Depth-Map and Albedo Estimation with Superior Information-Theoretic Performance

Adam P. Harrison[a] and Dileepan Joseph[a]

[a]University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

Lambertian photometric stereo (PS) is a seminal computer vision method. However, using depth maps in the image formation model, instead of surface normals as in PS, reduces model parameters by a third, making it preferred from an information-theoretic perspective. The Akaike information criterion (AIC) quantifies this trade-off between goodness of fit and overfitting. Obtaining superior AIC values requires an effective maximum-likelihood (ML) depth-map & albedo estimation method. Recently, the authors published an ML estimation method that uses a two-step approach based on PS. While effective, approximations of noise distributions and decoupling of depth-map & albedo estimation have limited its accuracy. Overcoming these limitations, this paper presents an ML method operating directly on images. The previous two-step ML method provides a robust initial solution, which kick starts a new nonlinear estimation process. An innovative formulation of the estimation task, including a separable nonlinear least-squares approach, reduces the computational burden of the optimization process. Experiments demonstrate visual improvements under noisy conditions by avoiding overfitting. As well, a comprehensive analysis shows that refined depth maps & albedos produce superior AIC metrics and enjoy better predictive accuracy than with literature methods. The results indicate that the new method is a promising means for depth-map & albedo estimation with superior information-theoretic performance.

Keywords: Depth-map & albedo estimation, image generation, photometric stereo, Akaike information criterion, maximum-likelihood estimation, separable nonlinear least-squares.

1. INTRODUCTION

Single-view photometric stereo (PS) addresses the problem of determining an object's surface normals and reflectance characteristics from a sequence of image observations from the same viewpoint, often under known light directions. A prominent goal of such an exercise is to later generate images from arbitrary light directions, *e.g.*, for virtual reflected-light microscopy.[1] Closely related, the goal of single-view depth-map estimation is to determine an object's height field. Such a 3D model can also be used for image generation purposes, but has value in and of itself. Most depth-map estimation techniques use PS as an initial step.

Thus, PS forms the foundation of two key computer vision tasks that use single-view images. When given enough observations, *i.e.*, images, PS-based techniques yield accurately generated images and depth-map & reflectance estimates, especially in Lambertian settings. Yet, even in this setting, as the number of observations approaches PS's 3-image minimum, overfitting becomes a significant issue, affecting both image generation and depth-map & albedo accuracy when imaging noise or physical deviations from the model are present.

Focusing on image generation, this problem can be viewed as a prediction task, *i.e.*, forecasting unseen images based on seen images. When one's goal is predictive accuracy, the question of model parsimony, or simplicity, looms large.[2] Simply put, the "best" model is not necessarily the one that maximizes likelihood, but instead one that best balances the tradeoff between goodness of fit and complexity.[3,4] Complexity can be thought of in many ways, but an important component is the number of parameters in the model. This is an information-theoretic approach to model selection, which can be seen as an extension to likelihood theory[3] and is exemplified most famously by the Akaike framework.[5] This framework's most famous tool is the Akaike information criterion (AIC), which assesses competing models based on how well they balance goodness of fit with parsimony.

Further author information: (Send correspondence to A.P.H.)

A.P.H.: E-mail: adam.p.harrison@gmail.com

D.J.: E-mail: dil.joseph@ualberta.ca

As a result, in the context of image generation, a more parsimonious model may well outperform PS, especially when observation counts are low. In fact, such a model is readily available. The depth map & albedo model is perfectly capable of representing image formation and possesses a parameter count roughly a third less than PS in Lambertian settings. However, to best use such a model one should determine an accurate ML estimate. Reaching high accuracy provides benefits outside of image generation, *i.e.*, estimating the depth-map itself. In cases where PS is susceptible to overfitting, the visual quality of any PS-based depth-map & albedo estimate will also suffer.

This work addresses key gaps in the state of the art in obtaining accurate ML depth-map & albedo estimates. Using the information-theoretic framework, this work motivates this pursuit under sound theoretical principles, even when PS provides excellent or near-perfect reconstruction of the observation images. In addition, this work outlines how the AIC provides a useful metric to gauge the predictive accuracy of the PS model or the more parsimonious depth-map & albedo model. In terms of obtaining accurate ML estimates, unlike the state of the art,[6–11] this work outlines a nonlinear minimization scheme that couples depth-map & albedo estimation together. To help accomplish this, ML estimation is cast as a separable nonlinear least-squares (SNLS) problem.[12,13] As well, boundary conditions are kept to Neumann boundary constraints coupled with two neighbouring Dirichlet constraints. Like Ikeda's approach,[10] this work uses the output from a PS-based technique as an initial solution. However, Harrison and Joseph's method[6] is used instead of Frankot and Chellapa's,[14] providing a more robust and accurate initial solution. These developments result in a practical and effective method to obtain depth map & albedo estimates. In concrete terms, this means at low observation counts this work's estimation technique leads to better predictive accuracy in image generation and improved depth-map reconstructions.

The following sections will outline the depth-map & albedo estimation technique explained in this work. Section 2 provides background information describing the image formation models, PS methods, overfitting, and the Akaike framework. This is followed by Section 3, which describes the depth-map & albedo estimation technique, including how it can be formulated as an SNLS problem. Section 4 outlines experiments demonstrating the efficacy of this work's approach in generating accurate images and depth-map & albedo estimates in challenging conditions with low observation counts. Section 5 discusses the state of the art, this work's significance, and future work. Finally, Section 6 concludes the paper.

2. CONCEPTUAL OVERVIEW

This section provides background and motivating concepts. Section 2.1 first outlines the Lambertian image formation model, highlighting the PS and depth-map & albedo generative models. Afterwards, Section 2.2 outlines the classic PS-based approach to image generation and depth-map & albedo estimation. Section 2.3 follows this up by illustrating how PS-based methods can be susceptible to overfitting, particularly when the number of observation images is low. Section 2.4 then casts these overfitting problems into the Akaike framework, showing how the AIC provides a theoretical foundation for determining the situations that favour the depth-map & albedo model over the PS model.

2.1 Lambertian Image Formation

In the context of this work, the Lambertian image formation model is used, which primarily assumes that an object enjoys ideal diffuse reflectance properties. Under additional assumptions, including orthographic cameras, a principal directional light source at infinity, a linear relationship between pixel intensity and irradiance, and no inter-reflections or cast shadows, the intensity of an image pixel at coordinates x and y can be expressed using a simple generative model:

$$I(x,y) = \boldsymbol{\ell}^{\mathsf{T}}\boldsymbol{\eta}(x,y) \cdot u(\boldsymbol{\ell}^{\mathsf{T}}\boldsymbol{\eta}(x,y)) + \epsilon(x,y), \tag{1}$$

$$\boldsymbol{\eta}(x,y) = \rho(x,y)\mathbf{n}(x,y), \tag{2}$$

where $\mathbf{n}(x,y) = (n_x, n_y, n_z)^{\mathsf{T}}$ represents surface normals, $\rho(x,y)$ is the surface albedo, $\boldsymbol{\ell} = (\ell_x, \ell_y, \ell_z)^{\mathsf{T}}$ is the light direction expressed as a unit vector, and $\epsilon(x,y)$ represents noise in the measurements. Here, $\boldsymbol{\eta}(x,y) = (\eta_x, \eta_y, \eta_z)^{\mathsf{T}}$ represents the surface normals multiplied by the albedo, and will be referred to as *weighted normals*. Attached

shadows are modelled by including $u(.)$, the unit step function, in the formulation. Image noise is assumed to be independently and identically distributed (IID) and zero-mean Gaussian.

Thus, given a set of weighted normals, one can generate an image from any light direction in the upper hemisphere using (1). Because there is a set of weighted normals for each pixel, image generation can be performed independently for each pixel. It should be noted that, apart from the attached-shadow term, the relationship is linear.

Yet, even though surface normals describe an object's depth map, $z(x, y)$, they still are not an explicit representation of it. If one desires a more explicit relationship between the depth map and image formation, a different formulation is needed. Dropping the x and y coordinates for notational simplicity, the image formation model can be expressed as

$$I = \rho \boldsymbol{\ell}^{\mathsf{T}} \mathbf{g}\left(p(z), q(z)\right) \cdot u(\boldsymbol{\ell}^{\mathsf{T}} \mathbf{g}(p(z), q(z))) + \epsilon(x, y), \tag{3}$$

$$\mathbf{g}(p(z), q(z)) = \frac{1}{\sqrt{p(z)^2 + q(z)^2 + 1}} \begin{pmatrix} -p(z) \\ -q(z) \\ 1 \end{pmatrix}, \tag{4}$$

$$\tag{5}$$

where $p(z)$ and $q(z)$ are the partial derivatives of z with respect to x and y, respectively. This formulation expresses an entirely nonlinear relationship.

If one has a representation of the depth map that allows analytic forms of the partial derivatives at each pixel location, then (3) can be applied directly. However, as is often the case in an imaging setting, the depth map is represented as a discrete array of values, $\mathbf{z}$. In this case, discrete approximations to the derivatives, $e.g.$, finite differences (FDs), must be applied. The surface gradients can be calculated using:

$$\begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} \simeq \begin{pmatrix} \mathbf{D}_x \\ \mathbf{D}_y \end{pmatrix} \mathbf{z}, \tag{6}$$

where $\mathbf{D}_x$ and $\mathbf{D}_y$ are FD approximations to the derivative operators. To model this relation using matrix notation, the depth map and its gradients have been represented in vector form, $e.g.$, using column- or row-major lexicographical order. The gradient values can then be substituted in (4). One of the implications of FD operators is that depth-map gradients are derived from a neighbourhood of values, linking pixel values calculated using (3) to their neighbours.

2.2 Photometric-Stereo Methods

While the above described the forward model of image generation, often the goal is to estimate the parameters themselves. PS is a seminal computer vision technique, first formulated by Woodham,[15] for such an estimation goal. Given a sequence of N single-view images illuminated under known light directions, PS is tasked with estimating the weighted normals, from which one can obtain both the albedo and surface normals. As such, PS is related to the image formation model of (1), rather than (3), and is faced with the following regression formulation:

$$\mathbf{i}(x, y) = \mathbf{L}\boldsymbol{\eta}(x, y) \cdot u(\mathbf{L}\boldsymbol{\eta}(x, y)) + \boldsymbol{\epsilon}(x, y), \tag{7}$$

where $\mathbf{L}$ is an $N \times 3$ matrix representing the light directions and the unit-step function's vector output is multiplied element-wise. Estimating the weighted normals can be done independently for each pixel. However, the attached shadows turns what would otherwise be a linear regression problem into a nonlinear one. Yet, if one can filter out the shadowed pixels so that only N_0 images remain, the problem becomes linear,

$$\mathbf{i}_0(x, y) = \mathbf{L_0}(x, y)\boldsymbol{\eta}(x, y) + \boldsymbol{\epsilon}(x, y), \tag{8}$$

where $\mathbf{L_0}(x, y)$ is an $N_0 \times 3$ matrix whose dependance on x and y reflects the dependance of attached shadows on pixel location. The weighted normals can then be estimated by solving (8) in the least-squares sense for each

pixel. More details can be found in Harrison and Joseph's paper, including a discussion of filtering tactics and how using the reduced linear regression form of (8) still keeps the result an ML estimate.[6]

With a set of weighted normals estimated for each pixel, images can be generated for any light direction using the model of (1). However, often the goal is to estimate the depth map itself, which is related to the PS output using the following expression for the gradients:

$$p = -\frac{\eta_x}{\eta_z}, \tag{9}$$

$$q = -\frac{\eta_y}{\eta_z}, \tag{10}$$

where again coordinates have been dropped for notational simplicity. Once this calculation has been done for each pixel, depth-map estimation can then be cast as a linear regression problem,

$$\begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} = \begin{pmatrix} \mathbf{D}_x \\ \mathbf{D}_y \end{pmatrix} \mathbf{z} + \boldsymbol{\epsilon}_{\mathbf{pq}}, \tag{11}$$

which differs from (6) in that an error term, $\boldsymbol{\epsilon}_{\mathbf{pq}}$, is included as part of the regression. Modelling the behaviour of $\boldsymbol{\epsilon}_{\mathbf{pq}}$ is challenging as its distribution is governed in part by the nonlinear operations of (9) and (10). However, Harrison and Joseph demonstrated that one can model the behaviour of gradient error using an anisotropic Gaussian distribution, provided regularity assumptions and asymptotic approximations are used.[6] Having such a model of gradient error in hand turns the task of ML estimation of depth maps into a generalized least-squares problem. This technique can be labelled *linearized ML (LML) depth-map estimation*, as linearized approximations to error distributions represents a core feature of the method.

LML depth-map estimation is a key component of the technique outlined in this work, with one important change. In the original work, an averaging strategy was used to overcome certain issues with standard FD operators.[6] However, as Horn details in Section 6 of his paper,[16] a more appropriate FD operator is a staggered grid. As Horn explains, if the image dimensions are $m_1 \times m_2$, this implies an $(m_1 + 1) \times (m_2 + 1)$ depth map. Using such a staggered FD operator obviates the need for Harrison and Joseph's averaging scheme. Yet, the staggered grids add an additional ambiguity to the unavoidable offset ambiguity in depth-map estimation. In practical terms, this is not a serious issue, as Dirichlet constraints can just be applied to two neighbouring pixels, instead of just one pixel. Thus, when executing LML depth-map estimation, this work uses Horn's staggered FD grids as $\mathbf{D}_x$ and $\mathbf{D}_y$.

2.3 Overfitting

PS, and algorithms based off of PS, can enjoy excellent performance and robustness. For instance, image generation[1] and LML depth-map estimation[6] can perform well in challenging conditions rife with violations of the Lambertian generative model. However, as the number of images approaches PS's three-image minimum, the challenges facing the technique increase. For one, in noisy conditions PS will tend to overfit to image noise. Moreover, it becomes less and less possible to filter out attached shadows. When only three images are available, PS is forced to forego filtering altogether to keep the linear sub-problem fully ranked, regardless of whether the pixels in question are under an attached shadow or not. Once shadowed pixels are included in the linear regression sub-problem of (8), PS may no longer produce an ML estimate and may exhibit major errors.

To illustrate this, the Lambertian image formation model of (1) was used to generate four images of Zhang *et al.*'s Mozart surface[17] along with a checkerboard albedo. Pixel values fell between 0 and 1 and image noise exhibited a standard deviation of 0.05, or 5% of the maximum intensity. The surface orientation needed for the weighted normals was calculated using Horn's staggered-grid FD operators.[16] After images were generated, PS was then executed to estimate the weighted normals from the noisy images.

As the second column of Figure 1 demonstrates, when generating an image originating from a light direction *not used* in the observations, PS produces a very noisy result that suffers from speckling. As well image details, such as the border between Mozart and the background, are obscured. The fact that PS faces difficulties is not surprising, as the four observation images used in the example do not allow PS very much leeway to mitigate noise effects and detect attached shadows.

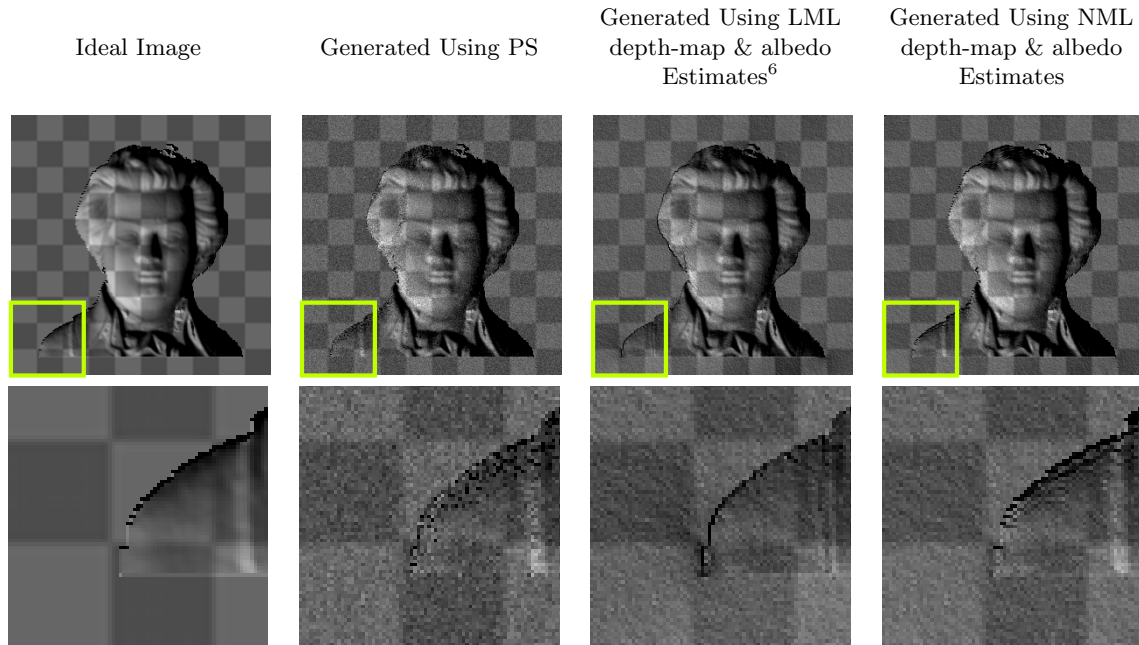| Ideal Image	Generated Using PS	Generated Using LML depth-map & albedo Estimates[6]	Generated Using NML depth-map & albedo Estimates

Figure 1. Images generated using the PS and depth-map & albedo models. The output of the PS, LML, and NML methods was used to generate an image from an *unseen* light direction with an elevation and azimuth angle of $(60°, 225°)$. The generated images can be compared with the ideal and noiseless version. The experimental setup used four observation images of the Mozart depth map, corrupted by Gaussian noise with 5% standard deviation and corresponding to illumination directions with elevation and azimuth angles of $(60°, 30°)$, $(45°, 150°)$, $(55°, 270°)$, and $(15°, 90°)$.

The low quality of images generated using PS can affect applications where fine examination is needed. For instance, when images are reconstructed for the purposes of virtual reflected-light microscopy, high image quality is required for specimen identification and examination tasks.[1] Moreover, the loss of important image details can also propagate to algorithms using PS output. The most obvious examples of this are depth-map estimation techniques that use the popular two-step approach of working off of the results of PS. Figure 2(b) illustrates how the errors inherent in PS can affect LML depth-map & albedo estimation, a leading two-step technique. As the figure demonstrates, the LML suffers from localized noisy effects. Moreover, the depth map suffers from global distortions, including a flattened look. Thus, if the goal is depth-map estimation, a technique other than a PS-based one should be adopted in these conditions.

Returning to the image formation task of Figure 1, a natural question is whether a more parsimonious model, *i.e.*, the formulation in (3) using a depth map & albedo, can avoid the overfitting issues encountered by PS and generate superior images. However, if the means to produce the depth-map estimate relies solely on the PS output, image generation will face continued problems. As the third column of Figure 1 demonstrates, which depicts images generated using the PS-based LML method, this is indeed the case. In general, while the LML method is able to generate images with less noise than the PS method, it is still unable to capture certain fine details. For instance, the extent of the raised ridge in the inset, corresponding to a wrinkle in Mozart's shirt, is made clear in the ideal image by lighter-coloured highlights. This is not captured by the LML method.

These examples suggest that the more parsimonious depth-map & albedo model can indeed help avoid overfitting to image noise. However, to avoid losing image detail, accurate depth-map & albedo estimates are needed. While the LML method produces high-quality estimates, it can struggle with image generation tasks when observation counts are low. For one, the LML method's asymptotic approximations lose their accuracy at low numbers of observation images, reducing its ability to model the stochastic behaviour of the surface gradients produced by PS. As well, the LML method is not designed to handle errors in filtering out attached shadows. For these reasons, to best compare the PS model with the depth-map & albedo model at low observation counts, one must employ a parameter estimate that works directly with image observations.

(a) True Depth Map   (b) LML Depth Map[6]   (c) NML Depth Map

Figure 2. Depth-map estimates under low observation counts. Using the experimental conditions of Figure 1, both LML and NML estimation were performed. As PS only estimates a representation of surface orientation, and not an explicit depth map, its results cannot be visually demonstrated. Differences in depth-map quality between the LML and NML methods can be significant.

Obtaining such a parameter estimate, one that avoids the pitfalls associated with PS while also enjoying the benefits of a more parsimonious model, is the goal of this work. In this work, the estimate is obtained by working directly with the challenging nonlinear formulation of (3). Consequently, the technique is labelled *nonlinear ML (NML) depth-map & albedo estimation.* As the fourth column of Figure 1 illustrates, NML estimation can produce image reconstructions that better capture detail, while still enjoying the benefits of a more parsimonious model, *i.e.,* exhibiting less susceptibility to noise. Moreover, as Figure 2(c) demonstrates, the NML technique can produce a depth-map estimate higher in quality than the LML method. The phenomena highlighted in Figures 1 and 2 motivates the NML estimation scheme outlined by this work.

2.4 Akaike Framework

The previous subsection used an example to illustrate how the balance between goodness of fit and model complexity affects the aims of computer vision tasks that use single-view image sequences. The questions surrounding this balance can be formalized using information-theoretic principles. The Akaike framework, one of the best known model-selection criteria, falls within the category. Broadly speaking, model selection under the Akaike framework attempts to maximize predictive accuracy.[2] It does so using the information-theoretic measure of Kullback-Liebler (KL) divergence, which quantifies the difference between the probability distributions of a model's hypothesis with the truth. The Akaike Information Criterion (AIC) provides a means to estimate differences between KL divergences of competing model hypotheses. What this boils down to in practical terms, is that the AIC, like many other criteria, including the Bayesian Information Criterion[18] and Minimum Description Length,[19] rewards goodness of fit and penalizes model complexity.

For general models, the AIC is defined as:[3]

$$AIC = -2\log(\mathcal{L}(\hat{\theta}|data)) + 2k, \tag{12}$$

where $\hat{\theta}$ is the set of parameters maximizing the model's likelihood and k is the number of such parameters. Two points are very important regarding the AIC. Firstly, to assess a model, the AIC value assumes that the ML estimate is available. However, depending on the model, the ML estimate may be difficult to determine or may suffer from inaccuracies. This will also affect the accuracy of the AIC value itself. Secondly, AIC values are never meant to be used absolutely, instead it is a relative measure that can be used to compare two or more models with each other. Models having smaller AIC values are considered to enjoy a greater degree of predictive accuracy.

Table 1. SSE and AIC_c values for the PS, LML, and NML methods under low observation counts. Using the experimental setup of Figure 1, PS posts a much better SSE value for the observation images than the LML method. However, its ability to forecast an *unseen* image is slightly worse. The NML method posts SSE values for the observation images roughly halfway between PS and LML, but is able to forecast the unseen image with markedly better accuracy. The AIC_c values align well with the PS and NML's predictive accuracy, but is less accurate for the LML method.

	PS	LML	NML
SSE Noisy Observation Images	230.3	787.8	446.1
SSE Unseen Ideal Image	245.2	238.0	135.0
AIC_c	$-470\,759.6$	$-1\,443\,426.4$	$-1\,644\,112.7$

When IID zero-mean Gaussian errors are assumed, (12) can be formulated as:

$$AIC = n\log(\hat{\sigma}^2) + 2k, \tag{13}$$

$$= n\log\left(\frac{SSE}{n}\right) + 2k, \tag{14}$$

where $\hat{\sigma}$ is the ML estimate of the variance, SSE is the sum-squared errors, n is the sample size, and k must also include the estimate of the variance within its parameter count. For small sample sizes, (14) can suffer from a large bias term, leading to the commonly-used corrected AIC (AIC_c) value:

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1}, \tag{15}$$

which was derived under the assumptions of a linear univariate model.[3]

In the context of this work, the two models in question are the PS model of (1) and the depth-map & albedo model of (3), both of which assume IID Gaussian noise. However, since both of these models do not exactly meet the criteria of (15), the implications should be addressed. Focusing on the PS model, technically (1) is a multivariate model, where each pixel provides its own set of discrete and independent observations. Nonetheless, since the error distribution is IID across all observations, specialized correction terms for multivariate models, such as Bedrick and Tsai's formulation,[20] are not needed. Moreover, the reduction of the problem to a reduced linear one, seen in (8), can also be replicated when measuring likelihood. Thus, (15) serves as an excellent measure for the PS model.

When focusing on the depth-map & albedo model of (3), the formulation is entirely nonlinear. Thus, the correction term in (15) will be a less accurate correction of bias. Nonetheless, the process by which (15) is derived is also valid for nonlinear models up to a first-order Taylor series approximation around the true parameter values. Thus, (15) is more accurate when the ML estimate is close to the true parameter values. Provided these caveats are understood, (15) can still serve as a useful approximate measure for nonlinear model selection. See Section 7.4 of Burnham and Anderson's book for more details on the derivation of AIC_c.[3]

Given an image with $m_1 \times m_2$ pixels, (1) employs $3m_1m_2$ parameters, whereas (3) employs $m_1m_2 + (m_1 + 1)(m_2 + 1)$ parameters if the FD operators described in Section 2.2 are used. Assuming N images are used, the sample size is Nm_1m_2.

Returning to the example in Section 2.3, Table 1 presents the SSE and AIC_c values for the PS, LML, and NML methods. Despite fitting the observation images considerably better than the LML method, the AIC_c value for the PS method is significantly inferior than that of the LML method. Thus, the AIC_c indicates that the depth-map & albedo model of (3) enjoys greater predictive accuracy. However, with the low observation counts and its reliance on PS output, the LML method's estimate is likely not able to realize these benefits. This conclusion is borne out when considering the LML method's SSE for the unseen ideal image, which despite what the AIC_c values predicted, posts a value close to that of the PS method. For this reason, ensuring that higher predictive accuracy is realized requires pursuing an estimation technique unencumbered by the LML method's approximations.

When this work's NML technique is used, the resulting depth-map & albedo estimate enjoys a lower $\mathrm{AIC_c}$ than either the PS or LML method. More importantly, the NML method's SSE for the unseen ideal image in Figure 1 is significantly lower than either the PS or LML methods. These quantitative results are reflected in concrete terms by the visual improvements, illustrated in Figures 1 and 2, garnered by NML estimation. The process by which an improved estimate for the depth map & albedo can be obtained is the focus of the next section.

3. METHOD

This section outlines how the depth-map & albedo can be estimated directly from image observations without using PS as an intermediate step. To help tackle this challenging nonlinear optimization, the problem can be formulated as an SNLS problem. Doing so reduces the number of optimization parameters by roughly a half, but complicates the calculation of the Jacobian. Details are provided on how to calculate this Jacobian.

NML estimation operates directly on the model in (3), by determining estimates of the depth map & albedo that minimize the SSE of the residuals for each pixel, $i.e.$,

$$\hat{\mathbf{z}}, \hat{\boldsymbol{\rho}} = \arg\min_{\mathbf{z}, \boldsymbol{\rho}} \sum_j \mathbf{r}_j(\mathbf{z}, \boldsymbol{\rho})^\mathsf{T} \mathbf{r}_j(\mathbf{z}, \boldsymbol{\rho}), \tag{16}$$

$$\mathbf{r}_j(\mathbf{z}, \boldsymbol{\rho}) = \mathbf{i}_j - \rho_j \mathbf{L}\, \mathbf{g}(p_j, q_j) \cdot u(\mathbf{L}\, \mathbf{g}(p_j, q_j)), \tag{17}$$

$$= \mathbf{i}_j - \rho_j \mathbf{h}(p_j, q_j), \tag{18}$$

where $\mathbf{g}(.,.)$ is defined as in (4), j indexes the image pixels, and p_j and q_j depend on $\mathbf{z}$, as described in Section 2.1. As well, $\mathbf{h}(p_j, q_j)$ in (18) is defined from (17).

The formulation in (18) is constructed such that if the parameters are considered in isolation, $\mathbf{z}$ can be viewed as a nonlinear parameter, whereas ρ_j is a linear one. This meets exactly the criteria for being an SNLS problem, which Golub and Pereyra identified as enjoying particular properties that can reduce the computational burden associated with optimization.[12, 13] Given an estimate of $\mathbf{z}$, the ρ_j that minimizes the SSE can be expressed using the Moore-Penrose pseudo-inverse operation for linear least-squares,

$$\hat{\rho}_j = \mathbf{h}_j^+ \mathbf{i}_j, \tag{19}$$

$$= \frac{\mathbf{h}_j^\mathsf{T} \mathbf{i}_j}{\mathbf{h}_j^\mathsf{T} \mathbf{h}_j}, \tag{20}$$

where $\mathbf{h}_j$ is used to denote $\mathbf{h}(p_j, q_j)$. The expression in (20) is included to demonstrate the inexpensiveness of the operation whereas, in actuality, the Moore-Penrose inverse need not be formulated explicitly to solve for $\hat{\rho}_j$.

Substituting (19) into (18) allows the residual to be expressed using only $\mathbf{z}$, $i.e.$,

$$\mathbf{r}_j(\mathbf{z}) = \mathbf{i}_j - \mathbf{h}_j \mathbf{h}_j^+ \mathbf{i}_j, \tag{21}$$

$$= \mathbf{i}_j - \mathbf{H}_j^\mathrm{P} \mathbf{i}_j, \tag{22}$$

$$= \mathbf{H}_j^\perp \mathbf{i}_j, \tag{23}$$

where $\mathbf{H}_j^\mathrm{P}$ denotes the projection operator corresponding to $\mathbf{h}_j$ and $\mathbf{H}_j^\perp$ denotes its complement. Thus, (21) reformulates the minimization problem to only incorporate $\mathbf{z}$, $i.e.$,

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \sum_j \mathbf{r}_j(\mathbf{z})^\mathsf{T} \mathbf{r}_j(\mathbf{z}), \tag{24}$$

reducing the optimization parameter count by roughly a half.

In principle, (24) is sufficient to determine the ML depth-map estimate (and subsequently the albedo estimate from it). However, most nonlinear least-squares algorithms require the Jacobian of the residuals to operate. To

avoid having to calculate this manually using FD operations, which is extremely costly, an expression for the Jacobian must be derived. This can be expressed as

$$\frac{\partial \mathbf{r}_j(\mathbf{z})}{\partial \mathbf{z}} = -\frac{\partial \mathbf{H}_j^{\mathrm{P}}}{\partial \mathbf{z}} \mathbf{i}_j, \tag{25}$$

The pseudo-inverse within the projection operator makes (25) non-trivial to formulate. Fortunately, Golub and Pereyra's SNLS method provides a means to express this derivative. In this case, the derivative is equivalent to

$$\frac{\partial \mathbf{H}_j^{P}}{\partial \mathbf{z}} = \mathbf{H}_j^{\perp} \frac{\partial \mathbf{h}_j}{\partial \mathbf{z}} (\mathbf{h}_j^+)^{\mathsf{T}} + \left(\mathbf{H}_j^{\perp} \frac{\partial \mathbf{h}_j}{\partial \mathbf{z}} (\mathbf{h}_j^+)^{\mathsf{T}} \right)^{\mathsf{T}}, \tag{26}$$

where $\partial \mathbf{h}_j / \partial \mathbf{z}$ can be considered as a sequence of vector slices, indexed by $\mathbf{z}$. Multiplication within (26) is performed "slice-wise", meaning across indices corresponding to $\mathbf{h}_j$ and repeated for each index within $\mathbf{z}$. More details on how to express the derivative of the projection operater can be found in Golub and Pereyra's original exposition.[12]

The crux of calculating (26) lies with calculating $\partial \mathbf{h}_j / \partial \mathbf{z}$, which can be expressed as

$$\frac{\partial \mathbf{h}_j}{\partial \mathbf{z}} = \mathbf{L} \frac{\partial \mathbf{g}(p_j, q_j)}{\partial \mathbf{z}} \cdot u(\mathbf{L}\,\mathbf{g}(p_j, q_j)) + \mathbf{L}\,\mathbf{g}(p_j, q_j) \cdot \frac{\partial u(\mathbf{L}\,\mathbf{g}(p_j, q_j))}{\partial \mathbf{z}}. \tag{27}$$

As the second term in (27) involves the derivative of the unit step function, *i.e.*, the Dirac delta function, it is zero except for when its arguments equal exactly 0. Assuming this does not occur, the second term can be dropped from consideration. With this assumption, the following expression represents the core of the partial derivative calculation:

$$\frac{\partial \mathbf{g}(p_j, q_j)}{\partial \mathbf{z}} = \frac{\partial \mathbf{g}(p_j, q_j)}{\partial p_j} \frac{\partial p_j}{\partial \mathbf{z}} + \frac{\partial \mathbf{g}(p_j, q_j)}{\partial q_j} \frac{\partial q_j}{\partial \mathbf{z}}. \tag{28}$$

Focusing on the first term, the partial derivatives in question can be expressed as

$$\frac{\partial \mathbf{g}(p_j, q_j)}{\partial p_j} = -\frac{p_j}{(p_j^2 + q_j^2 + 1)^{3/2}} \begin{pmatrix} -p_j \\ -q_j \\ 1 \end{pmatrix} - \frac{1}{\sqrt{p^2 + q^2 + 1}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{29}$$

$$\frac{\partial p_j}{\partial \mathbf{z}} = \mathbf{D}_x(j, :). \tag{30}$$

where $\mathbf{D}_x(j, :)$ denotes the jth row of the FD operator. Similarly, the second term in (28) can be expressed as

$$\frac{\partial \mathbf{g}(p_j, q_j)}{\partial q_j} = -\frac{q_j}{(p_j^2 + q_j^2 + 1)^{3/2}} \begin{pmatrix} -p_j \\ -q_j \\ 1 \end{pmatrix} - \frac{1}{\sqrt{p^2 + q^2 + 1}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{31}$$

$$\frac{\partial q_j}{\partial \mathbf{z}} = \mathbf{D}_y(j, :). \tag{32}$$

With the expressions of (29)–(32) in hand, (28) can be calculated, whose value can be cascaded all the way up to (25). Since the FD operators in (30) and (32) are very sparse, the Jacobian of the residuals is also very sparse, aiding the tractability of this large-scale minimization problem.

Like the LML method, two Dirichlet constraints are needed due to the nature of the staggered FD operators. As well, to aid convergence, Neumann boundary conditions are applied at the borders, *i.e.*, $\partial z / \partial \mathbf{n} = 0$, where $\mathbf{n}$ represents the normal to the depth-map boundary. Image sequences of an object resting on a flat surface satisfy these boundary conditions. However, if images do not satisfy these boundary constraints, padding can be applied to help constrain image sequences to these boundary conditions.

The expression for the residuals and its derivative can be provided to an established large-scale nonlinear least-squares routine. It should be noted that many such routines assume a smooth function, which the presence of the unit-step function in the residual violates. Nonetheless, in the authors' experience this has not been an issue. Should it become one, the unit-step function can always be replaced with a continuous approximation, along with the appropriate modifications to the partial derivative in (27).

Table 2. Illumination directions used for observation images. Each entry represents an elevation-azimuth angle pair. Experiments used an increasing number of observation images corresponding to the given light directions starting from the leftmost four entries.

Elevation:	60°	45°	55°	15°	75°	55°	30°	45°	20°	35°
Azimuth:	30°	150°	270°	90°	210°	300°	0°	60°	120°	240°

4. EXPERIMENTS

Experiments measured two main qualities. First, the predictive accuracy of the PS, LML, and NML methods was measured against different numbers of observation images. Second, the visual quality of depth-maps produced by the PS-based LML method was compared to the NML method. These two qualities gauge the effectiveness and benefits of executing NML estimation.

To measure these qualities, experiments used simulated images generated from the Mozart depth map, shown earlier, and the vase surface, generated using the formula described by Durou *et al.*[21] For the latter, a padding of 5 pixels was applied so that the surface satisfied Neumann boundary conditions. Tests employed MATLAB's large-scale nonlinear least squares routine, which is based on Coleman and Li's trust-region algorithm,[22] to determine an estimate of $\mathbf{z}$. Minimization was halted if the norm of any changes to $\mathbf{z}$ were less than $1e^{-2}$ or if the algorithm completed 200 iterations.

For both depth maps, experiments were performed on images corrupted with Gaussian noise with 5% standard deviation. Experiments then measured the effect of increasing the number of observation images on the performance of the PS, LML, and NML methods. This was repeated 10 times for each number of observation images, providing an ensemble of results with different realizations of the Gaussian noise. Table 2 provides the illumination directions used for the experiments.

Once observation images were generated, all three methods were executed. AIC_c values were then computed, using each method's SSE for the noisy observation images, to obtain an estimate of each method's predictive accuracy. Additionally, the output of all three methods was used to generate 72 images from roughly a hemisphere of mostly unseen light directions, corresponding to every pairing between elevation angles ranging from 0 to 75°, in increments of 15°, and azimuth angles ranging from 0 to 330°, in increments of 30°. The accuracy of these generated images was then measured by comparing them to ideal and noiseless images generated using the true depth maps. By providing a direct measure of predictive accuracy, this provides a means to gauge how well the AIC_c can assess predictive accuracy.

The top row of Figure 3 depicts the SSE values of each method for the noisy observation images. Except at the 6-images point for the vase surface, the PS method posts the best values for both surfaces at all numbers of observation images. This is expected, as the increased number of parameters in the PS model affords it greater ability to fit the observation images. However, as the middle row of the figure illustrates, the AIC_c value of the PS method is consistently worse than the LML and NML methods, suggesting that the more parsimonious depth-map & albedo model provides greater predictive accuracy. Nonetheless, as the number of observation images increases, differences in model parsimony begin to have less of an impact, and the AIC_c values of the PS method begin to converge to those of the LML and NML methods. These trend lines suggest that the NML method will enjoy the best predictive accuracy. The implications for the LML method are less clear, as the AIC_c is meant to gauge a model's predictive accuracy *given* an ML estimate and is not meant necessarily as a means to compare two estimates of the same model.

As the bottom row of Figure 3 demonstrates, the measured predictive accuracy confirms that the depth-map & albedo model, armed with the NML estimate, can generate significantly more accurate images than the PS model. On the other hand, despite its superior AIC_c values, the LML method underperforms the PS method in most situations. In fact, it only outperforms PS at 6 images or higher with the vase surface. Nonetheless, the predictive accuracy of all three models starts to converge as the number of observation images increases, which aligns with the AIC_c results. In particular, at the lowest number of observation counts, images generated using NML are roughly 40% more accurate than either the PS or LML methods for both surfaces. Thus, the NML method can generate significantly more accurate images at low observation counts. It should be noted, however, that these results are based upon the light directions in Table 2. It is expected that these results would
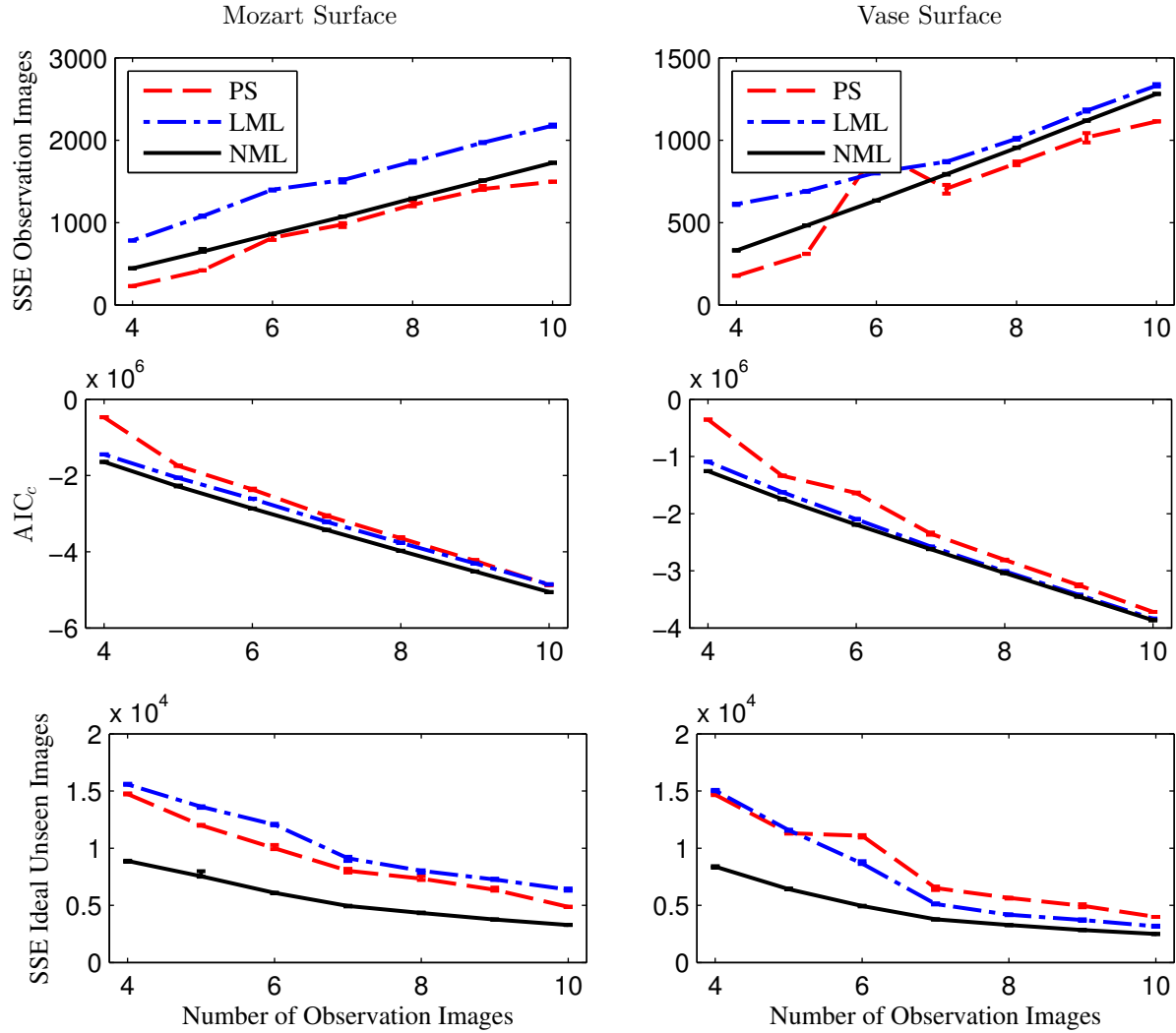
Figure 3. SSE, AIC$_c$ values, and predictive accuracy of the PS, LML, and NML methods for the Mozart and vase surfaces. Experiments ran 10 trials for each number of observation images, corresponding to the light directions given in Table 2. All trend lines depict median values, with error bars representing quartile values. The top row depicts the SSE for the noisy observation images, whereas the middle row depicts the AIC$_c$ values. The bottom row depicts the SSE between images from a hemisphere of light directions generated from the output of the three methods and ideal noiseless images generated from the *true* depth map.

generalize to any set of light directions, but this can be confirmed only with a lengthy set of experiments that run the gamut of the considerable permutations of different possible light directions.

Comparing the middle and bottom rows of Figure 3 suggests that the AIC$_c$ provides a valuable measurement of the predictive accuracy of the PS and NML estimates, which enjoy high levels of accuracy. The reliability of the AIC$_c$ value in assessing the LML's predictive accuracy is more questionable. Having said that, the AIC$_c$ still manages to opt for the depth-map & albedo *model* over the PS one, even when given the LML estimate. This suggests that the AIC$_c$ can provide a useful indicator of which *model* to use. However, anticipating the actual forecasting ability of the *estimates* requires accuracy commensurate with that of the PS and NML methods.

Apart from the significant boost to image generation, the NML method is also able to produce much more visually accurate depth-map estimates over the LML method. For instance, as Figures 4 and 5 demonstrate, the NML method is able to correct the flattened look of the LML's depth maps and suppress some of its noisy characteristics. Nonetheless, certain of the NML trials produced depth maps suffering from isolated spikes,

	Four Observations	Five Observations	Six Observations

(a)   (b)   (c)
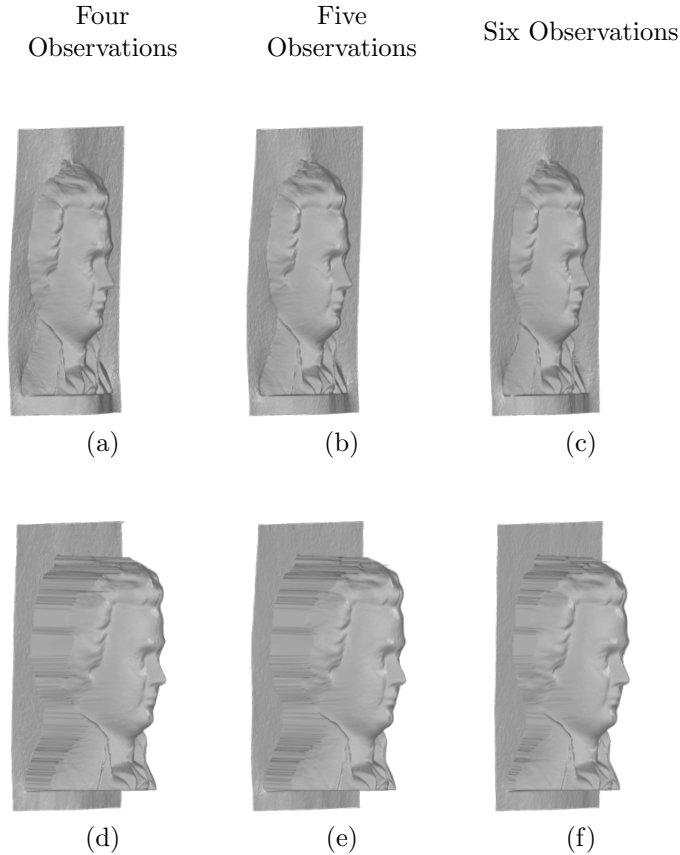
(d)   (e)   (f)

Figure 4. Visual results of the Mozart surface experiments. The depth-map estimates corresponding to the median SSE values for the unseen ideal images of the 10 trials for four, five, and six observation images are depicted. The top and bottom rows display the results for the LML and NML methods, respectively.

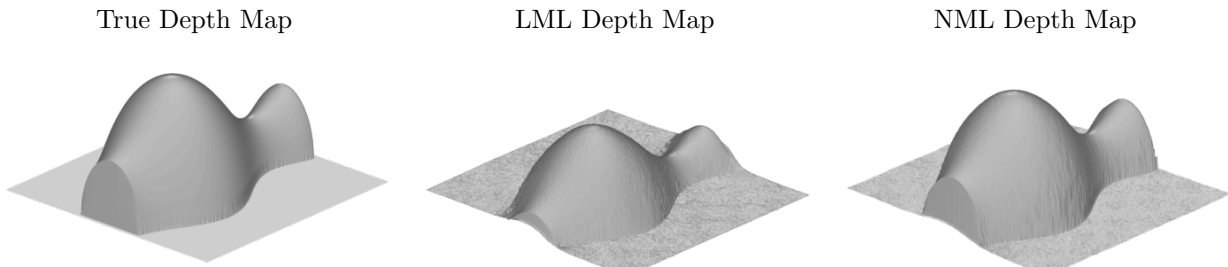True Depth Map   LML Depth Map   NML Depth Map

Figure 5. Visual results of the vase surface experiments. The LML and NML depth-map estimates corresponding to the median SSE values for the unseen ideal images of the 10 trials for *four* observation images are depicted, along with the true vase surface. Results for higher numbers of observation images are similar, except that both methods gradually correct asymmetries in the height of the background on each side of the vase surface. However, differences in scale between the two methods still persist.

indicating that the optimization step occasionally faced convergence issues. Even so, the visual Mozart and vase surface results indicate that in addition to providing an estimate with significantly greater predictive accuracy, the NML method can also produce depth-map estimates with superior appearance. Thus, the NML method is able to benefit both image generation and depth-map & albedo estimation tasks.

5. DISCUSSION

The results presented in this work demonstrate the superior information-theoretic performance of the depth-map & albedo model. As this work highlights, for low observation counts PS-based methods can suffer, motivating the NML method's approach of directly reconciling the depth map & albedo with image observations. Doing so results in more accurately generated images from unseen light directions and also produces depth maps closer to the ground truth. These benefits were anticipated by the Akaike framework, which values model parsimony along with goodness of fit. As a result, this work demonstrates the value of the Akaike framework, and related measures, for surface and reflectance estimation tasks.

The utility of an information-theoretic approach has been recognized in many other computer vision tasks. For instance, a recent book discusses information theory within the context of classification and image understanding tasks.[23] In terms of the AIC specifically, this measure has seen use in modelling the geometry of scenes[24, 25] and the distribution of anatomical structures in medical imaging,[26] and has also been used for tensor decomposition purposes.[27] However, when considering depth-map & reflectance estimation, most works consider only the visual quality of the estimates and not their ability to generate images from unseen light directions.

ML depth-map & albedo estimation has been approached previously by several authors. In the category of the popular PS-based two-step approach, which first conducts PS to estimate surface normals followed by a depth-map estimation step, most methods do not explicitly attempt to determine an ML estimate. Harrison and Joseph's method, *i.e.*, the LML method, is a notable exception, which qualifies as ML estimation under asymptotic conditions and IID Gaussian image noise.[6] While the asymptotic approximations provide for a very robust and effective technique, sample sizes must typically be very large for these approximations to take hold.[28] As well, while the LML method does filter out attached shadows in the PS step, at low observation counts filtering may no longer be possible, introducing errors outside of what Harrison and Joseph's technique is designed to handle. The NML method does not face these problems. In addition, LML estimation decouples albedo and depth-map estimation, when in actual fact both factors are tightly coupled in the image formation model. As a result, at low observation counts, improvements are possible to Harrison and Joseph's method, which this work highlighted.

The difficulties faced by the two-step approach to depth-map estimation motivates work that jettisons the intermediate PS step and instead directly estimates depth maps from image observations. Noakes and Kozera, and their coauthors, have published several works on this topic.[7–9] However, their approach is computationally taxing, leading them to pursue parallelization strategies.[9] Moreover, their work does not address attached shadows, which are typically unavoidable,[29] and employs the unrealistic assumption of uniform albedo. Nonlinearly estimating a depth-map directly from images has also been pursued by Ikeda in the form of a Jacobi iterative scheme.[10, 11] While Ikeda's stated motivations do not explicitly mention ML estimation, he does explicitly cite some of the difficulties PS faces with low observation image counts. Ikeda uses the output of Frankot and Chellapa's PS-based method[14] as an initial solution. However, as others have demonstrated, Frankot and Chellapa's method is susceptible to noise[6] and employs unrealistic periodicity assumptions.[30] Unlike Noakes and Kozera, Ikeda allows for non-constant albedo, but albedo estimation is performed prior to depth-map estimation, decoupling the two factors. In addition, Ikeda's generative model makes no allowance for attached shadows. In terms of convergence, Ikeda's method has only been tested using identical Dirichlet conditions along the border, and its performance on less stringent boundary conditions is unknown.

By coupling depth-map & albedo estimation together, explicitly modelling attached shadows, and using a robust initial solution, the NML method fills important gaps in the state of the art. A major aspect of future work is to reduce some of the convergence issues facing the nonlinear optimization step. This would ease the computational burden and would also help eliminate instances where the depth-map estimate suffers from isolated spikes. Exploring multi-level techniques is a promising direction for this goal. In addition, extending the method to models outside Lambertian image formation remains another important future goal.

6. CONCLUSION

This work outlined a depth-map & albedo estimation method able to produce an accurate ML estimate with superior predictive capabilities and visual appearance, particularly when the number of observation images is low.

The NML method explicated in this work is motivated using the information-theoretic Akaike framework, which provides a principled means to assess a model based on the tradeoff between goodness-of-fit and complexity. Thus, the value of determining an accurate ML depth-map & albedo estimate directly from image observations, even when PS posts low SSE values, is justified using sound theoretical principles. The NML method provides an effective means to attack this difficult problem, outlining how the challenging nonlinear formulation can be framed as an SNLS problem, halving the number of parameters in the optimization step.

The information-theoretic approach used in this work is confirmed by experimental results. When considering how accurately images can be generated using arbitrary light directions, the NML method outperforms PS and the LML by roughly 40% at low observation counts, which is a markedly superior performance and a boon to applications demanding accurately generated images. These results are in line with the AIC_c measure, which anticipated that the NML method would enjoy better predictive accuracy despite being less able to fit the observation images than PS. Focusing on the visual quality of the depth map itself, compared to the PS-based LML method, the NML method produces surfaces visually closer to the true depth map. These results demonstrate the importance of using a more parsimonious model, *i.e.*, obtaining an ML depth-map & albedo estimate directly from image observations, particularly at low observation counts.

The main constributions of this work are twofold. First, this work demonstrated that an information-theoretic approach is extremely valuable for surface and reflectance estimation. Secondly, this work presented a concrete means to realize the benefits of a more parsimonious model, by outlining an effective means to determine an accurate ML estimate using the depth-map & albedo model.

7. ACKNOWLEDGMENTS

REFERENCES

[1] Harrison, A. P., Wong, C., and Joseph, D., "Virtual Reflected-Light Microscopy," *Journal of Microscopy* **244**, 293–304 (2011).

[2] Forster, M. R., "The new science of simplicity," in [*Simplicity, Inference and Modeling: Keeping It Sophisticatedly Simple*], Zellner, A., Keuzenkamp, H. A., and McAleer, M., eds., Cambridge University Press (2002).

[3] Burnham, K. P. and Anderson, D. R., [*Model Selection and Multimodel Inference*], Springer-Verlag, New York, 2 ed. (2002).

[4] Burnham, K. P. and Anderson, D. R., "Multimodel Inference: Understanding AIC and BIC in Model Selection," *Sociological Methods & Research* **33**(2), 261–304 (2004).

[5] Sober, E., "Instrumentalism, Parsimony, and the Akaike Framework," *Philosophy of Science* **69**, S112–S123 (2002).

[6] Harrison, A. P. and Joseph, D., "Maximum Likelihood Estimation of Depth Maps Using Photometric Stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(7), 1368–1380 (2012).

[7] Noakes, L. and Kozera, R., "Nonlinearities and noise reduction in 3-source photometric stereo," *Journal of Mathematical Imaging and Vision* **18**(2), 119–127 (2003).

[8] Noakes, L. and Kozera, R., "Denoising images: Non-linear leap-frog for shape and light-source recovery," in [*Geometry, Morphology, and Computational Imaging*], *Lecture Notes in Computer Science* **2616**, 143–162, Springer Berlin / Heidelberg (2003).

[9] Cameron, T., Kozera, R., and Datta, A., "A parallel leap-frog algorithm for 3-source photometric stereo," in [*Computer Vision and Graphics: International Conference, ICCVG 2004*], *Computational Imaging and Vision*, 95–102, Springer (2006).

[10] Ikeda, O., "Synthetic Shape Reconstruction Combined with the FT-Based Method in Photometric Stereo," in [*ISVC 2010*], Bebis, G., Boyle, R. D., Parvin, B., Koracin, D., Chung, R., Hammoud, R. I., Hussain, M., Tan, K.-H., Crawfis, R., Thalmann, D., Kao, D., and Avila, L., eds., *Lecture Notes in Computer Science* **6453**, 678–687, Springer (2010).

[11] Ikeda, O., "Photometric Stereo Using Four Surface Normal Approximations and Optimal Normalization of Images," in [*Signal Processing and Information Technology, 2006 IEEE International Symposium on*], 672–679 (2006).

[12] Golub, G. and Pereyra, V., "The Differentation of Pseudo-Inverses and the Nonlinear Least Squares Problems Whose Variables Separate," *SIAM Journal on Numerical Analysis* **10**(2), 413–432 (1973).

[13] Golub, G. and Pereyra, V., "Separable nonlinear least squares: the variable projection method and its applications," *Inverse Problems* **19**, R1–R26 (2003).

[14] Frankot, R. T. and Chellappa, R., "A method for enforcing integrability in shape from shading algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **10**(4), 439–451 (1988).

[15] Woodham, R. J., "Photometric method for determining surface orientation from multiple images," *Optical Engineering* **19**(1), 139–144 (1980).

[16] Horn, B. K. P., "Height and Gradient from Shading," *International Journal of Computer Vision* **5**, 37–75 (1990).

[17] Zhang, R., Tsai, P., Cryer, J. E., and Shah, M., "Shape from Shading: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**, 690–706 (1999).

[18] Schwarz, G., "Estimating the Dimension of a Model," *The Annals of Statistics* **6**(2), 461–464 (1978).

[19] Grünwald, P. D., [*The Minimum Description Length Principle*], The MIT Press (2007).

[20] Bedrick, E. J. and Tsai, C. L., "Model Selection for Multivariate Regression in Small Samples," *Biometrics* **50**(1), 226–231 (1994).

[21] Durou, J.-D., Falcone, M., and Sagona, M., "Numerical Methods for Shape-from-shading: A New Survey with Benchmarks," *Computer Vision and Image Understanding* **109**(1), 22 – 43 (2008).

[22] Coleman, T. and Li, Y., "An interior, trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on Optimization* **6**, 418–445 (1996).

[23] Escolano, F., Suau, P., and Bonev, B., [*Information Theory in Computer Vision and Pattern Recognition*], Springer Verlag, London (2009).

[24] Kanatani, K., "Geometric Information Criterion for Model Selection," *International Journal of Computer Vision* **26**(3), 171–189 (1998).

[25] Kanatani, K., "Uncertainty Modeling and Model Selection for Geometric Inference," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(10), 1307–1319 (2004).

[26] Elnakib, A., Gimel'farb, G., Inanc, T., and El-Baz, A., "Modified Akaike information criterion for estimating the number of components in a probability mixture model," in [*Image Processing (ICIP), 2012 19th IEEE International Conference on*], 2497–2500 (2012).

[27] Tao, D., Sun, J., Wu, X., Li, X., Shen, J., Maybank, S., and Faloutsos, C., "Probabilistic tensor analysis with akaike and bayesian information criteria," in [*Neural Information Processing*], Ishikawa, M., Doya, K., Miyamoto, H., and Yamakawa, T., eds., *Lecture Notes in Computer Science* **4984**, 791–801, Springer Berlin Heidelberg (2008).

[28] Seber, G. A. and Wild, C. J., [*Nonlinear Regression*], John Wiley & Sons, Inc. (1989).

[29] Ramamoorthi, R., "Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(10), 1322–1333 (2002).

[30] Harker, M. and O'Leary, P., "Least squares surface reconstruction from measured gradient fields," in [*Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*], 1–7 (2008).