Sparse Representation of Cast Shadows via ℓ_1-Regularized Least Squares

Xue Mei
Center for automation Research
Electrical & Computer Engineering Dept.
University of Maryland, College Park, MD
xuemei@umiacs.umd.edu

Haibin Ling
Center for Information Science & Tech.
Computer & Information Science Dept.
Temple University, Philadelphia, PA
hbling@temple.edu

David W. Jacobs
Center for automation Research
Computer Science Dept.
University of Maryland, College Park, MD
djacobs@umiacs.umd.edu

Abstract

Scenes with cast shadows can produce complex sets of images. These images cannot be well approximated by low-dimensional linear subspaces. However, in this paper we show 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. 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 to several previously proposed methods.

1. Introduction

Dealing with shadows is an important and challenging problem in the study of illumination effects. This makes it difficult to recover illumination from a scene given a single input image. However, as observed in previous studies [23, 19, 26], images with cast shadows can often be sparsely represented, which is attractive since sparsity leads to efficient estimation, dimensionality reduction, and efficient modeling.

In this paper, we solve the problem of illumination recovery from a single image with cast shadows and show that the illumination can be well approximated by a combination of low frequency spherical harmonics and a sparse set of directional light sources. As in previous work, we recover lighting using a prior model of the scene that captures geometry and albedos [3, 2, 26, 24, 25, 19, 28]. It is pointed out in [12] that the assumption of known geometry is required by many illumination estimation methods. Illumination recovery with sparse light sources can be very helpful for many applications. 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.

We cast the problem of finding a sparse representation of directional sources as an ℓ_1-regularized least squares problem. This is partly motivated by recent advances in compressive sensing [4, 8]. 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. The proposed method is tested on synthetic and real images in which it outperforms other state-of-the-art approaches in both accuracy and speed.

The main contribution of this paper is an efficient solution using the ℓ_1-regularized least squares for sparse lighting representation with cast shadows. In our experiments, the proposed method compares favorably to previous approaches. We also feel that our result helps to understand other approaches to lighting recovery, by elucidating exactly which components of lighting are important in generating shadows.

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 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

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 in [27, 15] that, when ignoring all shadows, the images of a Lambertian scene lie in a three-dimensional linear subspace. 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 [3, 10] and analytically [2, 21]. The sparsity of cast shadows is recently studied in [23, 19]. The work in [23] 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. In [19], a sparse representation using a Haar wavelet basis is proposed to recover lighting in images with cast shadows.

The studies in [26, 24, 25, 19] are most 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 [19] a Haar wavelet basis is used to effectively capture lighting sparsely. Some work in precomputed radiance transfer and importance sampling in computer graphics is also very relevant [1, 18, 30, 31].

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 [17] 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 [11]. Much more accurate multiple illumination information is extracted from the shading of a sphere [33]. A signal-processing framework that describes the reflected light field as a convolution of the lighting and BRDF is introduced in [22]. This work suggests performing rendering using a combination of spherical harmonics and directional light sources with ray-tracing to check for shadows. However, our motivation and algorithms are quite different from theirs. Very recently, in [12], the number of point light sources and the reflectance property of an object are simultaneously estimated using the EM algorithm.

Our solution using ℓ_1-regularized least squares is motivated by recent advances in the field of compressed sensing [4, 8]. 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 [4, 8] 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 for problems such as face recognition [32], background subtraction [5], media recovery [9], texture segmentation and feature selection [14]. There is also work related to lighting in graphics [20, 29]. 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.

3. An Example

To strengthen our intuitions, we consider a very simple example of 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. A finite-dimensional linear subspace cannot capture any significant fraction of the effects of cast shadows.

But, let's 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)) we have the more shadows we have, but the lighter these shadows are. Therefore, while a sparse set of lights can produce strong cast shadows, many lights tend to wash out the effects of shadowing.

Now, suppose we approximate any possible image using one image of constant intensity, and a small number of images that are each produced by a directional source. If the actual image is produced by a small number of directional sources, we can represent its shadows exactly. If the image is produced by a large number of directional sources, we

(a) (b) (c)

Figure 1. A flagpole rendered with one directional source (a), two directional sources (b), and ten directional sources (c). The shadows are lighter as the number of directional sources increases.

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

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 (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, so we can render directional source images from it.

A scene is 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 amount) of image $I_{dir}(\theta)$.

For practical reason, integration over the continuous space $\mathcal{S}$ is replaced by a superposition over a large discrete set of lighting directions, 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),

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

It is known that in the absence of cast shadows, this lighting can be approximated using low frequency spherical harmonics [2, 21]. 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 [2]. 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., $||I_k||_2 = 1$.

The size of the image can be very large, which corresponds to a large linear system $Ax = \tilde{I}$. To reduce the dimensionality and speed up the computation, we apply PCA to the image set A. The standard PCA yields a projection

matrix $W \in \mathbb{R}^{m \times d}$ that consists of the m most important principal components of A.

Applying W to equation 6 yields:

$$\begin{aligned} & \arg\min_{\mathrm{x}} ||W(A\mathrm{x} - \mu) - W(\tilde{I} - \mu)||^2 \\ = \; & \arg\min_{\mathrm{x}} ||WA\mathrm{x} - W\tilde{I}||^2, \quad x_k \geq 0 \,, \end{aligned} \tag{7}$$

where μ is the mean vector of columns of A.

The dimension m is typically chosen to be much smaller than d. 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 ℓ_2 or Tikhonov regularization [16], which can be written as

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

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.

We are concerned with the problem of low-complexity recovery of the unknown vector x. Therefore, 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 sum of squares used in Tikhonov regularization:

$$\arg\min_{\mathrm{x}} ||WA\mathrm{x} - W\tilde{I}||^2 + \lambda ||\mathrm{x}||_1, \quad x_k \geq 0. \tag{9}$$

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 (9) can be reformulated as

$$\arg\min_{\mathrm{x}} ||WA\mathrm{x} - W\tilde{I}||^2 + \lambda \sum_{k=1}^{N} x_k, \quad x_k \geq 0. \tag{10}$$

Figure 2 left and right show 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-left. 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.

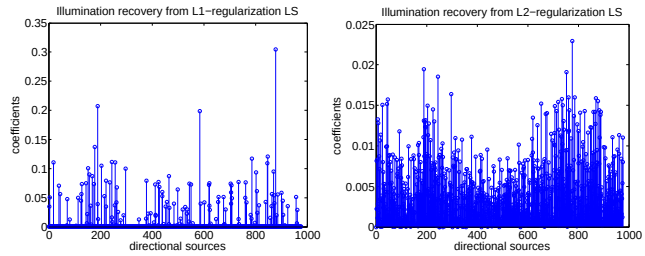


Figure 2. The recovered coefficients x from ℓ_1-Regularized LS (left) and ℓ_2-Regularized LS (right).

Algorithm 1 Sparse representation for inverse lighting

1: Obtain N directional source images 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: Project each directional source image I_k to the 9D spherical harmonic subspace and obtain the corresponding residual directional source image $\tilde{I}_k$.
3: Normalize $\tilde{I}_k$ such that $||\tilde{I}_k||_2 = 1$.
4: Generate matrix $A = [\tilde{I}_1 \tilde{I}_2 \cdots \tilde{I}_N]$.
5: Apply Principal Component Analysis to matrix A and obtain the projection matrix W by stacking the m most important principal components of A.
6: Project the query image I to the spherical harmonic subspace and obtain the residual image $\tilde{I}$.
7: Solve the ℓ_1-regularized least squares problem with nonnegativity constraints (10).
8: Render the scene with the spherical harmonic lighting plus the recovered sparse set of directional light sources.

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 [13]. 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 [6] for the minimization task in (10).

6. Experiments

In this section, we describe our experiments for illumination recovery on both synthetic and real data, in comparison with four other previous approaches.

6.1. Experimental Setup

6.1.1 Data

Both synthetic and real datasets are used in our experiments. The synthetic scene is composed of a coffee cup

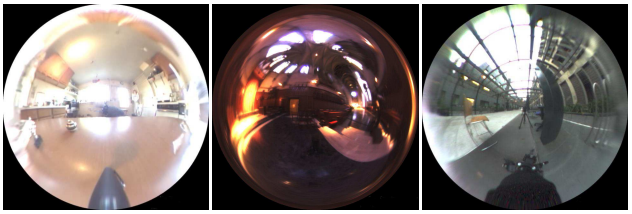


Figure 3. Light probes [7] used to generate our synthetic dataset: kitchen (left), grace (center), and building (right). The light probes are sphere maps and shown in low-dynamic range for display purposes.

and a spoon, with a plate underneath them (see Figure 4). Three synthetic images are obtained by rendering the scene with environment maps (namely *kitchen*, *grace*, and *building*, see Figure 3) provided by [7]. We considered a scene where the objects were placed on an infinite plane, so only lights coming from the upper hemisphere are taken into account.

For the real objects, we built CAD models of three objects (namely *chair1*, *chair2*, and *couch*, see Figures 5 and 6) and printed them with a 3D printer. The only difference between chair1 and chair2 is the number of backrest bars. These objects are placed under natural indoor illumination and images are taken by a Canon EOS Digital Rebel XT camera. The 3D models are registered to corresponding images by minimizing the distance between the feature points on the image and the corresponding feature points from the 3D model, projected onto the image.

One of our experiments involves recovering lighting from one object (chair1), and using it to render a model of a second object (chair2) [19]. For this reason, we take pictures of chair1 and chair2 in exactly the same illumination environment.

6.1.2 Methods for Comparison

We compare our proposed algorithm with Spherical Harmonics [2, 21], Non-Negative Least squares (NNL) [3], Semidefinite Programming (SDP) [28], and Haar wavelets [19] algorithms . NNL [3] finds the non-negative combination of directional sources that produce the best approximation to an image. To make this algorithm comparable, we use 100 directional sources to represent the lighting which is obtained from 977 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 (10) is set to 0.01 for all the experiments. We approximate the query image using 977 directional sources and pick the 100 directional sources which have the largest coefficients. SDP [28] is applied to perform a constrained optimization to quickly and accurately find the non-negative linear combination of spherical harmonics up to the order 10. The total harmonics used in SDP is $(10+1)^2 = 121$. It has been applied for specular object recognition on both synthetic and real data. This paper is the first to show that SDP can also be used to handle shadows. Haar wavelets [19] are used to recover illumination from cast shadows and are shown to be more reliable than spherical harmonics in reproducing cast shadows. We use the same procedure in [19] to estimate illumination using 102 Haar wavelet basis functions. To have a fair comparison with these methods, we use 100 directional sources for the illumination recovery from 977 possible directional sources.

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 [19]. Specifically, for an input image $I \in \mathbb{R}^d$ and its recovery $\hat{I} \in \mathbb{R}^d$, the RMS between them is defined as $r(I, \hat{I}) = ||I - \hat{I}||_2$.

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.[1] 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 Intel Core Duo at 1.73GHz with 2.0GB RAM laptop.

6.2. Experiments with Synthetic Data

In this section, we deal first with synthetic data, showing that illumination can be accurately recovered by using our proposed method. Comparison between our proposed method and other methods is made in terms of accuracy and speed.

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. Using these images, we numerically integrate to compute nine images of the scene, each with lighting consisting of a single spherical harmonic.

For this evaluation, we used synthetic images rendered with three environment maps provided by [7], from high

[1] The spherical harmonics and Haar wavelets also need these images for basis image estimation.

Table 1. RMS errors and average run time on the synthetic image dataset. Note: the running time does not include the preprocessing for generating "basis" images (same for Tables 2 and 3).

Method	Probe Kitchen	Probe Grace	Probe Building	Avg. Run Time (sec.)
Spherical Harmonics	8.00	12.23	12.21	0.01
NNL (100 DS)	55.74	17.31	39.41	9389.8
NNL (300 DS)	5.96	2.80	1.87	9389.8
SDP	3.21	4.11	3.48	10.9
Haar Wav. (102 basis)	3.42	3.12	1.61	1322.0
Our method (100 DS)	2.33	2.69	1.22	11.8

dynamic range light probe images, and recovered illumination from them. Figure 3 shows the sphere maps of three of the light probes used in our experiments. We considered a scene where the objects were placed on an infinite plane, so only lights coming from the upper hemisphere are taken into account.

Figure 4 shows the ground truth images (a)-(c) and images obtained by using the method based on spherical harmonics (d), NNL with 100 directional sources (DS) (e), NNL with 300 DS (f), SDP (g), Haar wavelets (h), and our method (i). We recovered illumination distributions from input images (a)-(c) rendered with the sphere maps of kitchen, grace, and building (Figure 3), respectively. The image approximated using spherical harmonics is obtained by projecting the image onto the harmonic subspace. It 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, from 977 possible ones. The reason is, as illustrated in Figure 4, NNL with 100 DS failed 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. We render the image with 102 Haar basis functions as in [19]. Both Haar wavelets and our method reproduce the shadows reliably.

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. 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. Therefore, our method works best for recovering illumination from cast shadows in terms of both accuracy and speed.

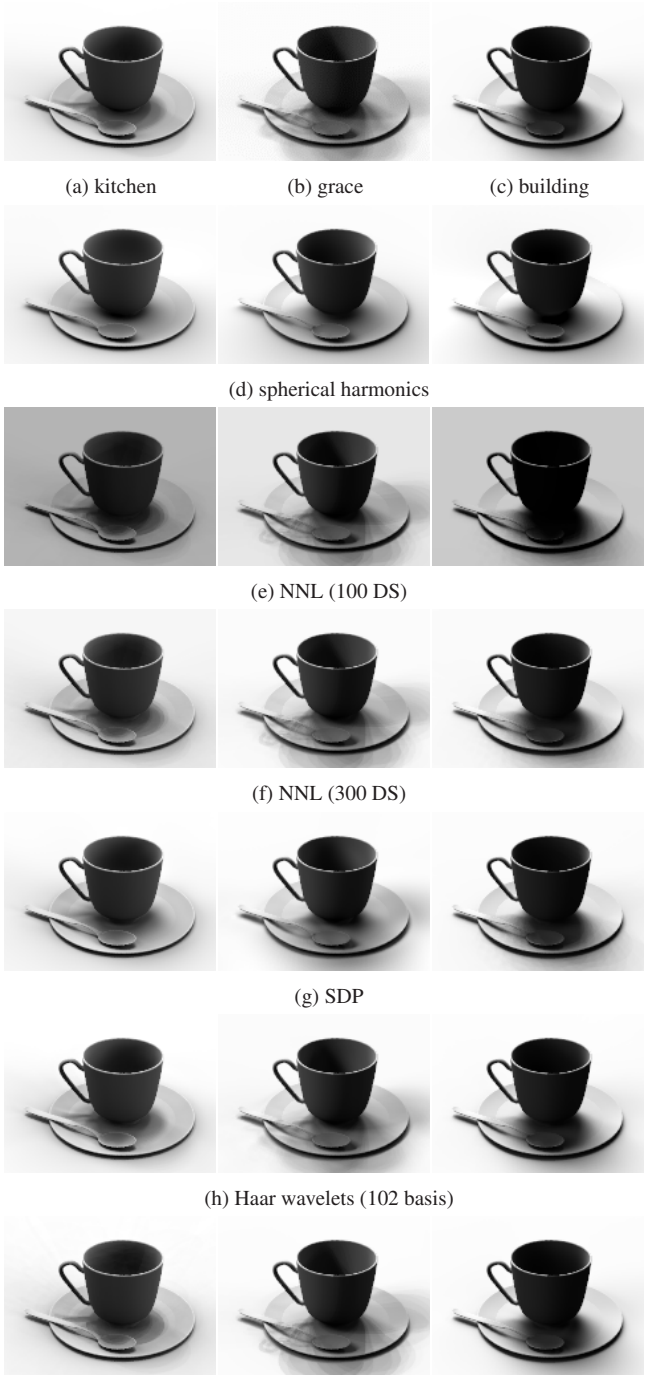


Figure 4. Experiments on synthetic images. (a)-(c): ground truth images from different lighting probes as indicated. (d)-(i) images recovered different approaches.

6.3. Experiments with Real Data

For real images, we conduct two kinds of experiments. First, all the algorithms are tested for illumination recovery tasks on *c*hair1 and *c*ouch. The results are shown in the left

Table 2. RMS errors and run times on the real images of chair1 and chair2. The RMS for chair1 is for lighting recovery (Fig. 5 left); while the RMS for chair2 is for lighting evaluation (Fig. 5 right).

Method	Chair1 RMS Estimation	Chair2 RMS Evaluation	Run time (sec.)
Spherical Harmonics	13.99	15.31	0.01
NNL (100 DS)	10.26	10.35	1854.89
SDP	9.38	9.40	10.88
Haar Wav. (102 basis)	10.75	11.02	1529.60
Our method (100 DS)	7.50	8.24	14.54

Table 3. RMS errors and run times on real images for the couch.

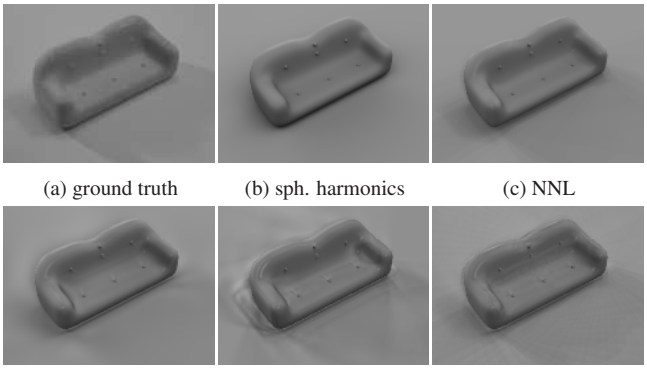
Method	RMS	Run time (sec.)
Spherical Harmonics	9.39	0.01
NNL (100 DS)	7.37	2050.22
SDP	7.01	14.62
Haar Wav. (102 basis)	7.84	1585.27
Our method (100 DS)	6.56	13.82

column of Figure 5 and in Figure 6. Second, we apply the recovered illumination from *c*hair1 to the model of *c*hair2 and then compare the results to the ground truth image. The second test is similar to those used for lighting recovery in [19]. The results are shown in the right column of Figure 5. The RMS errors and run time statistics are summarized in Tables 2 and 3.

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 shadows both on the seat and the ground. In addition, Table 2 shows the RMS error and run time for all the methods. Our method achieves the smallest error of all the methods in only tens of seconds run time. Figure 6 and Table 3 show the experimental results for the couch under natural indoor lighting. Again, our method achieves the best results in terms of speed and accuracy. Hence, it can be concluded that our method works reliably and accurately in recovering illumination and producing cast shadows for real images as well.

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 7 shows



(a) ground truth (b) sph. harmonics (c) NNL

(d) SDP (e) Wavelets (f) our method

Figure 6. (a) Ground truth image of couch. (b)-(f) show the image rendered with the lighting recovered from (a) using different approaches, where (c) and (3) use 100 directional sources (DS), and (e) uses 102 wavelet basis.

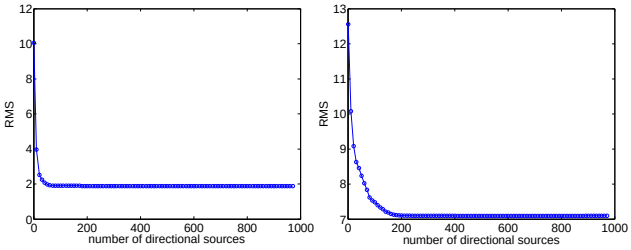
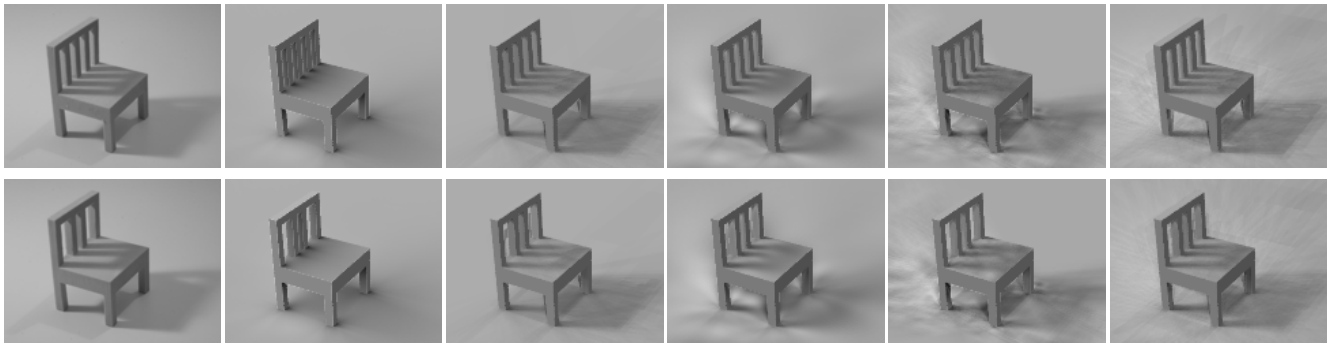


Figure 7. The improvement in accuracy by adding directional sources. RMS versus number of directional sources for a synthetic image rendered with grace light probe (left) and a real image in Figure 5 (a) (right) under natural indoor lighting.

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 images



(a) ground truth (b) spherical harmonics (c) NNL (d) SDP (e) Haar wavelets (f) our method

Figure 5. Column (a): Top - the image of chair 1. Bottom - the image of chair under the same lighting as chair 1. 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 souces.

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

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. Basri and D. Jacobs. "Lambertian Reflectances and Linear Subspaces", *PAMI*, 25(2):218-233, 2003. 1, 2, 3, 5

[3] P. Belhumeur and D. Kriegman. "What is the Set of Images of an Object Under All Possible Lighting Conditions?", *IJCV*, 28(3):245-260, 1998. 1, 2, 5

[4] 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, 2

[5] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa. "Compressive Sensing for Background Subtraction", *ECCV*, 2008. 2

[6] *http://www.stanford.edu/~boyd/l1_ls/*. 4

[7] 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. 5

[8] D. Donoho. "Compressed Sensing", *IEEE Trans. Inf. Theory*, 52(4):1289-1306, 2006. 1, 2

[9] J. Gu, S. Nayar, E. Grinspun, P. Belhumeur, and R. Ramamoorthi. "Compressive Structured Light for Recovering Inhomogeneous Participating Media", *ECCV*, 2008. 2

[10] P. Hallinan. "A Low-dimensional Representation of Human Faces for Arbitrary Lighting Conditions", *CVPR*, 995-999, 1994. 2

[11] K. Hara, K. Nishino, and K. Ikeuchi. "Determining reflectance and light position from a single image without distant illumination assumption", *ICCV*, 1:560-567, 2003. 2

[12] K. Hara, K. Nishino, and K. Ikeuchi. "Mixture of Spherical Distributions for Single-View Relighting", *PAMI*, 30(1):25-35, 2008. 1, 2

[13] 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. 4

[14] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. "Discriminative learned dictionaries for local image analysis", *CVPR*, 2008. 2

[15] Y. Moses. "Face Recognition: Generatlization to Novel Images", *PhD theis*, Weizmann Institute of Science, 1993. 2

[16] A. Neumaier."Solving ill-conditioned and singular linear systems: A tutorial on regularization", *SIAM Rev.*, 40(3):636666, 1998. 4

[17] K. Nishino, Z. Zhang, and K. Ikeuchi. "Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis", *ICCV*, 1:599-606, 2001. 2

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

[19] T. Okabe, I. Sato, and Y. Sato. "Spherical Harmonics vs. Haar Wavelets: Basis for Recovering Illumination from Cast Shadows", *CVPR*, 50-57, 2004. 1, 2, 5, 6, 7

[20] 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

[21] 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, 3, 5

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

[23] R. Ramamoorthi, M. Koudelka, and P. Belhumeur. "A Fourier Theory for Cast Shadows", *PAMI*, 24(2):288-295, 2005. 1, 2

[24] I. Sato, Y. Sato, and K. Ikeuchi. "Stability Issues in Recovering Illumination Distribution from Brightness in Shadows", *CVPR*, 400-407, 2001. 1, 2

[25] I. Sato, Y. Sato, and K. Ikeuchi. "Illumination Distribution from Shadows", *CVPR*, 306-312, 1999. 1, 2

[26] I. Sato, Y. Sato and K. Ikeuchi. "Illumination from Shadows", *PAMI*, 25(3):290-300, 2003. 1, 2

[27] A. Shashua. "On Photometric Issues in 3d Visual Recognition From a Single 2d Image", *IJCV*, 21(1-2):99-122, 1997. 2

[28] S. Shirdhonkar and D. Jacobs. "Non-Negative Lighting and Specular Object Recognition", *ICCV*, 1323-1330, 2005. 1, 5

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

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

[31] B. Walter, S. Fernandez, A. Arbree, K. Bala, M. Donikian and D. P. Greenberg. "Lightcuts: a scalable approach to illumination", *PAMI*, 24(3):1098-1107, 2005. 2

[32] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. "Robust Face Recognition via Sparse Representation", *PAMI*, 31(1):210-227, 2009. 2

[33] Y. Zhang and Y. Yang. "Illuminant direction determination for multiple light sources", *CVPR*, 1:269-276, 2000. 2