Illumination Recovery from Image with Cast Shadows via Sparse Representation

Xue Mei, Haibin Ling, *Member, IEEE*, and David W. Jacobs, *Member, IEEE*

Abstract—In this paper, we propose using sparse representation for recovering the illumination of a scene from a single image with cast shadows, given the geometry of the scene. The images with cast shadows can be quite complex and therefore cannot be well approximated by low-dimensional linear subspaces. However, it can be shown that the set of images produced by a Lambertian scene with cast shadows can be efficiently represented by a sparse set of images generated by directional light sources. We first model an image with cast shadows as composed of a diffusive part (without cast shadows) and a residual part that captures cast shadows. Then, we express the problem in an ℓ_1-regularized least squares formulation, with nonnegativity constraints (as light has to be nonnegative at any point in space). This sparse representation enjoys an effective and fast solution, thanks to recent advances in compressive sensing. In experiments on both synthetic and real data, our approach performs favorably in comparison with several previously proposed methods.

Index Terms—Illumination recovery, inverse lighting, sparse representation, compressive sensing, ℓ_1 minimization.

1 Introduction

Illumination recovery, also known as inverse lighting, is the problem of recovering an illumination distribution in a scene from the appearance of objects located in the scene. It is used for augmented reality, where the virtual objects match the existing image and cast convincing shadows on the real scene rendered with the recovered illumination. Shadows in a scene are caused by the occlusion of incoming light, and thus contain information about the lighting of the scene. Although shadows have been used widely in determining the 3D shape of the object that casts shadows onto the scene, few studies have been devoted to the illumination information provided by the shadows. This is mainly because that shadows bring great challenges to recover illumination from a scene given a single input image.

In this paper, we study the problem of illumination recovery from a single image with cast shadows. Instead of an explicit estimation of the illumination, the purpose of our method is to recover the lighting so that the image rendered with recovered lighting would look similar to the original image. A prior model about the scene, as in many previous works, is used that captures geometry and albedos [4], [3], [36], [34], [35], [27], [33], [13]. It is in fact pointed out in [13] that the assumption of known geometry is required by many illumination estimation methods.[1] Despite the challenges of the task, it has been observed [32], [27], [36] that images with cast shadows can often be sparsely represented (see Section 2 for details). Such a sparse representation is very attractive since sparsity leads to efficient estimation, dimensionality reduction, and efficient modeling. For example, by representing the lighting using a sparse set of directional sources, we can save a large amount of time in rendering very complex scenes while maintaining the quality of scene recovery.

Inspired by previous studies [32], [27], [36], we approximate an image with cast shadows by a combination of low frequency spherical harmonics and a sparse set of directional light sources. To find the sparse solution, we first approximate the diffusive part of the image with low dimensional harmonics. After that, the residual is modeled as a sparse combination of basis images representing different directional light sources. The sparse solution is then achieved by an ℓ_1-regularized least squares solver, which has been proved to be close to the ℓ_0 solutions under very flexible conditions [5], [9]. Compared to ℓ_2 minimization, ℓ_1 minimization tends to find the most significant directional sources and discard the insignificant ones. This is very suitable for our purpose in which we want to select a sparse representation from about one thousand directional sources. The solution to the ℓ_1-regularized least squares problem using the truncated Newton interior-point method is very fast and reliable, which enables our method to be used in many areas, such as lighting design. We tested the proposed method on synthetic and real images in comparison with serval state-of-the-art approaches. In all the experiments the proposed method demonstrates excellent performances in both recovery accuracy and run time efficiency.

The rest of the paper is organized as follows. Sec. 2 discusses related work. In Sec. 3, we show that the effects of cast shadows may not be well approximated by any low-dimensional representation. However, when only a few directional light sources illuminate a scene, they may be compactly represented. After that, the model for illumination recovery is described and analyzed in Sec. 4. Then, in Sec. 5, we propose a solution to find a sparse

1. However, in our experiments we use models that have uniform, white albedo, as these are easy to construct. With Lambertian objects, albedo linearly scales pixel intensities, and so it is easy to remove the effects of *known* albedo by simply scaling image pixels by corresponding albedos.

representation using the ℓ_1-regularized least squares. The experiments are described in Sec. 6, where the proposed approach demonstrates excellent performance in both accuracy and efficiency. Finally, we conclude the paper in Sec. 7.

2 Related Work

As one of the major challenges in illumination recovery, shadows have been studied for a long history. There has been a series of work aimed at understanding the complexity of the set of images produced by Lambertian objects lit by environment maps. It is shown by Shashua [38] that, when ignoring all shadows, the images of a Lambertian scene lie in a three-dimensional linear subspace. The result is used for rendering in [24]. When including attached shadows, the set of images produced by a Lambertian scene can be approximated by a low dimensional linear subspace. This is shown both empirically [4], [11] and analytically [3], [30]. Shadows in a real scene are areas where direct light from a light source cannot reach due to the occlusion by other objects and, thus can provide useful information in recovering the lighting [16], [28]. In Kim and Hong [16], a practical approach is proposed to estimate the illumination distribution from shadows cast on a textured, Lambertian surface. In Panagopoulos et al. [28], a graphical model is proposed to estimate the illumination which is modeled as a mixture of von Mises-Fisher distributions and detect the shadows of a scene with textured surfaces from a single image and only coarse 3D information. The sparsity of cast shadows is studied in [32], [27], [36]. A method is proposed by Sato et al. [36] to recover an illumination distribution of a scene from image brightness inside shadows cast by an object of known shape. It introduces an adaptive sampling framework for efficient estimation of illumination distribution using a smaller number of sampling directions. In Okabe et al. [27], a sparse representation using a Haar wavelet basis is proposed to recover lighting in images with cast shadows. The work by Ramamoorthi et al. [32] shows that, although the set of images produced by a scene with cast shadows can be of high dimension, empirically this dimension does not grow too rapidly.

Prior to our work, the idea that a complex lighting environment with cast shadows can be represented with a few point lights is known in previous work [42], [31], [1], [44]. In Ramamoorthi and Hanrahan [31], a signal-processing framework that describes the reflected light field as a convolution of the lighting and BRDF is introduced. This work suggests performing rendering using a combination of spherical harmonics and directional light sources with ray-tracing to check for shadows. Structured importance sampling is introduced by Agarwal et al. [1] to show how to break an environment map into a set of directional lights for image synthesis. It samples an environment map efficiently to render scenes illuminated by distant natural illumination. Walter et al. [44] propose a scalable framework for accurately approximating illumination from thousands or millions of point lights using a strongly sublinear algorithm. Our work share with the above work in sparse shadow representation, but differs in the motivation (Section 3) and recovery solutions.

The studies in [36], [34], [35], [27] are closely related to our work. These studies propose recovering lighting from cast shadows by a linear combination of basis elements that represent the light. Specifically, in [27] a Haar wavelet basis is used to effectively capture lighting sparsely. Given the compact support with various sizes of Haar wavelets, each incident direction of illumination is represented at different resolutions. The illumination is efficiently approximated using a small number of basis functions. In our method, we approximate the illumination using a small number of directional sources with nine spherical harmonics. Both methods are proposed to overcome the limitations from spherical harmonics for not capturing high frequency components in the lighting such as a point light or an area light with a sharp edge and the shadows cast by an object. In the experiments, we show that our method outperforms Haar wavelets in both accuracy and speed.

In the field of computer graphics, illumination information is necessary for augmented reality [2], where virtual objects are inserted and seamlessly integrated into a real scene. The recovered lighting can give real world appearance and make images look realistic. Illumination recovery has been of great interest to the graphics community and there has been significant previous research [39], [40], [41], [43], [1], [26], [44]. In [39], a method is proposed to divide the scene into cells and for each cell split the lights into important and unimportant lists with the latter very sparsely sampled. Precomputed radiance functions (PRT) [40] are used to represent low-frequency lighting environments that capture soft shadows, and interreflections. Cluster principal component analysis (CPCA) [41] is used to compress the high dimensional surface signal formed from per-point transfer matrices recorded by PRT. A new data representation and compression technique for precomputed radiance transfer is introduced in [43] and the light transfer functions and light sources are modeled with Spherical Radial Basis Functions (SRBFs). The environment map is approximated in a wavelet basis, and the light transport matrix is sparsely and accurately encoded in the same basis [26]. In [29], [37], a framework for capturing a sparse approximation of the light transport based on the theory of compressive sensing using a small set of illumination patterns is proposed.

There are many other methods to recover illumination distributions from images; though cast shadows are not handled specifically. The complexity of determining lighting grows dramatically when we must account for cast shadows. A framework is proposed in [25] to accomplish photo-realistic view-dependent image synthesis from a sparse image set and a geometric model. Two methods are presented for recovering the light source position from a single image without the distant illumination assumption [12]. Much more accurate multiple illumination information is extracted from the shading of a sphere [47]. In [22], a framework is proposed to automatically recover an object shape, reflectance properties and light sources from a set of

images. A unified framework to estimate both distant and point light sources is proposed in [48]. The number of point light sources and the reflectance property of an object are simultaneously estimated using the EM algorithm [13]. In [18], synthetic 3D objects are inserted into the scene based on the estimated illumination from a single outdoor image. The method relies on a combination of weak cues extracted from different portions of the image such as sky, vertical surfaces, and the ground.

Our solution to the sparse representation is motivated by recent advances in the field of compressed sensing [5], [9]. A goal of compressed sensing is to exploit the compressibility and sparsity of the true signal, which is an ℓ_0 minimization problem that is usually hard to solve. Recent studies [5], [9] show that, under very flexible conditions, the ℓ_0 minimization can be reduced to ℓ_1 minimization that further results in a convex optimization, which can be solved efficiently. The results from compressed sensing have been applied to different computer vision tasks [45] for problems such as face recognition [46], background subtraction [15], [6], media recovery [10], visual tracking [21], texture segmentation and feature selection [19]. In this work, we show that the number of directional sources needed to approximate the lighting is greatly compressible and the illumination recovery can be cast as an ℓ_1-regularized least squares problem.

A preliminary conference version of this paper appears in [20].

3 An Example and Motivation

In this section, we design a simple but effective example, a flag pole scene, to illustrate the intuition and motivation of our work. Although the set of images produced by a scene with cast shadows can be of high dimension, it has been shown [31] that the number of directional lights needed to approximate the lighting is highly compressible and the perceptual loss from the image constructed by the recovered lighting is hardly noticeable.

In the example, we consider a scene consisting of a flat playground with an infinitely thin flag pole. We view the scene from directly above, so that the playground is visible, but the flag pole appears only as a negligible point. Suppose the scene is illuminated by an arbitrary set of directional lights of equal intensity that each has an elevation of 45 degrees. In this case, the intensity of the lighting can be described as a one-dimensional function of azimuth. A single directional light illuminates the playground to constant intensity except for a thin, black shadow on it. The entire set of lights can cause shadows in multiple directions. None of these shadows overlap, because the pole is infinitely thin.

Now consider the linear subspace spanned by the images that this scene can produce. We first consider the set of images that are each produced by a single directional source. All images are nonnegative, linear combinations of these. We represent each image as a vector. By symmetry, the mean of these images will be the constant image produced in the absence of cast shadows. Subtracting the mean, each image is near zero, except for a large negative component at the shadow. All these images have equal magnitude, and are orthogonal to each other. Therefore, they span an infinite-dimensional space, and Principal Component Analysis (PCA) will produce an infinite number of equally significant components. In other words, a finite-dimensional linear subspace cannot capture any significant fraction of the effects of cast shadows.

However, we can look at the images of this scene differently. A single directional source produces a single, black shadow (Figure 1(a)). Two sources produce two shadows (Figure 1(b)), but each shadow has half the intensity of the rest of the playground, because each shadow is lit by one of the lights. The more lights (e.g., Figure 1(c,d,e,f)) we have the more shadows we have, but the lighter these shadows are. In Figure 1(f) in which there are 32 DS, the shadows are barely noticeable. Therefore, while a small set of lights can produce strong cast shadows, many lights tend to wash out the effects of shadowing.

The above observation implies that, to model a scene with cast shadows, perceptually only a limited number of directional light sources are needed plus a diffusive component where no shadow is considered. In fact, if a scene is indeed illuminated by a limited number of light sources, then its shadows can be represented exactly by these light sources. If a scene is otherwise illuminated by a large number of directional sources, we cannot represent the shadows well with a few sources, but we do not need to, because they have only a small effect and the image is approximately constant.

4 Modeling Images with Cast Shadows

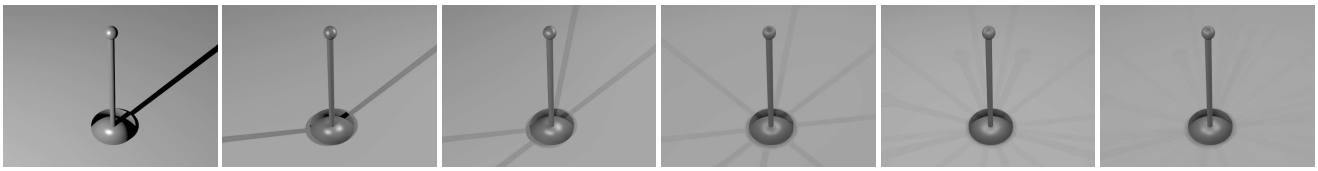
4.1 Lighting Model

We now model cast shadows in detail. We do not consider specular reflections, and in fact there is no reason to believe that sparse lighting can approximate the effects of full environment map lighting when there are significant specular reflections. For example, instead of our playground example, imagine images of a mirrored ball. Directional sources produce bright spots which do not get washed out as we add more directional sources. We also do not consider the effects of saturated pixels. We assume the geometry of the scene is given, as in many previous studies [4], [3], [36], [34], [35], [27], [33], [13], so we can render directional source images from it.

We model a scene as illuminated by light from all directions. Therefore, an image $I \in \mathbb{R}^d$ (we stack image columns to form a 1D vector) of a given scene has the following representation

$$I = \int_{\mathcal{S}} x(\theta) I_{dir}(\theta) d\theta \ , \quad x(\theta) \geq 0 \ , \tag{1}$$

where $\mathcal{S}$ is the unit sphere that contains all possible light directions, $I_{dir}(\theta)$ is the image generated by a directional light source with angle $\theta \in \mathcal{S}$, and $x(\theta)$ is the weight (or



(a) 1 DS (b) 2 DS (c) 4 DS (d) 8 DS (e) 16 DS (f) 32 DS

Fig. 1. A flagpole rendered with one directional source (a), two directional sources (b), four directional sources (c), eight directional sources (d), sixteen directional sources (e), and thirty-two directional sources (f). The shadows are lighter as the number of directional sources increases.

amount) of image $I_{dir}(\theta)$. Note that here, without loss of generality, an extremely simple camera model is assumed since our conclusion does not depend on camera models.

For practical reasons, integration over the continuous space $\mathcal{S}$ is replaced by a superposition over a large discrete set of samples, say $\{\theta_k\}_{k=1}^N$ with a large N. Denote the image generated by light from direction θ_k as $I_k = I_{dir}(\theta_k)$ and $x_k = x(\theta_k)$; we approximate with the discrete version of (1),

$$I = \sum_{k=1}^{N} x_k I_k \ , \quad x_k \geq 0. \tag{2}$$

4.2 Lighting Decomposition

It is known that in the absence of cast shadows, this lighting can be approximated using low frequency spherical harmonics [3], [30]. We use a nine-dimensional spherical harmonic subspace generated by rendering images of the scene, including their cast shadows, using lighting that consists of zero, first, and second order spherical harmonics. We will therefore divide the effects of these directional sources into low- and high-frequency components. We can then capture the low-frequency components exactly using a spherical harmonic basis. We will then approximate the high frequency components of the lighting using a sparse set of components that each represent the high frequency part of a single directional source.

We project the directional source image I_k onto the spherical harmonic subspace and it can be written as the sum of the projection image $\hat{I}_k$ and residual image $\tilde{I}_k$. Then Equation (2) can be written as:

$$I = \sum_{k=1}^{N} x_k (\hat{I}_k + \tilde{I}_k) \ , \quad x_k \geq 0. \tag{3}$$

We separate the low frequency component $\hat{I}_k$ from high frequency component $\tilde{I}_k$ and Equation 3 becomes:

$$I = \sum_{k=1}^{N} x_k \hat{I}_k + \sum_{k=1}^{N} x_k \tilde{I}_k \ , \quad x_k \geq 0. \tag{4}$$

We know that the low frequency component $\sum_{k=1}^{\infty} x_k \hat{I}_k$ lies in a low dimensional subspace and can be approximated using $\hat{I}$ by simply projecting I to the spherical harmonic subspace. Equation (4) can be written as:

$$I = \hat{I} + \sum_{k=1}^{N} x_k \tilde{I}_k \ , \quad x_k \geq 0. \tag{5}$$

$\hat{I}$ is simply the component of the image due to low-frequency lighting, where we solve for this component exactly using the method of [3]. We then approximate the high frequency components of the lighting using a sparse set of values for x_k. Note that these components will be reflected only in the cast shadows of the scene, and we expect that when these cast shadows are strong, a sparse approximation will be accurate.

Our problem is now reduced to finding a certain number of x_k's that best approximate the residual image $\tilde{I} = I - \hat{I}$. It can be addressed as a least squares (LS) problem with nonnegativity constraints:

$$\arg\min_{\mathrm{x}} ||A\mathrm{x} - \tilde{I}||^2, \quad x_k \geq 0 \ , \tag{6}$$

where $A = [\tilde{I}_1 \tilde{I}_2 \cdots \tilde{I}_N]$ and $\mathrm{x} = (x_1, ..., x_N)^\top$. To avoid ambiguity, we assume all the residual directional source images $\tilde{I}_k$ are normalized, i.e., $||I_k||_2 = 1$.

4.3 Low Dimensional Approximation

In practice, the size of the image can be very large, which corresponds to a large linear system $Ax = \tilde{I}$, which is very computationally expensive to solve. To alleviate the problem, we use low dimensional approximation by applying PCA to the image set A and $\tilde{I}$. The standard PCA yields from A a mean vector $\mu \in \mathbb{R}^d$ and a projection matrix $W \in \mathbb{R}^{m \times d}$, which consists of the m most important principal components of A.

Applying W and μ to equation 6, we have the following approximation:

$$\begin{aligned} \hat{x} &= \arg\min_{\mathrm{x}} ||W(A - \mu \otimes \mathbf{1}^\top)\mathrm{x} - W(\tilde{I} - \mu)||^2 \\ &\doteq \arg\min_{\mathrm{x}} ||A^*\mathrm{x} - \tilde{I}^*||^2, \quad x_k \geq 0 \ , \end{aligned} \tag{7}$$

where $\otimes$ denotes the cross product, $\mathbf{1} \in \mathbb{R}^{N \times 1}$ is a vector of 1's, $A^* \doteq W(A - \mu \otimes \mathbf{1}^\top)$ and $\tilde{I}^* \doteq W(\tilde{I} - \mu)$ denote the PCA transformation of A and $\tilde{I}$ respectively.

The dimension m is typically chosen to be much smaller than d. This is also a requirement for the sparse representation problem we introduce in the next section that the number of observations is smaller than the number of unknowns. The PCA plays an important role such that the system becomes underdetermined and hence satisfies the condition of the sparse representation problem. In this case, the system (7) is underdetermined in the unknown x and simple least squares regression leads to over-fitting.

5 ℓ_1-Regularized Least Squares

A standard technique to prevent over-fitting is the ℓ_2 or Tikhonov regularization [23], which can be written as

$$\arg\min_{\mathrm{x}} ||A^*x - \tilde{I}^*||^2 + \lambda||\mathrm{x}||_2, \quad x_k \geq 0. \tag{8}$$

where $||\mathrm{x}||_2 = (\sum_{k=1}^{N} x_k^2)^{1/2}$ denotes the ℓ_2 norm of x and $\lambda > 0$ is the regularization parameter.

The ℓ_2 regularization is known to be robust to Gaussian like errors. However, in our formulation, we are concerned with the problem of low-complexity recovery of the unknown vector x. This can be formulated with an ℓ_0-regularization,

$$\arg\min_{\mathrm{x}} ||A^*x - \tilde{I}^*||^2 + \lambda||\mathrm{x}||_0, \quad x_k \geq 0, \tag{9}$$

in which we are interested in minimizing the ℓ_0 norm that counts the number of non-zero elements in x. This formulation suffers unfortunately from impractical computational cost. Instead, recently study shows that it can be well approximated by using ℓ_1 minimizations under fairly flexible conditions [9], [5].

Motivated by the above observation, we exploit the compressibility in the transform domain by solving the problem as the ℓ_1-regularized least squares problem. We substitute a sum of absolute values for the ℓ_2 and ℓ_0 norms used in the above regularizations:

$$\arg\min_{\mathrm{x}} ||A^*\mathrm{x} - \tilde{I}^*||^2 + \lambda||\mathrm{x}||_1, \quad x_k \geq 0. \tag{10}$$

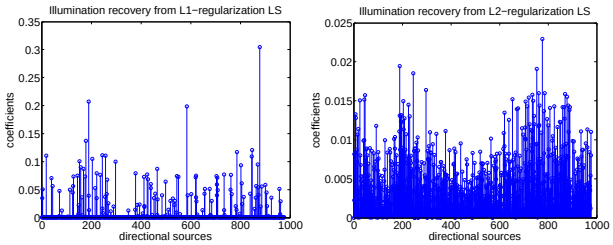
where $||\mathrm{x}||_1 = \sum_{k=1}^{N} |x_k|$ denotes the ℓ_1 norm of x and $\lambda > 0$ is the regularization parameter. This problem always has a solution, though not necessarily unique. ℓ_1-regularized least squares (LS) typically yields a sparse vector x, which has relatively few nonzero coefficients. In contrast, the solution to the Tikhonov regularization problem generally has all coefficients nonzero.

Since x is non-negative, the problem (10) can be reformulated as

$$\arg\min_{\mathrm{x}} ||A^*\mathrm{x} - \tilde{I}^*||^2 + \lambda \sum_{k=1}^{N} x_k, \quad x_k \geq 0. \tag{11}$$

Figure 2 shows the recovered coefficients x using ℓ_1-regularized LS and ℓ_2-regularized LS algorithms respectively for the synthetic image rendered with the light probe in Figure 3 (a). The query image is approximated using N=977 directional source images. The parameter λ's are tuned such that the two recoveries have similar errors. The results show that ℓ_1 regularization gives a much sparser representation, which fits our expectation.

Algorithm 1 summarizes the whole illumination recovery procedure. Our implementation solves the ℓ_1-regularized least squares problem via an interior-point method based on [17]. The method uses the preconditioned conjugate gradients (PCG) algorithm to compute the search direction and the run time is determined by the product of the total number of PCG steps required over all iterations and the cost of a PCG step. We use the code from [7] for the minimization task in (11).



(a) (b)

Fig. 2. The recovered coefficients x from ℓ_1-Regularized LS (a) and ℓ_2-Regularized LS (b).

Algorithm 1 Sparse representation for inverse lighting

Require: Input image I and the geometry of the scene I is generated from.

1: Obtain N directional source images $\{I_1, I_2, \ldots, I_N\}$ by rendering the scene with N directional light sources uniformly sampled from the upper hemisphere (object is put on a plane and there is no light coming from beneath).
2: Create the first nine spherical harmonic images by integrating the N directional source images.
3: Project each directional source image I_k to the 9D spherical harmonic subspace and obtain the corresponding residual directional source image $\{\tilde{I}_1, \tilde{I}_2, \ldots, \tilde{I}_N\}$.
4: Normalize $\tilde{I}_k$ such that $\|\tilde{I}_k\|_2 = 1$ and form matrix $A = [\tilde{I}_1 \tilde{I}_2 \cdots \tilde{I}_N]$.
5: Project the query image I to the spherical harmonic subspace and obtain the residual image $\tilde{I}$.
6: Apply PCA to A and obtain the projection matrix W and the mean vector μ.
7: $A^* \leftarrow W(A - \mu \otimes \mathbf{1}^\top)$ and $\tilde{I}^* \leftarrow W(\tilde{I} - \mu)$.
8: Solve the ℓ_1-regularized least squares problem with nonnegativity constraints (11).
9: Render the scene with the spherical harmonic lighting plus the recovered sparse set of directional light sources.

6 Experiments

In this section, we conduct two kinds of experiments: illumination recovery and illumination transfer on both synthetic and real data, in comparison with four other previous approaches. The illumination recovery experiments are recovering the lighting from the images and rendering the scene with the lighting recovered. In the illumination transfer experiments, we apply the recovered illumination from one scene to the other scene and vice versa.

6.1 Experimental Setup

6.1.1 Data

Both synthetic and real datasets are used in our experiments to validate the performance of our proposed method. For synthetics data, we use two synthetic scenes with four

different lighting environment maps. For real data, we use photos of three real objects printed by using a 3D printer from pre-designed CAD models. Details of these datasets are described in next subsections.

6.1.2 Methods for Comparison

We compare our proposed algorithms with four previous proposed methods: Spherical Harmonics [3], [30], Non-Negative Least squares (NNL) [4], Semidefinite Programming (SDP) [33], and Haar wavelets [27] algorithms. To make this algorithm comparable, we use about 100 directional sources to represent the lighting which is obtained from thousands of possible directional sources. The reason that we choose 100 sources is because by examining the coefficients for all the experiments, we find that the coefficients become zero or very small after 100 sources. The λ in Equation (11) is set to 0.01 for all the experiments. We experimentally choose 0.01 for λ that achieves the best approximation results.

In the following, we briefly review and analyze the four algorithms that are used for comparison.

Spherical Harmonics (SH): In [3], [30], it is shown that the image without cast shadows is well approximated by a linear combination of the first nine harmonics images.

Non-Negative Linear (NNL): In [4], it is shown that the set of images of an object produced by nonnegative lighting is a convex cone in the space of all possible images. Given an image I, the method attempts to minimize $\|Ha - I\|$ subject to $a >= 0$ where H is the matrix whose columns are directional source images. If we densely sample the illumination distribution, it makes the solution exceedingly expensive in terms of processing time and storage requirements because of the high dimensionality of the matrix H formed by point source images.

Semidefinite Programming (SDP): In [33], semidefinite programming is applied to perform a constrained optimization to quickly and accurately solve for the non-negative linear combination of spherical harmonics. It has been applied for specular object recognition on both synthetic and real data by better separating the correct and incorrect models. Our work is the first to show that SDP can also be used to handle shadows.

In [33], SDP is designed to approximate high frequency signals that cannot be captured by the 9D spherical harmonics. It works well on specular objects such as a shiny rubber ball and a ceramic shaker using harmonics up to 10th order. The total harmonics used in SDP is $(10+1)^2 = 121$. Since images with cast shadows generally have a lot of high frequency signals, it still misses a certain amount of information which is contained in higher order harmonics.

Haar Wavelets (Haar): In [27], spherical harmonics is compared to Haar wavelets as a basis for recovery from shadows. The Haar wavelets approach is similar since Haar wavelets also form an orthonormal basis. The advantages pointed out are the compact supports with various sizes which allow different resolutions in different regions.

In this method, the illumination distribution is mapped to a cube and two-dimensional Haar wavelet basis elements are used in each face of the cube. Similar to the spherical harmonics, the illumination distribution is represented by a linear combination of the wavelet basis functions $\Phi_i(\theta, \phi)$ and $\Psi_{ijkl}(\theta, \phi)$

$$L(\theta, \phi) = \sum_i (c_i \Phi_i(\theta, \phi) + \sum_{j,k,l} d_{i,j,k,l} \Psi_{ijkl}(\theta, \phi)) \quad (12)$$

where c_i and $d_{i,j,k,l}$ are coefficients of the corresponding basis functions. The coefficients are computed using a constrained least squares estimation which constrains the resulting distribution to be position.

6.1.3 Evaluation Criterion

Accuracy. To evaluate the accuracy of different algorithms, we use the Root-Mean-Square (RMS) errors of pixel values, which is also used in [27]. Specifically, for an input image $I \in \mathbb{R}^d$ and its recovery $I' \in \mathbb{R}^d$, the RMS between them is defined as $r(I, I') = \sqrt{\|I - I'\|_2^2 / d}$, where d is the number of pixels in the image.

Run Time. We divide the run time for illumination recovery into three parts: (1) preprocessing time, (2) time for solving the lighting recovery algorithm (e.g., solving the ℓ_1-regularized LS, SDP), and (3) rendering the scene with recovered lighting. First, part (1) can be done off-line and is actually similar for all methods. In fact, preprocessing time is dominated by the time for generating images using different directional light sources (via PovRay for all experiments). These images are pre-computed off-line, and are actually used in all methods[2]. Second, part (3) is usually much faster than the other two parts and therefore can be ignored. For the above reasons, in the paper we focus only on the part (2), which measures the time efficiency of different illumination recovery approaches.

All the algorithms were run in MATLAB 7.4.0. The computer used was a laptop with Intel Core Duo at 1.73GHz with 2.0GB RAM.

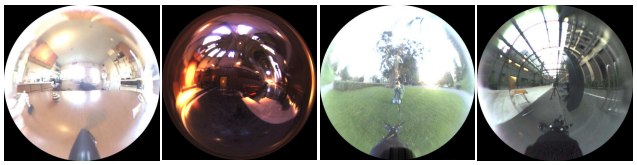
6.2 Experiments with Synthetic Data

6.2.1 Synthetic Images

There are two synthetic scenes used in our experiments. One is composed of a coffee cup and a spoon, with a plate underneath them (see Figure 5). The teacup and spoon will cast shadows on the plate when the light comes from certain directions. The other scene consists of one table and four chairs (see Figure 6). The legs of table and chairs are thin and will cast shadows on the ground which poses challenges for the illumination recovery.

Four synthetic images for each scene are obtained by rendering the scene with high dynamic environment maps (namely *kitchen*, *grace*, *campus*, and *building*, see Figure 3) provided by [8]. We consider a scene where the objects are placed on an infinite plane, so only lights coming from the upper hemisphere are taken into account.

2. The spherical harmonics and Haar wavelets also need these images for basis image estimation.



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

Fig. 3. Light probes [8] used to generate our synthetic dataset: (a) kitchen, (b) grace, (c) campus, and (d) building. The light probes are sphere maps and shown in low-dynamic range for display purposes.

6.2.2 Experiments

Using the POV-Ray ray tracer we generate directional images, each using a single directional light source. We obtain directions by uniformly sampling the upper hemisphere, and each direction represents a directional light. We generate directional source images each of which is rendered with one directional light. Using these images, we numerically integrate to compute nine harmonic images of the scene, each with lighting consisting of a single spherical harmonic. The coefficient of the integrated image is the intensity value in the spherical harmonic function at that direction.

Figure 4 shows the first nine harmonic images created from more than three thousand directional source images derived from a 3D model of one teacup (top), and a table with four chairs (bottom).

The recovery results of teacup images and table-chair images are shown in Figure 5 and Figure 6 respectively.

From the results we can see that spherical harmonics by itself fails to capture the apparent shadows cast on the plate and ground. For NNL, we tested two versions, using the 100 and 300 largest DS respectively. There are 977 DS used to approximate the lighting and the best 100 and 300 ones are picked, so these two versions have the same processing time. The reason to use 300 DS is, as illustrated in Figure 5, NNL with 100 DS fails to generate a reasonable result. This tells us the results of NNL is not sparse and require a large number of directional sources in order to produce good results. Comparing with spherical harmonics, SDP captures more details of the cast shadows, but the shadows are very fuzzy and the shadow boundaries are unclear. For the wavelet based method, we render the image with 102 Haar basis functions as in [27]. Both Haar wavelets and our method reproduce the shadows reliably.

In the illumination transfer experiments, we apply the recovered illumination from *teacup* to the model of *table-chair* and vice versa. Then we compare the results to the ground truth images. These tests are similar to those used for lighting recovery in [27]. The results are shown in Figure 7.

To quantitatively evaluate the performance of the methods in terms of speed and accuracy, we measure the quality of the approximation by looking at RMS and the speed by run time. The errors in pixel values and run time in seconds are shown in Table 1 and Table 2. One can find that the error in our method is the smallest of all the listed methods and the run time is much smaller than the Haar wavelets method which has comparable accuracy to our method. Table 3 shows the illumination transfer results. The ground truth images are rendered with high dynamic environment map *kitchen*. Our method has much smaller errors than the other methods. Therefore, our method works best for recovering illumination from cast shadows in terms of both accuracy and speed.

TABLE 1
RMS errors and average run time on the synthetic teacup images. Note: the running time does not include the preprocessing for generating "basis" images (same for all the rest tables).

Method	Probe Kitchen	Probe Grace	Probe Building	Probe Campus	Avg. Run Time (sec.)
SH	8.00	12.23	12.21	7.39	0.01
NNL (100 DS)	55.74	17.31	39.41	74.87	9389.8
NNL (300 DS)	5.96	2.80	1.87	12.50	9389.8
SDP	3.21	4.11	3.48	1.26	10.9
Haar Wav.	3.42	3.12	1.61	0.96	1322.0
Proposed	2.33	2.69	1.22	1.09	11.8

TABLE 2
RMS errors and average run time on the synthetic table-chair images.

Method	Probe Kitchen	Probe Grace	Probe Building	Probe Campus	Avg. Run Time (sec.)
SH	11.09	9.44	9.36	8.64	0.01
NNL (100 DS)	15.26	10.91	10.84	12.20	9410.5
NNL (300 DS)	8.72	10.90	10.73	9.51	9410.5
SDP	7.29	8.36	7.62	6.49	11.2
Haar Wav.	5.92	4.04	4.43	5.20	1332.0
Proposed	5.71	3.95	4.37	4.90	12.0

TABLE 3
Illumination transfer errors (RMS) on synthetic images.

Method	teacup to table-chair (top row Fig. 7)	table-chair to teacup (bottom row Fig. 7)
SH	43.09	52.25
NNL (100 DS)	57.21	18.93
NNL (300 DS)	15.42	13.07
SDP	18.89	59.47
Haar	18.09	15.93
Proposed	7.26	5.60

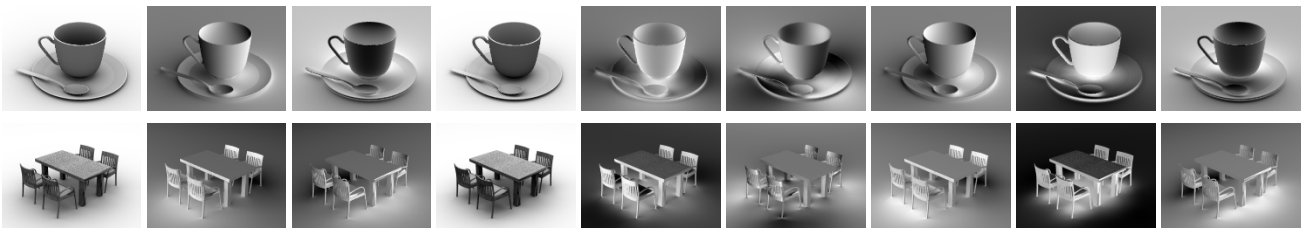


Fig. 4. The first nine harmonic images from a 3D model of one teacup (top) and a table with four chairs (bottom). The left most column contains the zeroth harmonic, the second to fourth columns contain the three first order harmonic images, and the rest are the images derived from the second harmonics (same for Figure 9). Image values are scaled to the full range of the intensity (same for Figure 9).

(a) Truth (b) SH (c) NNL1 (d) NNL2 (e) SDP (f) Haar (g) Proposed

Fig. 5. Experiments on synthetic images rendered from the teacup scene. The left column shows the ground truth images from different lighting environment probes shown in Figure 3 (*kitchen*, *grace*, *campus* and *building* from top to bottom, respectively). Columns (b) to (g) show results from different methods. Note: NNL1 uses 100 directional lighting sources and NNL2 uses 300.

6.3 Experiments with Real Data

6.3.1 Real Data

For the real objects, we built CAD models of three objects (namely *chair1*, *chair2*, and *couch*, see Figure 8 (a-b)) and printed them with a 3D printer (see Figure 8 (c) for results). The only difference between chair1 and chair2 is the number of backrest bars. The 3D printer we use is manufactured by Z Corporation with the model ZPrinter 310 Plus. We briefly explain how a 3D object is created from a 3D printer. First, a 3D CAD file is imported into the system software. The software slices the file into thin cross-sectional slices, which are fed to the 3D printer. Second, the printer creates the model one layer at a time by spreading a layer of powder and inkjet printing a binder in the cross-section of the part. Finally, the process is repeated until every layer is printed and the part is complete and ready to be removed.

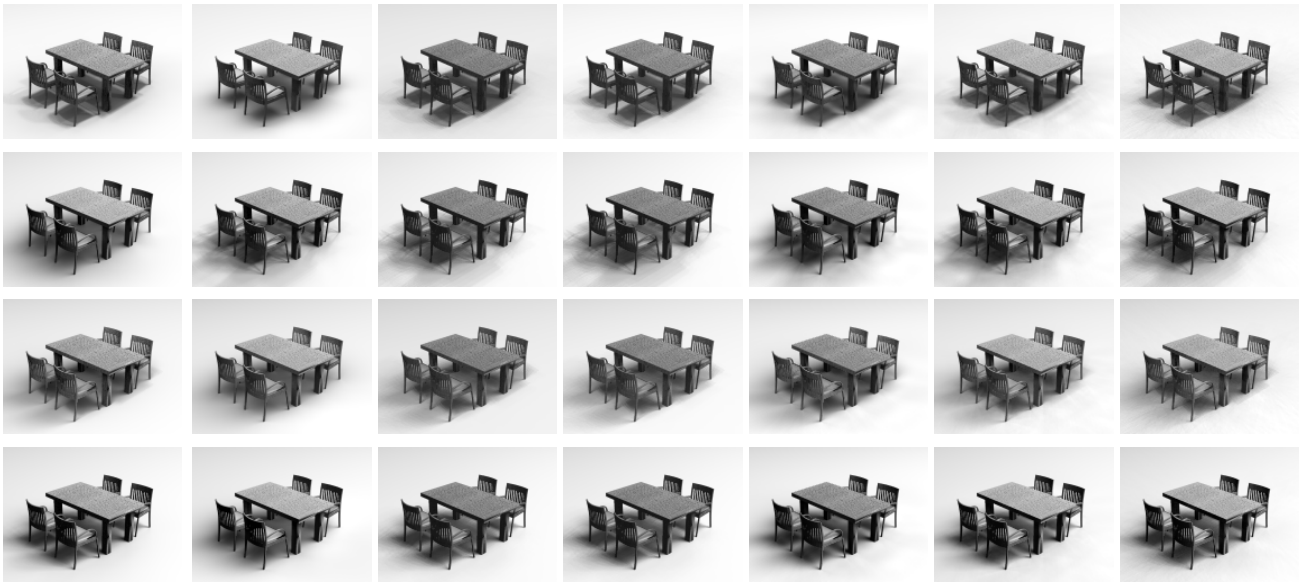
After the printing out the objects, we place them under natural indoor illumination and take pictures of them using a Canon EOS Digital Rebel XT camera. The images are then used in our experiments. Figures 10 and 11) show these images.

One of our experiments involves recovering lighting from one object (chair1), and using it to render a model of a second object (chair2) (as in [27]). For this reason, we take pictures of chair1 and chair2 in exactly the same illumination environment.

6.3.2 Registration

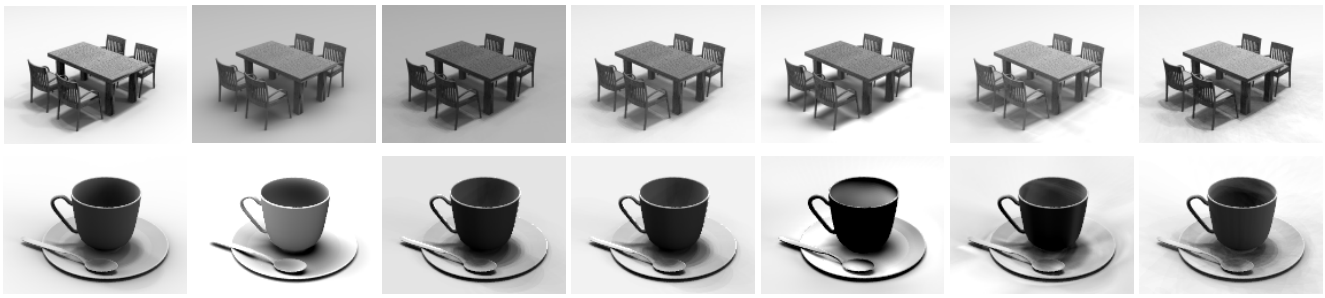
After we take pictures of the real objects, we need to register the object in the picture to the 3D model. We select the feature points (see Figure 8) and match them from the picture to the 3D model. For *chair1* and *chair2*, eight corner points are selected as feature points. For *couch*, there is no apparent corner points on the model. We add seven small balls with center on the couch whose center are used as feature points for the registration.

To do the registration, the object is first rotated and trans-



(a) Truth (b) SH (c) NNL1 (d) NNL2 (e) SDP (f) Haar (g) Proposed

Fig. 6. Experiments on synthetic images rendered from the table-chair scene. The left column shows the ground truth images from different lighting environment probes shown in Figure 3 (*kitchen*, *grace*, *campus* and *building* from top to bottom, respectively). Columns (b) to (g) show results from different methods. Note: NNL1 uses 100 directional lighting sources and NNL2 uses 300.



(a) Ground truth (b) SH (c) NNL1 (d) NNL2 (e) SDP (f) Haar (g) Proposed

Fig. 7. Illumination transfer experiments on synthetic images. The ground truth images are rendered with high dynamic environment map *kitchen*. Top row: transfer the illumination of *teacup* to *table-chair*. Bottom row: transfer the illumination of *table-chair* to *teacup*. Column (a): ground truth images. Columns (b–g): the images rendered with the lighting recovered using different methods. Note: NNL1 uses 100 directional lighting sources and NNL2 uses 300.

lated to the camera coordinate system and then projected onto the image plane. We use a simplified pinhole camera model [14] with six parameters to do the registration. The objects are registered to the images by minimizing the distance between the feature points on the image and the corresponding feature points from the 3D model.

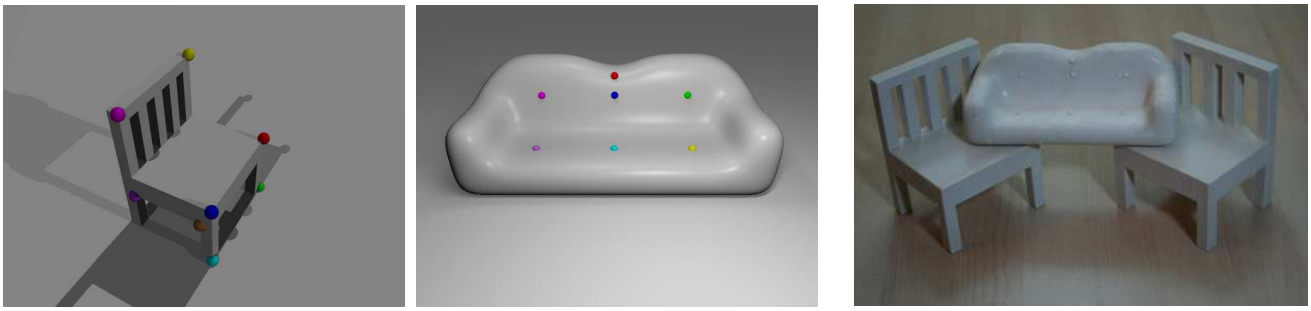
6.3.3 Experiments

In the illumination recovery experiments, we apply all the algorithms to real images of *chair1*, *chair 2* and *couch*. The recovered images are shown in Figure 10.

In the illumination transfer experiments, we apply the recovered illumination from *c*hair1 to the model of *c*hair2 and vice versa. Then we compare the results to the ground truth images. The results are shown in Figure 11.

Figure 9 show the first nine harmonic images created from more than three thousand directional source images derived from a 3D model of chair1.

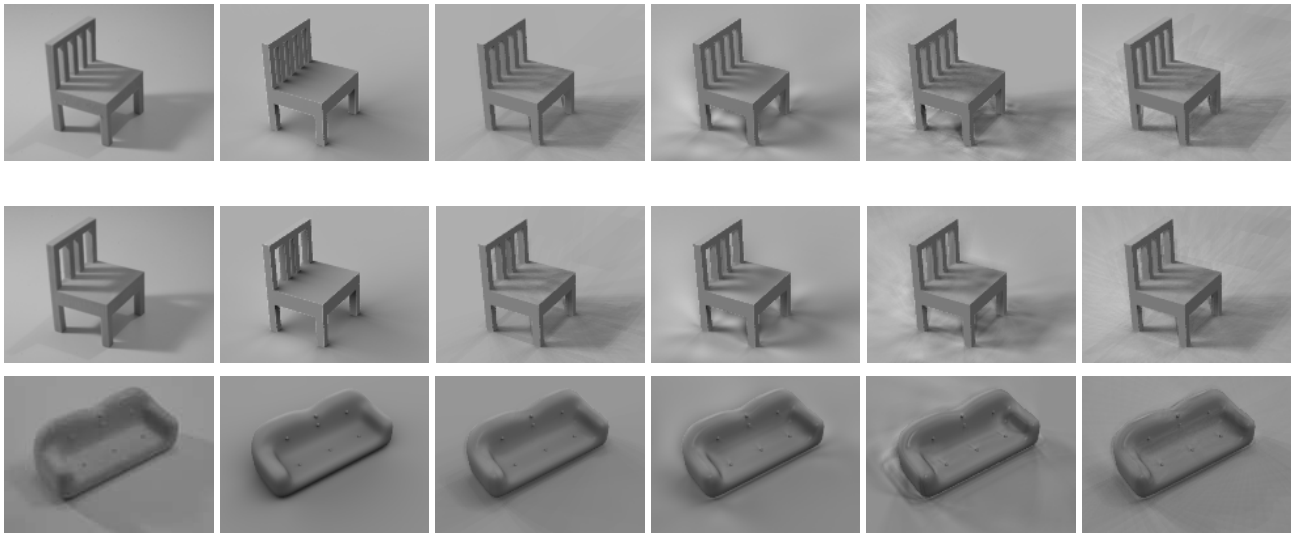
By visually checking the results in Figures 10 and 11, we can see that all these experiments show the superiority of our methods. Spherical harmonics fail to capture the apparent shadows cast on the seat of the chair and the ground. In comparison, SDP captures more details of the cast shadows, but the shadows are very fuzzy and there are some highlights on the ground. NNL can produce accurate shadows, but the shadows are intersecting and overlapping each other, causing the image to be unrealistic to the user. The Haar wavelets method produces accurate shadows, but there are some highlights on the ground. Our method generates visually realistic images and produces accurate



(a) (b) (c)

Fig. 8. (a-b): Two of the three 3D models used for generating real objects. The marked feature points in color are used for registration. (c): The three real objects printed out from 3D models by a 3D printer, from left to right: *chair1*, *couch* and *chair2*.

Fig. 9. The first nine harmonic images created from more than three thousand directional source images derived from a 3D model of the chair1.



(a) Ground truth (b) SH (c) NNL (d) SDP (e) Haar (f) Proposed

Fig. 10. Illumination recovery experiments on real images. (a) Ground truth images. (b)-(f) show the image rendered with the lighting recovered from (a) using different approaches, where (c) and (f) use 100 directional sources, and (e) uses 102 wavelet basis.

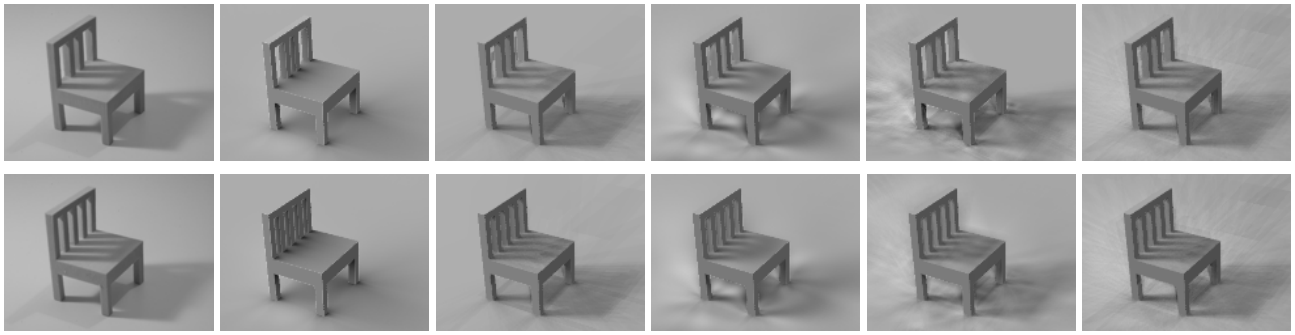
shadows both on the seat and the ground.

The quantitative evaluation, including the RMS and running times, is summarized in Table 4 for illumination recovery experiments and Table 5 for illumination transfer experiments. In both experiments, our method achieves the smallest error of all the methods in only tens of seconds run time.

6.3.4 Discussion

There is specular texture on the ground plane in the images. This is due to several reasons. The ground plane is a big white table. It may not be ideally Lambertian, and has specular component on the surface. Another reason is the object in the image to the 3D model registration. We use a simplified pinhole camera model with six parameters to do the registration. An ideal white Lambertian ground plane and more advanced registration method would help reduce the specular texture in the image. Note that these problems do not exist in synthetic experiments, where no specular texture on the ground is observed.

We also conducted experiments using random projection matrix as suggested in [46]. The entries of the projection matrix are independently sampled from a zero-mean normal distribution, and each row is normalized to unit length.



(a) Ground truth (b) SH (c) NNL (d) SDP (e) Haar (f) Proposed

Fig. 11. Illumination transfer experiments on real images. Top row: transfer the illumination of *chair1* to *chair2*. Bottom row: transfer the illumination of *chair2* to *chair1*. Column (a): ground truth images. Columns (b–f): the images rendered with the lighting recovered from chair 1 (top of (a)) using different methods. The method in (c,d,f) use 100 directional light sources, while that in (e) uses 102 sources.

TABLE 4
Illumination recovery errors (RMS) and run times (in seconds) on real images.

Method	*chair1* RMS	*chair1* run time	*chair2* RMS	*chair2* Run time	*couch* RMS	*couch* Run time
SH	13.99	0.01	13.77	0.01	9.39	0.01
NNL	10.26	1854.89	9.62	1826.67	7.37	2050.22
SDP	9.38	10.88	8.76	11.30	7.01	14.62
Haar	10.75	1529.60	7.87	1526.34	7.84	1585.27
Proposed	7.50	14.54	7.79	14.82	6.56	13.82

TABLE 5
Illumination transfer errors (RMS) on real images.

Method	Chair1 to Chair2 (top row Fig. 11)	Chair2 to chair 1 (bottom row Fig. 11)
SH	15.31	15.51
NNL	10.35	11.05
SDP	9.40	10.09
Haar	11.02	9.03
Proposed	8.24	8.58

(a) teacup (b) chair1

Fig. 12. Illumination recovery experiments using random projection matrix.

Figure 12 shows the illumination recovery results using random projection matrix. The RMS errors are 4.86 for the *teacup* and 10.43 for *chair1*, respectively. The results are less desirable than the ones using PCA.

6.4 Sparsity Evaluation

In the previous section, we argue that we can approximate the query image well using a sparse set of directional light sources. To justify our argument, we conduct experiments on synthetic and real images. Figure 13 shows the RMS versus number of possible directional sources for synthetic images rendered with the grace light probe (left) and a real image (right) under natural indoor lighting. The accuracy improves gradually as the number of directional sources increases. From the plots, we can see after a certain number of directional sources (≈ 50 for the left and ≈ 180 for the right), the error remains constant. It matches the argument that we can approximate the query image well enough using only a sparse set of directional sources and after a certain number of directional sources, increasing the number of directional sources does not improve the accuracy.

7 Conclusions

In this paper, we start from a simple example and explain that although the dimensionality of the subspace of images with cast shadows can go up to infinity, the illumination can still be well approximated by a sparse set of directional sources. Following this example, we derive a theoretical model and cast illumination recovery as an ℓ_1-regularized least squares problem. An efficient and fast solution is provided to find the most significant directional sources for the estimation. Experiments on both synthetic and real

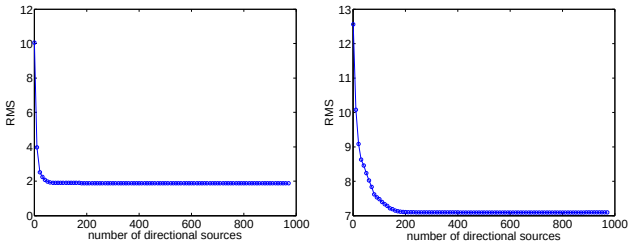


Fig. 13. The improvement in accuracy by adding directional sources. RMS versus number of directional sources for a synthetic image (*teacup*, second row of Figure 5) rendered with grace light probe (left figure) and a real image (*chair2*, middle left of Figure 10) under natural indoor lighting (right figure).

images have shown the effectiveness of our method in both accuracy and speed.

Acknowledgement: Haibin Ling is supported in part by the National Science Foundation (NSF) under the grants IIS-0916624 and IIS-1049032. David Jacobs is supported in part by the Office of Naval Research under the MURI Grant N00014-08-10638.

REFERENCES

[1] S. Agarwal, R. Ramamoorthi, S. Belongie, and H. W. Jensen. "Structured importance sampling of environment maps", *SIGGRAPH*, 22(3):605-612, 2003. 2

[2] R. Azuma. "A survey of augmented reality. In Com puter Graphics", *SIGGRAPH 95 Proceedings, Course Notes 9: Developing Advanced Virtual Reality Applications*, 1995. 2

[3] R. Basri and D. Jacobs. "Lambertian Reflectances and Linear Subspaces", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(2):218-233, 2003. 1, 2, 3, 4, 6

[4] P. Belhumeur and D. Kriegman. "What is the Set of Images of an Object Under All Possible Lighting Conditions?", *Int'l J. Computer Vision*, 28(3):245-260, 1998. 1, 2, 3, 6

[5] E. Candès, J. Romberg, and T. Tao. "Stable signal recovery from incomplete and inaccurate measurements", *Comm. on Pure and Applied Math*, 59(8):1207-1223, 2006. 1, 3, 5

[6] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa. "Compressive Sensing for Background Subtraction", *European Conf. Computer Vision*, 2008. 3

[7] *http://www.stanford.edu/˜boyd/l1_ls/*. 5

[8] P. Debevec. "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography", *SIGGRAPH*, 189-198, 1998. 6, 7

[9] D. Donoho. "Compressed Sensing", *IEEE Trans. Inf. Theory*, 52(4):1289-1306, 2006. 1, 3, 5

[10] J. Gu, S. Nayar, E. Grinspun, P. Belhumeur, and R. Ramamoorthi. "Compressive Structured Light for Recovering Inhomogeneous Participating Media", *European Conf. Computer Vision*, 845-858, 2008. 3

[11] P. Hallinan. "A Low-dimensional Representation of Human Faces for Arbitrary Lighting Conditions", *IEEE Conf. Computer Vision and Pattern Recognition*, 995-999, 1994. 2

[12] K. Hara, K. Nishino, and K. Ikeuchi. "Determining reflectance and light position from a single image without distant illumination assumption", *IEEE Int'l Conf. Computer Vision*, 1:560-567, 2003. 2

[13] K. Hara, K. Nishino, and K. Ikeuchi. "Mixture of Spherical Distributions for Single-View Relighting", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 30(1):25-35, 2008. 1, 3

[14] R. Hartley, and A. Zisserman. "Multiple View Geometry in Computer Vision", *Cambridge University Press*, March, 2004. 9

[15] J. Huang, X. Huang, and D. Metaxas. "Learning With Dynamic Group Sparsity", *IEEE Int'l Conf. Computer Vision*, 2009. 3

[16] T. Kim, and K. Hong. "A practical single image based approach for estimating illumination distribution from shadows", *IEEE Int'l Conf. Computer Vision*, 266-271, 2005. 2

[17] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. "A method for large-scale ℓ_1-regularized least squares", *IEEE J. on Selected Topics in Signal Processing*, 1(4):606-617, 2007. 5

[18] J. Lalonde, A. A. Efros, and S. G. Narasimhan. "Estimating Natural Illumination from a Single Outdoor Image", *IEEE Int'l Conf. Computer Vision*, 2009. 3

[19] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. "Discriminative learned dictionaries for local image analysis", *IEEE Conf. Computer Vision and Pattern Recognition*, 2008. 3

[20] X. Mei, H. Ling, and D.W. Jacobs. "Sparse Representation of Cast Shadows via ℓ_1-Regularized Least Squares", *IEEE Int'l Conf. Computer Vision*, 2009. 3

[21] X. Mei and H. Ling. "Robust Visual Tracking using ℓ_1 Minimization", in *IEEE Int'l Conf. Computer Vision*, 2009. 3

[22] B. Mercier, D. Meneveaux, and A. Fournier. "A Framework for Automatically Recovering Object Shape, Reflectance and Light Sources from Calibrated Images", *Int'l J. Computer Vision*, 73(1):77-93, 2007. 2

[23] A. Neumaier."Solving ill-conditioned and singular linear systems: A tutorial on regularization", *SIAM Rev.*, 40(3):636666, 1998. 5

[24] J. Nimeroff, E. Simoncelli, and J. Dorsey. "Efficient Re-rendering of Naturally Illuminated Environments", *Eurographics Workshop on Rendering*, 1994. 2

[25] K. Nishino, Z. Zhang, and K. Ikeuchi. "Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis", *IEEE Int'l Conf. Computer Vision*, 1:599-606, 2001. 2

[26] R. Ng, R. Ramamoorth, and P. Hanrahan. "All-Frequency Shadows Using Non-linear Wavelet Lighting Approximation", *SIGGRAPH*, 2003. 2

[27] T. Okabe, I. Sato, and Y. Sato. "Spherical Harmonics vs. Haar Wavelets: Basis for Recovering Illumination from Cast Shadows", *IEEE Conf. Computer Vision and Pattern Recognition*, 50-57, 2004. 1, 2, 3, 6, 7, 8

[28] A. Panagopoulos, D. Samaras, and N. Paragios. "Robust Shadow and Illumination Estimation Using a Mixture Model", *IEEE Conf. Computer Vision and Pattern Recognition*, 2009. 2

[29] P. Peers, D. Mahajan, B. Lamond, A. Ghosh, W. Matusik, R. Ramamoorthi, and P. Debevec. "Compressive light transport sensing", *ACM Trans. on Graphics*, 28(1), 2009. 2

[30] R. Ramamoorthi and P. Hanrahan. "On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object", *JOSA A*, 10:2448-2459, 2001. 2, 4, 6

[31] R. Ramamoorthi and P. Hanrahan. "A Signal-Processing Framework for Inverse Rendering", *SIGGRAPH*, 1:117-128, 2001. 2, 3

[32] R. Ramamoorthi, M. Koudelka, and P. Belhumeur. "A Fourier Theory for Cast Shadows", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(2):288-295, 2005. 1, 2

[33] S. Shirdhonkar and D. Jacobs. "Non-Negative Lighting and Specular Object Recognition", *IEEE Int'l Conf. Computer Vision*, 1323-1330, 2005. 1, 3, 6

[34] I. Sato, Y. Sato, and K. Ikeuchi. "Stability Issues in Recovering Illumination Distribution from Brightness in Shadows", *IEEE Conf. Computer Vision and Pattern Recognition*, 400-407, 2001. 1, 2, 3

[35] I. Sato, Y. Sato, and K. Ikeuchi. "Illumination Distribution from Shadows", *IEEE Conf. Computer Vision and Pattern Recognition*, 306-312, 1999. 1, 2, 3

[36] I. Sato, Y. Sato and K. Ikeuchi. "Illumination from Shadows", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(3):290-300, 2003. 1, 2, 3

[37] P. Sen and S. Darabi. "Compressive Dual Photography", *Eurographics*, 28(2), 2009. 2

[38] A. Shashua. "On Photometric Issues in 3d Visual Recognition From a Single 2d Image", *Int'l J. Computer Vision*, 21(1-2):99-122, 1997. 2

[39] P. Shirley, C. Wang, and K. Zimmerman. "Monte carlo techniques for direct lighting calculations", *ACM Trans. on Graphics*, 15(1):1-36, 1996. 2

[40] P. Sloan, J. Kautz, and J. Snyder. "Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments", *SIGGRAPH*, 2002. 2

[41] P. Sloan, J. Hall, J. Hart, and J. Snyder. “Clustered principal components for precomputed radiance transfer”, *SIGGRAPH*, 382-391, 2003. 2

[42] J. Snyder, P. Sloan, and Y. Ng. “Systems and methods for all-frequency relighting using spherical harmonics and point-light distributions”, US Patent 7,262,771. 2

[43] Y. Tsai, and Z. Shih. “All-Frequency Precomputed Radiance Transfer using Spherical Radial Basis Functions and Clustered Tensor Approximation”, *SIGGRAPH*, 25(3):967-976, 2006. 2

[44] B. Walter, S. Fernandez, A. Arbree, K. Bala, M. Donikian and D. P. Greenberg. “Lightcuts: a scalable approach to illumination”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(3):1098-1107, 2005. 2

[45] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T.S. Huang, and S. Yan. “Sparse Representation for Computer Vision and Pattern Recognition”, *Proceedings of the IEEE*, 98(6), 1031-1044, 2010. 3

[46] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. “Robust Face Recognition via Sparse Representation”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 31(1):210-227, 2009. 3, 10

[47] Y. Zhang and Y. Yang. “Illuminant direction determination for multiple light sources”, *IEEE Conf. Computer Vision and Pattern Recognition*, 1:269-276, 2000. 2

[48] W. Zhou and C. Kambhamettu. “A unified framework for scene illuminant estimation”, *Image and Vision Computing*, 26(3):415-429, 2008. 3