

# Photometric Stereo using Sparse Bayesian Regression for General Diffuse Surfaces

Satoshi Ikehata, *Student Member, IEEE*, David Wipf, *Member, IEEE*, Yasuyuki Matsushita, *Senior Member, IEEE* and Kiyoharu Aizawa, *Member, IEEE*

**Abstract**—Most conventional algorithms for non-Lambertian photometric stereo can be partitioned into two categories. The first category is built upon stable outlier rejection techniques while assuming a dense Lambertian structure for the inliers, and thus performance degrades when general diffuse regions are present. The second utilizes complex reflectance representations and non-linear optimization over pixels to handle non-Lambertian surfaces, but does not explicitly account for shadows or other forms of corrupting outliers. In this paper, we present a purely pixel-wise photometric stereo method that stably and efficiently handles various non-Lambertian effects by assuming that appearances can be decomposed into a sparse, non-diffuse component (*e.g.*, shadows, specularities, etc.) and a diffuse component represented by a monotonic function of the surface normal and lighting dot-product. This function is constructed using a piecewise linear approximation to the inverse diffuse model, leading to closed-form estimates of the surface normals and model parameters in the absence of non-diffuse corruptions. The latter are modeled as latent variables embedded within a hierarchical Bayesian model such that we may accurately compute the unknown surface normals while simultaneously separating diffuse from non-diffuse components. Extensive evaluations are performed that show state-of-the-art performance using both synthetic and real-world images.

**Index Terms**—Photometric Stereo, Sparse Regression, Piecewise Linear Regression, Sparse Bayesian Learning.



## 1 INTRODUCTION

PHOTOMETRIC stereo involves estimating the surface normals of an object given appearance variations in multiple images taken under different lighting conditions. While more than 20 years have passed since the original work on Lambertian photometric stereo [1], recovering the surface orientations of a natural scene from appearance variations remains challenging, largely because real-world objects exhibit diverse non-Lambertian effects. The design of practical photometric stereo algorithms involves complicated trade-offs related to robustness, efficiency, and versatility, with most approaches falling into one of two categories. First, there are many techniques that assume a basic Lambertian model but augmented with outlier detection for handling all non-Lambertian regions of the scene [2], [3], [4], [5], [6]. While this strategy is numerically stable and relatively insensitive to initializations, it may be computationally expensive since a large number of images is required for robust outlier rejection. Additionally, complex reflections such as rough specularities or non-Lambertian diffuse reflections can be highly disruptive. In contrast, a second basic approach relies on sophisticated reflectance models or the sharing of properties across materials to explicitly account for non-Lambertian reflections [7], [8], [9], [10], [11], [12]. While these methods are more capable of handling a wide variety of objects including rough surfaces, they are often more sensitive to initialization, numerical instabilities derived from the complex nonlinear optimization, and

the effects of shadows and other corrupting outliers. Further details regarding existing approaches will be discussed in Section 2.

This paper presents an alternative photometric stereo algorithm for stably and accurately estimating the surface normals of a scene in the presence of various non-Lambertian effects. For this purpose, a hierarchical Bayesian model is developed in Section 3 that automatically decomposes observed appearances into a continuous piecewise linear diffuse component and a sparse, non-diffuse component for capturing shadows, specularities, and other corruptions. Optimization and inference is accomplished using a robust majorization-minimization technique akin to the popular EM algorithm, with desirable convergence properties and quantifiable advantages over standard convex estimators.

For mathematical convenience, we develop our diffuse component by assuming, in the absence of non-diffuse corruptions, that pixelwise appearances are well-approximated by a monotonic (and therefore invertible) function of the dot-product between the surface normal and the lighting direction. We may then consider the inverse representation of the image formation process, where the unknown normal vector is now separated from the unknown monotonic inverse reflectance function. By parameterizing the latter using a piecewise linear approximation, we obtain a set of linear equations in both the surface normals and diffuse parameters, leading to simple, closed-form estimators. We note that the piecewise linear function has been widely used for regression purposes [13] because of its simplicity and representational capability. In the context of photometric stereo, this function is a the natural extension of the Lambertian reflectance model, which directly corresponds to the case where there is only one linear segment. Later we

• S. Ikehata and K. Aizawa are with University of Tokyo.

• D. Wipf and Y. Matsushita are with Microsoft Research Asia.

Manuscript received \*\*\*\*\* 2012; revised \*\*\*\*\* 2013.

will empirically demonstrate that additional piecewise linear diffuse segments (e.g., three) can effectively represent many complex non-Lambertian reflections.

Of course we do not know a priori where non-diffuse appearances will be located, and thus the overall estimation problem is underconstrained even with many available images. Thus, our hierarchical Bayesian model attempts to maximize the number of observations that can be adequately explained via the piecewise linear inverse diffuse function (the inlier model) while treating the remaining observations, including specularities, shadows, and sensor saturations, as non-diffuse elements with separate, unknown variances learned from the data (the outlier model). The partitioning between inliers and outliers is estimated simultaneously with the normal vectors and piecewise linear parameters using a principled variational Bayesian technique.

The proposed framework benefits from simple, efficient pixelwise optimization, which is easily amenable to parallel processing. Moreover it does not require the pre-processing of specularities/shadows, careful initialization strategies, or typical smoothness constraints for both object structure and reflectance, which can disrupt the recovery of fine details. Consequently, we do not suffer from numerical instability even with relatively few input images (e.g., 20) and we do not have difficulty handling spatially varying materials with high-frequency structures. Extensive experiments in Section 4 show that our implementation produces more stable, accurate, and efficient surface normal estimates than other robust algorithms such as [3], [6], even when applied to non-Lambertian materials that previously required complex model-based approaches such as [11]. A preliminary version of this work, without the piecewise linear diffuse model, many algorithmic details, or comprehensive empirical testing, has appeared in recent conference proceedings [14].

## 2 RELATED WORK

Since Woodham [1] first introduced photometric stereo for Lambertian scenes, the extension to non-Lambertian surfaces has drawn significant interest. These extensions are largely categorized by whether the non-Lambertian reflections are considered as outliers for removal, or as inliers using non-linear reflectance modeling.

In the former class, surface normals are recovered via a simple Lambertian reflectance model while non-Lambertian corruptions are robustly neglected. The most traditional yet still widely used technique is intensity thresholding for avoiding the effects of shadows and specularities because of its simplicity. To handle spatially varying albedo or sensor saturations, some recent approaches incorporate classification methods for detecting shadows and specularities [15], [16], [17], or exploit graph-based visibility estimation for shadow removal [18]. A color-based technique has also been used for specular removal [2]. More recently, techniques for handling other corruptions, such as inter-reflections [19] and calibration errors [20], have been proposed.

For robustly identifying outliers, the RANDOM SAMPLE CONSENSUS (RANSAC) scheme has been used in [3], [21], [5].

While RANSAC is effective when the number of observations and unknown inlier parameters are relatively small, it becomes computationally intractable with higher dimensionality. Alternatively, some works cast photometric stereo as a matrix decomposition or factorization problem considering non-Lambertian corruptions as missing entries. For example, Wu *et al.* [6] propose a rank-minimization approach that decomposes observations into a low-rank Lambertian structure and sparse non-Lambertian corruptions/outliers, while del Bue *et al.* [22] solve photometric stereo as a manifold constrained bilinear factorization problem in the presence of missing entries.

Spatial constraints are also exploited to effectively remove outliers. Tang *et al.* [23] use an Markov random field (MRF) for imposing discontinuity preserving smoothness using belief propagation. A similar setting has also been used in [24] where graph cuts is used for deriving the optimal solution. While the MRF-based approaches preserve discontinuities, they tend to over-smooth the surface normal map. Spatial information is also incorporated in filtering approaches [25], [4]. Miyazaki *et al.* [25] use a filter-based method where an inaccurate surface normal is refined via the median estimate of neighboring pixels. Yu *et al.* [4] propose a pixelwise scheme to find a maximum feasible subset of Lambertian observations via a Big-M algorithm.

In contrast to the former outlier removal techniques, a second class of methods treat non-Lambertian reflections as inliers using non-linear reflectance modeling. Various physically inspired parametric reflectance models of the bidirectional reflectance distribution function (BRDF) have been used in photometric stereo. For example, Torrance-Sparrow [26] and Ward [27] models have been used to account for specularities [7], [8]. To handle spatially varying materials, Goldman *et al.* [9] approximate basis materials by the Ward model, and reflectance parameters and surface normals are simultaneously estimated in a non-linear manner. More recently, Alldrin *et al.* [10] use a non-parametric representation of BRDFs for simultaneously recovering shape and BRDFs, which are represented by bi-variate form [28]. The model and optimization technique are further simplified in Shi *et al.* [11], where a compact biquadratic representation of isotropic BRDF is introduced.

Instead of explicitly modeling the reflectances, there are methods that use general reflectance properties that are shared among diverse real reflectances. The properties of isotropy and reciprocity symmetry [29], [30] have been exploited for surface normal reconstruction in [31], [32], [33]. The monotonicity property of general diffuse reflectances is also used in [33], [12]. Other than these properties, radiance similarity [34], image deviates [35] and diffuse maxima [36] are also used in the photometric stereo context. Example-based approaches [37], [38] also fall in this category, in which surface normals are estimated using a reference object whose shape is known and BRDF is the same as the target object.

Our method tries to capitalize on both of the above approaches in that it relaxes the often restrictive Lambertian reflectance model by using a piecewise linear representation of reflectances, while simultaneously performing robust estimation to avoid over-fitting. More specifically, our sparsity

penalty effectively rejects outliers while the inverse piecewise linear representation of diffuse reflectances enables us to handle a wide variety of materials.

### 3 PHOTOMETRIC STEREO VIA INVERSE PIECEWISE LINEAR SPARSE REGRESSION

In this section, we formulate the estimation of surface normals using photometric stereo as a piecewise-linear sparse Bayesian regression problem. Henceforth, we rely on the following assumptions:

- 1) Relative position between the camera and the object is fixed across all images.
- 2) Object is illuminated by a point light source at infinity from varying and known directions.
- 3) Camera view is orthographic, and the radiometric response function is linear.

#### 3.1 Problem Statement

Diverse appearances of real world objects can be encoded by a BRDF, which relates the observed intensity  $I$  at a given point on the object to the associated surface normal  $\mathbf{n} \in \mathbb{R}^3$ , the incoming lighting direction  $\mathbf{l} \in \mathbb{R}^3$ , and the outgoing viewing direction  $\mathbf{v} \in \mathbb{R}^3$  via

$$I = f(\mathbf{n}, \mathbf{l}, \mathbf{v}). \quad (1)$$

The dichromatic reflectance model [39] states that if the scene is illuminated by a single dominant point light source, the radiance is a linear combination of diffuse and specular reflectance via

$$I = f_d(\mathbf{n}, \mathbf{l}) + f_s(\mathbf{n}, \mathbf{l}, \mathbf{v}). \quad (2)$$

In practice, various additional effects are observed, including attached/cast shadows, image noise, inter-reflections, and so on. We can interpret these effects as additive corruptions  $e$  applied to the ideal scene leading to the image formation model giving

$$I = f_d(\mathbf{n}, \mathbf{l}) + f_s(\mathbf{n}, \mathbf{l}, \mathbf{v}) + e. \quad (3)$$

The photometric stereo is a problem to recover surface normal  $\mathbf{n}$  of a scene by inversely solving Eq. (3) from a collection of  $m$  observations under the unknown set of parameters  $(f_d, f_s, e)$ . Note that except for uncalibrated photometric stereo problems such as [36],  $\mathbf{l}$  and  $\mathbf{v}$  are usually known.

Early photometric stereo works assumed that the diffuse component of observation is represented by Lambertian reflectance model (*i.e.*,  $f_d = \rho \mathbf{n}^T \mathbf{l}$ , where  $\rho$  is a surface albedo) and discarded the non-Lambertian component  $f_s + e$  as outliers [3], [6], [14]. While the Lambertian assumption is effective for a certain materials, this strong assumption on the reflectance substantially limits the target objects. Instead, we introduce the general representation of the material diffuse function to handle non-Lambertian diffuse materials as follow,

$$f_d(\mathbf{n}, \mathbf{l}) = f(\mathbf{n}^T \mathbf{l}), \quad (4)$$

where we assume  $f(\mathbf{n}^T \mathbf{l})$  passes through the origin, that is  $f(0) = 0$ . Note that Lambertian image formation is a special

case when  $f$  is a linear function. Then, we also merge  $f_s$  and  $e$  as deviations from the diffuse reflection and use the following imaging model,

$$I = \tilde{I} + e = f(\mathbf{n}^T \mathbf{l}) + e. \quad (5)$$

Here,  $\tilde{I}$  is the diffuse component of  $I$ . Our goal is to recover unknown surface normal  $\mathbf{n}$ , diffuse reflectance function  $f$  and non-diffuse corruptions  $e$  from a collection of lighting directions  $\mathbf{l}$  and associated appearances  $I$ . However, there are two critical issues which must be solved: (a) the coincidence of unknown parameters  $\mathbf{n}$  and  $f$  in the same term, and (b) an under-constrained problem since the number of unknowns (equal to  $m + 2$  plus however many degrees of freedom are needed to describe  $f$ ) always exceeds the number of equations (equal to the number of images  $m$ ). We overcome these difficulties by combining a convenient, piece-wise linear inverse representation of the imaging model with a sparsity penalty applied through latent variables embedded in a robust hierarchical Bayesian framework.

#### 3.2 Inverse Diffuse Reflectance Model

For simplicity, we first neglect the non-diffuse corruptions  $e$  in Eq. (5). Then, we assume the monotonicity of the diffuse reflectance function like [33], [12] which provides the following constraint on the function under two different lighting directions  $\mathbf{l}_1$  and  $\mathbf{l}_2$ ,

$$\mathbf{n}^T \mathbf{l}_i > \mathbf{n}^T \mathbf{l}_j \leftrightarrow f(\mathbf{n}^T \mathbf{l}_i) > f(\mathbf{n}^T \mathbf{l}_j). \quad (6)$$

Under this assumption, the unique existence of the inverse function of  $f$  is guaranteed giving

$$f^{-1}(\tilde{I}) = g(\tilde{I}) = \mathbf{n}^T \mathbf{l}. \quad (7)$$

Now that we assume only diffuse reflections appear in the scene, *i.e.*,  $\tilde{I} = I$ , the following equation is acquired,

$$\mathbf{n}^T \mathbf{l} = g(I). \quad (8)$$

We call Eq. (8) the *inverse diffuse reflectance model*. The fundamental advantage of this model is that an unknown function  $g(I)$  and a surface normal  $\mathbf{n}$  are separated, which contributes to the simplicity of the problem. Eq. (8) suggests that the per-pixel collection of 2-D plot  $(I_i, \mathbf{n}^T \mathbf{l}_i)$  corresponding to  $i$ -th image must draw a monotonic inverse diffuse function  $g(I)$ . While this relationship limits the solution space of both  $\mathbf{n}$  and  $g$ , the problem is that there are still multiple feasible solutions of a pair of  $\mathbf{n}$  and  $g(I)$  as illustrated in Fig. 1, especially when  $m$  is small. To reduce inherent ambiguity of the problem, we further assume a parametric model of the general inverse diffuse function  $g(I)$ .

Given that the left-hand-side of Eq. (8) is linear in the unknown normal vector  $\mathbf{n}$ , for computational simplicity we would like to impose similar linearity on the right-hand-side in our parameterized representation of  $g(I)$ . For this purpose, we then choose to express  $g(I)$  as a summation over  $p$  fixed and known, non-linear basis functions  $g_k(I)$  weighted by an unknown coefficient vector  $\mathbf{a}$ , leading to the representation

$$g(I) = \sum_{k=1}^p a_k g_k(I). \quad (9)$$

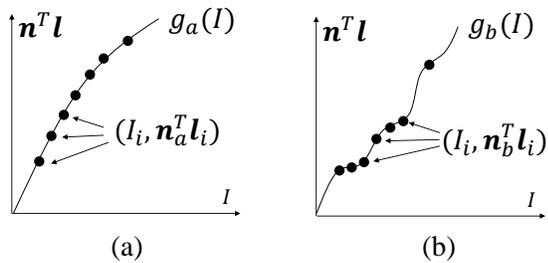


Fig. 1. We can draw different monotonic curves which interpolate 2-D plots  $(I_i, \mathbf{n}^T \mathbf{l}_i)$  derived from  $(\mathbf{n}_a)$  in (a) and  $(\mathbf{n}_b)$  in (b), which illustrates that there are multiple feasible solutions of a pair of  $\mathbf{n}$  and  $g(I)$  which satisfy the inverse diffuse reflectance model in Eq. (8).

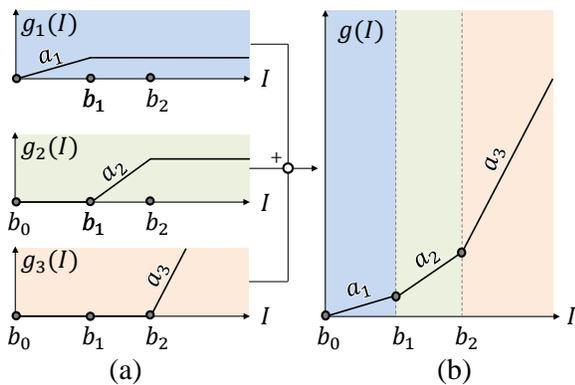


Fig. 2. The illustration of the piecewise linear function. We show the case when the number of basis functions is three. (a) Each basis function is defined as a polylinear function which has a breaking point, and (b) the piecewise linear function is defined as the summation of these basis functions.

While non-linear in  $I$ ,  $g(I)$  is clearly linear in  $\mathbf{a} = [a_1, \dots, a_p]^T$ . Choices for each  $g_k$  include polynomial, Gaussian, logistic, and spline functions as well as many others.

We choose to adopt a piecewise linear representation [13] which is composed of multiple polylinear functions of the form

$$g_k(I) = \begin{cases} 0 & (0 \leq I < b_{k-1}) \\ I - b_{k-1} & (b_{k-1} \leq I < b_k) \\ b_k - b_{k-1} & (b_k \leq I) \end{cases} \quad (10)$$

where each  $b_{k-1}$  denotes the point where the  $(k-1)$ -th linear segment ends and the  $k$ -th segment begins. See Fig. 2 for details. By construction,  $g(I)$  will be a *continuous* piecewise linear function, meaning each adjoining linear segment connects to one another at the corresponding point  $b_{k-1}$ , regardless of the coefficients  $\mathbf{a}$ . We also assume that  $b_0 = 0$ , meaning  $g(I)$  will intersect the origin. Remaining values of  $b_{k-1}$  are chosen such that each piecewise linear segment spans an equal-sized bin over the range of  $I$ .

This choice of basis function leads to a natural generalization of the Lambertian reflectance model, which is obtained

when all elements of  $\mathbf{a}$  are set to the same positive value. We thus preserve preferable properties of the Lambertian model where appropriate, while still allowing us to handle more general non-linear diffuse reflections where needed. In fact, even with  $p$  small, we may nonetheless approximate a wide variety of non-linear functionals, with monotonicity ensured whenever  $\mathbf{a}$  has all non-negative elements (while we do not strictly enforce non-negativity, our learning procedure described in the following sections strongly disfavors any  $a_k < 0$ ). Here we should note that our model relates to the recent bilinear BRDF model proposed by Romeiro and Zickler [40], which is represented as a linear combination of non-negative basis functions learned through non-negative matrix factorization of 100 materials in the MERL BRDF database [41]. However, there are two different points. Firstly, our basis functions are learned directly from the data itself (*i.e.*, linear segment ends in Eq. (10) are decided from samples at each pixel) while [40] learns basis functions from the external database. More importantly, we only model the diffusive component by a simple linear function which contributes to the computationally efficiency and stability, while [40] adapts basis functions represented as discretized two-dimensional matrices.

By substituting Eq. (9) into the inverse diffuse reflectance model, Eq. (8) becomes

$$\mathbf{n}^T \mathbf{l} = \sum_{k=1}^p a_k g_k(I). \quad (11)$$

Collecting variations of observation at the same pixel under different lighting directions, the equations can be merged into following linear problem,

$$\tilde{A} \mathbf{x} = 0, \quad (12)$$

where  $\mathbf{x} \triangleq [n_x, n_y, n_z, a_1, a_2, \dots, a_p]^T \in \mathbb{R}^{p+3}$ , and  $n_x, n_y, n_z$  are the three elements of the surface normal.  $\tilde{A} \in \mathbb{R}^{m \times (3+p)}$  is a data matrix whose  $j$ -th row is given by

$$\tilde{A}_j = [-l_x^j, -l_y^j, -l_z^j, b_1 - b_0, \dots, b_{k-1} - b_{k-2}, I_j - b_{k-1}, 0, \dots, 0]. \quad (13)$$

Here we assume  $b_{k-1} \leq I_j < b_k$  and  $l_x^j, l_y^j, l_z^j$  are three elements of the  $j$ -th lighting direction.

Without loss of generality, we may avoid the degenerate  $\mathbf{x} = 0$  solution to Eq. (12) by constraining  $\sum_{k=1}^p a_k = 1$ . For this purpose we replace  $\tilde{A}$  by  $A$  where  $[0, 0, 0, 1, \dots, 1]$  is appended as the  $(m+1)$ -th row of  $\tilde{A}$  and we define  $\mathbf{y} \in \mathbb{R}^{m+1}$  as a vector of all zeros except for a one as the last element. Given the appearance variations  $(I_1, I_2, \dots, I_m)$  under different known lighting conditions  $(\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_m)$ , the optimal surface normal  $(\mathbf{n})$  and model parameters  $(a_1, a_2, \dots, a_p)$  are recovered by solving the linear problem

$$\min_{\mathbf{x}} \|W(A\mathbf{x} - \mathbf{y})\|_2^2. \quad (14)$$

Here  $W \triangleq \text{diag}[1, \dots, 1, \infty] \in \mathbb{R}^{(m+1) \times (m+1)}$  is a diagonal weighting matrix designed to strictly enforce the constraint  $\sum_{k=1}^p a_k = 1$ . Note that  $2 + p$  linearly independent images are sufficient for producing a unique solution to Eq. (14). We will refer to this photometric stereo method as *piecewise linear*

*least-squares regression (PL-LS)*. In spite of its simplicity, PL-LS works for a wide variety of non-Lambertian diffuse materials (see our experimental results in Section 4.3). The problem of course is that real images are frequently contaminated with various non-diffuse effects as modeled in Eq. (5). The next section will focus on how to handle these corruptions within our inverse diffuse reflectance model.

### 3.3 Piecewise Linear Sparse Regression with Non-Diffuse Corruptions

By introducing additive non-diffuse corruptions  $e'$  in Eq. (11), the image formation model becomes

$$\mathbf{n}^T \mathbf{l} = \sum_{k=1}^p a_k g_k(I) + e'. \quad (15)$$

Note that although the value of  $e'$  in Eq. (15) is different from the  $e$  in Eq. (5), both factors are directed at the same observations and thus serve the same overall purpose (see Fig. 3). The standard least-squares based photometric stereo problem corresponding to Eq. (15) would involve solving

$$\min_{\mathbf{x}, e'} \|W(A\mathbf{x} + e' - \mathbf{y})\|_2^2 \quad \text{s.t.} \quad e'_{(m+1)} = 0, \quad (16)$$

where  $e' \triangleq [e'_1, e'_2, \dots, e'_{m+1}]^T \in \mathbb{R}^{m+1}$ . Note that  $e'_{(m+1)}$  must be set to zero to maintain the constraint  $\sum_{k=1}^p a_k = 1$  (see details in Section 3.4).

Given observations  $I$  and lighting directions  $\mathbf{l}$ , the ultimate goal is to recover surface normals and model parameters  $\mathbf{x}$  and non-diffuse corruptions  $e'$ . However, this is an under-constrained problem since the number of unknowns  $p+2+m$  will always exceeds the number of independent equations  $m$ .

One solution to this ambiguity is to apply simple shadow/specular thresholding [9], [11] or a color channel transformation [2] as a preprocessing step, to obtain an estimate of  $e'$  and/or discard outliers. However, these types of heuristics may discard useful information at times and come with an additional computational expense. Moreover, graph-based approaches [24], [23], [42] and robust algorithms [3], [6] do not naturally embed within our framework since they may conflict with our inverse piecewise diffuse model and/or degrade the numerical stability.

In this paper, we instead introduce a special sparsity penalty applied to  $e'$ , whose minimization disambiguates the infinity of feasible solutions to Eq. (15). This penalty quantifies the reasonable observation that objects in the natural world exhibit dominant diffuse reflections while non-diffuse effects emerge primarily in limited areas of its appearance. For example, specularities surround the spot where the surface normal is oriented halfway between lighting and viewing directions, while shadows are created only when  $\mathbf{l}^T \mathbf{n} \leq 0$  (attached shadow) or when a non-convex surface blocks the light (cast shadow). Strictly speaking, we assume that the optimal feasible solution to Eq. (16) is acquired when the largest possible number of observations are lying on the piecewise linear diffuse reflectance function. Reflecting this assumption, our estimation problem can be formulated as

$$\min_{\mathbf{x}, e'} \|e'\|_0 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x} + e', \quad e'_{(m+1)} = 0. \quad (17)$$

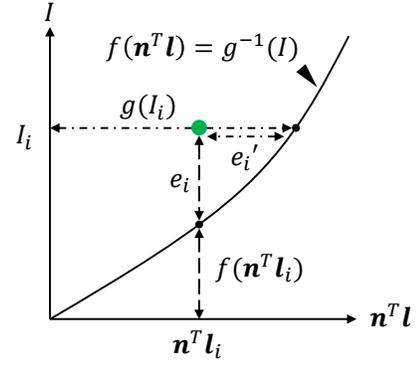


Fig. 3. A 2-D point  $(\mathbf{n}^T \mathbf{l}_i, I_i)$  can be represented by both the forward image formation model ( $I_i = f(\mathbf{n}^T \mathbf{l}_i) + e_i$ ) and the inverse image formation model ( $\mathbf{n}^T \mathbf{l}_i = g(I_i) + e'_i$ ). The illustration indicates that they are uniquely convertible if the reflectance function is monotonic.

Here,  $\|\cdot\|_0$  represents the  $\ell_0$ -norm, which counts the number of non-zero entries in a vector. To reiterate, Eq. (17) builds on the assumption that images are captured under known lighting conditions and any non-diffuse corruptions have sparse structure. If these assumptions are not true (e.g., because of imperfect lighting calibration, no dominant diffuse structure, etc.), then the hard constraint in Eq. (17) is no longer appropriate. To compensate for more modeling errors, we relax the hard constraint via an additional model mismatch penalty giving

$$\min_{\mathbf{x}, e'} \|W(A\mathbf{x} + e' - \mathbf{y})\|_2 + \lambda \|e'\|_0 \quad \text{s.t.} \quad e'_{(m+1)} = 0, \quad (18)$$

where  $\lambda$  is a nonnegative trade-off parameter balancing data fit with sparsity. Note that in the limit as  $\lambda \rightarrow 0$ , problems (17) and (18) are equivalent (the limit must be taken outside of the minimization). Eq. (18) entails a difficult, combinatorial optimization problem that must be efficiently solved at every pixel. Here we consider two alternatives to brute force exhaustive search. First, in the machine learning and statistics literature, it is common to replace the discontinuous, non-convex  $\ell_0$  norm with the convex surrogate  $\ell_1$  norm. The  $\ell_1$  norm of a vector  $\mathbf{z}$  is given by  $\sum_i |z_i|$ , which constitutes the tightest convex approximation to the  $\ell_0$  norm. In certain situations the resulting estimate will closely match the solution to Eq. (18); however, in the context of photometric stereo this substitution may not always be adequate (see Section 3.5 for more details). Secondly, we can apply a simple hierarchical Bayesian approximation to estimate  $\mathbf{x}$  while simultaneously accounting for  $e'$ . This formulation, a generalized version of sparse Bayesian learning (SBL) [43], is described in detail next.

### 3.4 Recovery of Normals and Corruptions Via SBL

We assume the standard Gaussian likelihood function for the first-level, diffuse errors giving

$$p(\mathbf{y}|\mathbf{x}, e') = N(\mathbf{y}; A\mathbf{x} + e', \lambda W^{-1}), \quad (19)$$

Note that we define  $W^{-1} \triangleq \text{diag}[1, \dots, 1, 0] \in \mathbb{R}^{(m+1) \times (m+1)}$ . We next apply an independent, zero-mean Gaussian prior distributions on both  $\mathbf{x}$  and  $\mathbf{e}'$ :

$$p(\mathbf{x}) = N(\mathbf{x}; \mathbf{0}, \Sigma_x), \quad p(\mathbf{e}') = N(\mathbf{e}'; \mathbf{0}, \Gamma). \quad (20)$$

$\Sigma_x$  describes the prior variance of the unknown normal vector  $\mathbf{n}$  and model parameters  $\mathbf{a}$  as  $\Sigma_x = \text{diag}(\sigma_n^2 I_{(3)}, \sigma_a^2 I_{(p)})$  where  $I_{(k)} \in \mathbb{R}^{k \times k}$  is the identity matrix; they are fixed to convey our lack of *a priori* certainty about  $\mathbf{x}$ . Thus the prior on  $\mathbf{x}$  will be relatively uninformative. The values of  $\sigma_n^2$  and  $\sigma_a^2$  will be discussed further below. In contrast,  $\Gamma \triangleq \text{diag}[\boldsymbol{\gamma}]$  is a fully-parameterized, diagonal matrix, where  $\boldsymbol{\gamma} \triangleq [\gamma_1, \dots, \gamma_{m+1}]^T$  is a non-negative vector of variances in one-to-one correspondence with elements of  $\mathbf{e}'$ . A large variance  $\gamma_i$  indicates that the corresponding  $e'_i$  is free to reflect the data, compensating for non-diffuse effects (outliers), while a small or zero-valued variance implies that the associated error term is constrained near zero (inliers). While we are ignorant of which observations are outliers,  $\gamma_{(m+1)}$  is fixed to be zero because of the constraint regarding  $e'_{(m+1)}$  in Eq. (18).

Combining the likelihood and prior using Bayes' rule leads to the posterior distribution  $p(\mathbf{x}, \mathbf{e}' | \mathbf{y}) \propto p(\mathbf{y} | \mathbf{x}, \mathbf{e}') p(\mathbf{x}) p(\mathbf{e}')$ . To estimate  $\mathbf{x}$ , we may further marginalize over  $\mathbf{e}'$  to give

$$p(\mathbf{x} | \mathbf{y}) = \int p(\mathbf{x}, \mathbf{e}' | \mathbf{y}) d\mathbf{e}' = N(\mathbf{x}; \boldsymbol{\mu}, \Sigma), \quad (21)$$

with mean and covariance defined as

$$\begin{aligned} \boldsymbol{\mu} &= \Sigma A^T (\Gamma + \lambda W^{-1})^{-1} \mathbf{y}, \\ \Sigma &= \left[ \Sigma_x^{-1} + A^T (\Gamma + \lambda W^{-1})^{-1} A \right]^{-1}. \end{aligned} \quad (22)$$

We now have a convenient closed-form estimator for  $\mathbf{x}$  given by the posterior mean. The only issue then is the values for the unknown parameters  $\Gamma$ . Without prior knowledge as to the locations of the sparse errors, the empirical Bayesian approach to learning  $\Gamma$  is to marginalize the full joint distribution over all unobserved random variables, in this case  $\mathbf{x}$  and  $\mathbf{e}'$ , and then maximize the resulting likelihood function with respect to  $\Gamma$  [43]. Equivalently, we will minimize

$$\begin{aligned} L(\Gamma) &\triangleq -\log \int p(\mathbf{y} | \mathbf{x}, \mathbf{e}') p(\mathbf{x}) p(\mathbf{e}') d\mathbf{x} d\mathbf{e}' \\ &\equiv \log |\Sigma_y| + \mathbf{y}^T \Sigma_y^{-1} \mathbf{y} \\ &\quad \text{with } \Sigma_y \triangleq A \Sigma_x A^T + \Gamma + \lambda W^{-1}, \end{aligned} \quad (23)$$

with respect to  $\Gamma$ . While  $L(\Gamma)$  is non-convex, optimization can be accomplished by adapting a majorization-minimization approach from [44] to the photometric stereo problem. This technique essentially involves the construction of rigorous upper bounds (see below) on each of the two terms in Eq. (23) using auxiliary variables  $\mathbf{z} \triangleq [z_1, \dots, z_{m+1}]^T$  and  $\mathbf{u} \triangleq [u_1, \dots, u_{m+1}]^T$  (the EM algorithm can be viewed as a special case). For fixed values of  $\mathbf{z}$  and  $\mathbf{u}$ , a closed form solution for  $\Gamma$  exists. Likewise, for a fixed value of  $\Gamma$ , the auxiliary variables can be updated in closed form to tighten the upper bound around the current  $\Gamma$  estimate.

While some details are omitted for brevity, using results from convex analysis it can be shown that for all  $\mathbf{u} \geq 0$ ,

$$\begin{aligned} \log |\Sigma_y| &= \log |\Gamma| + \log |A \Sigma_x A^T + \lambda W^{-1}| \\ &\quad + \log \left| \Gamma^{-1} + (A \Sigma_x A^T + \lambda W^{-1})^{-1} \right| \\ &\leq \log |\Gamma| + \log |A \Sigma_x A^T + \lambda W^{-1}| \\ &\quad + \sum_i \frac{u_i}{\gamma_i} - h^*(\mathbf{u}) \\ &\equiv \sum_i \left( \frac{u_i}{\gamma_i} + \log \gamma_i \right) - h^*(\mathbf{u}), \end{aligned} \quad (24)$$

where  $h^*(\mathbf{z})$  denotes the concave conjugate function [45] of  $h(\boldsymbol{\beta}) \triangleq \log \left| \text{diag}[\boldsymbol{\beta}] + (A \Sigma_x A^T + \lambda W^{-1})^{-1} \right|$  and we have removed irrelevant factors independent from  $\mathbf{u}$  or  $\boldsymbol{\gamma}$ . It can be shown that equality (and therefore the minimum of the right-hand side) is obtained in Eq. (24) if and only if

$$\mathbf{u} = \text{diag} \left[ \Gamma^{-1} + (A \Sigma_x A^T + \lambda W^{-1})^{-1} \right]^{-1}. \quad (25)$$

In a somewhat related fashion, the second term in  $L(\Gamma)$  can be upper-bounded via

$$\begin{aligned} \mathbf{y}^T \Sigma_y^{-1} \mathbf{y} &\leq (\mathbf{y} - \mathbf{z})^T (A \Sigma_x A^T + \lambda W^{-1})^{-1} (\mathbf{y} - \mathbf{z}) \\ &\quad + \sum_i \frac{z_i^2}{\gamma_i}, \end{aligned} \quad (26)$$

which holds for all  $\mathbf{z}$ , with equality if and only if

$$\mathbf{z} = \Gamma \Sigma_y^{-1} \mathbf{y}. \quad (27)$$

Now with  $\mathbf{z}$  and  $\mathbf{u}$  fixed, the overall upper bound on  $L(\Gamma)$  decouples and we can solve for each  $\gamma_i$  individually by collecting the  $\boldsymbol{\gamma}$ -dependent terms from Eq. (24) and Eq. (26), leading to the problem

$$\min_{\gamma_i \geq 0} \frac{z_i^2 + u_i}{\gamma_i} + \log \gamma_i, \quad (28)$$

which has a simple closed-form solution.

Combining all of the above, and using matrix inversion formulae to produce numerically efficient computations, produces update rules for the  $(k+1)$ -th iteration given by

$$\begin{aligned} \gamma_i^{(k+1)} &\leftarrow \left( z_i^{(k)} \right)^2 + u_i^{(k)}, \forall i, \quad \Gamma^{(k+1)} = \text{diag}[\boldsymbol{\gamma}^{(k+1)}] \\ \mathbf{z}^{(k+1)} &\leftarrow \Gamma^{(k+1)} \left( S^{(k+1)} \right)^{-1} \mathbf{y} \\ \mathbf{u}^{(k+1)} &\leftarrow \text{diag} \left[ \Gamma^{(k+1)} - \left( \Gamma^{(k+1)} \right)^2 \left( S^{(k+1)} \right)^{-1} \right], \end{aligned} \quad (29)$$

where  $S^{(k+1)}$  is computed via

$$\begin{aligned} S^{(k+1)} &= D - DA \left[ \Sigma_x^{-1} + A^T DA \right]^{-1} A^T D \\ &\quad \text{and } D \triangleq (\Gamma^{(k+1)} + \lambda W^{-1})^{-1}. \end{aligned} \quad (30)$$

These expressions only require  $O(n)$  computations and are guaranteed to reduce  $L(\Gamma)$  until a fixed point  $\Gamma_*$  is reached. This value can then be plugged into Eq. (22) to estimate the unknown normal vector and model parameters. We denote this point estimator as  $\mathbf{x}_{sbl}$ . If the variances  $\Gamma_*$  reflect the true profile of the sparse errors, then  $\mathbf{x}_{sbl}$  will closely approximate

the true surface normal. This claim will be quantified more explicitly in the next section.

We have thus far omitted details regarding the choice of  $\lambda$ ,  $\sigma_n^2$  and  $\sigma_a^2$ . The first one can be reasonably set according to our prior expectations regarding the magnitudes of diffuse modeling errors, but in practice there is considerable flexibility here since some diffuse errors will be absorbed into  $e'$ . In contrast, we can realistically set  $\sigma_n^2 \rightarrow \infty$ , which implies zero precision or no prior information about the normal vectors and yet still leads to stable, convergent update rules. However we have observed that on certain problems a smaller selection for  $\sigma_n^2$  can lead to a modest improvement in performance, presumably because it has a regularizing effect that smoothes out the cost function and improves the convergence path of the update rules from Eq. (29). Interestingly and perhaps counter-intuitively, in certain situations it does not alter the globally optimal solution as discussed below. It is also possible to learn  $\sigma_n^2$  using similar updates to those used for  $\Gamma$ , but this introduces additional complexity and does not improve performance.

Finally, we discuss the effects of  $\sigma_a^2$  on our algorithm. Since the value of  $\sigma_a^2$  regularizes the shape of the piecewise linear function Eq. (10), smaller values will prevent the reverse of the sign of each linear segment and enforce the monotonicity of the function. While too small of a  $\sigma_a^2$  may limit the generality of  $g(I)$ , it can be proven that as  $\sigma_a^2 \rightarrow 0$  the model naturally reduces to a simple, Lambertian form for the diffuse component. We will empirically determine an appropriate value for  $\sigma_a^2$  in Section 4.3.

### 3.5 Analytical Evaluation

Previously we discussed two tractable methods for solving Eq. (18): a convex  $\ell_1$ -norm-based relaxation and a hierarchical Bayesian model called SBL. This section briefly discusses comparative theoretical properties of these approaches relevant to the photometric stereo problem. To facilitate the analysis, here we consider the idealized case where there are no diffuse modeling errors, or that  $\lambda$  is small. In this situation, the basic estimation problem reverts back to Eq. (17).

If the lighting directions and sparse errors are in general position, meaning they are not arranged in an adversarial configuration with zero Lebesgue measure, then it can be shown that the minimizer of Eq. (17), denoted  $\mathbf{x}_0$ , is guaranteed to be the correct normal vector as long as the associated feasible error component  $e' = \mathbf{y} - A\mathbf{x}_0$  satisfies  $\|e'\|_0 < m - p - 2$ . Therefore, a relevant benchmark for comparing photometric stereo algorithms involves quantifying conditions whereby a candidate algorithm can correctly compute  $\mathbf{x}_0$ .

In this context, recent theoretical results have demonstrated that any minimizer  $\mathbf{x}_1$  of the  $\ell_1$  relaxation approach will equivalently be a minimizer of Eq. (17) provided  $\|e'\|_0$  is sufficiently small relative to a measure of the structure in columns of the matrix  $A$  [46]. Unfortunately however, for typical photometric stereo problems the requisite equivalency conditions often do not hold (*i.e.*,  $\|e'\|_0$  is required to be prohibitively small) both because of structure imposed by the lighting geometry and implicit structure that emerges from the

relatively small dimensionality of the problem (meaning we do not benefit from asymptotic large deviation bounds that apply as  $m$  becomes large). Fortunately, SBL offers the potential for improvement over  $\ell_1$  via the following result.

**Theorem:** For all  $\sigma_n^2, \sigma_a^2 > 0$  (and assuming  $\lambda \rightarrow 0$ ), if  $\Gamma_*$  is a global minimum of Eq. (23), then the associated estimator  $\mathbf{x}_{sbl}$  will be a global minimum of Eq. (17). Moreover, for  $\sigma_n^2$  and  $\sigma_a^2$  sufficiently large it follows that: (i) Any analogous locally minimizing SBL solution is achieved at an estimate  $\mathbf{x}_{sbl}$  satisfying  $\|\mathbf{y} - A\mathbf{x}_{sbl}\|_0 \leq m - p - 2$ , (ii) SBL can be implemented with a tractable decent method such that convergence to a minimum (possibly local) that produces an  $\mathbf{x}_{sbl}$  estimator as good or better than the global  $\ell_1$  solution is guaranteed, meaning  $\|\mathbf{y} - A\mathbf{x}_{sbl}\|_0 \leq \|\mathbf{y} - A\mathbf{x}_1\|_0$ .

The proof, which has been omitted for brevity, is relatively straightforward and uses block-matrix inverse and determinant identities, as well as ideas from [46], to extend SBL properties derived in [47] to problems in the form of Eq. (17). We may thus conclude that SBL can enjoy the same theoretical guarantees as the  $\ell_1$  solution yet boosted by a huge potential advantage assuming that we are able to find the global minimum of Eq. (23), which will always produce an  $\mathbf{x}_{sbl} = \mathbf{x}_0$ , unlike  $\ell_1$ .

There are at least two reasons why we might expect this to be possible based on insights drawn from [47]. First, as discussed previously,  $A$  will necessarily have some structure unlike, for example, high dimensional random matrices. In this environment SBL performance is often vastly superior to  $\ell_1$  because it can be shown to be implicitly based on an  $A$ -dependent sparsity penalty that can compensate, at least in part, for structure in  $A$ . Secondly, the sparse errors  $e'$  will likely have substantially different magnitudes depending on image and object properties (meaning the non-zero elements of  $e'$  will not all have the same magnitude), and it has been shown that in this condition SBL is more likely to converge to the global minimum [47].

## 4 EXPERIMENTAL RESULTS

In this section, we quantitatively evaluate our method on synthetic and real image data. All experiments were performed on an Intel Core2 Duo E6400 (2.13GHz, single thread) machine with 4GB RAM and were implemented in MATLAB. For the SBL- and  $\ell_1$ -based methods we used  $\lambda = 1.0^{-6}$  in the synthetic experiments with no additive image noise and perfect lighting calibrations (Section 4.2 and Section 4.3), and  $\lambda = 10^{-2}$  for the other cases (Section 4.4). We set  $\sigma_n^2 = 10^6$  for all experiments. As for  $\sigma_a^2$ , which can affect the solution when  $p > 1$ , we experimentally find the optimal value in Section 4.3-(a).

### 4.1 Datasets

For quantitatively evaluating our method, four different datasets are used. We generate 32-bit HDR images of two target scenes, *Bunny* ( $256 \times 256$ ) and *Caesar* ( $300 \times 400$ ), illuminated under random directional lightings. We use a few

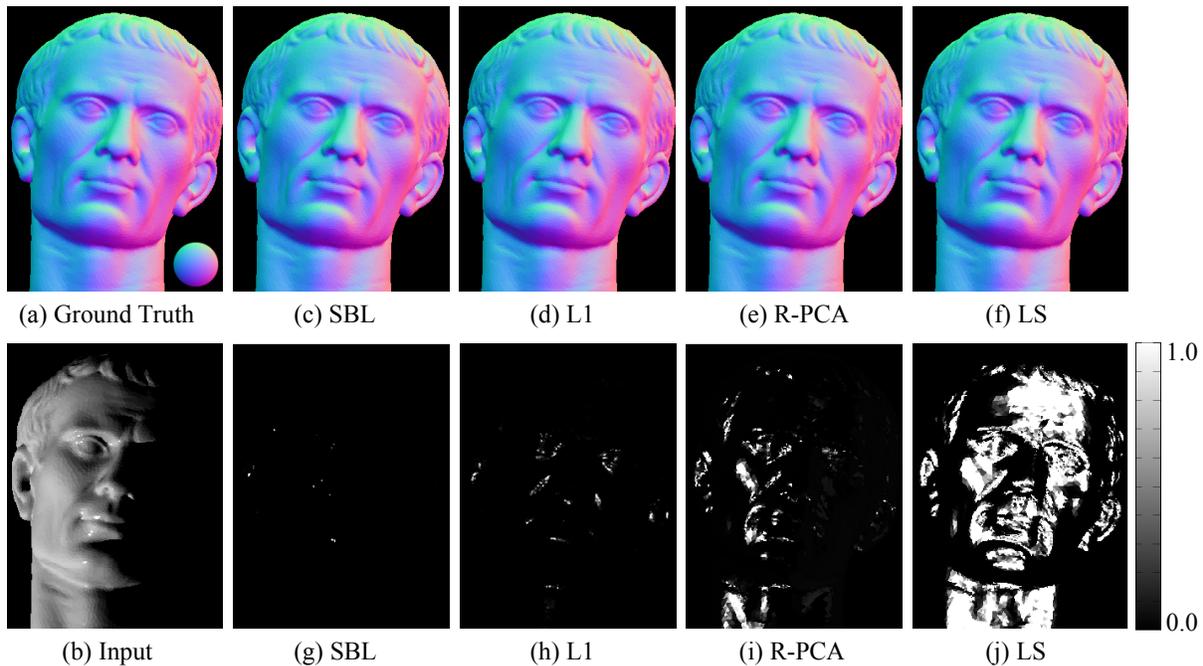


Fig. 4. Recovery of surface normals from 40 images of *Caesar* ( $300 \times 400$ ) in dataset (A) with explicit shadow removal. (a) Input, (b) Ground truth, (c)-(j) Recovered surface normals and Error maps (in degrees).

TABLE 1  
Experimental results of *Bunny* (left) / *Caesar* (right) in dataset (A) with varying number of images.

No. of images	Mean error (in degrees)				Median error (in degrees)				Elapsed time (sec)			
	SBL	L1	R-PCA	LS	SBL	L1	R-PCA	LS	SBL	L1	R-PCA	LS
5	<b>6.0</b>	6.0	15.3	7.0	4.7	<b>4.3</b>	10.7	4.8	46.5	13.6	15.7	<b>4.6</b>
10	<b>0.09</b>	0.61	3.8	1.9	<b>0.27</b>	0.58	0.81	1.8	36.3	13.6	37.8	<b>5.9</b>
15	<b>0.076</b>	0.16	0.21	1.6	<b>0.052</b>	0.13	0.19	1.6	26.8	13.1	55.1	<b>6.3</b>
20	<b>0.033</b>	0.080	0.11	1.6	<b>0.022</b>	0.078	0.11	1.6	24.2	13.5	70.5	<b>6.9</b>
25	<b>0.018</b>	0.055	0.084	1.6	<b>0.010</b>	0.048	0.069	1.6	23.1	14.1	86.0	<b>7.6</b>
30	<b>0.012</b>	0.037	0.080	1.7	<b>0.0048</b>	0.032	0.065	1.7	23.1	14.2	121.0	<b>8.4</b>
35	<b>0.0057</b>	0.023	0.098	1.6	<b>0.0029</b>	0.019	0.093	1.6	22.7	14.6	161.3	<b>8.5</b>
40	<b>0.0039</b>	0.019	0.12	1.6	<b>0.0020</b>	0.015	0.12	1.6	22.6	15.0	200.7	<b>9.4</b>

different BRDF settings for rendering; (A) combination of Lambertian diffuse reflection and specularity of the Cook-Torrance reflectance model [48], (B) combination of Lafortune diffuse reflection [49] and Cook-Torrance specularity, and (C) MERL BRDF database [41]. Additionally, as the fourth dataset, denoted (D), we record real images for qualitatively evaluating our method in a practical scenario. For the datasets (A), (B), and (C), both cast shadows and attached shadows are also rendered. Note that though light calibration is noise-free, and shadowed pixel intensities are exactly zero in the synthesized datasets, some calibration errors and non-zero shadowed pixels exist in the real data (D).

#### 4.2 Performance Evaluation with $p = 1$

We begin with the simplest case where the number of basis functions in the piecewise linear model is one *i.e.*, the diffuse function is assumed to be Lambertian. We evaluate performance using the synthesized images in dataset (A), which has Lambertian diffuse reflections along with non-Lambertian

specularities and shadows. Here we compare our SBL-based and convex  $\ell_1$ -norm-based sparse regression models from Sections 3.3 and 3.4 with the recent R-PCA-based method proposed by Wu *et al.* [6], which attempts the simultaneous recovery of a low-rank Lambertian factor and sparse corruptions. We also compare with the standard least squares (LS) Lambertian photometric stereo estimator [1], which is equivalent to solving Eq. (14) with  $p = 1$ .

We change experimental conditions with regard to the number of images, surface roughness (*i.e.*, the ratio of specularities), shadow removal (*i.e.*, whether or not a shadow mask is used to remove zero-valued elements from the observed images), and the presence of additional Gaussian noise. Note that when in use as defined for each experiment, the shadow mask is applied equivalently to all algorithms. To increase statistical reliability, all experimental results are averaged over 20 different sets of 40 input images. While the amount of cast/attached shadows are fixed to 24.0% and 27.8% of the number of all the pixels, respectively, the

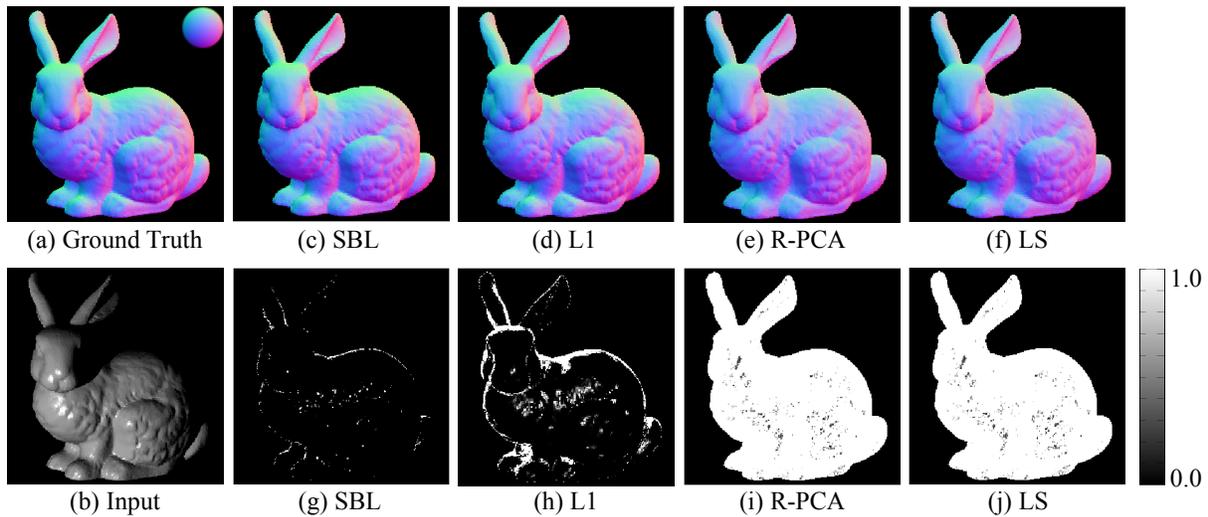


Fig. 5. Recovery of surface normals from 40 images of *Bunny*( $256 \times 256$ ) in dataset (A) without explicit shadow removal. (a) Input, (b) Ground truth, (c)-(j) Recovered surface normals and Error maps (in degrees).

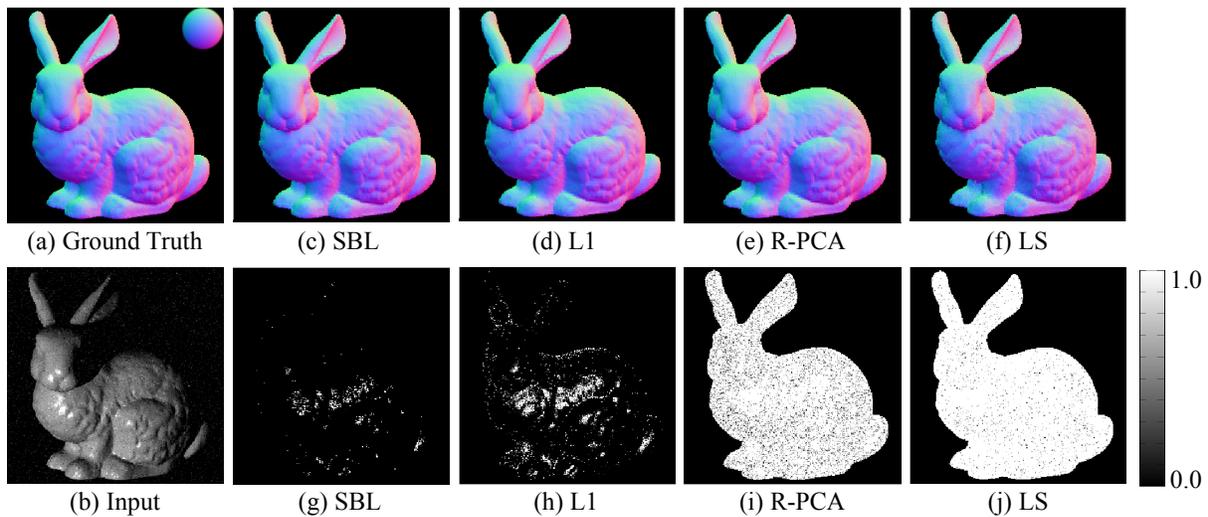


Fig. 6. Recovery of surface normals from 40 images of *Bunny* in dataset (A) with explicit shadow removal and additive Gaussian noises (30%). (a) Input, (b) Ground truth, (c)-(j) Recovered surface normals and Error maps (in degrees).

percentage of specularities depends on the surface roughness parameter that is independently selected in each experiment. To quantitatively evaluate the performance, we compute the angular error between the recovered normal map and the ground truth.

#### Valid number of images for efficient recovery in the presence of specularities

In this experiment, we vary the number of images to estimate the minimum number required for effective recovery when using the shadow mask with fixed surface roughness. The percentage of specular pixels in *Bunny* and *Caesar* are 8.4% and 11.6%, respectively.

Once 40 images are generated for each dataset, the image subset is randomly subsampled without replacement. The results are displayed in Table 1 and Fig. 4. We observe that the sparse-regression-based methods are significantly more

accurate than both R-PCA and LS. It is also clear that SBL is more accurate than  $\ell_1$ , although somewhat more expensive computationally.<sup>1</sup> Note that, although not feasible in general, when the number of images is only 5, the most accurate and efficient implementation for regression could be to just systematically test every subset of 3 images (*i.e.*, brute force search only requires 10 iterations at each pixel).

#### Robustness to shadows and image noise

We now evaluate the robustness of our method against corruptions; shadows and image noise. We set two conditions for evaluating the effects of (i) shadows (fixed specularities, no shadow removal, no image noise), (ii) additive Gaussian image

1. The SBL convergence rate can be slower with fewer images because of the increased problem difficulty. This explains why the computation time may actually be shorter with more images.

TABLE 2

Results of *Bunny* in dataset (A) without shadow removal.

No. of images	Mean error (in degrees)				Median error (in degrees)				Elapsed time (sec)			
	SBL	L1	R-PCA	LS	SBL	L1	R-PCA	LS	SBL	L1	R-PCA	LS
5	<b>5.2</b>	11.9	12.1	12.1	<b>5.0</b>	12.3	12.5	12.5	213.0	37.0	45.8	<b>5.1</b>
10	<b>2.8</b>	5.6	10.9	10.9	<b>2.3</b>	5.6	11.3	11.3	98.9	33.0	93.7	<b>6.0</b>
15	<b>1.9</b>	4.0	9.9	10.0	<b>2.3</b>	4.0	10.1	10.2	66.8	32.5	153.3	<b>7.4</b>
20	<b>1.2</b>	2.7	9.4	9.5	<b>1.0</b>	2.7	9.6	9.6	52.9	30.0	177.5	<b>7.6</b>
25	<b>0.81</b>	1.9	8.9	9.0	<b>0.69</b>	1.8	8.9	9.0	46.2	31.0	196.9	<b>9.1</b>
30	<b>0.62</b>	1.6	9.0	9.1	<b>0.61</b>	1.5	8.9	8.9	41.1	32.0	231.7	<b>9.4</b>
35	<b>0.59</b>	1.5	9.1	9.1	<b>0.58</b>	1.4	9.3	9.3	41.1	34.4	259.2	<b>11.0</b>
40	<b>0.53</b>	1.2	8.8	8.9	<b>0.58</b>	1.2	9.0	9.1	39.4	33.3	281.2	<b>10.7</b>

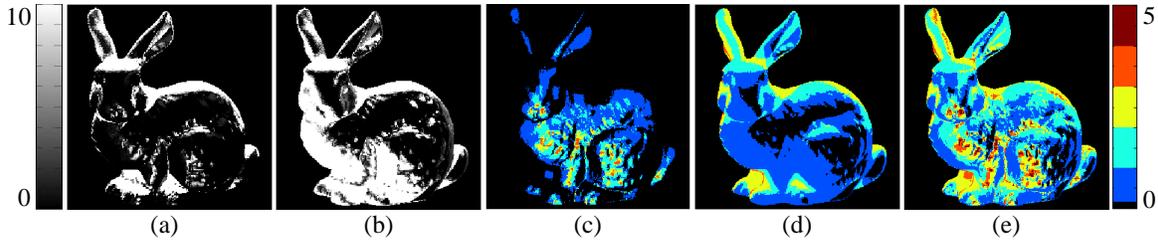


Fig. 7. Comparison between SBL and  $\ell_1$ -based method. Error maps of (a) SBL and (b) L1 (in degrees). The per-pixel number of (c) specularities, (d) shadows, (e) corruptions (The maximum is 5). These maps are illustrated based on one of twenty datasets.

noise (fixed specularities, explicit shadow removal, varying amount of image noise). The ability to estimate surface normals without an explicit shadow mask is important, since in practical situations shadow locations are not always easy to be determined a priori. The number of images is 40 in (ii) and varying from 5 to 40 in (i). We use *Bunny* for evaluation, and the ratio of specularities is 8.4% in (i) and (ii), and image noise is 10% to 50% in (ii). Image noise obeys a zero-mean iid Gaussian distribution with  $\sigma^2 = 0.1$ .

The results are illustrated in Fig. 5, Fig. 6, Table 2 and Table 3. While performance of each method degrades when additional corrupted pixels (outliers) contaminate the estimation process, our sparse regression methods outperform both R-PCA and LS in accuracy and outperform R-PCA in efficiency in the presence of shadows and noise. We observe that the performance of R-PCA degrades when a shadow mask is unavailable (*i.e.*, the positions of missing entries are unknown) while our sparse-regression based method automatically compensates for missing entries in the estimation process.

We also observe that SBL is more accurate than  $\ell_1$  in all conditions with slight more expense of computational effort. We further compare the results of SBL and  $\ell_1$  using the case where the number of images is 5 without removing shadows. The error maps and the numbers of corruptions per-pixel are displayed in Fig. 7. We observe that the  $\ell_1$  method typically suffer from shadows while SBL can find the correct solution in most pixels as long as the number of corruptions is less than 3, which is the theoretical limit when only 5 observations are given.

### Sparseness of corruptions

While it does not often occur, our assumption that corruptions

TABLE 3

Experimental results of *Bunny* in dataset (A) with varying amount of additive Gaussian noises.

Dens. of noises (%)	Mean error (in degrees)				Median error (in degrees)			
	SBL	L1	R-PCA	LS	SBL	L1	R-PCA	LS
10	<b>0.0079</b>	0.040	0.16	3.3	<b>0.0060</b>	0.039	0.16	3.3
20	<b>0.021</b>	0.11	0.79	4.4	<b>0.019</b>	0.099	0.80	4.3
30	<b>0.068</b>	0.29	3.6	5.3	<b>0.060</b>	0.25	3.2	5.2
40	<b>0.21</b>	0.70	9.8	6.2	<b>0.18</b>	0.63	9.9	6.1
50	<b>0.58</b>	1.5	11.7	7.0	<b>0.53</b>	1.4	11.7	6.9

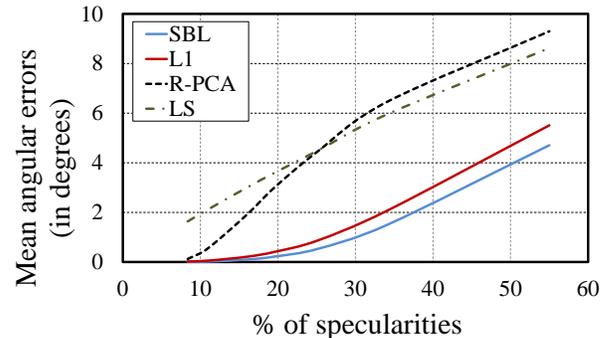


Fig. 8. Experimental results of *Bunny* in dataset (A) with varying amount of specularities. The x-axis and y-axis indicate the ratio of specularities and the mean angular error of normal map.

appear sparsely over images can be violated in some situations, *e.g.*, observations include wide-lobe specularities. In this experiment, we examine how our approach is affected by the sparseness of corruptions by changing the width of the specular lobe (*i.e.*, percentage of specularities appeared in observations). Here, we use *Bunny* dataset whose amount of specularities is varying from 10% to 60%, respectively, by appropriately choosing the model parameter of Cook-Torrance reflectance model. In this experiment, we remove shadows explicitly to factor out their effect.

The result is illustrated in Fig. 8. As expected, performance of both SBL and  $\ell_1$  degrades as the ratio of specular corruptions increases. However, even when the sparseness of outliers is difficult to be held, we observe that our

sparse-regression based method still outperforms R-PCA and LS.

### Comparison with RANSAC-based approaches

We also compared our approach with two different RANSAC-based methods. Here we use 40 images of *Bunny* dataset whose percentage of specular pixels is 8.4%. Shadows are removed and images noises are not added in this experiment.

The first one is a photometric linearization approach [3], which converts an input image into the ideal Lambertian image. It uses RANSAC for robustly identifying the basis images in the presence of non-Lambertian corruptions. Once linearized images are acquired, the standard Lambertian-based approach [1] is applied with known lightings to estimate surface normals. The second one uses RANSAC more directly in Lambertian photometric stereo method like [21]. In this approach, three images are randomly sampled for each position independently, and the surface normal and albedo are estimated from them using Lambertian photometric stereo [1]. Then the Euclidean distances between observations and predicted intensities are computed, and the number of *inliers* whose distances are less than a threshold are counted. After the sampling process, the surface normal and albedo with the maximum number of inliers are adopted.

The results are summarized in Table 4 and Table 5, respectively. We have also included standard deviations in the tables for observing the estimation stability. We observe in Table 4 that although we use a large number of samples for RANSAC (e.g., 2000) for the photometric linearization, this approach cannot always stably find the solution especially when the number of images is large. On the other hand, our method is able to successfully and efficiently finds the solution. We also observe in Table 5 that a large number of samples (e.g., 200) gives very accurate estimation while it takes much time to compute it. The smaller number of samples gives more efficient but less accurate estimation. When the number of samples is 20, the computational cost is almost same as our SBL-based method, but the mean error is 50 times larger than ours, which also demonstrates the effectiveness of the proposed method.

It is also crucial to emphasize that the RANSAC algorithm scales very poorly as the number of inlier model parameters is increased. Simply put, a larger number of inlier random samples is required to robustly estimate a larger number of model parameters, and obtaining such inlier samples becomes combinatorially more difficult in higher dimensions. Consequently, while RANSAC may work reasonably well here recovering surface normals under a Lambertian diffusive model (where the number of unknowns is effectively only 3), it will become intractable when using more complex reflectance models, which is a central concern herein. In contrast, our method scales linearly in the number of inlier parameters and can therefore be robustly expanded to handle non-Lambertian parameterized reflectance functions with outliers, e.g., the  $p > 1$  model detailed in Section 4.3 or other more sophisticated extensions.

TABLE 4

Comparison with RANSAC based approach [3] with *Bunny* in dataset (A) (Number of samples is 2000).

No. of Images	Mean error		Median error		Standard deviation		Elapsed time	
	SBL	[3]	SBL	[3]	SBL	[3]	SBL	[3]
5	<b>6.0</b>	6.7	<b>4.7</b>	5.4	<b>4.1</b>	4.4	<b>46.5</b>	52.3
10	<b>0.09</b>	0.74	<b>0.27</b>	0.38	<b>0.35</b>	1.4	<b>36.3</b>	544.0
15	<b>0.076</b>	0.61	<b>0.052</b>	0.12	<b>0.059</b>	1.9	<b>26.8</b>	958.4
20	<b>0.033</b>	0.70	<b>0.022</b>	0.058	<b>0.027</b>	1.2	<b>24.2</b>	1048.9
25	<b>0.018</b>	1.0	<b>0.010</b>	0.063	<b>0.020</b>	2.6	<b>23.1</b>	1141.8
30	<b>0.012</b>	2.3	<b>0.0048</b>	0.042	<b>0.012</b>	4.0	<b>23.1</b>	1227.8
35	<b>0.0057</b>	3.2	<b>0.0029</b>	0.051	<b>0.0064</b>	9.2	<b>22.7</b>	1327.9
40	<b>0.0039</b>	2.1	<b>0.0020</b>	0.046	<b>0.0048</b>	4.3	<b>22.6</b>	1430.4

TABLE 5

Performance of RANSAC based approach [21] with *Bunny* in dataset (A) (Number of images is 40).

No. of Samples	Mean error	Median error	Standard deviation	Elapsed time
200	0.000019	7.9E-06	0.000027	283.0
100	0.0035	7.6E-06	0.31	142.4
80	0.0099	7.6E-06	0.70	115.2
60	0.019	7.6E-06	0.98	86.0
40	0.034	7.3E-06	1.2	57.9
20	0.22	7.4E-06	3.4	29.8
10	0.51	7.5E-06	5.0	15.6

### 4.3 Performance Evaluation with $p > 1$

If the majority of observations are represented by non-linear diffuse reflections, then piecewise linear sparse regression with  $p > 1$  basis functions or segments is expected to be considerably more effective. In this experiment, we evaluate the  $p$ -functions piecewise sparse linear regression to those complex objects by using the dataset (B) rendered with non-linear Lafortune diffuse reflectance model [49] and Cook-Torrance Reflectance model, and (C) rendered with one hundred BRDF functions from the MERL BRDF database [41].

We generate 32-bit HDR images of *Bunny* ( $256 \times 256$ ) scene with Lafortune diffuse reflectance and Torrance sparrow specularities under 40 different lightings in dataset (B), and 100 different lightings for 100 different materials in dataset (C). Note that the average percentage of specularities in dataset (B) is 8.4% and 24.0% of cast/attached shadows are rendered in both dataset (B) and dataset (C).

Here, in addition to the SBL-based piecewise linear sparse regression (PL-SBL), we further implement the method which is solving Eq. (14) (PL-LS). We compare our methods with the R-PCA-based method [6] and a recent parametric non-Lambertian photometric stereo method with biquadratic reflectance model [11], which reasonably represents the low-frequency of non-linear reflectance though shadows and high-frequency observations must be removed in advance. The

biquadratic model is a approximated BRDF as follow,

$$I = \alpha_1(\mathbf{n}^T \mathbf{l})^2 (\mathbf{l}^T \mathbf{h})^2 + \alpha_2(\mathbf{n}^T \mathbf{l})^2 (\mathbf{l}^T \mathbf{h}) + \alpha_3(\mathbf{n}^T \mathbf{l})^2 + \alpha_4(\mathbf{n}^T \mathbf{l}) (\mathbf{l}^T \mathbf{h})^2 + \alpha_5(\mathbf{n}^T \mathbf{l}) (\mathbf{l}^T \mathbf{h}) + \alpha_6(\mathbf{n}^T \mathbf{l}) + \alpha_7(\mathbf{l}^T \mathbf{h})^2 + \alpha_8(\mathbf{l}^T \mathbf{h}) + \alpha_9, \quad (31)$$

where  $\mathbf{h}$  is a half vector directed at  $\mathbf{l} + \mathbf{v}$  and  $\alpha \triangleq [\alpha_1, \alpha_2, \dots, \alpha_9]$  are model parameters. In the implementation of [11], we used the non-shadowed pixels whose intensities are ranked below the 25%, and both surface normals and model parameters are iteratively updated initialized by the Lambertian photometric stereo method. Note that, shadowed pixels are also rejected, but all observations are taken into account in other methods.

### Choosing $\sigma_a^2$ and $p$

Here we examine the choices for  $\sigma_a^2$  and for the number of basis functions  $p$  in piecewise linear model by using dataset (C). Note that if  $p$  is large enough, we can essentially represent any complex non-linear function, although we run the risk of over-fitting when  $m$  is too small. However, we can always compensate to some extent via  $\sigma_a^2$ , since regardless of  $p$ , for  $\sigma_a^2$  sufficiently small we approximate the standard Lambertian model with all linear segments having equal slope, and therefore equal diffuse albedo.

In this experiment, we vary individually  $\sigma_a^2$  and  $p$  to find the optimal parameters. The results are illustrated in Fig. 9. Here, average angular errors in each normal map are further averaged over 100 materials. We observe that as expected, too small  $\sigma_a^2$  deteriorates the performance, yet does not affect the performance when the value is sufficiently large. Therefore, in the following experiments, we fix  $\sigma_a^2$  by 1.0. As for the number of basis functions, it appears that  $p = 3$  is optimal in the case of dataset (C).

We also illustrate per-material angular errors with different  $p$  in Fig. 10. We observe that piecewise linear function with many basis functions, *e.g.*, 6, works very well under the material with narrow specular peaks, *e.g.*, for (1) specular-white-phenolic and (4) gray-plastic, while it suffers from over-fitting in polished metal with the broad specularity, *e.g.*, for (82) silver-metallic paint or materials with complex 2-lobe BRDF, *e.g.*, (53) natural-209 and (83) ipswich-pine-221), where we can hardly see the underlying monotone diffuse structure in observations.

### Quantitative comparison with other methods

We evaluate of the performance of our method by a numerical comparison with other methods using datasets (B) and (C). Here, our methods (PL-SBL, PL-LS) are compared with the biquadratic photometric stereo method by Shi *et al.* [11] (Biquadratic), LS and R-PCA (R-PCA is included only for dataset (C)).

First, we use dataset (B) to verify that our inverse piecewise linear diffuse model reasonably handles non-Lambertian diffuse reflections. In the Lafortune model [49], a general rotationally symmetric diffuse component  $\rho_d$  is written as  $\rho_d = (\mathbf{n}^T \mathbf{l})^k (\mathbf{n}^T \mathbf{v})^k$ . Here  $k$  is a model parameter which is fixed to 3.0 in our experiment. Note that specularities of

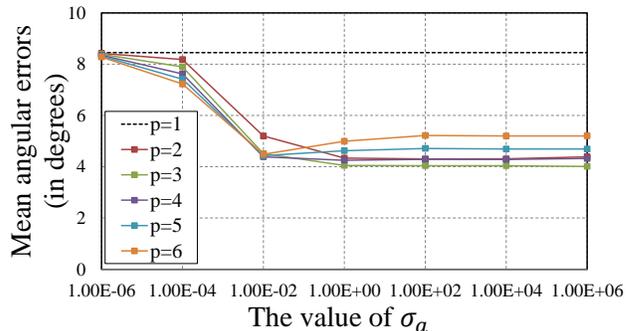


Fig. 9. Experimental results of dataset (B) with different values of  $p$  and  $\sigma_a^2$ . The results are averaged over 100 different materials.

dataset (B) are also rendered with Cook-Torrance model in the same manner with dataset (A). Therefore we can consider dataset (B) as a combination of dominant general diffuse reflections and sparse specular reflections.

The results are illustrated in Fig. 11 and Table 6. We present the results of PL-SBL with  $p = 1$  and  $p = 3$ , and PL-LS with  $p = 3$ . Note that PL-LS with  $p = 1$  is exactly same with LS. We observe that while non-Lambertian diffuse reflections degrade the performances of LS and PL-SBL( $p = 1$ ), Biquadratic, PL-LS( $p = 3$ ) and PL-SBL( $p = 3$ ) work better since they are potentially capable of non-Lambertian reflections. We also observe that PL-SBL( $p = 3$ ) performs best since the estimation of Biquadratic and PL-LS are disrupted by specularities which are not included in each model while the sparsity penalty in PL-SBL( $p = 3$ ) reasonably neglects them as the model outlier.

Secondly, we evaluated the performance of our method by using dataset (C). The results are illustrated in Fig. 12. We observe that while SBL( $p = 1$ ) outperforms R-PCA and LS in most of materials, the average angular errors are large for various complex materials. In contrast, PL-SBL( $p = 3$ ) works better for those kind of objects since the piecewise linear function with several basis functions can capture the nonlinearity of the non-Lambertian diffuse structure. We also observe that the Biquadratic model is the most effective method of all for dataset (C), however we emphasize that our method achieves competitive performance even without heuristically filtering high-frequency specularities. Consequently, on real images our approach produces qualitatively superior results (see Section 4.4 below). Finally, we can see from the comparison between PL-LS( $p = 3$ ) and PL-SBL( $p = 3$ ) that the sparsity penalty is also helpful for large  $p$  in our method, which can reject specularities efficiently.

We also compare our method with others in the presence of shadows and image noises. The results are illustrated in Table 7. We observe that our SBL-based methods still work under attached/cast shadows and image noises in contrast that Biquadratic and least-squares-based methods are easily disrupted by those corruptions.

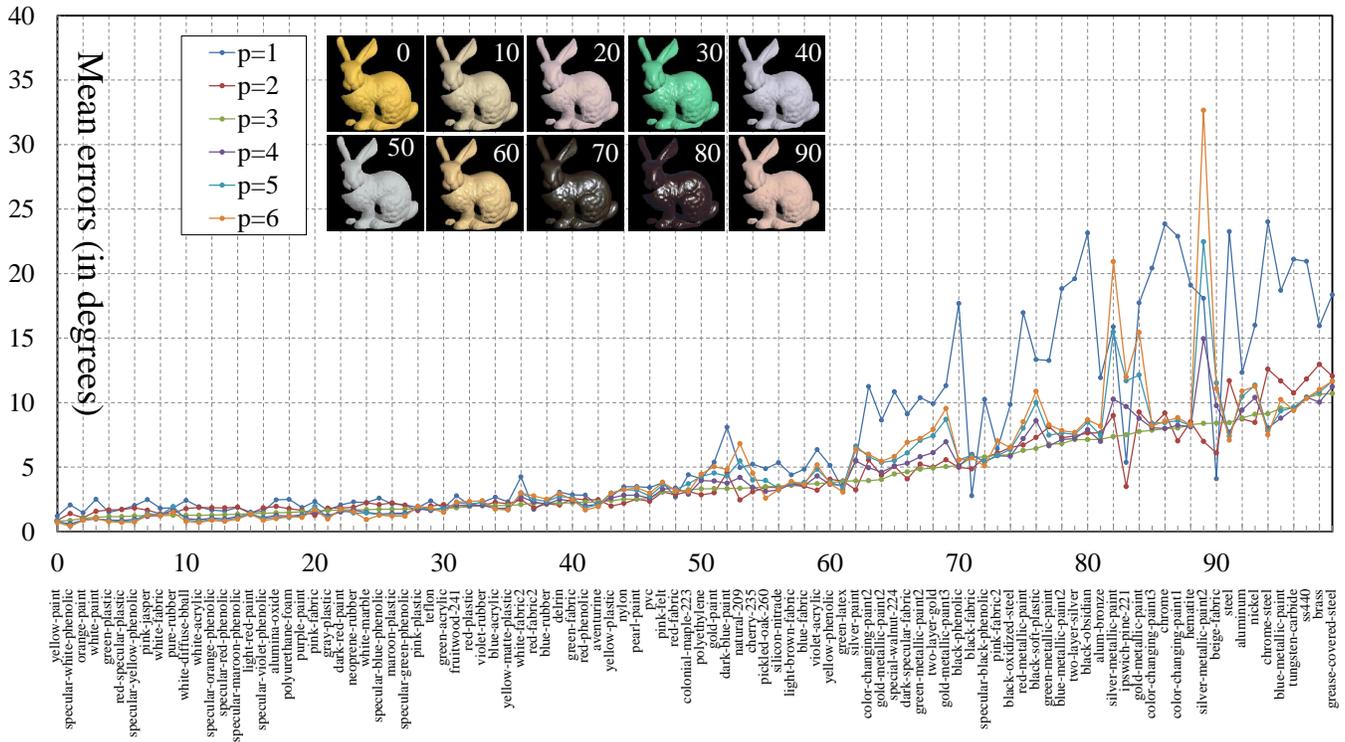


Fig. 10. Experimental results of dataset (C) under the different number of basis functions [41]. We aligned results in ascending order of mean angular error of PL-SBL( $p = 3$ ). Some rendered images corresponding to each material ID are also illustrated.

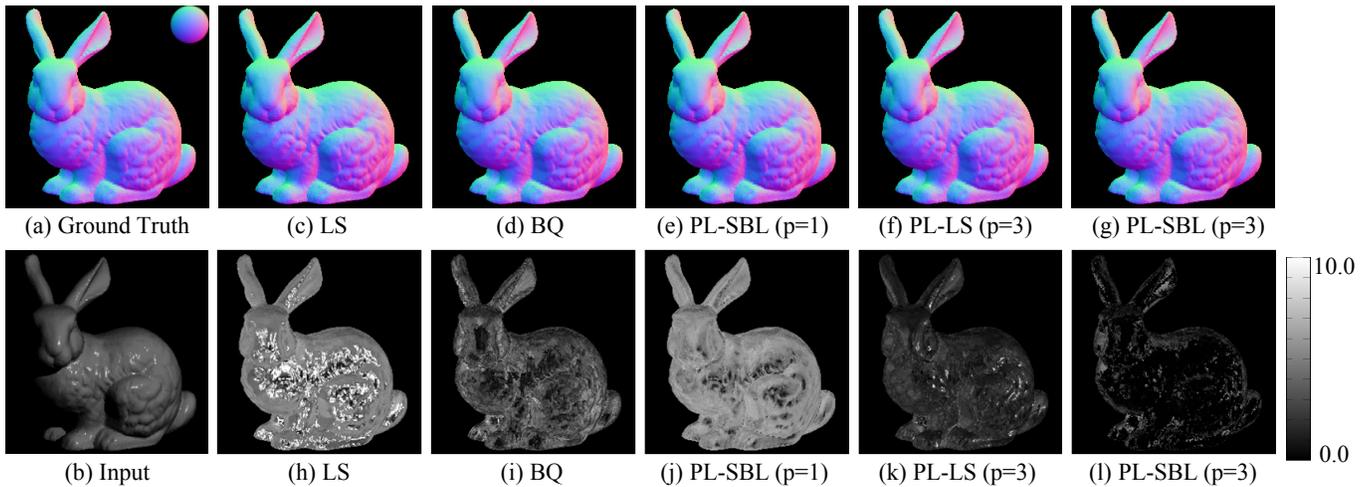


Fig. 11. Recovery of surface normals from 40 images of *Bunny* in dataset (B) with explicit shadow removal. (a) Input, (b) Ground truth, (c)-(l) Recovered surface normals and Error maps (in degrees).

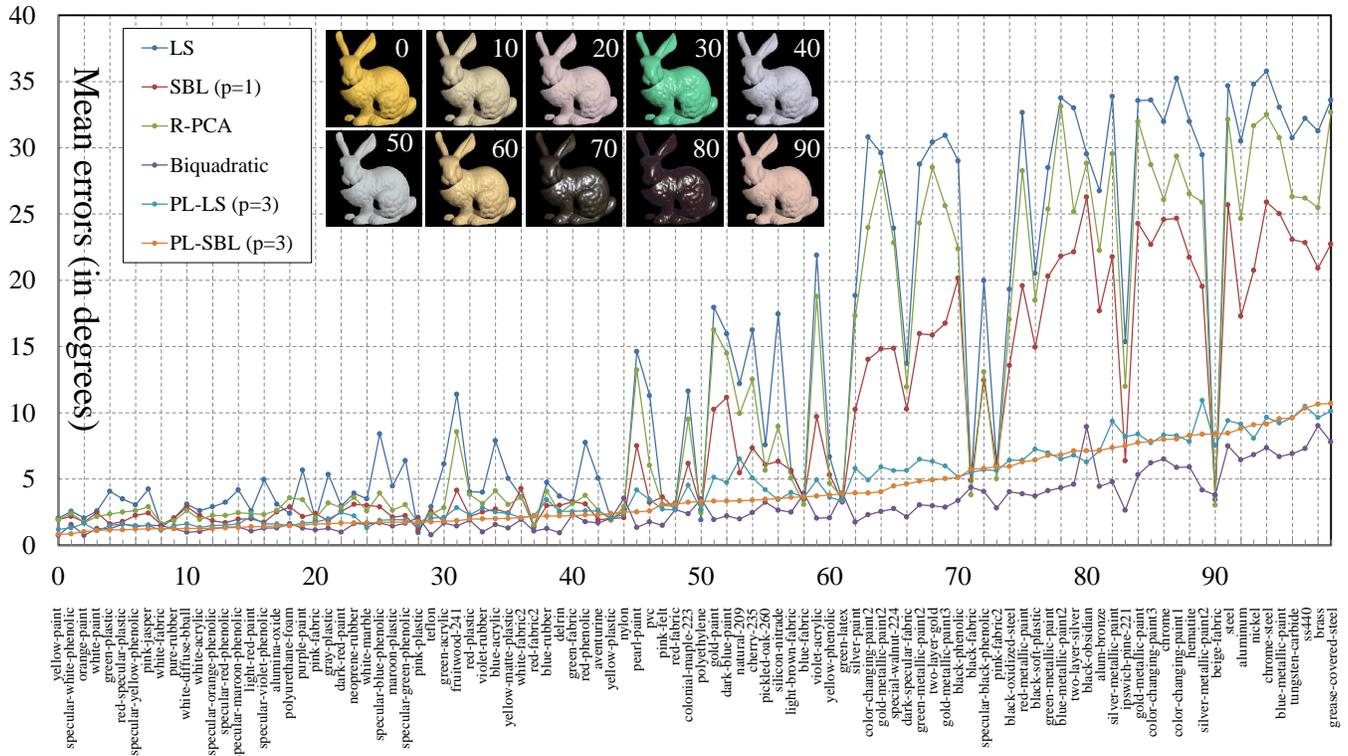


Fig. 12. Comparison among different methods with dataset (C). We aligned results in ascending order of mean angular error of PL-SBL( $p = 3$ ). Some rendered images corresponding to each material ID are also illustrated.

TABLE 6

Comparison among different methods with dataset (B).

No. of Images	Mean error (in degrees)				
	LS	BQ	PL-SBL (p=1)	PL-LS (p=3)	PL-SBL (p=3)
5	<b>4.9</b>	6.9	5.4	25.1	6.0
10	3.7	4.4	3.3	2.6	<b>1.5</b>
15	3.7	3.5	3.5	1.4	<b>0.80</b>
20	3.8	2.8	3.6	1.3	<b>0.67</b>
25	3.8	2.2	3.6	1.3	<b>0.58</b>
30	3.7	1.9	3.5	1.2	<b>0.49</b>
35	3.8	1.7	3.6	1.2	<b>0.47</b>
40	3.8	1.6	3.7	1.2	<b>0.49</b>

TABLE 7

Experimental results of dataset (C) under the different kind of corruptions. This table illustrates average angular errors of surface normals for 100 materials.

	Mean error (in degrees)				
	LS	SBL (p=1)	Biquadratic	PL-LS (p=3)	PL-SBL (p=3)
w/o shadow and noise	13.9	8.9	<b>2.9</b>	4.4	4.1
w/ shadow	14.3	<b>7.7</b>	44.0	28.0	10.5
w/ noise	35.8	8.2	11.0	14.1	<b>4.8</b>

#### 4.4 Qualitative Evaluation with Real Images

We also evaluate our algorithm (PL-SBL) using real images (dataset (D)). We captured RAW images without gamma correction by Canon 30D camera with a 200[mm] tele-photo lens and set it 1.5[m] far from target object. Lighting conditions are randomly selected from a hemisphere whose radius is 1.5[m] with the object placed at the center. For calibrating light sources, a glossy sphere was placed in the scene. We use a set of 25 images of *Chocolate bear* ( $261 \times 421$ ), and 40 images each of *Doraemon* ( $269 \times 420$ ) and *Fat guy* ( $293 \times 344$ ). Note that in this experiment, we did not remove shadows from images by zero-intensity thresholding since even shadowed pixels have non-zero values due to the presence of slight ambient illumination, sensor saturations, low signal-to-noise ratio, inter-reflections between the object and the floor and so on. Therefore, we sort all observations in increasing order and use lowest 25% of observations for the low-intensity condition of Biquadratic (low). We also applied the biquadratic reflectance model using all observations denoted as Biquadratic (all). We evaluate the performance of PL-SBL by visual inspection of the output normal maps and recovered surface meshes by a poisson solver [50], which recovers surface meshes from a gradient map.

The results are illustrated in Fig. 13. First, we observe that PL-SBL( $p = 1, p = 3$ ) succeeds in efficiently rejecting specularities and estimating smoother and more reasonable normal maps. While Biquadratic (low) has produced the highest performance in Section 4.4(b), it has substantial difficulty with

real images where calibration errors, shadowed pixels, and sensor saturations are often mis-classified as low-frequency reflectance, leading to unpredictable errors in practice (see Table 7). We note that increasing the threshold for shadow removal might solve this problem, but the optimal threshold selection remains hard and scene-dependent. Although including higher frequency observations may sometimes give better results by diluting low-frequency corruptions with a larger number of samples (see Biquadratic (all)), accurate estimation still remains hard since the simple biquadratic model has difficulty representing complex, non-linear high-frequency observations (as is mentioned in [11]).

Second, we observe that reconstructed surface meshes by PL-SBL( $p = 3$ ) are more reasonable than PL-SBL( $p = 1$ ) *e.g.*, the stomach of *fatguy* recovered by PL-SBL( $p = 1$ ) is shaper than that of PL-SBL( $p = 3$ ), though we can not make the further quantitative comparison due to the lack of the ground truth. From those observations, we can say that our method is effective, especially when  $p > 1$ , in the practical situation.

## 5 CONCLUSION

In this paper, we have proposed the inverse piecewise linear sparse regression based photometric stereo method which works under various kind of non-diffuse corruptions. At first, we presented a special case of Lambertian-based sparse regression which consider all kind of non-Lambertian corruptions as outliers. Our detailed discussion and experimental results has shown that our method are not only robust to those kind of corruptions, but also computationally more efficient and stable comparing than other robust approaches. However, since Lambertian sparse regression could not handle non-Lambertian materials, we have secondly discussed the general case where the piecewise linear function has several basis functions. Our extensive evaluation with both synthetic and real images indicated that our method works not only for various non-diffuse corruptions, but also real images where the shadow removal is not easy.

Our current limitation is that we assume the existence of the unique dominant diffuse structure in observations. Therefore our method cannot potentially handle materials with unknown dominant structure *e.g.*, materials with unknown BRDF where dominant structure is specularities or 2-lobe BRDF which is represented by a sum of several functions whose dominant directions are unknown *polyethylene and natural ones like fabrics*. We are very interested in extending our work to handle those complex materials. The other interesting direction of our future work is applying our method to the uncalibrated case since we currently assume that the lighting environment is calibrated. Since the simple structure of our image formation model have the possibility to be combined with other known constraints such as *isotropy symmetry*, it seems that the re-alization of the uncalibrated photometric stereo method for general diffuse materials is not impossible.

## ACKNOWLEDGMENTS

This work was supported by the Grants-in-Aid for JSPS Fellows (248615).

## REFERENCES

- [1] P. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Engg.*, vol. 19, no. 1, pp. 139–144, 1980.
- [2] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur, "Beyond lambert: reconstructing specular surfaces using color," *Proc. CVPR*, 2005.
- [3] Y. Mukaigawa, Y. Ishii, and T. Shakunaga, "Analysis of photometric factors based on photometric linearization," *Journal of the Optical Society of America*, vol. 24, no. 10, pp. 3326–3334, 2007.
- [4] C. Yu, Y. Seo, and S. Lee, "Photometric stereo from maximum feasible lambertian reflections," in *Proc. ECCV*, 2010.
- [5] K. Sunkavalli, T. Zickler, and H. Pfister, "Visibility subspaces: Uncalibrated photometric stereo with shadows," *Proc. ECCV*, 2010.
- [6] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," in *Proc. ACCV*, 2010.
- [7] A. Georgiades, "Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo," *Proc. ICCV*, 2003.
- [8] H. Chung and J. Jia, "Efficient photometric stereo on glossy surfaces with wide specular lobes," *Proc. CVPR*, 2008.
- [9] D. Goldman, B. Curless, A. Hertzmann, and S. Seitz, "Shape and spatially-varying brdfs from photometric stereo," in *Proc. ICCV*, October 2005.
- [10] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proc. CVPR*, 2008.
- [11] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "A biquadratic reflectance model for radiometric image analysis," *Proc. CVPR*, 2012.
- [12] —, "Elevation angle from reflectance monotonicity," in *Proc. ECCV*, 2012.
- [13] E. Vieth, "Fitting piecewise linear regression functions to biological responses," *Journal of Applied Physiology*, vol. 67, no. 1, pp. 390–396, 1989.
- [14] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," *Proc. CVPR*, 2012.
- [15] S. Barsky, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1239–1252, 2003.
- [16] V. Argyriou and M. Petrou, "Recursive photometric stereo when multiple shadows and highlights are present," *Proc. CVPR*, 2008.
- [17] J. Ackermann, F. Langguth, S. Fuhrmann, and M. Goesele, "Photometric stereo for outdoor webcams," *Proc. CVPR*, 2012.
- [18] M. Chandraker, S. Agarwal, and D. Kriegman, "Shadowcuts: Photometric stereo with shadows," *Proc. CVPR*, 2007.
- [19] M. Liao, X. Huang, and R. Yang, "Interreflection removal for photometric stereo by using spectrum-dependent albedo," *Proc. CVPR*, 2011.
- [20] T. Okatani and K. Deguchi, "Optimal integration of photometric and geometric surface measurements using inaccurate reflectance/illumination knowledge," *Proc. CVPR*, 2012.
- [21] C. Hernández, G. Vogiatzis, and R. Cipolla, "Multi-view photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 3, pp. 548–554, 2008.
- [22] A. del Bue, J. Xavier, L. Agapito, and M. Paladini, "Bilinear modeling via augmented lagrange multipliers (balm)," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1496–1508, 2012.
- [23] K. Tang, C. Tang, and T. Wong, "Dense photometric stereo using tensorial belief propagation," in *Proc. CVPR*, 2005.
- [24] T. Wu and C. Tang, "Dense photometric stereo using a mirror sphere and graph cut," *Proc. CVPR*, 2005.
- [25] D. Miyazaki, K. Hara, and K. Ikeuchi, "Median photometric stereo as applied to the segonko tumulus and museum objects," *International Journal of Computer Vision*, vol. 86, no. 2, pp. 229–242, 2010.
- [26] K. Torrance and E. Sparrow, "Theory for off-specular reflection from roughened surfaces," *Journal of the Optical Society of America*, vol. 57, no. 9, pp. 1105–1112, 1967.
- [27] G. Ward, "Measuring and modeling anisotropic reflection," *Computer Graphics*, vol. 26, no. 2, pp. 265–272, 1992.
- [28] S. Rusinkiewicz, "A new change of variables for efficient brdf representation," in *Proc. of Eurographics Symposium on Rendering*, 1998.
- [29] P. Tan, L. Quan, and T. Zickler, "The geometry of reflectance symmetries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2506–2520, 2011.
- [30] M. Chandraker and R. Ramamoorthi, "What an image reveals about material reflectance," *Proc. ICCV*, 2012.
- [31] P. Tan, S. P. Mallick, L. Quan, D. Kriegman, and T. Zickler, "Proc. CVPR," *Isotropy, Reciprocity and the Generalized Bas-Relief Ambiguity*, 2007.

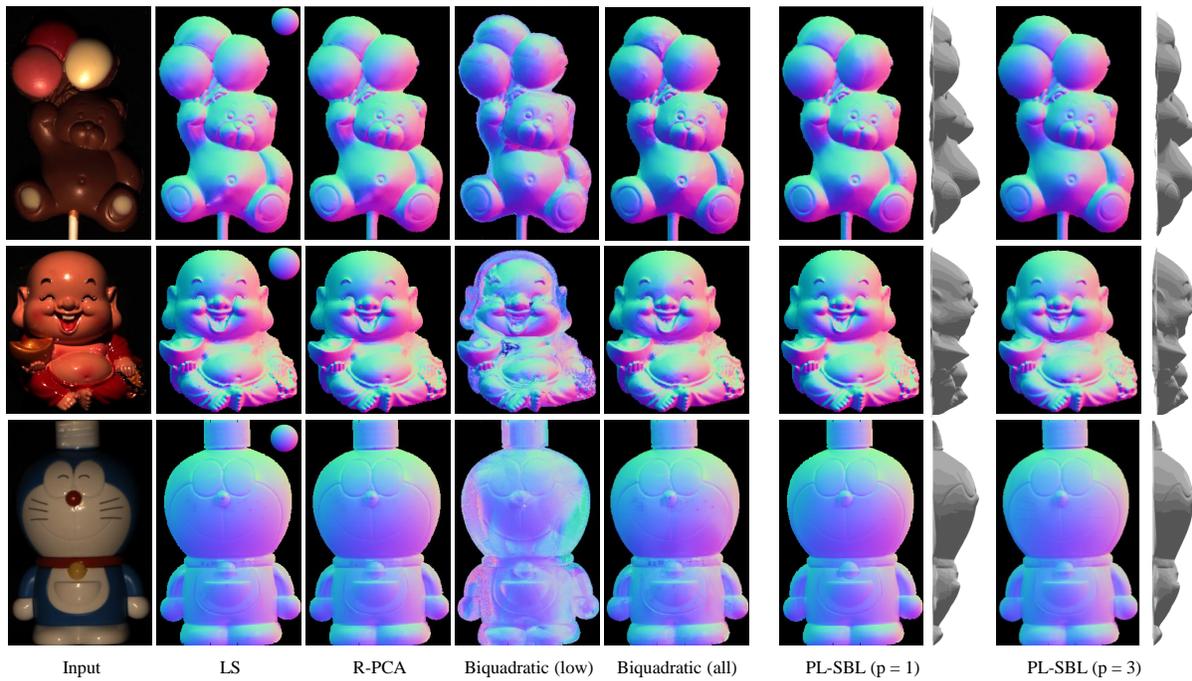


Fig. 13. Experimental results with real datasets (D). We used three kind of datasets called *Chocolate bear* (25 images with  $261 \times 421$ ), *Fat guy* (40 images with  $293 \times 344$ ) and *Doraemon* (40 images with  $269 \times 420$ ). We show example of input images, recovered surface normals and surface meshes (only for PL-SBL( $p = 3$ )).

[32] N. Alldrin, "Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach," *Proc. ICCV*, 2007.

[33] T. Higo, Y. Matsushita, and K. Ikeuchi, "Consensus photometric stereo," *Proc. CVPR*, 2010.

[34] I. Sato, T. Okabe, Q. Yu, and Y. Sato, "Shape reconstruction based on similarity in radiance changes under varying illumination," *Proc. ICCV*, 2007.

[35] M. Chandraker, J. Bai, and R. Ramamoorthi, "A theory of differential photometric stereo for unknown brdfs," *Proc. CVPR*, 2011.

[36] P. Favaro and T. Papadhimetri, "A closed-form solution to uncalibrated photometric stereo via diffuse maxima," *Proc. CVPR*, 2012.

[37] W. M. Silver, *Determining shape and reflectance using multiple images*. Master's thesis, MIT, 1980.

[38] A. Hertzmann and S. M. Seitz, "Example-based photometric stereo: shape reconstruction with general, varying brdfs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1254–1264, 2005.

[39] S. Shafer, "Using color to separate reflection components," *COLOR research and application*, vol. 10, no. 4, pp. 210–218, 1985.

[40] F. Romeiro and T. Zickler, "Blind reflectometry," in *Proc. ECCV*, 2010.

[41] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Trans. on Graph.*, vol. 22, no. 3, pp. 759–769, 2003.

[42] F. Verbiest and L. V. Gool, "Photometric stereo with coherent outlier handling and confidence estimation," *Proc. CVPR*, 2008.

[43] M. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

[44] D. Wipf and S. Nagarajan, "Iterative reweighted  $\ell_1$  and  $\ell_2$  methods for finding sparse solutions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 317–329, 2010.

[45] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[46] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[47] D. Wipf, B. D. Rao, and S. Nagarajan, "Latent variable bayesian models for promoting sparsity," *IEEE Transactions on Information Theory*, vol. 57, no. 9, 2011.

[48] R. Cook and K. Torrance, "A reflectance model for computer graphics," *ACM Trans. on Graph.*, vol. 15, no. 4, pp. 307–316, 1981.

[49] E. Laforune, S.-C. Foo, K. Torrance, and D. Greenberg, "Non-linear approximation of reflectance functions," in *Proc. ACM SIGGRAPH*, 1997.

[50] A. Agrawal, R. Raskar, and R. Chellappa, "What is the range of surface reconstructions from a gradient field ?" *Proc. ECCV*, 2006.



Human Cognition and Perception.

**Satoshi Ikehata** received the B.A. of Psychology in 2009 and the M.S. of Information Studies in 2011 from the University of Tokyo respectively and am working for Ph.D at the University of Tokyo supervised by Prof. Kiyoharu Aizawa. He worked as an intern at the Visual Computing Group at Microsoft Research Asia as a visiting student in 2011. His main interests lie in Image-based 3D Scene Reconstruction (photometric stereo, reflectance recovery, multi-view stereo), Time-of-flight Imaging, Image Super-resolution,



ing Group at Microsoft Research Asia.

**David Wipf** received the B.S. degree in electrical engineering from the University of Virginia, Charlottesville, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, San Diego. He was later an NIH Postdoctoral Fellow in the Biomagnetic Imaging Lab at the University of California, San Francisco where his research involved the development and analysis of Bayesian learning algorithms for functional brain imaging and sparse coding. Currently he is with the Visual Computing Group at Microsoft Research Asia.



**Yasuyuki Matsushita** received his B.S., M.S. and Ph.D. degrees in EECS from the University of Tokyo in 1998, 2000, and 2003, respectively. He joined Microsoft Research Asia in April 2003. He is a Senior Researcher in Visual Computing Group. His areas of research are computer vision (photometric techniques, such as radiometric calibration, photometric stereo, shape-from-shading), computer graphics (image relighting, video analysis and synthesis). He is on the editorial board member of IEEE Transactions on

Pattern Analysis and Machine Intelligence (TPAMI), International Journal of Computer Vision (IJCV), IPSJ Journal of Computer Vision and Applications (CVA), The Visual Computer Journal, and Encyclopedia of Computer Vision. He served/is serving as a Program Co-Chair of PSIVT 2010, 3DIMPVT 2011, and ACCV 2012 and a General Co-Chair for ACCV 2014. He is appointed as a Guest Associate Professor at Osaka University (April 2010-), Visiting Associate Professor at National Institute of Informatics (April 2011-) and Tohoku University (April 2012 - ), Japan. He is a senior member of IEEE.



**Kiyoharu Aizawa** received the B.E., the M.E., and the Dr.Eng. degrees in Electrical Engineering all from the University of Tokyo, in 1983, 1985, 1988, respectively. He is currently a Professor at the Department of Information and Communication Engineering of the University of Tokyo. He was a Visiting Assistant Professor at University of Illinois from 1990 to 1992. His current research interest is in image processing and multimedia applications, and he is currently engaged in multimedia life log and 3D video etc.

He received the 1987 Young Engineer Award and the 1990, 1998 Best Paper Awards, the 1991 Achievement Award, 1999 Electronics Society Award from IEICE Japan, and the 1998 Fujio Frontier Award, the 2002 and 2009 Best Paper Award from ITE Japan. He received the IBM Japan Science Prize in 2002. He is currently on Editorial Board of ACM TOMCCAP, Journal of Visual Communications and Image Processing and APSIPA Transactions on Signal and Information Processing. He served as the Editor in Chief of Journal of ITE Japan, an Associate Editor of IEEE Trans. Image Processing, IEEE Trans. CSVT and IEEE Trans. Multimedia. He has served a number of international and domestic conferences; he was a General co-Chair of ACM Multimedia 2012 and IEEE VCIP2012.