Thirteenth Eurographics Workshop on Rendering (2002)
P. Debevec and S. Gibson (Editors)

Accurate Image Based Relighting through Optimization

Pieter Peers   Philip Dutré[†]

Department of Computer Science, K.U.Leuven, Belgium

Abstract

In this paper we present a new relighting technique that, for a single viewpoint, accurately captures the reflectance field of real and virtual objects, without restrictions on their geometrical complexity or material properties. As a result the objects can be relit under arbitrary lighting conditions. To capture the reflectance field, we take photographs of an object lit by several lighting patterns on a surrounding discretized hemicube. The illumination of each pixel due to emitted radiance from each discrete patch on the hemicube is approximated by a reflection coefficient and a rectangular support which are found through an optimization procedure. The discretization of the hemicube ensures sufficient angular sampling to capture diffuse material properties. The use of a sub-patch support accommodates small solid angles of incoming light important for specular materials. To relight the object, the target illumination, a high-dynamic range environment map, is averaged over the support and multiplied by the reflection coefficient, per pixel and per discrete patch on the hemicube. The results obtained show accurate relighting of diffuse, glossy and specular objects.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three dimensional graphics and realism I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture

1. Introduction

Image-based relighting represents a class of techniques that apply new lighting conditions to a scene, given a set of basis images. Possible applications range from lighting design to augmented reality. Lighting design can be a tedious task, since moving objects and light sources is not always feasible. Objects can be very fragile (e.g. archaeological artifacts), the illumination is not controllable (e.g. the sun through a sky-window) or an object is too massive to move around (e.g. a statue). Augmented reality applications are constantly evolving and pushing the limits of widely accepted techniques. Placing real objects in virtual environments influences the appearance of the objects. It is not always possible to approximate the desired illumination at the moment an object is captured (e.g. special effects in movies or computer games which combine real and virtual environments). Relighting is an important part of computer graphics applicable to a large range of application.

In this paper we develop a relighting technique that combines strengths from several techniques, the Light Stage[2, 5] and environment matting[1, 14], into a single framework. This allows us to capture complex geometrical objects without placing restrictions on material properties.

To capture the reflectance field a hemicube around the object is discretized in several light patches. Each patch emits a series of illumination patterns and for each pattern a high-dynamic range photograph[3] is recorded. Currently we use a calibrated CRT monitor per hemicube side to emit the illumination patterns (figure 1). *For each pixel* and *for each light patch* the received pixel radiance is then approximated by a reflection coefficient and a support area on the light patch. This support area approximates the area on the light patch that is important for the pixel under consideration. A least square minimization procedure on the pixel radiance over different illumination patterns is used to determine the best approximation for the support. To relight the scene with an arbitrary high-dynamic range light map, the average illumination emitted over the support area by the light map is computed and multiplied by the reflectance coefficient. This is repeated for each pixel and each light patch. The results show that the technique is able to accurately relight complex geometry with diffuse, glossy or specular material properties.

† {pieterp, phil}@cs.kuleuven.ac.be

The paper is structured as follows: In the next section we discuss previous work. In section 3, a general overview of the technique is given, and is presented as a three step algorithm in sections 4, 5 and 6. The accuracy is verified in section 7 and the relation with the Light Stage and environment matting is further explored in section 8. Some results obtained with our relighting technique are found in section 9. Finally directions for future work and a conclusion is given in section 10.

2. Previous Work

In this section we give an overview of work related to image based relighting.

A first class of techniques requires or reconstructs a geometrical model of the scene. Loscos et al.[8, 9] reconstruct a simplified geometrical model, approximate the illumination and calculate unoccluded illumination textures with an adapted radiosity algorithm. Relighting is performed by modifying light intensities. Yu et al.[12, 13] focus mainly on recovering the reflectance properties of materials by using an iterative optimization procedure to find parameters to represent the reflectance properties of surfaces. The scene with the recovered material properties is re-rendered using a global illumination rendering system. Reconstruction of geometry is advantageous when visualization from multiple viewpoints is desired, but is a disadvantage when a scene consists out of geometrical complex objects.

Another class of techniques circumvents the need for complex geometrical models by directly manipulating images. Nimeroff et al.[10] introduced the concept of linearly weighting and combining basis images of a scene. The lighting condition in these basis images are chosen in such a way that any desired novel lighting condition can be constructed by linearly combining and weighting these basis images. Wong et al.[11] based their relighting method on light field rendering[4, 6]. By sampling the scene under different viewpoints and illumination an apparent BRDF for each surface pixel is found. Lin et al.[7] define a reflected irradiance field, the dual of a light field, and derive upper sampling bounds for each BRDF to reproduce it truthfully.

The Light Stage[2, 5] uses basis images of a scene lit by lights regularly spaced on a sphere surrounding the scene. A light map is applied by linearly combining these basis images into a single resulting image.

Environment matting, originally introduced by Zongker et al.[14] and extended by Chuang et al.[1], capture the reflectance properties of specular and refractive materials by lighting them with different illumination patterns. The direction and solid angle important for the illumination of a pixel can be derived from this information in the form of a reflection coefficient and a support on the back or side-drops.

Our technique combines the strength and features of the Light Stage and of environment matting. Similar to the Light Stage incoming light directions are sufficiently discretized to approximate the reflectance properties of diffuse materials, where as the use of illumination patterns, similar to environment matting, makes it possible to find small solid angles important to approximate specular reflectance properties.

3. Overview of the method

The goal of our relighting method is to apply a new light map (in the form of an environment map) to an existing scene, photographed from a fixed viewpoint. To do this, we need to know, for each direction in the light map, how much it contributes to the radiance seen through a pixel. To find these contributions, the scene can be lit by light sources located around the object on a bounding volume and the contribution to the illumination through a pixel can be recorded. We can now formulate the goal of our relighting method: for each pixel, the goal is to compute the weight (contribution to the total radiance seen through the pixel) of each point on the bounding volume for an existing scene. With this it is possible to relight the image using any environment map. The bounding volume is discretized in a number of light patches that can be lit independently of other light patches.

Consider for a moment that only direct illumination emitted from a patch contributes to the radiance seen through a pixel. If the reflection to the eye is according to a diffuse BRDF, then all direction in the patch will be equally significant. If on the other hand the reflection to the eye is according to a specular BRDF, then only a few directions are significant (located around the perfect reflected direction). The set points on the light patch that are significant for the contribution to the pixel radiance are called the support of that light patch.

In this paper we model the reflectance for each pixel and each light patch by a reflection coefficient and a support. The refection coefficient represents the global influence of the light patch to the observed pixel radiance, where as the support represents the subset of important directions within a light patch. This can be written as:

$$I(p) = \sum_k R_{p,k} \times < S_{p,k}, L_k >$$

The intensity I in pixel p equals the sum over all light patches k of the reflection coefficient $R_{p,k}$ multiplied by the average of the incoming radiance L_k over the support $S_{p,k}$.

To relight a pixel, the illumination values in the light map need to be averaged over the support. We can generalize this averaging as a filter operation. We call this filtering the 'evaluation' of the support. We used a box filter in the text above, but in principle any filter can be used.

We will now develop a method to capture, compute and use these sets of reflection coefficients and supports to relight each pixel. First a number of basis images are recorded (section 4). From this gathered input data the reflection coefficients and supports are extracted using a least square minimization algorithm (section 5) and stored for relighting afterwards. This relighting (section 6) can be done an arbitrary number of times with different light maps and requires minimal processing time.

4. Basis Image recording

4.1. General overview

We measure the reflectance fields as follows: a hemicube is assumed around the object for which the reflectance field will be captured. Each side of the hemicube is subdivided in a regular grid of connecting light patches. From each of these light patches a series of illumination patterns are emitted (while emitting no radiance from other light patches) and a high-dynamic range photograph[3] is recorded. If n illumination patterns are emitted from each light patch and there are m light patches then a total of $n \times m$ high-dynamic range photographs will be recorded.

4.2. Practical setup

In our experiments we used a CRT monitor for each side of the hemicube (figure 1).

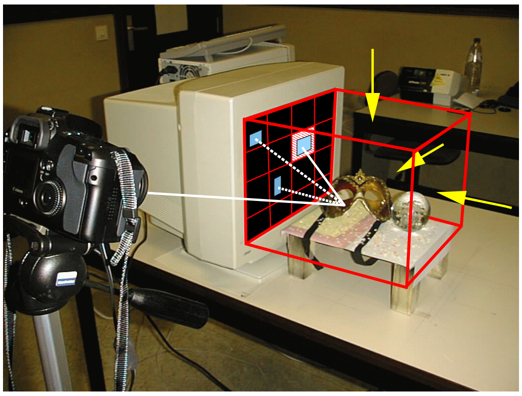

Figure 1: *A CRT monitor is used for each side of the hemicube. A 4×4 grid of light patches emits a series of illumination patterns. For each light patch a reflection coefficient and support are computed. Our current setup limits monitor placement to 4 sides, excluding the side where the camera is located.*

The CRT monitor needs to be radiance calibrated, because the mapping between pixel values and emitted radiance is a non-linear function for most CRT monitors. To calibrate the CRT monitor, a high-dynamic range photograph is recorded of the CRT monitor displaying a calibration pattern. All our illumination patterns are transformed using the non-linear mapping between radiance and pixel values.

In our experiments we use a 4×4 grid on each hemicube side. This choice is motivated by the fact that it produces good results for a reasonable amount of recorded photographs. Due to practical limitations the side facing the camera was omitted.

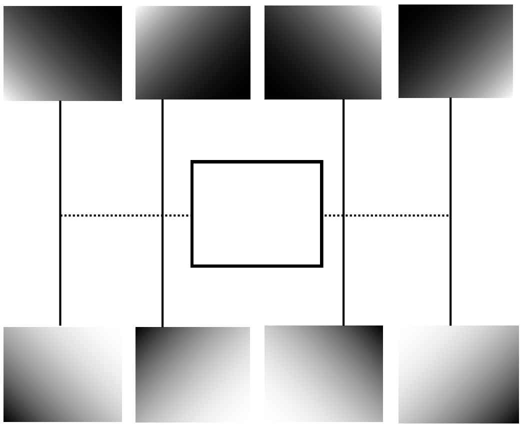

Figure 2: *The illumination patterns used to record the basis images. The solid white pattern is displayed in the center, while inverse patterns are displayed on opposite sides.*

4.3. Illumination Patterns

For each light patch a reflection coefficient and support needs to be calculated for a pixel. The illumination patterns need to provide enough information to extract this data.

The choice of which illumination patterns depends on the choice of the specific filter used and on the representation of the supports. In this paper use an axis-aligned support and a box filter.

It is possible to calculate the reflection coefficient using just one illumination pattern. The reflection coefficient is defined in section 3 as the global influence of a light patch to the pixel radiance. By emitting a solid white pattern, the reflection coefficient can be directly computed from the observed pixel radiance.

There are many kinds of illumination patterns that can be used to compute a support. In environment matting were three different kinds of illumination patterns used: horizontal and vertical stripes (combined with a box filter), Gaussian bands sweeping over the screen at 0, 45, 90 and 135 degrees (combined with a elliptical Gaussian filter) and a single colored gradient (restricted to colorless materials).

We used a variation on gradients, and use 4 quadratic (gray-scale) gradients oriented at $45 + (n \times 90)$ degrees with $n = [0..3]$ (figure 2).

Since dark regions in a high-dynamic range photograph are prone to noise, we also used 4 inverse patterns, because the inverse patterns are bright where the normal patterns are dark and vice versa. This way we are sure that the effects of noise are minimized.

A detailed analysis of the choice of supports is beyond the scope of this paper.

5. Calculation of the reflection coefficients and supports

Once the basis images are recorded, the reflectance coefficient and its support is found for each pixel and for each light patch through an optimization procedure. A pixel consists out of a number of channels (e.g. RGB) and two approaches are possible. The first possibility is to treat each channel independently and calculate a support and reflection coefficient for each channel. The second possibility is to use the same support for each channel and treat the reflection coefficient as a vector of reflection coefficients for each channel. We used the latter approach.

As said, the reflection coefficient can be directly derived from the basis image lit with the solid white pattern. The evaluation of any support over a solid white illumination pattern equals a constant factor. The observed pixel value lit by the solid white illumination pattern equals the reflection coefficient multiplied by the constant factor.

To find the support within a light patch, a least square optimization is used. We use a steepest decent optimization algorithm (algorithm 1).

```
Algorithm 1:

1) Select an initial support
2) Calculate the least-squared error
3) while not minimized
   a) Test all expansion possibilities
             of the support
   b) Calculate least-squared
             errors and select minimum error
   c) if minimum error < error then
             update support
      else minimized = true
```

There are a number of possible initial supports. In our implementation we use the following: if the reflection coefficient of the previous pixel is almost equal as the reflection coefficient of the current pixel then the initial support is set equal to the support of the previous pixel. If the reflection coefficients differ too much then the initial support is set equal to the entire light patch. The idea is that if the reflection coefficients are almost equal then it is likely that the material properties of these pixels are almost the same and thus have a simular behavior with respect to supports.

To compute the least square error in each optimization step, a trial support is evaluated for each illumination pattern and scaled by the known reflection coefficient. The computed value will differ from the observed pixel value in the basis image lit by the same illumination pattern. This error

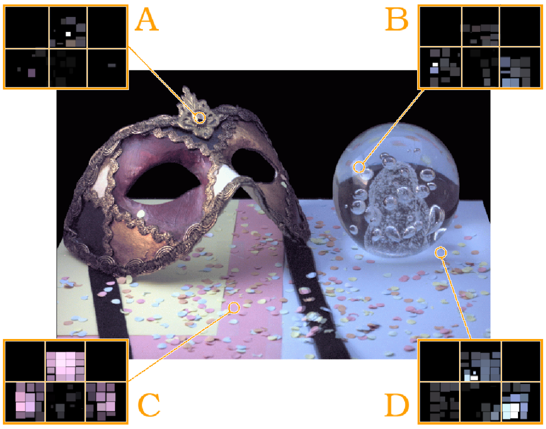

Figure 3: *The reflection coefficients and supports of 4 different pixels. The size of each square represents the support and the intensity the reflection coefficient (a non-linear mapping was applied to the intensities). The upside down T represents the unfolded hemicube (top; left, back, right).*

is squared and summed for the trail support over all illumination patterns.

$$SquaredError(S) = \sum_j (R_p \times < P_j, S > - I(p, P_j))^2$$

The summed squared error is calculated over the illumination patterns P_j for a trial support S scaled by a reflection coefficient R_p. $I(p, P_j)$ denotes the observed pixel p lit by the illumination pattern P_k.

The use of a box filter and an axis-aligned support makes it possible to use summed area tables on the illumination patterns which significantly speed up calculation of the squared error.

In figure 3 the reflection coefficients and supports are plotted for a number of different pixels. The size of the squares represents the size of the different supports and the intensity represents the magnitude of the reflection coefficients. Pixel A represents a specular surface. The bright white support is around the prefect reflected direction. The other supports visible for this pixel are less then 0.5 percent in intensity and are due to noise in the recorded basis images. Pixel B represents a transparent surface. One can clearly see the two distinct sets of supports representing the influence from the reflection (left) and refraction (right). Pixel C represents a diffuse surface. All supports are almost maximal in size. The reason that there are almost no supports visible for the backside is due to occlusion by the mask. Finally, pixel D represents a diffuse surface at the base of the glass sphere. The reader can clearly see the effects of direct (upper hemicube side) and indirect illumination (right hemicube side) due to refraction through the glass sphere in the different supports.

6. Relighting

Once the reflection coefficients and supports are calculated for each pixel and each light patch it is easy to apply a new light map to the scene. To relight the scene using an arbitrary high-dynamic range environment map, the support is evaluated over the new illumination and scaled by the reflection coefficient, for each light patch and for each pixel. The results from the different supports are summed for each pixel, resulting in the total observed radiance for that pixel. In case a box filter was used over the different supports a significant speedup can be obtained by using summed area tables for the light-map, because the supports differ for each pixel and thus for each pixel, the illumination pattern has to be evaluated over the support.

7. Verification

Different degrees of accuracy are possible with the proposed technique. The discretization rate of the hemicube and the number of illumination patterns directly influence the accuracy. In this section we verify the accuracy. We use an artificial scene containing 4 spheres, each with different material properties, placed on a diffuse surface (figure 4) and lit by a light map placed at the righthand side of the scene. A reference image of this scene was rendered using a global illumination renderer (figure 4, left picture). The basis images lit with the illumination patterns were also rendered using the same global illumination renderer. The reflection coefficients and supports were calculated with our method using these basis images. Finally the same light map that was used in the reference image was used in the relighting step of our technique (figure 4, right image). No difference is visually noticeable, soft shadows are reproduced accurately (figure 4 zoom-ins on the lefthand side) and glossy/specular reflection are reproduced without visual difference (figure 4 zoom-ins on the righthand side). The average relative error over all pixel confirms the visual verification and was less than 1 percent. The verification shows that our method can accurately relight diffuse, glossy and specular materials. Soft shadows are also reproduced faithfully.

8. Discussion

In this section we discuss the implications of our developed method. One can see that the proposed technique is a superset of methods like the Light Stage and environment matting. The discretization of the incoming light directions directly maps to the regular sampling on a sphere surrounding the scene by the Light Stage. The Light Stage is good in reproducing diffuse material properties, but fails for specular materials because the coarse angular sampling is not sufficient to capture the high frequency spikes of specular reflection functions. Illumination patterns to determine a support are also used in environment matting. This method is very good in reproducing specular and transparent material properties, because it is able to find a single small solid angle important for the reflection function of specular materials. Environment matting however fails for diffuse materials because the support would cover the entire hemisphere and a single reflection coefficient (combined with a simple filter) is unable to approximate the reflection function sufficiently.

A straightforward combination of the Light Stage and environment matting (proposed by Debevec et al.[2]), where a two step data acquisition stage first preforms a step simular to the Light Stage and then as a second step preforms an environment matting step, would only be able to correctly handle diffuse and specular materials. Glossy materials would still be difficult to capture because their angular frequency could be too small for the Light Stage but still be too large for environment matting to represent it by a single reflection coefficient and a support (and a simple filter). Using more complex filters (e.g. oriented Gaussian filters as proposed by Chuang et al.[1]) would minimize this problem, but not completely solve it. Certain materials are almost impossible to represent by a single reflection coefficient and a filter.

The proposed method differs from the straightforward combination due fact that for each discrete part of incoming light directions a support and a reflection coefficient are calculated. Glossy materials can still be approximated with simple filters and with sufficient accuracy because they still can span several supports and reflection coefficients, but aren't bound by the discretization rate of the incoming light directions. Using more complex filters (e.g. elliptical Gaussian filters) would allow even more accurate approximations at the cost of extra illumination patterns. The fact that everything is on a per-pixel-basis and that a different support is possible per discrete set of incoming light directions makes it possible to use a much coarser subdivision of the incoming light directions.

The proposed method combines the strengths of the Light Stage and environment matting in an unified framework. The results in section 9 show that the method is indeed able to capture a wide range of materials even with a limited number of photographs.

9. Results

In this section two results (figure 5 and figure 6) are discussed. Both images show a wooden toy car placed on a diffuse surface.

Figure 5 is relit using a light map recorded inside the Uffizi Galleria, where most of the light comes from directly above. The zoom-in on the bumper shows the elongated glossy highlight caused by the left to right orientation of the illumination in the light map. The zoom-in on the back of the toy car clearly show the relatively hard shadows caused by the illumination from above.

Figure 6 is relit using an artificial light map, where each side of the hemicube emits a different color. One can clearly distinguish the four sides of the hemicube in the reflection

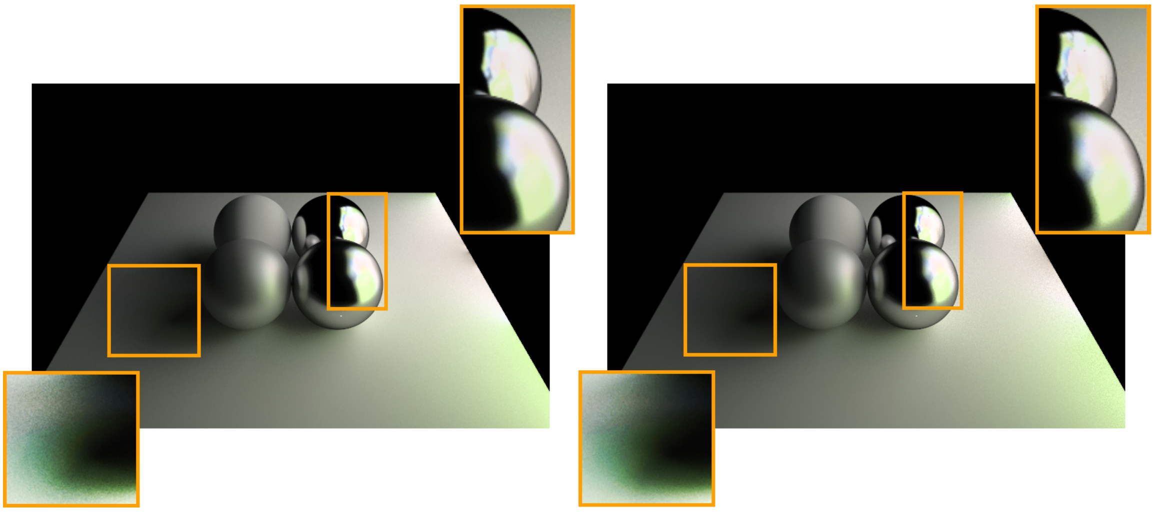

Figure 4: *A comparison of a rendered (left) and relit image (right). Both use the same light map placed at the righthand side of the scene.*

in the glossy bumper and the zoom-in on the back of the car shows almost no shadow because the radiance emitted from each side of the light map is approximately the same.

To capture the basis images for these scenes we used a 19” Iiyama CRT monitor. We moved this monitor around for each side of the hemicube. By using only one monitor we avoid subtle differences between different monitors. The photographs were recorded using a Canon EOS D30 camera. We automated the process of capturing and, as mentioned before, only the monitor had to moved around. The recording time was about 10 hours (of which at least 6 hours were due to overhead of the slow speed of the camera). Recording could be sped up by using a digital video camera and using different F-stop filters to acquire high-dynamic range images.

The processing took 20 hours on an Athlon 1.2Ghz system with 512meg ram. This comes down to approximately 1 pixel per second per 200Mhz of cpu power. Since each pixel can be processed independently, it is possible to subdivide the processing time among different computers without loss in performance.

The relighting can be done in less than a minute and could be done in real-time if all reflection coefficients and supports would fit in the main memory ($16 \times 4 \times NrPixels$ supports).

When high-dynamic range environment maps are of much larger dynamic range than the range of the monitor, visual artifacts appear. Neighboring pixels can have relatively different reflection coefficients due to noise in the recorded high-dynamic range photographs, combine this with large intensities from the environment map for a certain support, and large differences become visible for neighboring pixels. This problem could be reduced by applying noise reduction filters as preprocessing on each recorded basis image.

10. Conclusions

In this paper we developed a new and practical technique for capturing the reflectance field of objects for a fixed viewpoint, but without placing restrictions on material properties or geometrical complexity. The method scales very well in terms of accuracy versus number of basis images. Higher accuracy can be obtained by using more complex filters and/or a finer discretization. The practical implementation can be done with off the shelf materials. Diffuse, glossy and specular materials are reproduced faithfully, as are more complex illumination effects (such as caustics).

A whole new range of applications is possible: the digital capture and display of objects with a mix of diffuse and specular appearance (e.g. jewelry or sensitive archaeological objects), augmented reality applications, (interactive) relighting of fully virtual scenes, ...

Future research will focus on faster and easier capturing of the reflectance field, the use of other projection devices such as VR-caves that can accommodate larger objects, reducing the number necessary photographs by using better patterns (e.g. oriented box filters), and better optimization procedures.

More results in color can be found at the dedicated web page: http://www.cs.kuleuven.ac.be/∼graphics/CGRG.PUBLICATIONS/ARTO/

Acknowledgments

We would like to thank Frank Suykens, Vincent Masselus and Karl vom Berge for proof-reading the paper and many constructive discussions. Furthermore we would like to thank Paul Debevec for his high quality environment maps (http://www.debevec.org/Probes). All artificial scenes were rendered using RenderPark (http://www.renderpark.be).

References

1. Yung-Yu Chuang, Douglas E. Zongker, Joel Hindorff, Brian Curless, David H. Salesin, and Richard Szeliski. Environment matting extensions: Towards higher accuracy and real-time capture. In Kurt Akeley, editor, *Siggraph 2000, Computer Graphics Proceedings*, Annual Conference Series, Los Angeles, 2000. ACM Siggraph, Addison Wesley Longman.

2. Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In Kurt Akeley, editor, *Siggraph 2000, Computer Graphics Proceedings*, Annual Conference Series, pages 145–156, Los Angeles, 2000. ACM Siggraph, Addison Wesley Longman.

3. Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In Turner Whitted, editor, *Siggraph 97 Conference Proceedings*, Annual Conference Series, pages 369–378. ACM Siggraph, Addison Wesley, August 1997.

4. Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. In Holly Rushmeier, editor, *Siggraph 96 Conference Proceedings*, Annual Conference Series, pages 43–54. ACM Siggraph, Addison Wesley, August 1996.

5. Tim Hawkins, Jonathan Cohen, and Paul Debevec. A photometric approach to digitizing cultural artifacts. In *In 2nd International Symposium on Virtual Reality, Archaeology, and Cultural Heritage, Glyfada, Greece, November 2001*.

6. Marc Levoy and Pat Hanrahan. Light field rendering. In Holly Rushmeier, editor, *Siggraph 96 Conference Proceedings*, Annual Conference Series, pages 31–42. ACM Siggraph, Addison Wesley, August 1996.

7. Zhouchen Lin, Tien-Tsin Wong, and Heung Yeung Shum. Relighting with the reflected irradiance field: Representation, sampling and reconstruction. In *Proceedings of IEEE Computer Vision and Pattern Recognition*, December 2001.

8. Céline Loscos and George Drettakis. Low-cost photometric calibration for interactive relighting. In *Proceedings of the First French-British International Workshop on Virtual Reality*, Brest, France, July 2000.

9. Céline Loscos, George Drettakis, and Luc Robert. Interactive virtual relighting of real scenes. In *IEEE Transactions on Visualization and Computer Graphics*, volume 6(4), pages 289–305. IEEE Computer Society, 2000.

10. Jeffry S. Nimeroff, Eero Simoncelli, and Julie Dorsey. Efficient Re-rendering of Naturally Illuminated Environments. In *Fifth Eurographics Workshop on Rendering*, pages 359–373, Darmstadt, Germany, June 1994.

11. Tien-Tsin Wong, Pheng-Ann Heng, Siu-Hang Or, and Wai-Yin Ng. Image-based rendering with controllable illumination. In Julie Dorsey and Philipp Slusallek, editors, *Eurographics Rendering Workshop 1997*, pages 13–22, New York City, NY, June 1997. Eurographics, Springer Wien.

12. Yizhou Yu, Paul Debevec, Jitendra Malik, and Tim Hawkins. Inverse global illumination: Recovering reflectance models of real scenes from photographs from. In Alyn Rockwood, editor, *Siggraph 99, Annual Conference Series*, Annual Conference Series, pages 215–224, Los Angeles, 1999. ACM Siggraph, Addison Wesley Longman.

13. Yizhou Yu and Jitendra Malik. Recovering photometric properties of architectural scenes from photographs. In Michael Cohen, editor, *Siggraph 98 Conference Proceedings*, Annual Conference Series, pages 207–218. ACM Siggraph, Addison Wesley, July 1998.

14. Douglas E. Zongker, Dawn M. Werner, Brian Curless, and David H. Salesin. Environment matting and compositing. In Alyn Rockwood, editor, *Siggraph 1999, Computer Graphics Proceedings*, Annual Conference Series, pages 205–214, Los Angeles, 1999. ACM Siggraph, Addison Wesley Longman.

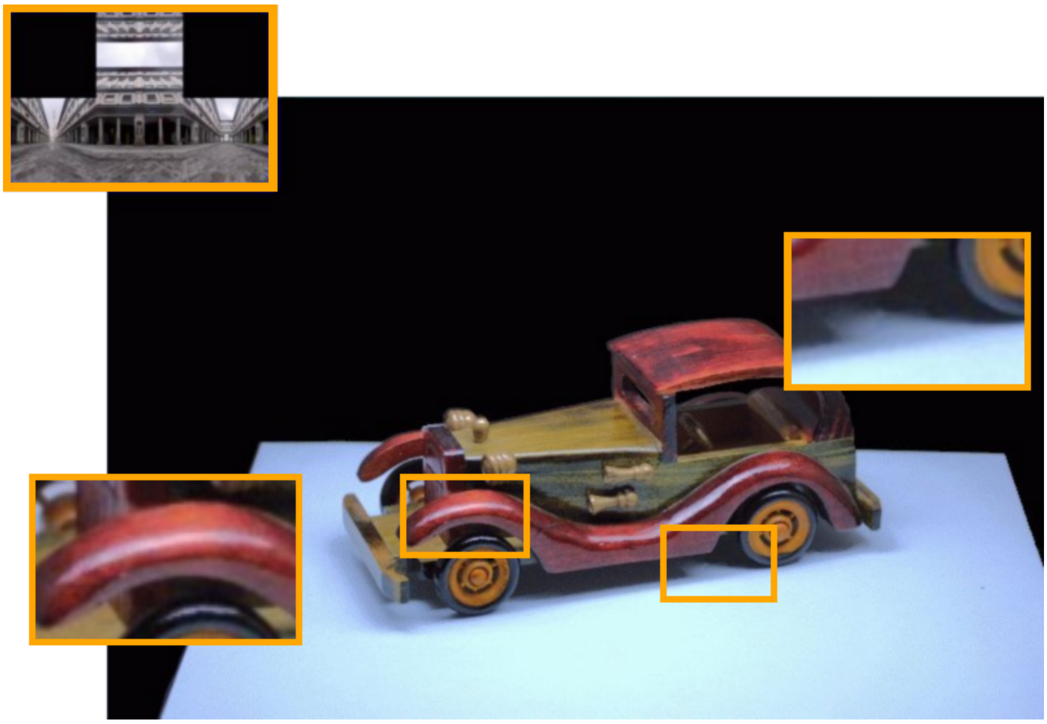

Figure 5: *A wooden toy car placed on a diffuse surface, relit with a light map recorded in the Uffizi Galleria. Note the glossy highlight and the soft shadows.*

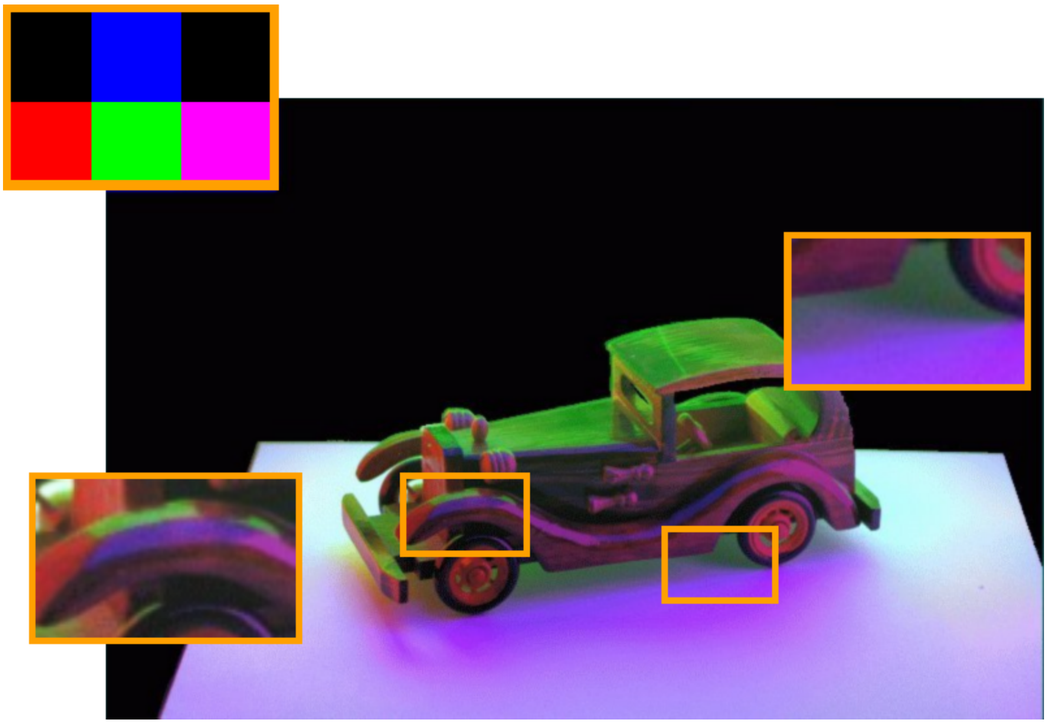

Figure 6: *A wooden toy car placed on a diffuse surface, relit with an artifical light map.*