

Illumination Decomposition for Material Recoloring with Consistent Interreflections

Robert Carroll

Ravi Ramamoorthi
University of California, Berkeley

Maneesh Agrawala

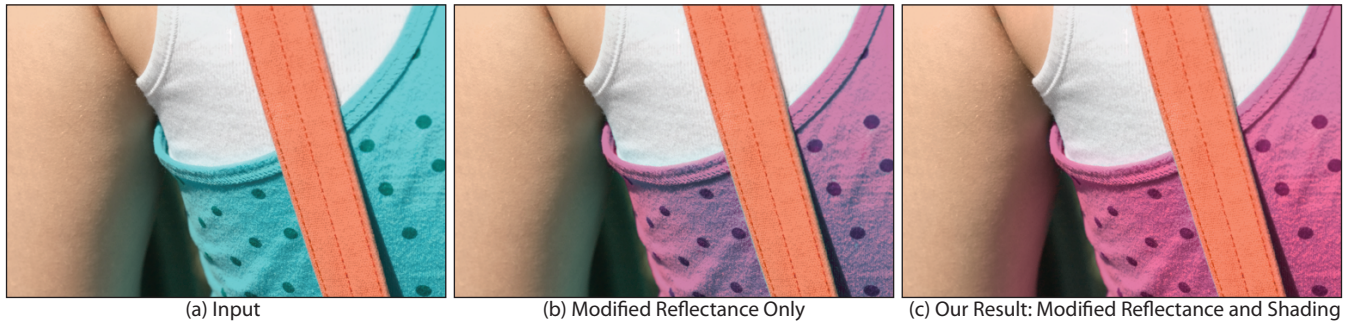


Figure 1: We seek to recolor the input image (a). However, changing the color (reflectance) of the shirt alone, without modifying the illumination, does not account for the correct diffuse reflection on the girl's arm or interreflections in the fine texture of the shirt (b). Indeed, the image in (b) still has bluish reflections on the arm and a purple color shift on the shirt. Our user-assisted decomposition (Figure 2) lets us modify indirect illumination to match the modified shirt color (c), leading to a much more consistent and natural looking recoloring.

Abstract

Changing the color of an object is a basic image editing operation, but a high quality result must also preserve natural shading. A common approach is to first compute reflectance and illumination intrinsic images. Reflectances can then be edited independently, and recomposed with the illumination. However, manipulating only the reflectance color does not account for diffuse interreflections, and can result in inconsistent shading in the edited image. We propose an approach for further decomposing illumination into direct lighting, and indirect diffuse illumination from each material. This decomposition allows us to change indirect illumination from an individual material independently, so it matches the modified reflectance color. To address the underconstrained problem of decomposing illumination into multiple components, we take advantage of its smooth nature, as well as user-provided constraints. We demonstrate our approach on a number of examples, where we consistently edit material colors and the associated interreflections.

Links: DL PDF WEB

1 Introduction

Adjusting the color of an object is a common photo editing task. Yet, the color we see at each pixel is the result of complex interactions between the lighting and the reflectance of materials in the scene. A promising approach for recoloring objects is to first estimate *intrinsic images* [Barrow and Tenenbaum 1978] that separate each image pixel into a component due to illumination or shad-

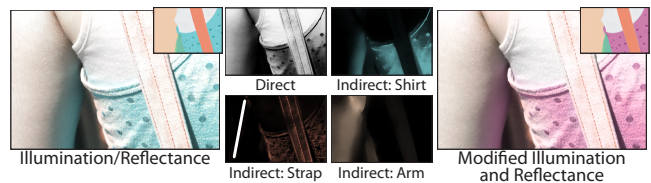


Figure 2: Our technique takes an illumination/reflectance intrinsic image pair and further factors the illumination into contributions due to direct lighting and indirect lighting from various material colors. The user can add strokes to the decomposition to locally remove the contribution of individual sources. In this example, the user added a long stroke across the arm to prevent attributing indirect illumination to the strap. With the decomposition we can individually modify the indirect illumination components to match the modified reflectance colors.

ing and a component due to reflectance. Users can then modify the reflected color of an object independently from the shading and recombine the two to produce the recolored image [Weiss 2001; Bousseau et al. 2009].

However, editing the reflectance image alone does not properly account for diffuse interreflections. Such interreflections are subtle, but visually important features of natural photographs. Photographs appear visually incorrect when the colors of the diffuse interreflections are inconsistent with the colors of the materials. In Figure 1b for example, we altered the color of the shirt from blue to pink, but the reflection on the arm remains blue. In addition, the shirt appears purple rather than pink because the blue interreflections caused by the fine texture of the shirt are unchanged.

In this paper we propose a user-assisted method for further separating the illumination image into direct and multiple indirect components (Figure 2). With this decomposition users can recolor materials, and our system updates the colors of the diffuse interreflections accordingly. Figure 1c shows that our approach is able to recolor the interreflections to match the modified color, leading to a more consistent and natural looking image. It would be very difficult to properly modify these interreflections using traditional image editing tools.

ACM Reference Format

Carroll, R., Ramamoorthi, R., Agrawala, M. 2011. Illumination Decomposition for Material Recoloring with Consistent Interreflections. *ACM Trans. Graph.* 30, 4, Article 43 (July 2011), 9 pages. DOI = 10.1145/1964921.1964938 <http://doi.acm.org/10.1145/1964921.1964938>.

Copyright Notice

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, fax +1 (212) 869-0481, or permissions@acm.org.
© 2011 ACM 0730-0301/2011/07-ART43 \$10.00 DOI 10.1145/1964921.1964938
<http://doi.acm.org/10.1145/1964921.1964938>



Figure 3: Our color editing system has 3 stages. We first compute intrinsic images using the approach of Bousseau et al., and sample the reflectance image to obtain basis sources. In the second stage we compute the illumination decomposition, with the user iteratively adding strokes to refine the result. In the final stage we modify the reflectance image and obtain a new set of basis sources, which is used with the decomposition to construct a new illumination image.

The illumination at any point in a scene is an aggregate of light that travels directly from the light sources and light that first bounces off other surfaces, possibly multiple times. When light bounces off a diffuse material it is modulated by that material’s color, resulting in a colored interreflection. Our approach to factoring the illumination is to model indirect illumination as a single bounce off a set of material colors, which we sample from the reflectance intrinsic image. We then decompose the illumination at each pixel into a linear combination of the light color and the light color modulated by each material color. Since there are only three color channels in the illumination image, finding the linear combination that explains the illumination is ill-posed whenever the scene contains more than two sources of indirect lighting (due to interreflections from two or more materials) in addition to the direct lighting.

The key idea of our approach is to address this ambiguity by taking advantage of the low-frequency variation of illumination on surfaces [Land et al. 1971; Ramamoorthi and Hanrahan 2001]. In particular we apply a smoothness prior to the decomposition that is based on the Total-Variation (TV) of individual illumination components. We demonstrate that TV is particularly well suited for our decomposition because it preserves edges and it promotes contributions from a sparse set of sources at each pixel (Section 4.6). We also allow users to mark strokes that locally constrain the contribution of each illumination component. The effects of these user-specified constraints are propagated through the decomposition via the smoothness prior.

Our system does rely on a few assumptions about the scene. It works best when illumination is locally a mix of a small set of independent colors, although we mitigate this limitation through user interaction and the smoothness prior (Figure 6). We assume there is only a single bounce of indirect illumination. In principle our approach could be extended to multiple bounces by considering the pairwise products of the material colors, but this would greatly increase the number of variables and make the separation more ambiguous, for little practical benefit. Similarly, we assume a single direct light color, although our model extends directly to multiple lights. Finally, as is common in many image editing techniques we assume the scene is Lambertian. Although we make simplifying assumptions about the scene, the goal of our system is not to exactly invert the image formation process, but to enable plausible editing of material colors and interreflections. Our decomposition is often sufficient for recoloring scenes that do not strictly adhere to these assumptions.

We demonstrate the applicability of our approach on a number of examples where modifying reflectance alone results in inconsistencies (figures 8, 9, 11, 12). We show that our decomposition is sufficient for plausibly manipulating indirect illumination. We also compare our decomposition to a computationally simpler alternative and show that it is superior (Section 4.6).

2 Related Work

Our illumination decomposition builds upon the intrinsic image decomposition [Barrow and Tenenbaum 1978]. Researchers have proposed a number of techniques for automatically computing intrinsic images from a single image [Horn 1986; Sinha and Adelson 1993; Tappen et al. 2005; Shen et al. 2008], from multiple images [Weiss 2001], and from a single image with user assistance [Bousseau et al. 2009]. While many of these approaches are designed to recover monochromatic illumination, our system requires colored illumination as input.

Factoring images into different components is a general problem that arises in a number of contexts, including matting [Chuang et al. 2001; Wang and Cohen 2007; Levin et al. 2008], transparency/reflection separation [Levin and Weiss 2007], and haze removal [Fattal 2008]. Hsu et al. [2008] estimate the relative contribution of two light sources at each point in the image in order to white balance them independently. Treating each material as a source of indirect illumination, we are left with a similar estimation problem. However, unlike all of these previous decomposition techniques which generate two components, we focus on factoring illumination into more than two components.

There are a number of techniques for automatically obtaining illumination decompositions using active lighting under controlled conditions. Seitz et al. [2005] propose an inverse light transport method, later refined by Bai et al. [2010], that separates an image into a sum of components for each light bounce. This method uses a large number of images, corresponding to separately illuminating each point in the scene. Nayar et al. [2006] propose a faster method that uses only a few images to separate direct and global illumination components. Unlike these systems, our approach does not require active controlled illumination, but instead incorporates some user interaction.

Researchers have proposed many approaches for editing materials and lighting in images. Oh et al. [2001] propose an image based modeling approach with applications to texture editing and relighting. Fang and Hart [2004] use shape from shading to synthesize textures with proper foreshortening. Khan et al. [2006] use very rough depth and lighting estimation to replace an object’s BRDF and even add transparency. Extending this work, Gutierrez et al. [2008] add caustics formed by the transparent objects. A number of authors have proposed methods for editing shadows in images [Finlayson et al. 2006; Mohan et al. 2007; Shor and Lischinski 2008], which can be considered an effect of global illumination. Our work is complementary to these techniques, as it is the only one to address diffuse reflection.

Editing diffuse interreflections has a long history for synthetic scenes. A good overview of radiosity techniques is given by Cohen and Wallace [1993], and we relate our decomposition to the radiosity form factors in the Appendix. Our decomposition also bears

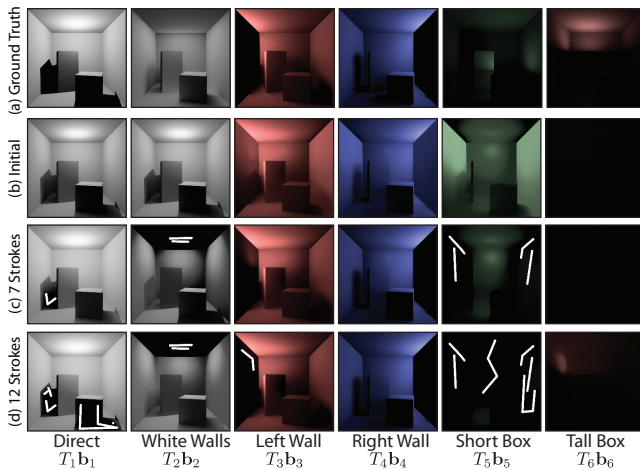


Figure 4: The illumination at each pixel is a linear combination of basis sources accounting for direct illumination and indirect illumination off each material in the scene. Indirect illumination from the white walls are combined into a single source, but are separate from direct lighting. The ground truth illumination decomposition of a Cornell box was rendered with a modified path tracer (a). In our system, users add strokes to remove contributions of the various indirect illumination sources to refine the decomposition (b-d). (Best viewed on a monitor)

resemblance to methods for precomputed radiance transfer, where a low dimensional representation of light transport is used to efficiently render interreflections under varying illumination [Hašan et al. 2006]. Similar to our stroke-based interface, researchers have also developed rendering systems that adjust light source intensities [Schoeneman et al. 1993] and light transport itself [Obert et al. 2007] to match user-specified target illumination.

Sparse gradient priors, like the one used by our system, have previously been used for image restoration. Such priors are typically motivated by their robustness to outliers compared to the L2 norm, allowing them to better reconstruct sharp edges. Rudin et al. [1992] introduced Total Variation, which measures the L1 norm of gradient magnitudes, as a regularization for image denoising. The importance of the L1 norm for sparse signals has also been highlighted by recent work on compressive sensing [Candes et al. 2006; Donoho 2006]. More recently, “sparser” hyper-Laplacian priors have been applied to super-resolution [Tappen et al. 2003] and deblurring [Levin et al. 2007], as well as denoising [Joshi et al. 2009]. As we will show, we use the Total-Variation prior not only because it preserves strong edges, but also because it enforces smoothness within individual illumination components without influencing the distribution of gradients between components.

3 Overview

Our approach to modifying indirect illumination is comprised of three stages (Figure 3). In the setup stage we first white balance the input image by dividing out the color of a user marked white object. We then assume a white light source throughout the rest of the decomposition process. Next we split the image into illumination and reflectance components. In practice we use the approach of Bousseau et al. [2009], though any source of intrinsic images with colored illumination would work with our system. The user selects a subset of materials from the reflectance image that account for indirect illumination in the scene. Our system combines these reflectance colors with white direct illumination into a set of *basis sources*.

In the illumination decomposition stage we use a nonlinear energy

minimization to compute a per-pixel linear combination of basis sources that explains the illumination image. The user can refine the decomposition by drawing strokes to specify that a basis source cannot contribute to the illumination in a particular area. Our system then re-optimizes the decomposition subject to the constraints. Users can iteratively add more strokes until they are satisfied with the decomposition. Computing this decomposition is our main algorithmic contribution.

In the third stage of our approach we recombine a modified reflectance image with the appropriately modulated illumination. Users modify the reflectance image using existing image editing tools, which is relatively simple because the reflectance image is made of piecewise constant colors. We substitute the modified reflectances for our basis sources to construct the modified illumination image. Multiplying the modified reflectance and illumination images gives us our final result.

Although our technique is designed to work with photographs, throughout Section 4 we will illustrate our approach using a synthetic rendering of a Cornell box. Using synthetic data allows us to evaluate our algorithm with accurate intrinsic images, and compare against a ground truth decomposition and recoloring.

4 Illumination Decomposition

Our illumination decomposition technique relies on a Lambertian scene assumption. We first describe our image formation model under this assumption, and then show how we use the low-frequency nature of illumination to guide our decomposition.

4.1 Image Formation Model

For Lambertian scenes, the observed color at each pixel is given by

$$\mathbf{I}(x, y) = \mathbf{S}(x, y)\mathbf{R}(x, y) \quad (1)$$

where \mathbf{I} is the observed image, \mathbf{S} is the shading or illumination intrinsic image, and \mathbf{R} is the reflectance intrinsic image. The bold variables indicate RGB vectors and multiplication is component-wise. We assume \mathbf{S} and \mathbf{R} are given as input from the setup stage.

Our illumination model is based on the assumption that there is only one bounce of indirect illumination. The light hitting any point in the scene either follows a direct path from the emitter, or it first bounces off one other surface and is modulated by the material color of that surface. We combine the emitted light color¹ and the emitted light color modulated by each material color into a set of basis sources² that account for the illumination at each pixel (Figure 4). The illumination at each pixel is then

$$\mathbf{S}(x, y) = \sum_{i=1}^n T_i(x, y)\mathbf{b}_i \quad (2)$$

where \mathbf{S} is the RGB illumination image, T_i are transport terms containing the scalar contribution of basis source \mathbf{b}_i at each pixel, and n is the number of sources. T_i is related to light transport in the standard global illumination model, and we also derive Eqn. 2 from the radiosity equation in the Appendix. For physically valid illumination T_i must be non-negative. Solving for the T_i is usually underconstrained because there are 3 equations and n unknowns at every pixel.

¹In practice we white balance the input image, so we can assume the emitted light is white. However, our model works with any light color.

²Basis sources model indirect illumination from different materials and are not light sources in the typical sense of emitting light. A similar concept of virtual point lights has been used in the context of global illumination rendering [Keller 1997; Walter et al. 2005]. However, our basis sources encompass the effects of all surfaces with a given material color, rather than forming a single point light.

We modified the PBRT rendering engine [Pharr and Humphreys 2004] to directly generate T_i , \mathbf{R} , and \mathbf{S} for a Cornell box. \mathbf{I} , \mathbf{S} , and \mathbf{R} are shown on the left side of Figure 3. The ground truth decomposition in Figure 4a shows six illumination components for direct illumination and indirect illumination from the white walls, the red left wall, the blue right wall, the short green box, and the tall red box.

4.2 Energy Minimization

The key observation that guides our decomposition is that illumination tends to be smooth across smooth surfaces. For Lambertian scenes we only expect to see strong gradients around occlusion boundaries, sharp changes in the surface geometry, and boundaries of cast shadows from direct illumination [Land et al. 1971]. We formulate our decomposition as an energy minimization based on this smooth variation property, along with terms accounting for fidelity to the input illumination, non-negativity, and a regularization on the contribution of each source.

Data Fidelity. To ensure that we obtain a valid decomposition of the illumination we turn equation 2 into an error metric using the squared L2 norm

$$E_{df} = \left\| \mathbf{S} - \sum_i T_i \mathbf{b}_i \right\|^2. \quad (3)$$

Non-negativity. For physically valid illumination, all the coefficients in T_i must be non-negative. We enforce non-negativity with a penalty function

$$E_{nn} = \sum_i \sum_{x,y} \max(-T_i(x,y), 0) \quad (4)$$

which is zero for $T_i \geq 0$ and increases linearly for $T_i < 0$.

Smoothness. We encode our key assumption of smooth illumination with a Total Variational (TV) smoothness term, which measures the L1 norm of gradient magnitudes

$$E_{sm} = \sum_i \|T_i\|_{TV} = \sum_i \sum_{x,y} \|\nabla T_i(x,y)\|, \quad (5)$$

where $\|\cdot\|$ is the standard Euclidean magnitude. In conjunction with a data term, the L1 prior is known to encourage sparsity. For our application the L1 prior on gradient magnitudes has two related effects; it encourages illumination components to be smooth with some sharp edges, and it tends to attribute gradients in the illumination image to a few rather than all of the components. We demonstrate the benefit of using this prior in Section 4.6.

Sparsity. Our final energy term is designed to encourage sparsity in the absolute contribution of each source, i.e. that only a few sources contribute to the illumination at each pixel. The main purpose of the regularization is to help guide the decomposition when the other terms do not have a clear minimum, and also to enforce user specified constraints on the contribution of each source as described in Section 4.4. The sparsity energy is given by

$$E_{sp} = \sum_i \|W_i T_i\|_1 = \sum_i \sum_{x,y} w_i(x,y) |T_i(x,y)| \quad (6)$$

where W_i is a diagonal weighting matrix for each component of the illumination with $w_i(x,y)$ on the diagonal. By default we set $w_i(x,y)$ to 1.

4.3 Optimization

To decompose illumination we minimize a weighted sum of the individual error terms

$$E = E_{df} + \lambda_{nn} E_{nn} + \lambda_{sm} E_{sm} + \lambda_{sp} E_{sp}, \quad (7)$$

where the λ variables determine the relative weights of the various energy terms. Intuitively we set these weights so that non-negativity is a very strong constraint, smoothness has less of an effect than data fidelity, and sparsity only has an effect when the decomposition otherwise does not have a strong minimum. We experimentally found the weights $\lambda_{nn} = 10^3$, $\lambda_{sm} = 10^{-4}$, and $\lambda_{sp} = 10^{-10}$ to work well in practice.

We optimize E using iteratively reweighted least-squares (IRLS) [Holland and Welsch 1977], which has previously been applied to sparse gradient priors in images [Levin and Weiss 2007; Levin et al. 2007; Joshi et al. 2009]. The IRLS technique is general enough to handle all of our energy terms, including non-negativity. It works by solving a series of weighted least-squares problems, where weights are determined by a previous iteration. We substitute the following weighted least squares energies into Eqn. 7:

$$E'_{sm} = \sum_i \sum_{x,y} \psi_{sm,i}(x,y) \|\nabla T_i(x,y)\|^2 \quad (8)$$

$$E'_{sp} = \sum_i \sum_{x,y} \psi_{sp,i}(x,y) w_i(x,y) T_i(x,y)^2 \quad (9)$$

$$E'_{nn} = \sum_i \sum_{x,y} \psi_{nn,i}(x,y) T_i(x,y)^2 \quad (10)$$

where

$$\psi_{sm,i}(x,y) = (\|\nabla T_i(x,y)\| + \epsilon)^{-1}$$

$$\psi_{sp,i}(x,y) = (|T_i(x,y)| + \epsilon)^{-1}$$

$$\psi_{nn,i}(x,y) = \begin{cases} 0 & \text{for } T_i(x,y) > 0 \\ (|T_i(x,y)| + \epsilon)^{-1} & \text{otherwise} \end{cases}$$

are weighting terms and ϵ is a parameter that smooths the objective function and prevents division by zero. As ϵ approaches zero, the modified energy terms revert to the original non-smooth versions. We set $\epsilon = 10^{-6}$ for all our results. The IRLS optimization alternates between calculating the weights and solving the least-squares system with fixed weights. The weights are initialized to 1 and updated on each iteration using the current values of T_i .

We use the synthetic Cornell box rendering to evaluate our algorithm. Our initial decomposition is shown in Figure 4b and can be compared to the ground truth decomposition in 4a. We are able to capture indirect illumination from the left and right wall fairly well. However, we overestimate green indirect illumination from the small box and do not accurately separate indirect white illumination from the direct lighting. To address these issues we incorporate user-specified constraints.

4.4 User Constraints

Users can specify strokes in our interface to remove the contribution of particular sources in areas where the contribution is overestimated. The strokes simply increase the weights $w_i(x,y)$ in Eqn. 6. We set the weights to 10^{10} for all pixels covered by a stroke. Increasing the weights for one component forces the other components to explain the illumination in the affected area. The smoothness and fidelity terms propagate the effect of the strokes throughout the illumination decomposition.

Adding strokes to the Cornell box example improves the decomposition significantly. In Figure 4c the user adds strokes to remove

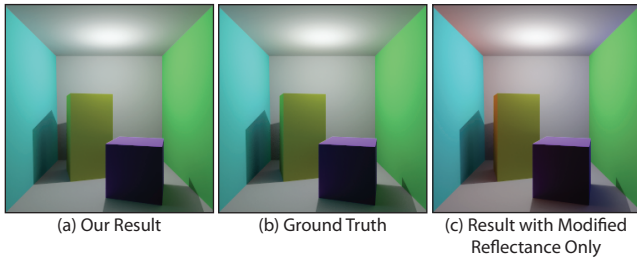


Figure 5: Our recoloring result of the Cornell box (a) uses the decomposition from Figure 4d, and is visually indistinguishable from the ground truth recoloring (b). Accounting for indirect illumination when recoloring leads to much more consistent results than the approach of modifying only reflectance (c). (Best viewed on a monitor)

indirect white illumination on the ceiling. Although the ground truth decomposition contains some indirect white illumination on the ceiling, it is negligible compared to direct illumination there. The user also adds strokes to remove direct illumination in the cast shadow of the tall box, and indirect illumination from the short box onto the left and right walls. After this first set of strokes there is still too much direct illumination in the cast shadows and green illumination from the tall box on the back wall. In Figure 4d the user adds a second set of strokes to fix these issues, as well as a stroke to remove some illumination on the left wall that is attributed to the left wall itself. After incorporating user-strokes, our decomposition matches the ground truth much more closely.

4.5 Color Editing

With a complete decomposition we are able to change any of the basis sources and modulate the illumination accordingly. We modify the reflectance image using the quick selection tool in Photoshop and use the Robust Matting software of Wang et al. [2007] for objects that have soft boundaries in the reflectance image. We substitute the modified basis sources \mathbf{b}' into Eqn. 2 to obtain the recolored illumination image \mathbf{S}' and generate our final image \mathbf{I}' using Eqn. 1.

In Figure 5a we use our decomposition from Figure 4d to recolor the Cornell box, changing all material colors except the white walls. The result is nearly impossible to visually distinguish from the ground truth image rendered in PBRT (Figure 5b). The benefit of modifying indirect illumination is especially clear in comparison to modifying the reflectance channel alone and combining it with the unmodified illumination (Figure 5c). Although indirect illumination may seem subtle at first, the effect of inconsistent diffuse reflections can be striking.

Note that because data fidelity is just one of our energy terms, our decomposition is not guaranteed to sum back to the input illumination. In particular, for real-world scenes that may partially violate our assumptions the reconstruction deviates when illumination is not spanned by a non-negative combination of the basis sources. Moreover, the tradeoff between fidelity and smoothness also leads to discrepancies in the reconstruction. We can optionally add the residual $\mathbf{S} - \sum_i T_i \mathbf{b}_i$ into \mathbf{S}' to retain illumination not captured by our decomposition. We demonstrate the effect of the residual in Figure 9.

4.6 Discussion

Our technique takes advantage of two assumptions: 1) that illumination is locally a mix of independent colors, and 2) that the illumination is smooth. We consider how our decomposition is affected by these assumptions.

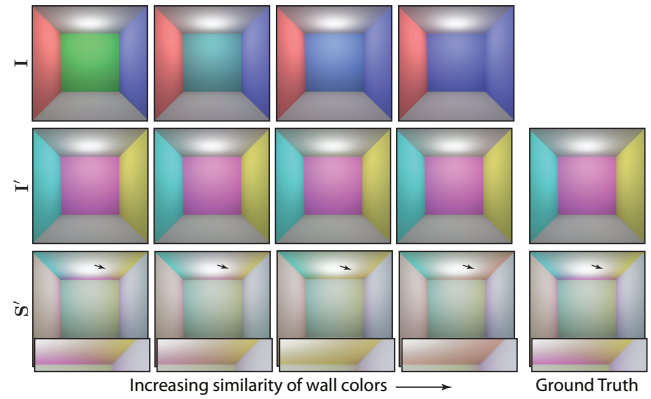


Figure 6: We decompose illumination from rendered Cornell boxes with increasingly dissimilar wall colors. For each variation of input reflectances we render the input shading image, decompose it with our system and use the decomposition to create modified shading images \mathbf{S}' . For inputs with similar wall colors \mathbf{S}' closely matches the ground truth, but as the input wall colors become more similar the indirect illumination from those walls become hard to separate and the modified interreflections blend together. See Section 4.6 for details. (Best viewed on a monitor)

Basis Source Ambiguity. Our decomposition works best when the illumination at each pixel is formed by a subset of independent basis sources. When this assumption is violated we rely heavily on the smoothness prior. In Figure 6 we demonstrate how the quality of the decomposition and subsequent material edits degrades as we progressively shift the color of the back wall to match the right wall. In this example we combine direct illumination and indirect illumination from the white surfaces into a single basis source. We also use a fixed set of strokes for all input images to prevent self-interreflections on the walls. The first row of Figure 6 shows the input images \mathbf{I} . Each input image is the product of the rendered reflectance \mathbf{R} and illumination \mathbf{S} (not shown). Since the geometry is the same for each box, they all have the same transport terms T_i . For space we omit showing all of these images, but they are available in supplemental material.

The second row of Figure 6 shows the recolored image \mathbf{I}' constructed from our estimated decomposition of each input, as well as the ground truth rendering. Differences are subtle, but we can see them more clearly in the recolored illumination images \mathbf{S}' . We enhanced the saturation of the illumination images to better reveal the differences (unmodified versions are available in supplemental material). The main differences show up on the ceiling in the diffuse reflections of the back and right walls. When the colors of the walls are very different, we are able to estimate the transport terms well and \mathbf{S}' closely matches the ground truth. However, as the colors of the walls become more similar, it is harder to disambiguate their diffuse reflections on the ceiling, and our decomposition allocates some of the indirect lighting to the wrong wall. When the walls are the same color (column 4) this misallocation results in a large red reflection instead of distinct pink and yellow reflections on the ceiling. Despite inaccuracies in \mathbf{S}' , the differences between the final recolored images \mathbf{I}' and the ground truth are very subtle. Closely examining the boundary between the walls and ceiling reveals small differences, but none of the results have inconsistencies that visually stand out.

Total-Variational Smoothness. The Total-Variation prior is typically motivated by its robustness to large gradients as compared to the L2 norm. We distinguish two effects of this robustness for our application. Just as is useful in denoising and deblurring applications, Total-Variation prevents sharp edges from being blurred, be-

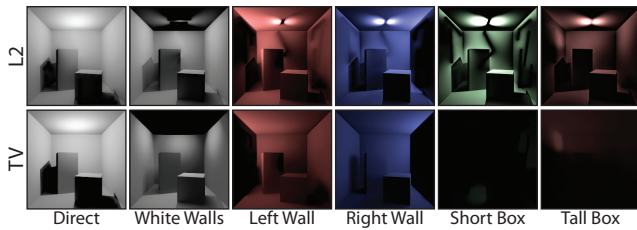


Figure 7: Comparison of Total Variational and L2 smoothness priors on a decomposition of the Cornell box rendering. The L2 prior has a strong preference to explain the bright highlight on the ceiling as a combination of the light sources, which interferes with the user strokes propagating across the surface. Under the Total Variational prior a single large gradient is equally likely as many smaller gradients. (Best viewed on a monitor)

cause one large gradient is penalized the same as many small gradients. For the same reason, and perhaps more importantly for our application, TV is better able to attribute a gradient in the illumination to one basis source rather than a combination of all the sources. Although computationally simpler, a least-squares smoothness prior favors attributing the gradient to many small contributions from all sources. We demonstrate the advantage of the TV regularization over L2 in Figure 7. User strokes (the same used in Figure 4-bottom) are not propagated well in the least-squares result, because explaining the highlight on the ceiling as a combination of sources has a much lower energy under the L2 norm. In contrast, our result using the TV norm attributes the highlight largely to the direct illumination.

4.7 Implementation Details

We implemented our IRLS optimization in MATLAB using a sparse direct solver to solve the weighted linear system in each iteration. We fix the number of iterations to 10, which was sufficient for the decomposition to visually converge in our experiments. Runtime varies based on image size and the number of illumination components. Decomposing the 250x250 Cornell box illumination (Figure 4) into 6 components takes approximately 3 minutes per set of user constraints. While the computation time is a limiting factor in making our interactive system practical for high-resolution images, there are a number of possibilities for increasing performance. For the IRLS iterations, using an iterative method such as conjugate gradient with an appropriate preconditioner could reduce the running time, but finding an effective preconditioner for weighted least squares systems is non-trivial. Another approach is to replace IRLS with an alternative nonlinear optimization strategy; recent iterative shrinkage-thresholding algorithms have been used to efficiently minimize similar error functions [Bioucas-Dias and Figueiredo 2007]. Finally, a number of techniques could aid in applying the technique to higher resolution images, such as, coarse-to-fine methods, upsampling [Kopf et al. 2007], or image-space dimensionality reduction [Fattal et al. 2009]. We have experimented with a simple method for creating higher resolution results by decomposing the input illumination into base and detail layers. We downsample the base layer and modify it using our system, and then add the detail layer into the upsampled result. We include these high-resolution images in the supplemental material.

5 Results

We demonstrate the results of our approach with several example photographs in figures 1, 8, 9, 11, and 12. *Many of the results are subtle and we encourage readers to view images on a monitor to zoom in, and to look at the supplemental material.*

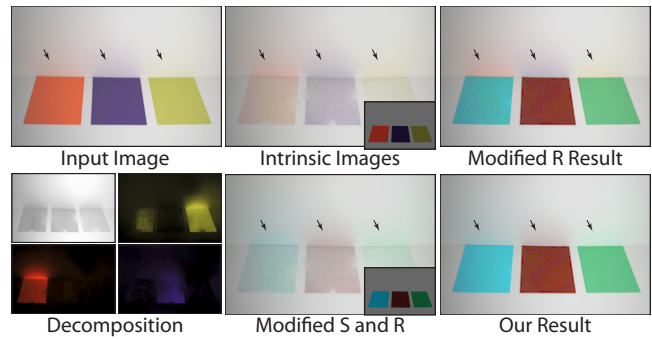


Figure 8: Each of the colored sheets of paper casts a diffuse reflection onto the wall, as well as interreflection within the rough surface of the paper. We decompose the illumination into a component for each sheet, as well as a combined component due to direct illumination and indirect illumination from the white surfaces. We are able to modify the colors of each of the sheets, and modify the diffuse reflections to match. (Best viewed on a monitor)

Figure 8 shows a photo taken in a laboratory setting. We placed three pieces of differently colored paper at the corner of two white surfaces, resulting in diffuse reflections. The illumination image **S** contains color bleeding from the papers onto the vertical wall. **S** also contains some colored illumination on the papers themselves, which can be explained by concavities in the rough surface of the paper. We set the basis sources to the colors of the three pieces of paper and white. We combine the direct lighting and indirect white illumination into a single source because they are difficult to disambiguate. Because of the simple nature of this scene, our algorithm obtains a clean decomposition without any user constraints. The only noticeable artifact is a small amount of illumination on the orange paper attributed to indirect illumination from the yellow paper. Nevertheless, using our decomposition to modify the reflectance and illumination results in consistent color bleeding. In contrast, the result of modifying the reflectance alone, without the illumination, results in an inconsistent image.

In the photo in Figure 1 the woman's shirt casts blue illumination on her arm. As shown in the illumination image in Figure 2, the fine texture of the shirt also results in self-interreflection. Making the shirt pink without modifying the shading produces inconsistent results. We decompose the illumination into components due to the direct lighting and indirect lighting from the skin, shirt, and orange strap. We also use a single user stroke to prevent illumination on the arm from being attributed to the strap. After modifying indirect illumination from the shirt to match the modified reflectance, our result looks much more consistent. The diffuse reflection on the arm changes to pink, and the color shift goes away.

In Figure 9 we demonstrate our technique for recoloring the cabinets in a kitchen. The bright pink surfaces cast strong indirect illumination over a large portion of the image. The user marks the cabinets and beige walls as basis sources, for a total of three when combined with direct lighting. With the pink illumination separated, we are able to modify the indirect illumination to match the recolored cabinets, resulting in blue illumination on the floor and under the table. For this result we also add the residual illumination, which helps capture the glossy illumination on the ceramic jar and the bright colors in the curtains that were misattributed to illumination in the intrinsic image decomposition. Since we only use three basis sources for this decomposition, it is possible to solve Eqn. 2 with a simple inversion, but this method can result in negative illumination contributions. Furthermore, without the smoothness term the decomposition is sensitive to noise, which is noticeable in individ-

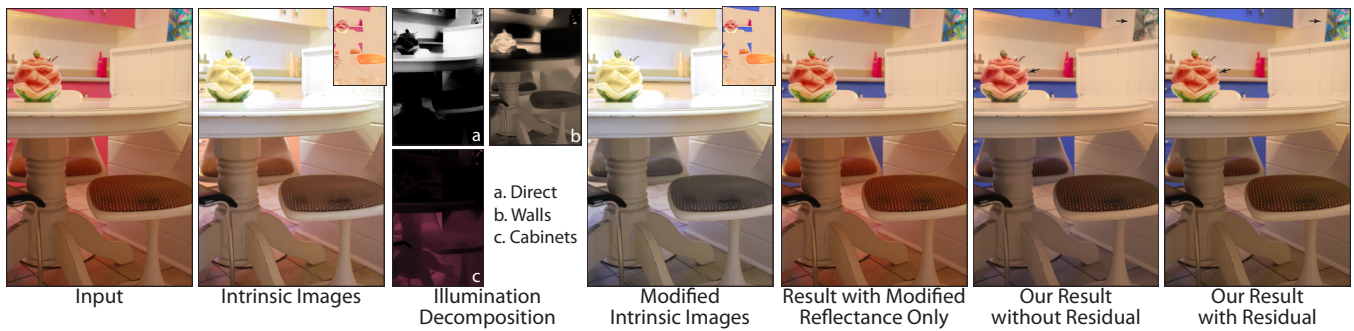


Figure 9: Indirect lighting from the cabinets in this example results in pink illumination throughout much of the kitchen. Our decomposition contains illumination components due to direct white light, indirect lighting from the beige paint, and the indirect lighting from the cabinets. We recolor the cabinets to be blue, and convincingly modify the illumination to match. (Best viewed on a monitor)

ual illumination components even where the original illumination image appears smooth (Figure 10). The recoloring using a simple inversion leads to artifacts in places where illumination is poorly explained by our basis sources, for example the yellow lighting under the cabinets and the ceramic jar. Our method also does a better job at recoloring all of the pink lighting under the table.

Figure 11 shows an example of changing the color of a pillow and the wall, along with subtle interreflections. In the input illumination image we see that the wall casts brown illumination over much of the image, especially on the bib. Our system initially overestimated the contribution of light from the wall hitting the baby’s head, which the user fixed with a few strokes. The blue pillow causes a diffuse reflection onto the adjacent gray pillow and the baby’s legs and arm, which we were able to separate. The user added some additional strokes to remove contribution from the blue pillow elsewhere in the image. In our result the user changes the blue pillow to be green, and our system updates the interreflection accordingly. We also change the illumination coming off the back wall from blue in order to match the modified reflectance. Because they are reflective and not modeled well under the Lambertian assumption, the intrinsic image separation attributes some color from the eyes to the shading image. Our decomposition attributed the eye color to the blue pillow, resulting in green eyes in our result.

Our final example in Figure 12 shows the result of changing the color of a girl’s shirt. The girl is looking down causing red illumination from her shirt to reflect on her face, which can be seen clearly in the illumination image. We are able to factor the illumination without any user constraints because there are only a few materials contributing to indirect illumination. We modify the reflectance of the shirt to be green, and obtain a matching green interreflection on the girl’s face.

6 Limitations

Intrinsic Image Quality. Our system requires intrinsic images as input and any artifacts (e.g. texture variation attributed to illumination) present in the intrinsic images are propagated into our decomposition. Generating intrinsic images remains a challenging problem, and even state-of-the-art approaches cannot fully separate illumination and reflectance. However, these imperfect intrinsic images are often sufficient for plausibly modifying reflectance using our illumination decomposition system. In addition, our system works independently of the particular intrinsic image technique and will benefit from any future advances in that area.

The intrinsic image decomposition of Bousseau et al. can vary significantly depending on the parameters and strokes provided to the system. We found that our tool was able to provide a good decom-

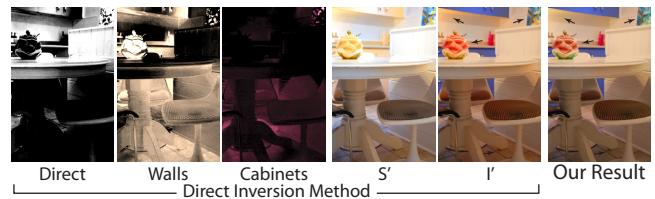


Figure 10: We model the illumination in Figure 9 with only 3 basis sources, so the linear system in Eqn. 2 can be solved directly for each pixel. However, because real images can deviate from our image formation model in many ways, this direct decomposition can contain noise and negative contributions. In contrast, our result (Figure 9) is nonnegative and smoother. (Best viewed on a monitor)

position of the illumination given a plausible intrinsic image pair, and was not very sensitive to the particular intrinsic image separation.

User Interaction. As with many stroke based image editing tools, users must learn how to draw effective strokes. In our system the user modifies the decomposition by marking areas in which to remove contributions from particular basis sources. One alternative approach is to use strokes to increase the contribution of a particular source; however, it is much less clear how to implement such strokes. The strokes in our system increase the weight of the sparsity constraint for the stroked pixels in the chosen illumination component. However, decreasing the weights would not cause much of a change, because the sparsity term already has a small effect with the default weights. We could also use strokes to indicate the (non-zero) absolute contribution of a source at a particular location, but this value would be more difficult for the user to specify. Although it does have a learning curve, our approach of using strokes to remove illumination contributions leads to both a simple implementation and an interaction that users can learn relatively quickly.

7 Conclusion

Changing the color of an object is one of the most basic image editing tasks, but doing so also requires modifying the shading, including diffuse interreflections. In this paper, we have presented a way to decompose an input image into multiple basis sources, corresponding to direct lighting and indirect illumination from each material. The decompositions produced by our system may not be physically accurate—indeed, computing the decomposition is an underconstrained problem for which we introduce a principled optimization framework. However, we demonstrate in a number of examples that we do obtain a good decomposition, that is sufficient to create visually plausible edits, with consistent interreflections.

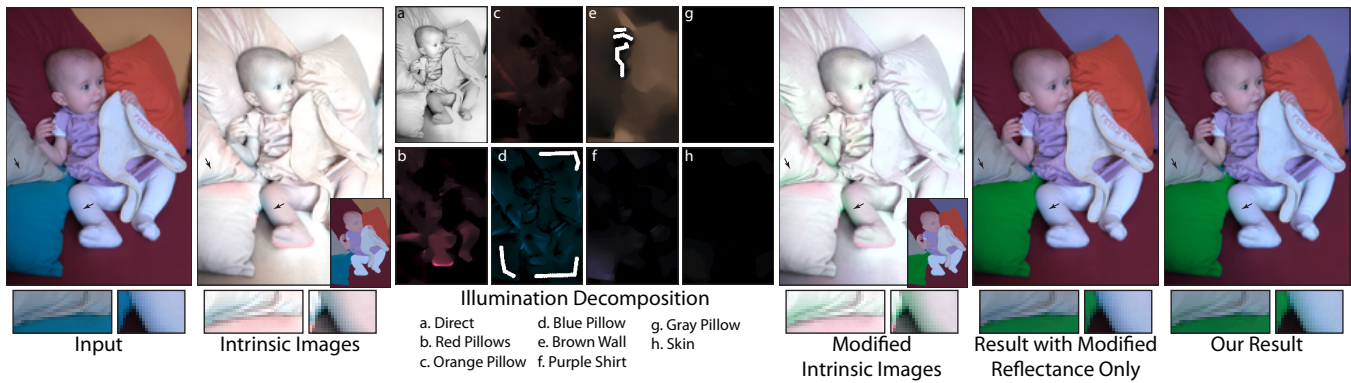


Figure 11: In the input image, the blue pillow casts indirect illumination on the gray pillow, and the baby’s arm and leg. We capture this indirect lighting in our decomposition, and change its color along with the pillow. We also modify the color of the wall to be blue along with indirect illumination it casts across much of the image, most notable on the baby’s bib. (Best viewed on a monitor)

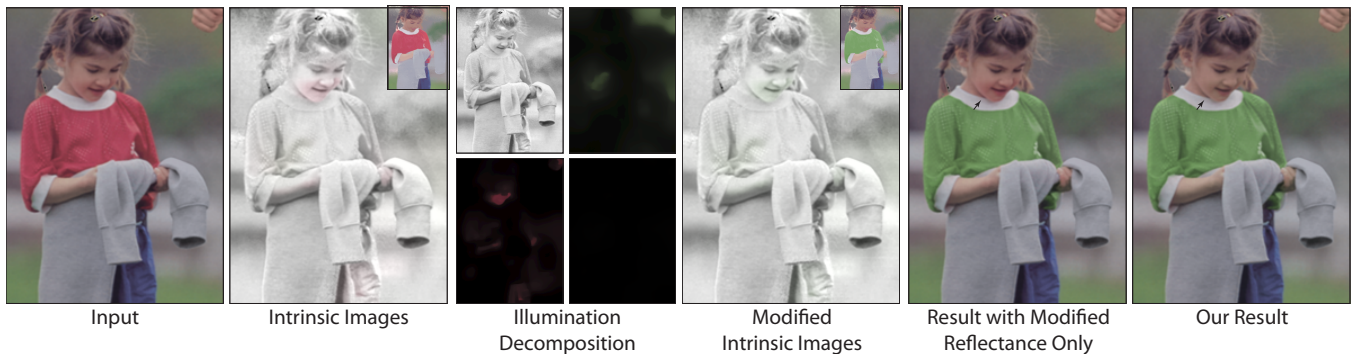


Figure 12: In this example the girl’s shirt casts red indirect illumination on her face. We decompose the illumination into components due to direct light, and indirect illumination from the red shirt, green grass, and blue pants. We then make the shirt green, and appropriately modify the interreflection to match. (Best viewed on a monitor)

In summary, we have shown that some insight into the physics of light transport, and new optimization methods for finding plausible solutions to ill-posed inverse problems, can lead to novel approaches to image editing. We believe such techniques can lead to a richer suite of editing tools, that better account for the physical aspects of real world scenes.

Acknowledgments. We would like to thank Adrien Bousseau for providing his intrinsic images code and invaluable guidance. We also thank Michael Lustig for helpful discussions regarding optimization, and reviewers for their feedback. This work was supported in part by ONR PECASE grant N00014-09-1-0741, NSF CAREER grant 09-24968, NSF grant CCF-0643552 and generous equipment and support from Intel, NVIDIA, Adobe and Pixar. We would also like to thank Flickr users kimhas7cats and msjacoby for their images.

References

- BAI, J., CHANDRAKER, M., NG, T.-T., AND RAMAMOORTHY, R. 2010. A dual theory of inverse and forward light transport. In *ECCV '10*, 294–307.
- BARROW, H., AND TENENBAUM, J. 1978. Recovering intrinsic scene characteristics from images. *Computer Vision Systems* 27, 9, 3–26.
- BIUCAS-DIAS, J. M., AND FIGUEIREDO, M. A. T. 2007. A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. on Image Processing* 16, 12 (dec.), 2992–3004.
- BOUSSEAU, A., PARIS, S., AND DURAND, F. 2009. User-assisted intrinsic images. *ACM Trans. Graph.* 28, 130:1–130:10.
- CANDES, E., ROMBERG, J., AND TAO, T. 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory* 52, 2, 489–509.
- CHUANG, Y.-Y., CURLESS, B., SALESIN, D. H., AND SZELISKI, R. 2001. A bayesian approach to digital matting. In *CVPR '01*, vol. 2, 264–271.
- COHEN, M. F., AND WALLACE, J. 1993. *Radiosity and realistic image synthesis*. Academic Press Professional, Inc., San Diego, CA.
- DONOHOO, D. 2006. Compressed sensing. *IEEE Trans. on Inform. Theory* 52, 4, 1289–1306.
- FANG, H., AND HART, J. C. 2004. Textureshop: texture synthesis as a photograph editing tool. *ACM Trans. Graph.* 23, 354–359.
- FATTAL, R., CARROLL, R., AND AGRAWALA, M. 2009. Edge-based image coarsening. *ACM Trans. Graph.* 29, 6:1–6:11.
- FATTAL, R. 2008. Single image dehazing. *ACM Trans. Graph.* 27, 72:1–72:9.
- FINLAYSON, G., HORDLEY, S., LU, C., AND DREW, M. 2006. On the removal of shadows from images. *IEEE Trans. PAMI* 28, 1, 59–68.
- GUTIERREZ, D., SERON, F. J., LOPEZ-MORENO, J., SANCHEZ,

- M. P., FANDOS, J., AND REINHARD, E. 2008. Depicting procedural caustics in single images. *ACM Trans. Graph.* 27, 120:1–120:9.
- HAŠAN, M., PELLACINI, F., AND BALA, K. 2006. Direct-to-indirect transfer for cinematic relighting. *ACM Trans. Graph.* 25, 1089–1097.
- HOLLAND, P. W., AND WELSCH, R. E. 1977. Robust regression using iteratively reweighted least-squares. *Communications in Statistics - Theory and Methods* 6 (September), 813–827.
- HORN, B. K. P. 1986. *Robot Vision*. The MIT Press, March.
- HSU, E., MERTENS, T., PARIS, S., AVIDAN, S., AND DURAND, F. 2008. Light mixture estimation for spatially varying white balance. *ACM Trans. Graph.* 27, 70:1–70:7.
- JOSHI, N., ZITNICK, C., SZELISKI, R., AND KRIEGMAN, D. 2009. Image deblurring and denoising using color priors. In *CVPR '09*, 1550–1557.
- KELLER, A. 1997. Instant radiosity. In *Proc. SIGGRAPH*, 49–56.
- KHAN, E. A., REINHARD, E., FLEMING, R. W., AND BÜLTHOFF, H. H. 2006. Image-based material editing. *ACM Trans. Graph.* 25, 654–663.
- KOPF, J., COHEN, M. F., LISCHINSKI, D., AND UYTENDAELE, M. 2007. Joint bilateral upsampling. *ACM Trans. Graph.* 26.
- LAND, E. H., JOHN, AND MCCANN, J. 1971. Lightness and retinex theory. *Journal of the Optical Society of America*, 1–11.
- LEVIN, A., AND WEISS, Y. 2007. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Trans. PAMI* 29, 9, 1647–1654.
- LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. T. 2007. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.* 26.
- LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2008. A closed-form solution to natural image matting. *IEEE Trans. PAMI* 30 (February), 228–242.
- MOHAN, A., TUMBLIN, J., AND CHOUDHURY, P. 2007. Editing soft shadows in a digital photograph. *IEEE Comput. Graph. Appl.* 27 (March), 23–31.
- NAYAR, S. K., KRISHNAN, G., GROSSBERG, M. D., AND RASKAR, R. 2006. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph.* 25, 935–944.
- OBERT, J., KŘIVÁNEK, J., SÝKORA, D., AND PATTANAIK, S. 2007. Interactive light transport editing for flexible global illumination. In *ACM SIGGRAPH 2007 sketches*.
- OH, B. M., CHEN, M., DORSEY, J., AND DURAND, F. 2001. Image-based modeling and photo editing. In *Proc. SIGGRAPH*, 433–442.
- PHARR, M., AND HUMPHREYS, G. 2004. *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann.
- RAMAMOORTHY, R., AND HANRAHAN, P. 2001. A signal-processing framework for inverse rendering. In *Proc. SIGGRAPH*, 117–128.
- RUDIN, L. I., OSHER, S., AND FATEMI, E. 1992. Nonlinear total variation based noise removal algorithms. *Phys. D* 60 (November), 259–268.
- SCHOENEMAN, C., DORSEY, J., SMITS, B., ARVO, J., AND GREENBERG, D. 1993. Painting with light. In *Proc. SIGGRAPH*, 143–146.
- SEITZ, S., MATSUSHITA, Y., AND KUTULAKOS, K. 2005. A theory of inverse light transport. In *ICCV '05*, vol. 2, 1440–1447.
- SHEN, L., TAN, P., AND LIN, S. 2008. Intrinsic image decomposition with non-local texture cues. In *CVPR '08*, 1–7.
- SHOR, Y., AND LISCHINSKI, D. 2008. The shadow meets the mask: Pyramid-based shadow removal. *Comput. Graph. Forum* 27, 2, 577–586.
- SINHA, P., AND ADELSON, E. 1993. Recovering reflectance and illumination in a world of painted polyhedra. In *ICCV '93*, 156–163.
- TAPPEN, M. F., RUSSELL, B. C., AND FREEMAN, W. T. 2003. Exploiting the sparse derivative prior for super-resolution and image demosaicing. In *IEEE Workshop on Stat. and Comput. Theories of Vision*.
- TAPPEN, M. F., FREEMAN, W. T., AND ADELSON, E. H. 2005. Recovering intrinsic images from a single image. *IEEE Trans. PAMI* 27, 9, 1459–1472.
- WALTER, B., FERNANDEZ, S., ARBREE, A., BALA, K., DONIKIAN, M., AND GREENBERG, D. P. 2005. Lightcuts: a scalable approach to illumination. *ACM Trans. Graph.* 24, 1098–1107.
- WANG, J., AND COHEN, M. 2007. Optimized color sampling for robust matting. In *CVPR '07*, 1–8.
- WEISS, Y. 2001. Deriving intrinsic images from image sequences. In *ICCV '01*, vol. 2, 68–75.

Appendix: Derivation of Illumination Model

Assuming that the scene is Lambertian, we can relate our illumination model to the radiosity equation. We begin with the standard Neumann series expansion of the radiosity equation truncated to 2 bounces

$$\mathbf{B}_i = \mathbf{E}_i + \rho_i \sum_j F_{ij} \mathbf{E}_j + \rho_i \sum_j \sum_k \rho_j F_{ij} F_{jk} \mathbf{E}_k + \dots$$

where \mathbf{B}_i is the radiosity at surface patch i , \mathbf{E} is the emission, ρ is the reflectance or albedo, and F_{ij} is the form factor describing the light transfer from patch i to patch j . Bold variables indicate RGB vectors. We assume that the albedo at each patch is a weighted sum of a set of basis albedos $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$, such that we have $\rho_j = \sum_l c_{jl} \mathbf{b}_l$. If we assume the light sources are white (or a white balanced input image) we can substitute $e_i \mathbf{1}$ for \mathbf{E}_i , where e_i is a scalar. Then,

$$\mathbf{B}_i \approx e_i \mathbf{1} + \rho_i \sum_j F_{ij} e_j + \rho_i \sum_j \sum_k \sum_l c_{jl} \mathbf{b}_l F_{ij} F_{jk} e_k$$

Substituting $T_i = \sum_j F_{ij} e_j$ and $T_{il} = \sum_j \sum_k c_{jl} F_{ij} F_{jk} e_k$ and assuming light sources aren't visible (i.e. $e_i = 0$) we get

$$\mathbf{S}_i = \frac{\mathbf{B}_i}{\rho_i} \approx T_i + \sum_l T_{il} \mathbf{b}_l, \quad (11)$$

where \mathbf{S}_i is the illumination at patch i . This equation is equivalent to our shading decomposition equation (Eqn. 2), except it uses indexed patches instead of pixels, and the distinction between direct and indirect illumination is made explicit.

