(To appear in ICCV '99)

Coupled Lighting Direction and Shape Estimation from Single Images

Dimitrios Samaras and Dimitris Metaxas
Vision, Analysis and Simulation Technologies Laboratory
Department of Computer & Information Science
University of Pennsylvania
Philadelphia PA 19104-6389

Abstract

This paper presents a new method for the simultaneous estimation of lighting direction and shape from shading. The method estimates the shape and the lighting direction using a two step iterative process. We assume an initial (possibly incorrect) estimate of the lighting position. A stiff deformable model is then fitted to the image, assuming this lighting position. Next, a least-squares estimate of the lighting position is derived from the model using the Levenberg-Marquart method. The two steps — model fitting and lighting-position estimation — are iterated. Once the light direction has converged to a stable solution the deformable model stiffness is lowered and the model fits accurately given the lighting model. In addition, we show how the method can be used with either orthographic or perspective projection assumptions. In a variety of experiments on real and synthetic data, the method is robust to errors both to the initial light position and shape estimates.

1 Introduction

We present a new method for the simultaneous estimation of lighting direction and shape from shading. The difficulty of the problem stems from the nonlinear relationship between the lighting direction and the surface properties, regardless of the lighting model used. Most of the previous work in this area has focused on either shape from shading or on light direction estimation. There have been attempts to address both problems, but they are limited to either assuming a Lambertian lighting model or making local surface assumptions. In addition, they do not provide a tight coupling between shape from shading and the lighting direction estimation.

The method we present in this paper couples the shape from shading and the lighting direction estimation problems using a deformable model framework. Our approach requires the use of a lighting model that is differentiable with respect to both the surface parameters and the lighting direction. Hence, a Lambertian or a more sophisticated lighting model can be used. We use deformable models as shape primitives due to their wide coverage of shapes. A deformable grid is used, based on both global and local parameters.

The method estimates the shape and the lighting direction using a coupled two step iterative process. We first assume an initial (possibly incorrect) estimate of the lighting position. A stiff deformable model is then fitted to the image [19], assuming this lighting position. In the beginning of the fit, only gross-scale shape features will emerge. These features will be roughly correct under most light source configurations, so they will be an improved estimate for the correct light source as well. Next, a least-squares estimate of the lighting position is derived from the model using the Levenberg-Marquart method. The two steps — model fitting and lighting-position estimation — are iterated. Once the light direction has converged to a stable solution the deformable model stiffness is lowered and the model fits accurately given the lighting model.

We aim to solve two tightly coupled problems: a) to improve the fit of the model, and b) to gain a more accurate estimate of the light position. The quality of a model's fit to shading data strongly depends on knowledge of the lighting conditions. Fitting to an incorrect light will either cause the fitting process to not converge or will introduce additional error in the form of excessive wrinkling. Fig. 1 shows two fitted models. In (b), the model is fitted with fixed light position, as estimated by the method in [22]. While the method of [22] gives an approximate estimate of the light position, it is clearly not sufficient for accurate model fitting. In (c), the model is fitted using the method described in this paper; re-estimation of the light position gives a clear improvement, and the error of fit reduces by 11.2 percent. We performed multiple experiments of the method on two synthetic data sets, illuminated under three different lighting conditions. In all cases the lighting direction converged to within 10 degrees of the true direction; in five out of six cases the error was 5.6 degrees or less. In the two cases where the light direction estimate in [22] was rather poor, our method showed some striking improvements. These experiments showed improvements in the light direction estimation error from 42 degrees to 2.2 degrees and 83 degrees to 5.6 degrees. A number of experiments exploring the importance of the initial estimate of the light position demonstrated that the convergence of the light estimation method is almost independent from it. Convergence was within 5 degrees of the true light vector in all of the examples where the initial light estimate was within 30 degrees of the true position and for most of the examples where the light vector was initialized 45 to 90 degrees off. In an example performed on a real image of a face, the final estimate of the light was not affected by such violations of the lighting model as specularities, or varying albedo. Finally, we show how the method can be used with either orthographic or perspective projection assumptions.

The paper is organized as follows. Section 2 describes previous work, and the integration of SFS in deformable models. Section 3 describes the simultaneous estimation of light position and model shape. Section 4 describes two generalizations of the model: the incorporation of perspective projection, and the use of singular points information. Section 5 gives results and section 6 gives the conclusions.

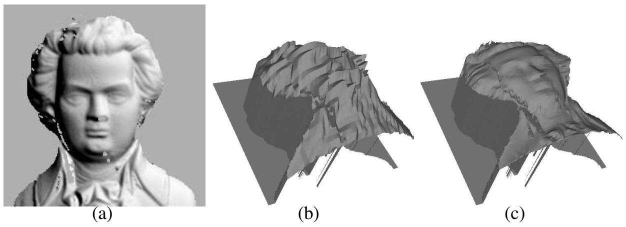


Figure 1: The effects of inaccurate light source position on shape estimation are evident in (b). (a) is the original image; (b) is the fitted model without light position re-estimation; (c) is the fitted model with the light position re-estimation method described in this paper.

2 Background

2.1 Previous work

This work is based on an approach that integrates shape from shading (SFS) within the deformable models framework. Most of the earlier work on SFS has been compiled in [7]; the first comprehensive comparative study of a number of SFS algorithms is [21]. Most of the methods use a regularization approach combined with some additional constraints [3, 8, 11, 9, 22, 12]. [9, 4] combine stereo and shading; [4] handles perspective projection in their stereo and shading mesh. Other approaches are described in [6, 16, 10, 20].

A useful discussion of the ambiguities involved in light source estimation can be found in [2]. A number of researchers have proposed methods for the estimation of the light source direction. [3] proposes an iterative method that updates both the shape and the illuminant direction at every iteration. To avoid local minima a good initial state is often necessary, and furthermore, the requirement for a light source vector of unit length is not enforced. [10] used a Gaussian sphere model for the surface normal distribution and local spherical patches, but did not take shadowing effects into account. [22] presents the most sophisticated of the image based methods, considering shadowing effects and using a uniform distribution of the tilt and slant angles of surface normals (instead of a Gaussian sphere). They still assume local spherical patches and their algorithm suffers on surfaces that deviate significantly from this assumption. [9] derives accurate light source information from stereo data.

In our approach we use the method of [22] to obtain an initial estimate which we then refine using our two step iterative process within our deformable model framework. Our process makes no assumptions about the surface shape, instead it uses the already recovered shape information.

2.2 Deformable Models

In this section we review briefly the general formulation of deformable models [14, 15].

Geometrically, the models used in this paper are parameterized surfaces in space whose intrinsic parameters are $\mathbf{u} = (u, v)$, defined on a domain Ω. The positions $\mathbf{x}(\mathbf{u}, t)$ of points on the model relative to an inertial frame of reference Φ in space are given by $\mathbf{x} = \mathbf{c} + \mathbf{R}\mathbf{p}$, where $\mathbf{c}(t)$ is the origin of the model frame, ϕ, and $\mathbf{R}(t)$ is the rotation matrix expressing the orientation of ϕ. $\mathbf{p}(\mathbf{u}, t)$ denotes the positions of points on the model relative to the model frame. We introduce global and local deformations, by further expressing $\mathbf{p}$ as the sum of a reference shape $\mathbf{s}(\mathbf{u}, t)$ and a displacement function $\mathbf{d}(\mathbf{u}, t)$, i.e., $\mathbf{p} = \mathbf{s} + \mathbf{d}$.

For the applications in this paper, we define $\mathbf{s}$ as a geometric primitive parameterized in uv space (see [14, 15] for details and formulas). Local displacements $\mathbf{d}$ are computed based on the use of triangular finite elements [14, 15], which provide a tessellation of the deformable model's surface. Associated with every finite element node i is a nodal vector variable $\mathbf{q}_{d,i}$. We collect all the nodal variables into a vector of local degrees of freedom $\mathbf{q}_d = (\ldots, \mathbf{q}_{d,i}^\top, \ldots)^\top$, and we compute the local displacement $\mathbf{d}$ based on finite element theory as $\mathbf{d} = \mathbf{S}\mathbf{q}_d$. $\mathbf{S}$ is the shape matrix whose entries are the finite element shape functions.

$$\mathbf{q} = (\mathbf{q}_c^\top, \mathbf{q}_\theta^\top, \mathbf{q}_s^\top, \mathbf{q}_d^\top)^\top \tag{1}$$

are the model parameters with $\mathbf{q}_c = \mathbf{c}$ and $\mathbf{q}_\theta$ is the model's rotational degrees of freedom expressed as a quaternion.

Our goal when fitting the model to visual data is to recover the vector $\mathbf{q}$ which expresses the model's degrees of freedom. Based on Lagrangian dynamics we make our model dynamic in $\mathbf{q}$, and we arrive at the motion equations:

$$\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}_q, \tag{2}$$

where $\mathbf{K}$ is the stiffness matrix, (see [15] for definitions). $\mathbf{f}_q$ are the generalized external forces computed from the 3D forces which in our application will be computed from edge information.

2.3 Integration of Illumination Constraints in Deformable Models

We now briefly summarize the method for integration of lighting constraints in deformable models, as presented in [19], where it was applied to two lighting models; the simple Lambertian model, also briefly described below, and the more accurate and complex model proposed in [17].

In the Lambertian model, if we assume a point light source at infinity, the scene radiance is expressed as

$$I_L = B\frac{\rho}{\pi}\cos\theta = B\rho'\,\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}, \tag{3}$$

where θ is the angle between the surface unit normal vector, $\hat{\mathbf{n}} = \mathbf{n}/\|\mathbf{n}\|$, and the unit light source direction vector $\hat{\mathbf{s}}$. B is the strength of the light source and ρ the constant albedo of the surface. From (3) we get the brightness constraint: $C = \hat{\mathbf{s}}\cdot\mathbf{n} - I'_L\|\mathbf{n}\| = 0$, where $I'_L = I_L/(B\rho')$, and the values of I'_L range between 0 and 1.We want to recover the shape parameters based on the intensity information at n points and therefore we will have an n-component constraint vector $\mathbf{C}$.

We use Lagrange multipliers to incorporate the constraint vector $\mathbf{C}$ in a deformable model framework, as a *nonlinear holonomic constraint* $\dot{\mathbf{C}}(\mathbf{q},t) = 0$. (2) then becomes:

$$\dot{\mathbf{q}} = \mathbf{f} - \mathbf{K}_d\mathbf{q} - \mathbf{C}_{\mathbf{q}}^{\top}\boldsymbol{\lambda}, \tag{4}$$

where $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_n]^{\top}$ are the Lagrange multipliers, $\mathbf{f}$ are generalized edge based forces and $\mathbf{C}_{\mathbf{q}}$ is the Jacobian matrix of the constraints $\mathbf{C}$ w.r.t. the shape parameters. We can consider $-\mathbf{C}_{\mathbf{q}}^{\top}\boldsymbol{\lambda}$ to be the vector of generalized forces on the model parameters due to the lighting constraint.

To obtain the additional necessary equations we differentiate the constraint equation w.r.t. time $\dot{\mathbf{C}}(\mathbf{q},t) = 0$ yielding $\mathbf{C}_{\mathbf{q}}\dot{\mathbf{q}} + \mathbf{C}_t = 0$. In our application $\mathbf{C}$ is not directly dependent on time, therefore $\mathbf{C}_t = 0$. Since we will be fitting a deformable model to the data, the constraints will be far from being satisfied initially, so we will use Baumgarte's [1] stabilization method [14], to replace $\dot{\mathbf{C}}(\mathbf{q},t) = 0$ with the following constraint equation

$$\dot{\mathbf{C}} + \alpha\mathbf{C} = 0, \tag{5}$$

where α is a stabilization factor. Combining (4), (5) we can solve for the generalized constraint forces $-\mathbf{C}_{\mathbf{q}}^{\top}\boldsymbol{\lambda}$ and obtain

$$\mathbf{C}_{\mathbf{q}}^{\top}\boldsymbol{\lambda} = \mathbf{C}_{\mathbf{q}}^{+}(\alpha\mathbf{C} + \mathbf{C}_{\mathbf{q}}(\mathbf{f} - \mathbf{K}_d\mathbf{q})). \tag{6}$$

The matrix $\mathbf{C}_{\mathbf{q}}^{+}$ is the pseudo-inverse of the matrix $\mathbf{C}_{\mathbf{q}}$, and if we define the vector of all non-constraint (generalized) forces, as $\mathbf{b} = \mathbf{f} - \mathbf{K}_d\mathbf{q}$, (4) becomes

$$\dot{\mathbf{q}} = \mathbf{b} - \mathbf{C}_{\mathbf{q}}^{+}(\alpha\mathbf{C} + \mathbf{C}_{\mathbf{q}}\mathbf{b}) = -\mathbf{C}_{\mathbf{q}}^{+}\alpha\mathbf{C} + (\mathbf{I} - \mathbf{C}_{\mathbf{q}}^{+}\mathbf{C}_{\mathbf{q}})\mathbf{b}. \tag{7}$$

The first term $-\mathbf{C}_{\mathbf{q}}^{+}\alpha\mathbf{C}$ in (7) is a model-based least-squares solution to the Baumgarte lighting constraint equations, while the second term $(\mathbf{I} - \mathbf{C}_{\mathbf{q}}^{+}\mathbf{C}_{\mathbf{q}})\mathbf{b}$ projects the forces $\mathbf{b}$ (in this case due to the smoothness of the model) to the space of the constraint forces, thus canceling the part that violates the lighting constraint.

3 Coupling Light and SFS Estimation

In this section we will describe our method that allows the simultaneous estimation of the light source direction and the object shape from shading information. This is done by iteratively improving the estimates of the shape and of the light vector.

3.1 Overview

In monocular orthographic images of Lambertian surfaces, there is inherent ambiguity in the configuration of light and surface shape for the generation of images; in the most general case there is an infinite number of combinations of surface shape and albedo on one hand and light position and intensity on the other [2]. However, considering only surfaces of constant and known albedo, it has been shown that for every image point there are two possible configurations for the normal and the light source [3, 2]. More specifically, if the constraint equations $\mathbf{C}$ are satisfied by shape *A* with normals $\hat{\mathbf{n}}_A$ under light source $\hat{\mathbf{s}}_A$, then they are also satisfied by a dual shape *B* with normals $\hat{\mathbf{n}}_B$ under light source $\hat{\mathbf{s}}_B$ with $\hat{\mathbf{n}}_B = 2\hat{\mathbf{v}}(\hat{\mathbf{n}}_A\hat{\mathbf{v}}) - \hat{\mathbf{n}}_A$ and $\hat{\mathbf{s}}_B = 2\hat{\mathbf{v}}(\hat{\mathbf{s}}_A\hat{\mathbf{v}}) - \hat{\mathbf{s}}_A$, where $\hat{\mathbf{v}}$ is the viewer direction. When the viewing direction is perpendicular to the image plane, this relation preserves unit lengths. This is the standard *in-out* ambiguity which apart from the case of the source being collinear with the viewer, can be easily resolved by taking shadow information into account. Our goal is to estimate either the pair $(\hat{\mathbf{n}}_A, \hat{\mathbf{s}}_A)$ or the pair $(\hat{\mathbf{n}}_B, \hat{\mathbf{s}}_B)$.

Our method is based on the following observation: The integrability constraint on the shape of the model surface, (i.e., the requirement for a physically plausible shape), which is implicitly imposed by our deformable model and the constant albedo assumption, make the image plausible under only one light vector (and its dual)[2]. Therefore, any fitting process, unless we use the correct light vector (or its dual), will not satisfy the constraint by modifying only the surface normals. This is due to the two sources of error: the incorrect normals and the incorrect light position.

In the beginning of the fitting process, while the model surface is still smooth and the model is stiff, only gross-scale shape features will emerge. These features (or their duals) will be roughly correct under most light source configurations (or their duals), so they will result into an improved estimate for the correct light source. That increases our hope that if we re-estimate the light at this point, we will move closer to the true light vector. We can then fit the model to the new light source, re-estimate the light position and iterate this process until convergence.

We therefore introduce the following process which consists of an initial step and an iterative two step procedure.

3.2 Initial light estimate

In the initial step we fit the model under a crude estimate of the light source which we obtain using the image based method described in [22]. This method estimates the light source based on image statistics, under certain assumptions on the probability distribution of the surface normals. These assumptions are general and do not take into account any already existing knowledge about the model shape. Therefore we cannot use it to take advantage of any improvement in the model shape estimates. However it is useful in the initial step given the absence of any shape information; in most of the images we tested it on, it estimated a light direction vector forming an angle of less than $\pi/4$ from the correct solution. As will be seen in the experiments section this is good enough for our algorithm. Since we want the light source vector to be of unit length, we only need to estimate the slant γ (the angle between the illuminant and the positive z axis) and the tilt τ (the angle between the illuminant and the $x-y$ plane). Therefore $\hat{\mathbf{s}} = (\cos\tau\sin\gamma, \sin\tau\sin\gamma, \cos\gamma)$.

For the estimation of τ [22] uses the formulas

$$\tau = \arctan\left(\frac{E_{x,y}\left\{\frac{\tilde{y}_L}{\sqrt{\tilde{x}_L^2+\tilde{y}_L^2}}\right\}}{E_{x,y}\left\{\frac{\tilde{x}_L}{\sqrt{\tilde{x}_L^2+\tilde{y}_L^2}}\right\}}\right), \tag{8}$$

where $E\{\}$ are ensemble averages over the local estimates $\tilde{x}_L, \tilde{y}_L$ are the x and y components of the local estimate of τ, which can be computed by:

$$\begin{bmatrix} \tilde{x}_L \\ \tilde{y}_L \end{bmatrix} = \mathbf{B}^+ \begin{bmatrix} \delta I_1 \\ \vdots \\ \delta I_N \end{bmatrix}, \text{ with } \mathbf{B} = \begin{bmatrix} \delta x_1 & \delta y_1 \\ \vdots & \vdots \\ \delta x_N & \delta y_N \end{bmatrix}, \tag{9}$$

where $\mathbf{B}^+$ the pseudoinverse of $\mathbf{B}$. N is the number of local image derivatives δI.

The slant γ can be determined numerically from the monotonically decreasing function

$$f_3(\gamma) = \frac{E_{x,y}\{I\}}{E_{x,y}\{I^2\}} \tag{10}$$

3.3 Two Step Iteration

In the first iteration of the iterative process we fit the model to the image data using the light source direction estimated in the above initial step. Using the method described in [19], with high stiffness coefficients ($w_0 = w_1 = 0.05$ in our experiments) we compute an initial fit of the shape. The model parameters are updated using (7).

Based on the fitted shape, we can now re-estimate the light direction. This time we want to estimate the source direction, taking into account the surface shape. We want the closest estimate to the source direction which, given the current shape would produce an image as close as possible to the original. We define "closeness" in a least squares sense, in the absence of any general knowledge of which parts of the image are more conforming to the illumination model than others. However one could envision more powerful error metrics for applications where more information is available. We cannot use linear least squares to estimate the three components of the light source independently, because the resulting vector would not be of unit length. For this reason we need the non-linear (τ, γ) parameterization. We use the Levenberg-Marquart method [13, 18], a standard non-linear least-squares estimation technique, to update the values of τ and γ. Levenberg-Marquart smoothly combines an inverse-Hessian method and a steepest descent method, with the objective of minimizing all the jacobians of the error w.r.t. the model's parameters. The method approximates its solution rapidly, but then can wander around the true minimum. So we stop iterating when the error decreases minimally (less than 10^{-3}). In our experiments this typically happens within a few iterations.

Once we have re-estimated the model's parameters, we repeat the model fitting stage under the new light source. We continue iterating between these two steps until the light stabilizes. At that point we decrease (in decrements of 10 percent) the model's stiffness parameters, allowing it to minimize the lighting constraint error.

3.4 Another formulation

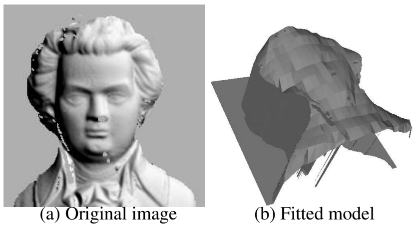
Another formulation of the light and state estimation problem is to include the light parameters in the deformable model framework. Although the light source is physically independent from changes in the deformable model's shape parameters, (they do not appear in the formalism of equations (1) and (2)) they can still be estimated using the parameter update rule of (4) through the constraint Jacobian matrix $\mathbf{C_q}$. We can include the illuminant direction parameters in the parameter vector $\mathbf{q}$ to obtain $\mathbf{q}\prime = (\mathbf{q}^\top, \tau, \gamma)^\top = (\mathbf{q}_c^\top, \mathbf{q}_\theta^\top, \mathbf{q}_s^\top, \mathbf{q}_d^\top, \tau, \gamma)^\top$ and proceed in the same manner as in section (2.3). Experiments show that the above two methods are equivalent in terms of results. However, the first approach is faster, since the light parameters are not updated at every step. An added advantage of the first approach is that this procedure can be used with any iterative shape estimation method.

4 Further Generalizations of the Approach

In this section we present two further generalizations of the approach to account for perspective projection and incorporate singular point information.

4.1 Perspective Projection

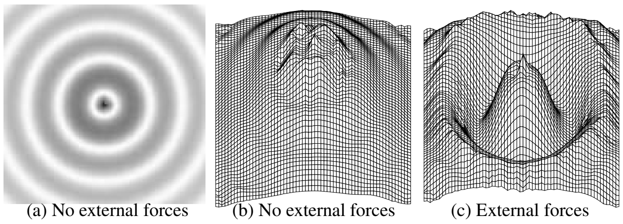
A common assumption in shape estimation from shading is that of orthographic projection. Perspective projection can be easily handled in the deformable model formulation since the reconstruction takes place in 3D space and so the mathematical treatment does not change. There are two differences which affect the reconstruction process.



(a) Original image (b) Fitted model

Figure 2: A model fitted under perspective projection, with a low-resolution model

An image point $\mathbf{P} = (X, Y, -f)$, where f is the focal length of the projection will correspond to model point $\mathbf{p} = (x, y, z) = (-z * X/f, -z * Y/f, z)$. That means that as z changes the model point that corresponds to $\mathbf{P}$ changes too. Points that were visible might become occluded, and visibility checking would be required after every iteration. Furthermore, the surface cannot be canonically sampled in u, v anymore and the local deformation updates have to be computed simultaneously and not per element. That increases substantially the size of the pseudoinverse $\mathbf{C}_{\mathbf{q}}^{+}$ that needs to be computed with SVD.



(a) No external forces (b) No external forces (c) External forces

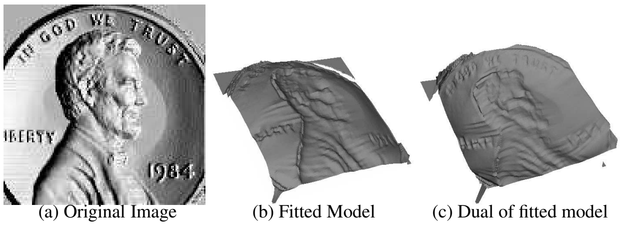
Figure 3: In the case of the sombrero lit from the top, there is convex/concave ambiguity. Knowledge about how to resolve the ambiguity, (eg. singular points) can be applied to the model in the form of external forces. Here a few points in the inner lower circle were pulled down.

4.2 Singular Points Information

It has been shown [16], that the known ambiguities in shape from shading estimation can be resolved with the knowledge of singular points in the image. Here we show how we can incorporate such biases in the deformable model formulation. In the sombrero image viewed head-on in figure 3, there is a convex-concave ambiguity. Our deformable model fits the convex solution only. If we know that we want the inner lower circle to be lowered down, we can simply apply external forces $\mathbf{f}_e$ in (4) on a few of the points on that circle (in the example presented here just five). Then the whole model will converge to the desired solution, in order to satisfy the illumination constraint. Each shading ambiguity has to be resolved independently through the introduction of the proper forces; no forces were applied to the tip of the sombrero which still points upwards.

5 Results

We now present experiments of coupled estimation of light direction and SFS on real and synthetic data. We get substantially improved results in light estimation than previous methods, and consequently in the shape estimation.



(a) Original Image (b) Fitted Model (c) Dual of fitted model

Figure 5: In the case of the penny lit from [1,0,1], the correct shape (c) is the dual of the fitted model (b).

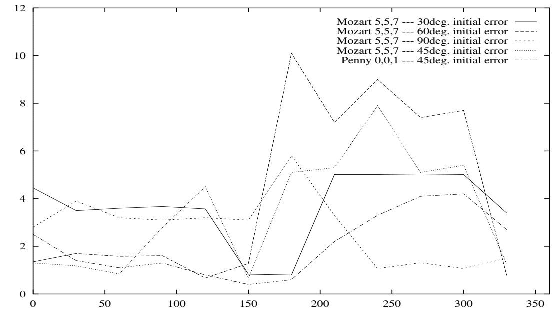


Figure 6: Light direction error analysis w.r.t. initial estimate. We plot the error in the final estimate when the direction of the initial estimate is 30,45,60 and 90 degrees off the true direction. Each initial lighting direction off the true light forms a cone w.r.t. the true light axis. We sample the cone at intervals of 30 degrees. Four sequences at different angles were generated for the Mozart image with true light direction (5,5,7) and one sequence for Penny illuminated from above (0,0,1).

5.1 Synthetic Data

We first present results of experiments on synthetic data sets where we had ground truth for the shape and the generating light source direction. Each data set was illuminated under three different light source directions. For each image we first estimated the light source directions using the method in [22]. This method gave accurate results for the images that were lit head-on (with direction (0,0,1)). In these cases our method obviously gave no improvement (although it did not decrease accuracy — the lighting position remained stable). In all other cases there was a marked im-

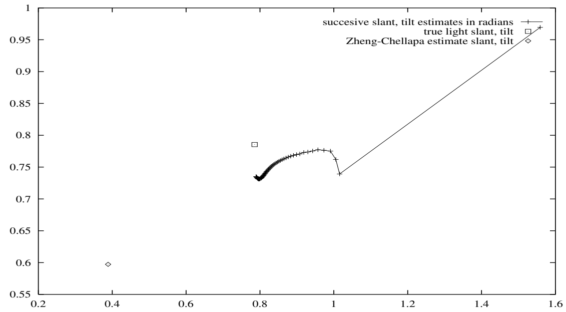
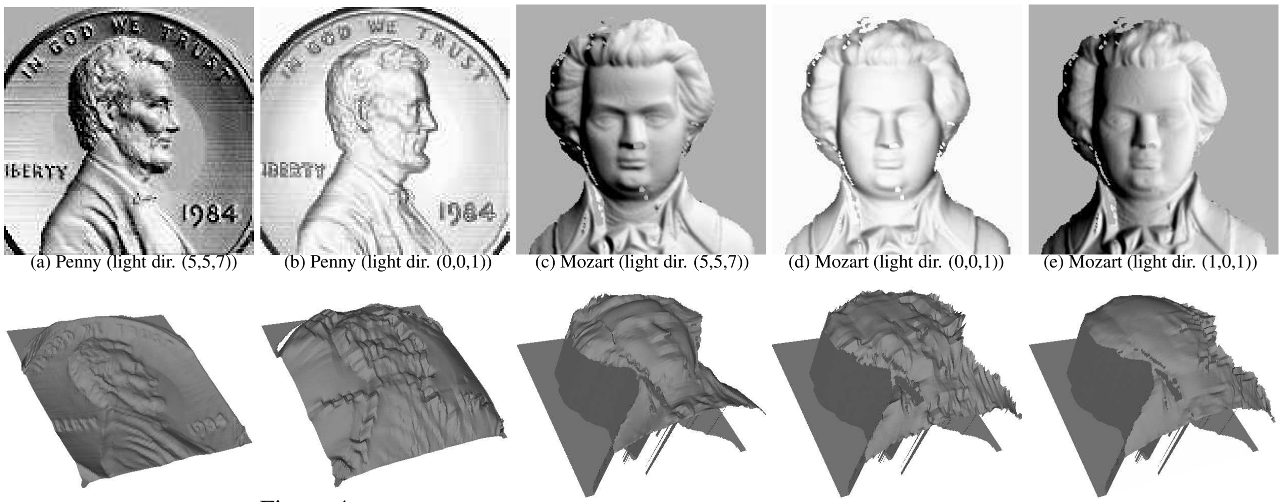


Figure 7: Light estimation results for Mozart lit from [5,5,7]. Initial estimate at top right was 45 degrees off the true light, denoted by a square. Crosses are the estimates of our algorithm connected by a line in the order that the algorithm successively generates them. The final estimate was 2 degrees off the truth. The circle at the bottom left is the estimate generated by the algorithm in [22] which was 23.7 degrees off the true light.



(a) Penny (light dir. (5,5,7)) (b) Penny (light dir. (0,0,1)) (c) Mozart (light dir. (5,5,7)) (d) Mozart (light dir. (0,0,1)) (e) Mozart (light dir. (1,0,1))

Figure 4: Top: Synthetic Images and Light Directions. Bottom: Fitted Models.

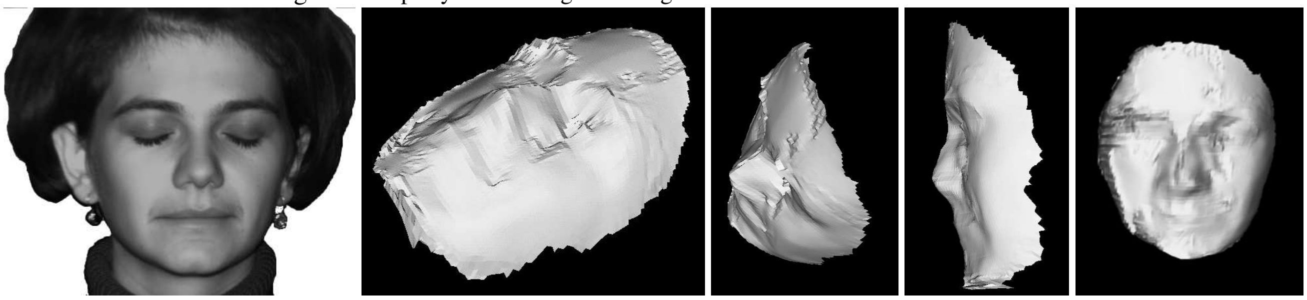


Figure 8: Experiments on real data. Initial image on the left and 4 views of the recovered model

provement over [22].

To test the robustness of our method to an initially poor lighting-direction estimate, we introduced to the images obtained with lighting direction (0,0,1) an initial lighting direction error of 45 degrees. In the mozart example the light estimate converged to 4.8 degrees from the true direction vector and the in the penny example it converged to 3.1 degrees. The recovered shapes can be seen in fig. 4.

The examples with light source direction (5,5,7) (or $\tau = \gamma = \pi/4$) gave the best results. The initial estimate from the [22] method for Mozart was 23.7 degrees off the true vector and our method recovered the illumination direction with a 2 degrees error. The initial estimate for the penny was 83 degrees off. Our model converged to 5.6 degrees of the true light vector.

The examples lit sideways (with $\tau = 0$, direction (1,0,1)) were the ones with the highest errors. The method in [22] gave an estimate for mozart that was 18.9 degrees off from the correct direction. Our method reduced the error to 9.7 degrees. Although a significant improvement, it still wasn't close enough to avoid visible errors in the reconstruction. As for the penny example, the initial estimate was 83 degrees off from the true light but only 7 from its dual (direction (-1,0,1). As expected, our method converged to the dual with an error of 5 degrees. As can be seen in fig. 5 the fitting process gave us the dual shape, as described earlier. We converted this to the correct fit by taking the complement of the normals w.r.t the viewing direction.

In fig. 2 we show the results of fitting to a perspective projection image. The focal length was assumed to be 2 and the distance from the focal point was assumed to be 4, so the perspective distortion effects on the image are quite noticeable. The light source direction was (5,5,7). Our light source estimate was 2.5 degrees off. We show the coarse model that we used while still estimating the light source. Once the light stabilizes we increase the model's resolution and fit it more closely to the lighting constraints.

In our experiments, we found the Levenberg-Marquart method typically to converge very quickly (within three to five iterations) for the estimates of τ and γ.

We analyzed how the choice of the initial lighting estimate affects the error of the recovered light direction. The initial estimates were chosen in the following way: We selected an initial estimate at 30, 45, 60, 90 degrees off the true vector. Then we rotated each estimate around the true vector and formed a cone. We sampled the cone at intervals of 30 degrees. This gave us 12 different estimates for each fixed angle off the true vector, which we used as initial estimates to our method. We see that for all the experiments where the initial estimate was no more than 30 degrees off the true vector, our algorithm gave a lighting position within 5 degrees from the correct answer (fig. 6).

The successive light estimates returned in one of those experiments (initial estimate is 45 degrees off from the truth) are shown in fig. 7, as they quickly converge towards the the true light source direction.

Furthermore, we tested how the initial shape affects the light estimation. The Mozart (5,5,7) image was fitted using a sphere, an ellipsoid and a planar mesh. All three converged to 2.2, 3.5 and 4.6 degrees, respectively. So, based on our experiments, we observe that the method can estimate the light direction, provided that the initial shape has the ability to deform to fit the data. It can take significantly longer to converge if the initial shape is far from the true data.

The method typically converges to the accurate light estimate quickly, in less than 20 iterations of a low resolution model for the Penny experiments.

5.2 Real Data

Fig 8 shows the fitting results to a real image of a face. We compare the reconstruction to data obtained by stereo using the method in [5]. The camera pose and the perspective projection parameters were supplied with the stereo data. We compared the recovered light vector, to the light vector that can be estimated from the stereo data, and the two estimates are 2.6 degrees off. The recovered depth is on average 1.34 percent off from the true data. There are visible errors in the reconstruction of the mouth and the eyes due to the albedo changes and at the tip of the nose due to specularities that violate the Lambertian assumption. These are typical shortcomings when SFS alone is used on real data. However, the light estimation was not adversely affected by the violations of the shading assumptions.

6 Conclusion

We have presented an approach for the simultaneous shape from shading and lighting direction estimation using a deformable model framework. As opposed to previous approaches the method only requires that the lighting constraint used is differentiable with respect to the lighting direction parameters and the deformable model parameters. We have shown that we can solve the above two nonlinearly coupled problems under both orthographic and perspective projections, based on a two step iterative process . Our experiments demonstrate that even if the initial estimate of the lighting direction is very crude, the correct shape and lighting direction are estimated at the end of our iterative approach. Future work includes integrating this methodology with other cues such as optical flow. We plan to experiment with other lighting models, but this is beyond the scope of this paper due to the higher number of parameters they have.

Acknowledgments

The test data sets and images were obtained from the University of Central Florida Computer Vision lab. Stereo data used as ground truth for the face image were supplied by Pascal Fua. Michael Collins and Doug De Carlo provided useful comments. This research has been partially supported by grants from NSF (IRI-97-01803, IRI-9624604, EIA98-09209), ARO and ONR-YIP N00014-97-1-0817.

References

[1] J. Baumgarte. Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 1:1-16, 1972.

[2] P Belhumeur, D Kriegman, A Yuille. The Bas-Relief Ambiguity. In *CVPR 1997* (1060-1066)

[3] M.J. Brooks and B.K.P. Horn. Shape and source from shading. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 932-936, 1985.

[4] P. Fua and Y.G. Leclerc. Object-Centered Surface Reconstruction: Combining Multi-Image Stereo and Shading, IJCV(16), September 1995.

[5] , P. Fua and C. Miccio, From Regular Images to Animated Heads: A Least Squares Approach, ECCV 1998, (188-202)

[6] B.K.P Horn. 1970. Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View. PhD thesis, MIT, 1970.

[7] B.K.P. Horn and M.J. Brooks. eds. Shape from Shading. Cambridge, MA: MIT Press, 1989

[8] K. Ikeuchi and B.K.P Horn. Numerical Shape from shading and occluding boundaries. *Artificial Intelligence*, 17(1-3):141-184, 1981.

[9] Y.G. Leclerc and A.F. Bobick. The Direct Computation of Height from Shading. In *CVPR 1991*, pages 552-558,1991.

[10] C.H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence*, 26:125-143, 1985.

[11] K.M. Lee and C.C. Kuo. Shape from Shading with a Linear Triangular Element Surface Model. *PAMI* 15(8) Aug. 1993

[12] S.H. Lai and B.C. Vemuri. Physically-Based Adaptive Preconditioning for Early Vision, PAMI(19), June 1997.

[13] D.W. Marquart *SIAM* 11:431-441, 1963

[14] D. Metaxas and D. Terzopoulos. Shape and Nonrigid Motion Estimation Through Physics-based Synthesis. *PAMI–15*(6), pp. 580–591, June 1993.

[15] D. Metaxas, Physics-Based Deformable Models: Applications to Computer Vision, Graphics and Medical Imaging, Kluwer-Academic Publishers, November 1996.

[16] J. Oliensis and P. Dupois. A global algorithm for shape from shading. In *ICCV 1993*, pages 692-701, 1993.

[17] M. Oren and S.K. Nayar. Diffuse reflectance from rough surfaces. In *ICCV 1993* pages 763-764, 1993.

[18] W.H. Press, S.A. Teukolsky, W.T. Vettering and B.P. Flannery. *Numerical Recipes in C*, Cambridge University Press, 1992.

[19] D. Samaras and D. Metaxas Incorporating Illumination Constraints in Deformable Models In CVPR '98 (322-329)

[20] P.S. Tsai and M. Shah. A simple shape from shading algorithm. In *CVPR 1992* pages 734-736, 1992.

[21] R. Zhang, P. Tsai, J.E. Cryer, M. Shah. Analysis of shape from shading techniques. In *CVPR 1994* pages 377-384, 1994.

[22] Zheng, Chellappa Estimation of illumination direction, albedo, and shape from shading. *PAMI* 13(7):680-702, 1991.