- *BPY* module, built-in into *Blender*, to generate and manipulate the 3D scene
- *PIL* module, to save read the images resulting from the rendering process
- *NUMPY* module, to perform generic operations on matrices
- *SCIPY* module, to perform advanced operations on matrices, such as optimization

The first basic scene shown in Fig.3 is composed of a cube, a sphere and a cylinder; the light source is a white point light and the materials used are a red diffusive surface for the cube, a blue one for the sphere and a green one for the cylinder. Once this first step is completed the next step involves the substitution of the original materials with a flat grey material and the re-positioning of the point light-source; the complexity of the image is then calculated and its value stored into a file that is subsequently loaded to produce the plots of the cost function. A single render takes about half a second, so the whole process takes about $(0.5*360*90)/3600 = 4.5$ hours to complete.

The recovery of the reflectance, however, does not entail the exploration of the whole (ψ,ϕ) domain, for it is the result of a minimization process. To this end we employed the Nelder-Mead method implemented in the *optimize* package of *SciPy*. For the scene of Fig. 3, Nelder-Mead took an average of 100 steps to converge to the right solution. Fig. 7 shows the step-by-step error between the correct solution and the current one.

A more realistic scene is depicted in Fig. 8. This scene includes all the characteristics we added so far to our basic scene, that are: cast shadows, textured materials, sun lamp

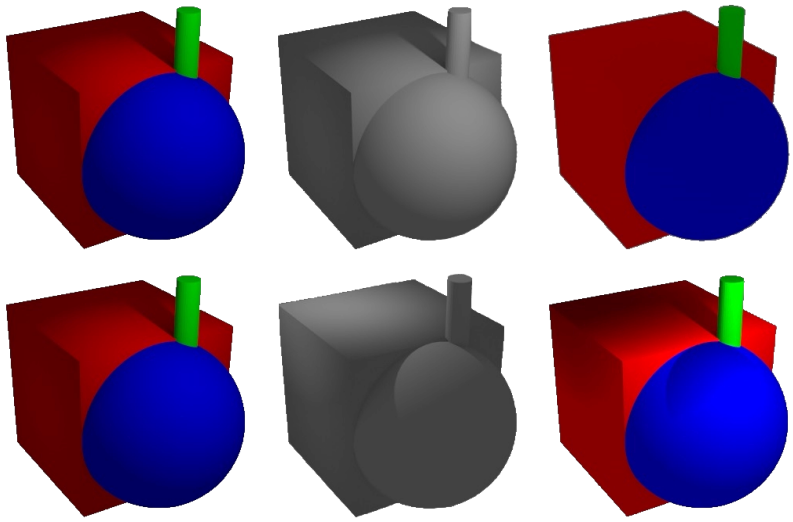

Figure 3: *An example of a simple rendered scene; Top: the correct shading and correct reflectance images. Bottom a wrong decomposition, where the resulting reflectance image contains shading artifacts.*

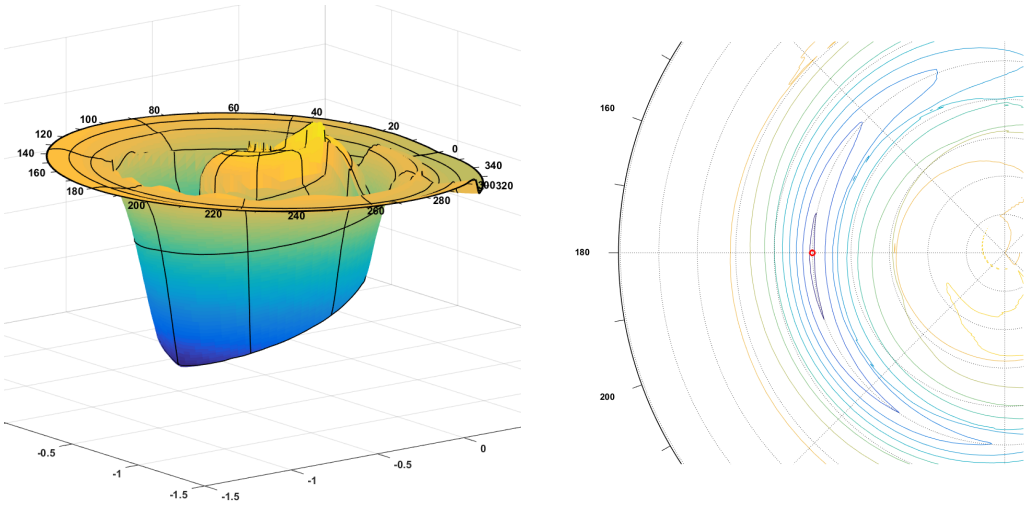

Figure 4: *Plot of complexity values obtained moving the light-source in a hemisphere, and its relative isolines, for the scene of Fig. 3. The red circle indicates the position of the ground truth global minimum.*

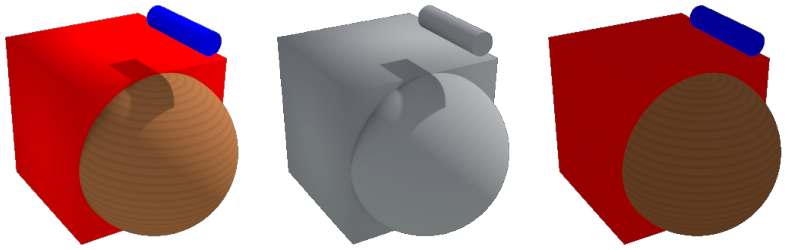

Figure 5: *Same scene as in Fig. 3, but with sun lamp, texture and Ambient Occlusion enabled.*

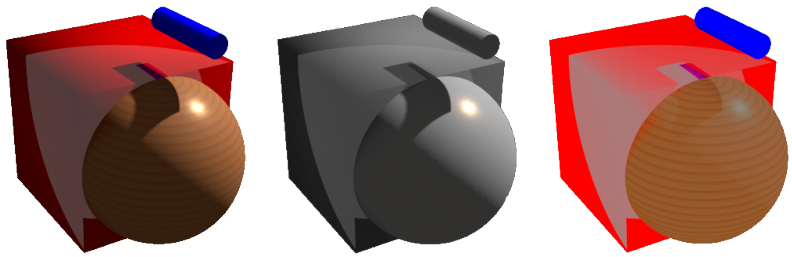

Figure 6: *Similar scene as in previous experiment, now with added reflective component to materials.*

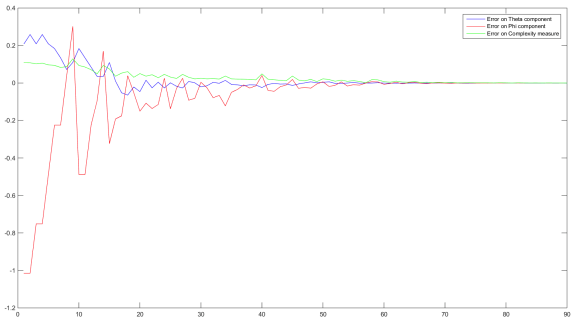

Figure 7: *Evolution of the errors relative to* ψ *(blue curve),* ϕ *(red curve), and the complexity (green curve), using Nelder-Mead method.*

and ambient occlusion, giving a realistic look to the whole scene. Please note that the sky in Blender is modelled so as to yield a coloured light, hence the S image is in colour as well.

The shading component is obtained by substituting all the materials with a grey diffusive one, and the final reflectance component is then calculated as the ratio between the target image and the shading one.

Figure 9 shows the plot of the cost function (complexity of a tentative reflectance) for this experiment. The shape is different (and there are fewer samples) than before, but still there is a distinct global minimum in the correct position $(\psi_o, \phi_o) = (45^\circ, 180^\circ)$, which the optimization procedure duly reaches.

In the last experiment (Fig. 10) we wanted to test how the algorithm would behave with a full global illumination: this means that the rendering engines calculates up to 128 bounces for a single light ray, and (potentially) all the surfaces of the scene contribute with their colour to determine the illumination.

The reflectance image (Fig. 10 right) is mainly correct, but an artifact appears in the area shown in detail in Fig. 11 (between the two roofs) that has a *reddish* colour, due to the fact that the red roof contributes to its illumination.

This problem is more severe in darker areas, where the indirect illumination prevail, and the colour of the illuminant is significantly different to that obtained in the grey scene used to compute the shading, where the information about the colour of the surfaces is not taken into account. Further investigation will be aimed at solving this problem. The working idea is that the image I can be used as a proxy of the unknown R in order to assign a colour to the light (while the surfaces remains grey) during the computation of S with global lighting.

Figure 12 shows the plot of the cost function for this experiment, with a distinct global minimum in the correct position $(\psi_o, \phi_o) = (45^\circ, 180^\circ)$.

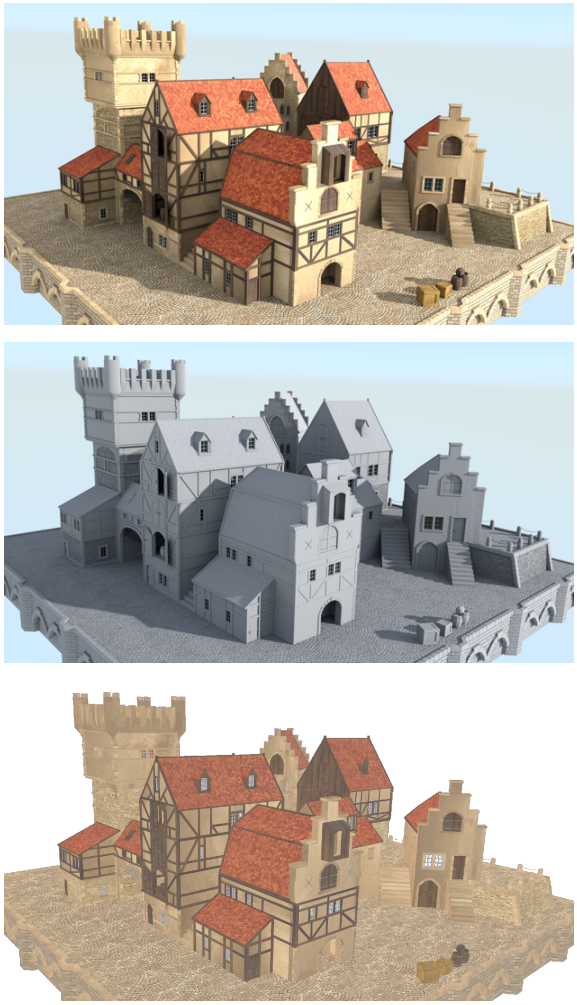

Figure 8: *From top to bottom: to right: original image, representing a small town with a tower, the shading component and the reflectance recovered by our method.*

Discussion

In summary, the algorithm proposed here does not require any type of classification on derivatives or learning process, but relies only on the correct determination of the light source position knowing the 3D model of the scene and knowing the position of the camera. Then, the intrinsic reflectance component is automatically determined, provided that the illumination model and the material properties describe sufficiently well the image formation process.

This represents the first step of a broader program, aimed at producing textured models from multiple images, where the texture is *intrinsic*, hence it does not carry with itself the illumination artifacts. We believe that the idea of identifying the correct reflectance image as the simplest one among the many images that can be obtained by rendering a synthetic shading image (suitably parameterized) is effective, and this preliminary experiments support the claim.

Future work will be aimed at (possibly) improving the complexity measure, coping with glossy and specular surfaces and with global illumination effects, with the aim of demonstrating the method on real images in conjunction with a structure-from-motion pipeline.

Acknowledgements

The 3D models of Fig. 1, 8, and 10 have been downloaded from the following web sites, respectively: www.blendswap.com/blends/view/67533, www.blendernation.com/2012/08/06/model-medieval-kind-of-seaport, www.blendswap.com/blends/view/26264.

References

[Abr] ABRASH M.: Quake's lighting model: Surface caching. www.bluesnews.com. Retrieved 2015-09-07. 1

[BBS14] BELL S., BALA K., SNAVELY N.: Intrinsic images in the wild. *ACM Trans. on Graphics 33*, 4 (July 2014), 159:1–159:12. 3

[BM11] BARRON J., MALIK J.: High-frequency shape and albedo from shading using natural image statistics. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* (June 2011), pp. 2521–2528. 3

[BM12a] BARRON J., MALIK J.: Color constancy, intrinsic images, and shape estimation. In *Proc. of the European Conf. on Computer Vision*, vol. 7575 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 57–70. 3

[BM12b] BARRON J., MALIK J.: Shape, albedo, and illumination from a single image of an unknown object. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* (June 2012), pp. 334–341. 3

[BM15] BARRON J., MALIK J.: Shape, illumination, and reflectance from shading. *IEEE Trans. on Pattern Analysis and Machine Intelligence 37*, 8 (Aug 2015), 1670–1687. 3

[BPD09] BOUSSEAU A., PARIS S., DURAND F.: User-assisted intrinsic images. *ACM Trans. on Graphics 28*, 5 (December 2009), 130:1–130:10. 2

[BT78] BARROW H., TENENBAUM J.: Recovering intrinsic scene characteristics from images. In *Computer Vision Systems* (1978), pp. 3–26. 2

[FDL04] FINLAYSON G., DREW M., LU C.: Intrinsic images by entropy minimization. In *Proc. of the European Conf. on Computer Vision*, vol. 3023 of *Lecture Notes in Computer Science*. 2004, pp. 582–595. 2

[FF07] FARENZENA M., FUSIELLO A.: Recovering intrinsic images using an illumination invariant image. In *Proc. of the IEEE International Conf. on Image Processing* (2007), vol. 3. 2

[FH01] FINLAYSON G. D., HORDLEY S. D.: Color constancy at a pixel. *Journal of the Optical Society of America A 18*, 2 (Feb 2001), 253–264. 2

[FHD02a] FINLAYSON G. D., HORDLEY S. D., DREW M.: Removing shadows from images using retinex. *Color and Imaging Conf.*, 1 (2002), 73–79. 2

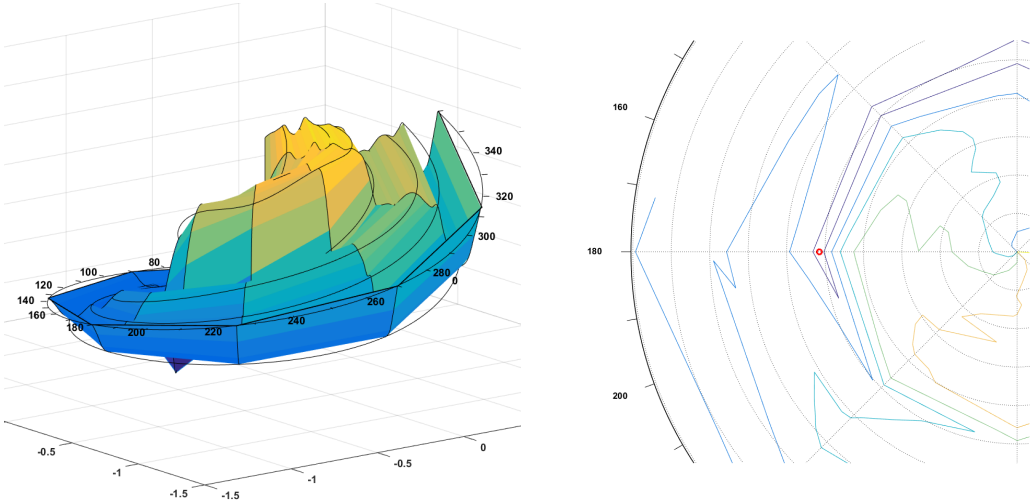

Figure 9: *Plot of complexity values calculated as the mean of gradient magnitudes of tentative reflectance image, and relative isolines, for the scene of Fig. 8. The red circle indicates the position of the ground truth.*

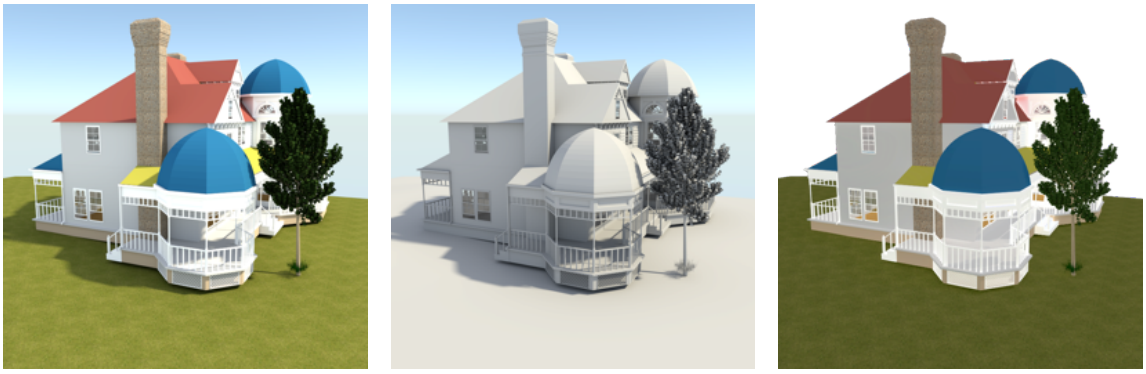

Figure 10: *Left to right: original image, representing a house, its shading and reflectance components, as retrieved by our algorithm.*

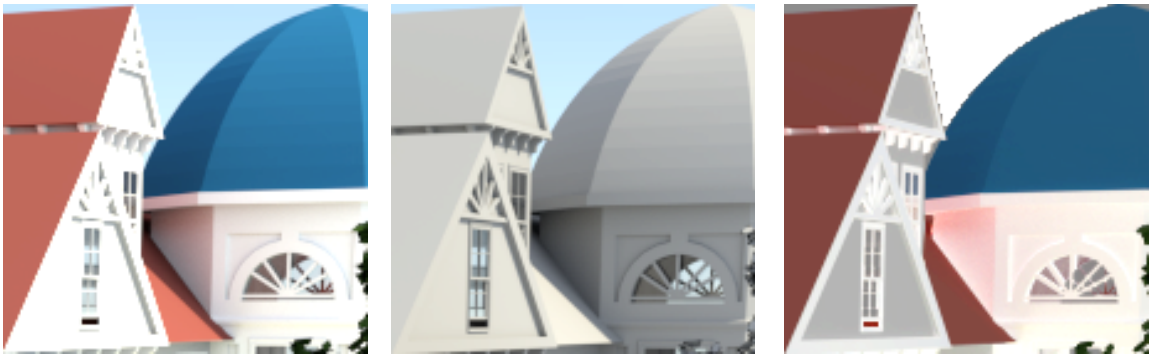

Figure 11: *A crop of Figure 10, where an artifact due to global illumination is noticeable.*

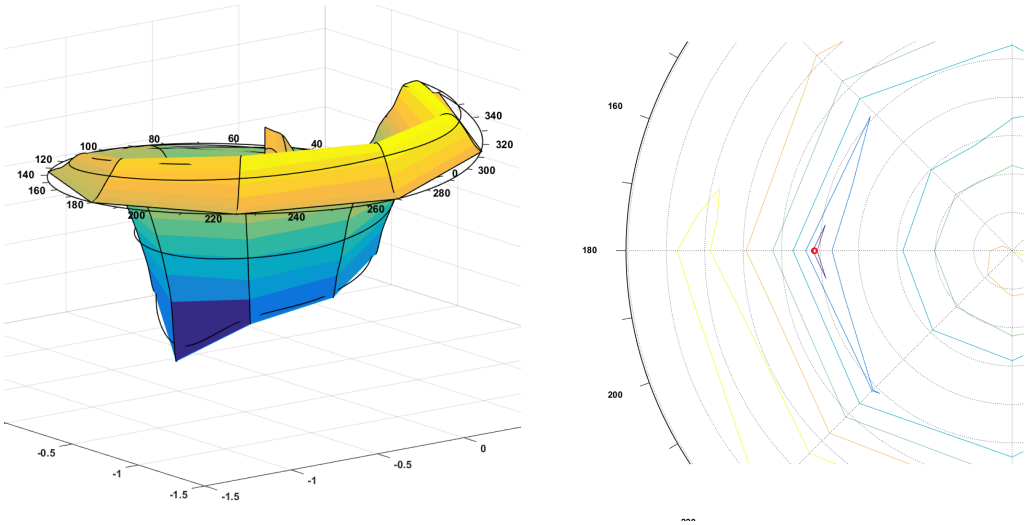

Figure 12: *Plot of complexity values calculated as the mean of gradient magnitudes of tentative reflectance image, and relative isolines, for the scene of Fig. 10. The red circle indicates the position of the ground truth.*

[FHD02b] FINLAYSON G. D., HORDLEY S. D., DREW M. S.: Removing shadows from images. In *Proc. of the European Conf. on Computer Vision* (2002), Springer-Verlag, pp. 823–836. 2

[GJAF09] GROSSE R., JOHNSON M. K., ADELSON E. H., FREEMAN W. T.: Ground-truth dataset and baseline evaluations for intrinsic image algorithms. In *Proc. of the International Conf. on Computer Vision* (2009), pp. 2335–2342. 3

[GMLMG12] GARCES E., MUNOZ A., LOPEZ-MORENO J., GUTIERREZ D.: Intrinsic images by clustering. *Computer Graphics Forum (Proc. EGSR 2012) 31*, 4 (2012). 3

[LBP*12] LAFFONT P.-Y., BOUSSEAU A., PARIS S., DURAND F., DRETTAKIS G.: Coherent intrinsic images from photo collections. *ACM Trans. on Graphics 31*, 6 (November 2012), 202:1–202:11. 3

[MLKS04] MATSUSHITA Y., LIN S., KANG S. B., SHUM H.-Y.: Estimating intrinsic images from image sequences with biased illumination. In *Proc. of the European Conf. on Computer Vision* (May 2004), pp. 274–286. 3

[MNIS03] MATSUSHITA Y., NISHINO K., IKEUCHI K., SAKAUCHI M.: Illumination normalization with time-dependent intrinsic images for video surveillance. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* (2003), vol. 1. 3

[RKZ*11] ROTHER C., KIEFEL M., ZHANG L., SCHÃŰLKOPF B., GEHLER P. V.: Recovering intrinsic images with a global sparsity prior on reflectance. In *Advances in Neural Information Processing Systems* (2011), pp. 765–773. 3

[SA93] SINHA P., ADELSON E.: Recovering reflectance and illumination in a world of painted polyhedra. In *Proc. of the International Conf. on Computer Vision* (1993), pp. 156–163. 2

[SL03] STAINVAS I., LOWE D.: A generative model for separating illumination and reflectance from images. *Journal of Machine Learning Research 4* (December 2003), 1499–1519. 3

[SLX09] SHI B., LI Y., XU C.: Intrinsic image decomposition using color invariant edge. In *Fifth International Conf. on Image and Graphics* (Sept 2009), pp. 307–312. 2

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

[SY11] SHEN L., YEO C.: Intrinsic images decomposition using a local and global sparse representation of reflectance. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* (June 2011), pp. 697–704. 3

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

[TAF06] TAPPEN M., ADELSON E., FREEMAN W.: Estimating intrinsic component images using non-linear regression. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* (2006), vol. 2, pp. 1992–1999. 2

[TFA05] TAPPEN M., FREEMAN W., ADELSON E.: Recovering intrinsic images from a single image. *IEEE Trans. on Pattern Analysis and Machine Intelligence 27*, 9 (2005), 1459–1472. 2

[TFG*13] TOLDO R., FANTINI F., GIONA L., FANTONI S., FUSIELLO A.: Accurate multiview stereo reconstruction with fast visibility integration and tight disparity bounding. In *Proceedings of the Workshop: 3D Virtual Reconstruction and Visualization of Complex Architectures (3D-ARCH)* (Trento, Italy, 2013), vol. XL-5/W1, pp. 243–249. 2

[TGFF15] TOLDO R., GHERARDI R., FARENZENA M., FUSIELLO A.: Hierarchical structure-and-motion recovery from uncalibrated images. *Computer Vision and Image Understanding 140* (November 2015), 127 – 143. 2

[Wei01] WEISS Y.: Deriving intrinsic images from image sequences. In *Proc. of the International Conf. on Computer Vision* (2001), pp. 68–75 vol.2. 3

[YGL*14] YE G., GARCES E., LIU Y., DAI Q., GUTIERREZ D.: Intrinsic video and applications. *ACM Trans. on Graphics 33*, 4 (July 2014), 80:1–80:11. 3

[YW13] YU H., WINKLER S.: Image complexity and spatial information. In *Fifth International Workshop on Quality of Multimedia Experience* (July 2013), pp. 12–17. 3