Shape and Albedo from Multiple Images using Integrability.

A. L. Yuille and D. Snow *

December 22, 2002

Abstract

Previous work [5], [2] have developed an approach for estimating shape and albedo from multiple images assuming Lambertian reflectance with single light sources. The main contributions of this paper are: (i) to show how the approach can be generalized to include ambient background illumination, (ii) to demonstrate the use of the integrability constraint for solving this problem, and (iii) an iterative algorithm which is able to improve the analysis by finding shadows and rejecting them.

* *Smith-Kettlewell Eye Research Institute, San Francisco, CA 94115.*

In Proceedings CVPR'97. Puerto Rico.

1 Introduction

In recent years there has been growing interest in understanding realistic imaging models for objects. This stems from the realization that accurate imaging (or lighting) models are needed to design vision systems which can recognize objects in complex lighting situations, and for the related reconstruction problem of photometric stereo. In both problems – learning object models and photometric stereo – the input is a set of images of the object, or scene, taken under different lighting conditions. The task is to estimate the shape of the object or scene, its reflection function, and its albedo.

Previous work has shown the use of Singular Value Decomposition (SVD) as a tool to solve this problem provided that the reflectance function is Lambertian [5],[2]. SVD, however, is only able to solve the problem up to an unknown constant linear transform in the shape and albedo of the viewed object(s) and further assumptions must be made to resolve this ambiguity. In addition, the SVD approach ignores shadows, background ambient illumination, and specularities.

The main contributions of this paper are: (i) to show how the approach can be generalized to include ambient background illumination, (ii) to clarify the nature of the linear transformation ambiguity and demonstrate ways to solve it using the surface consistency condition, and (iii) to demonstrate an iterative algorithm which is able to improve the analysis by finding and rejecting shadows.

2 Previous Work

There have been many important papers on photometric stereo [11], [12],[8] and on learning lighting models for objects [3], [9]. In this paper we are mainly concerned with the work of [5] and [2].

Suppose we have a set of images generated by a Lambertian model where the lighting conditions vary. We use $\mathbf{x}$ to label positions in the image plane Ω and let $|\Omega|$ be the number of these positions (we assume a finite grid). The light source directions are *unknown* and are labeled by $\mu = 1, ..., M$. This gives us a set of M images:

$$I(\mathbf{x}, \mu) = a(\mathbf{x})\mathbf{n}(\mathbf{x}) \cdot \mathbf{s}(\mu) \equiv \mathbf{b}(\mathbf{x}) \cdot \mathbf{s}, \tag{1}$$

where $a(\mathbf{x})$ is the albedo of the object, $\mathbf{n}(\mathbf{x})$ is its surface normal, $\mathbf{b}(\mathbf{x}) \equiv a(\mathbf{x})\mathbf{n}(\mathbf{x})$ (observe that $a(\mathbf{x}) = |\mathbf{b}(\mathbf{x})|$ and $\mathbf{n}(\mathbf{x}) = \widehat{\mathbf{b}(\mathbf{x})}$), and $\mathbf{s}(\mu)$ is the light source direction. We will typically work with $\mathbf{b}(\mathbf{x})$ instead of $a(\mathbf{x})$ and $\mathbf{n}(\mathbf{x})$. Equation (1), however, has several limitations. It ignores shadows, ambient illumination, and specularities.

We wish to solve equation (1) for the albedo, shape, and light source directions. To do this, we define a least squares cost function:

$$E[\mathbf{b}, \mathbf{s}] = \sum_{\mu, \mathbf{x}} \{I(\mathbf{x}, \mu) - \sum_{i=1}^{3} b_i(\mathbf{x}) s_i(\mu)\}^2 \tag{2}$$

It is possible to minimize this cost function to solve for $\mathbf{b}(\mathbf{x})$ and $\mathbf{s}(\mu)$ up to a constant linear transform using Singular Value Decomposition (SVD) [5],[2].

To see this, observe that the intensities $\{I(\mathbf{x}, \mu)\}$ can be expressed as a $M \times |\Omega|$ matrix $\mathbf{J}$ where M is the number of images (light sources) and $|\Omega|$ is the number of points $\mathbf{x}$.. Similarly we can express the surface properties $\{b_i(\mathbf{x})\}$ as a $|\Omega| \times 3$ matrix $\mathbf{B}$ and the light sources $\{s_i(\mu)\}$ as a $3 \times M$ matrix $\mathbf{S}$. SVD implies that we can write $\mathbf{J}$ as:

$$\mathbf{J} = \mathbf{U}\mathbf{D}\mathbf{V}^T, \tag{3}$$

where $\mathbf{D}$ is a diagonal matrix whose elements are the square roots of the eigenvalues of $\mathbf{J}\mathbf{J}^T$. The columns of $\mathbf{U}$ correspond to the normalized eigenvectors of the matrix $\mathbf{J}^T\mathbf{J}$. The ordering of these columns corresponds to the ordering of the eigenvalues in $\mathbf{D}$. Similarly, the columns of $\mathbf{V}$ correspond to the eigenvectors of $\mathbf{J}\mathbf{J}^T$.

If our image formation model is correct then there will only be three nonzero eigenvalues of $\mathbf{J}\mathbf{J}^T$ and so $\mathbf{D}$ will have only three nonzero elements. We do not expect this to be true for our dataset because of shadows, ambient background, specularities, and noise. But SVD is guaranteed to gives us the best least squares solution in any case. Thus the biggest three eigenvalues of Σ, and the corresponding columns of $\mathbf{U}$ and $\mathbf{V}$ represent the Lambertian part of the reflectance function of these objects. We define the vectors $\{\mathbf{f}(\mu) : \mu = 1, ..., M\}$ to be the first three columns of $\mathbf{U}$ and the $\{\mathbf{e}(\mathbf{x})\}$ to be the first three columns of $\mathbf{V}$.

This assumption enables us to use SVD to solve for $\mathbf{b}$ and $\mathbf{s}$ up to a linear transformation. The solution is:

$$\begin{aligned} \mathbf{b}(\mathbf{x}) &= \mathbf{P}_3\mathbf{e}(\mathbf{x}),\ \forall\ \mathbf{x}, \\ \mathbf{s}(\mu) &= \mathbf{Q}_3\mathbf{f}(\mu),\ \forall\ \mu, \end{aligned} \tag{4}$$

where $\mathbf{P}_3$ and $\mathbf{Q}_3$ are 3×3 matrices which are constrained to satisfy $\mathbf{P}_3^T\mathbf{Q}_3 = \mathbf{D}_3$, where $\mathbf{D}_3$ is the 3×3 diagonal matrix containing the square roots of the biggest three eigenvalues of $\mathbf{J}\mathbf{J}^T$. There is an ambiguity $\mathbf{P}_3 \mapsto \mathbf{A}\mathbf{P}_3$, $\mathbf{Q}_3 \mapsto \mathbf{A}^{\mathbf{-1}\ \mathbf{T}}\mathbf{Q}_3$ where $\mathbf{A}$ is an arbitrary invertible matrix.

This ambiguity is inherent in the original Lambertian equation (1), where it is clear that the equation is invariant to the transformation $\mathbf{b}(\mathbf{x}) \mapsto \mathbf{A}\mathbf{b}(\mathbf{x})$ and $\mathbf{s}(\mu) \mapsto \mathbf{A}^{-1^T}\mathbf{s}(\mu)$.

Hayakawa [5] proposed assuming that the light source had constant magnitude for six or more images. By equations (4), we have $\mathbf{s}(\mu) \cdot \mathbf{s}(\mu) = \mathbf{f}^T(\mu)\mathbf{Q}_3^T\mathbf{Q}_3\mathbf{f}(\mu)$. Imposing the constraint that this is constant only allows us to solve for $\mathbf{Q}_3^T\mathbf{Q}_3$ and hence there is a further ambiguity $\mathbf{Q}_3 \mapsto \mathbf{R}\mathbf{Q}_3$, where $\mathbf{R}$ is a rotation matrix satisfying $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. Hayakawa assumes that this rotation $\mathbf{R}$ is the identity, which as is shown in [13] is equivalent to making assumptions of symmetry about the dataset which are only appropriate for very special situations. An alternative assumption suggested by Hayakawa – that the magnitude of the surface reflectance was known for six or more surface points – would lead to knowledge of $\mathbf{P}_3^T \mathbf{P}_3$ and hence to an equivalent ambiguity $\mathbf{P}_3 \mapsto \mathbf{R}\mathbf{P}_3$.

Epstein, Yuille and Belhumeur [2] described several techniques of solving for the linear transformation some of which assume prior knowledge of the object class. More importantly, they argued for use of the surface integrability constraint as a way of constraining the linear transform. They showed that, in theory, the constraint restricted the ambiguity to a three-dimensional family of transformations which send $b_1(\mathbf{x}) \mapsto \lambda b_1(\mathbf{x}) + \alpha b_3(\mathbf{x})$, $b_2(\mathbf{x}) \mapsto \lambda b_2(\mathbf{x}) + \beta b_3(\mathbf{x})$, and $b_3(\mathbf{x}) \mapsto \tau b_3(\mathbf{x})$. This transformation, *the generalized bas relief transform* (GBR), has been further investigated [1] and shown to apply even when shadows are present. In our previous work, however, it was only possible to check surface integrability rather than directly using it [2].

3 Generalization to include ambient illumination

In this paper we generalize the model to include a background ambient illumination term. This will mean that we can obtain the albedo, shape, and ambient term even if the ambient term is highly complex such as projecting another image onto the object, see figure (1).

This means we modify the equations to be:

$$I(\mathbf{x}, \mu) = \mathbf{b}(\mathbf{x}) \cdot \mathbf{s}(\mu) + \tilde{a}(\mathbf{x}), \tag{5}$$

where $\tilde{a}(\mathbf{x})$ is the ambient illumination which we assume to be independent of μ. (I.e. we assume that the ambient illumination stays constant while the Lambertian light sources vary).

We define a cost function for estimating $\mathbf{b}$, $\mathbf{s}$, and $\tilde{a}$:

$$E[\mathbf{b}, \mathbf{s}, \tilde{a}] = \sum_{\mathbf{x}, \mu} \{I(\mathbf{x}, \mu) - (\mathbf{b}(\mathbf{x}) \cdot \mathbf{s}(\mu) + \tilde{a}(\mathbf{x}))\}^2. \tag{6}$$

It is straightforward to generalize our previous approach and apply SVD to estimate $\mathbf{b}$, $\mathbf{s}$, and $\tilde{a}$. The important difference is that we now rely on the first four eigenvalues of $\mathbf{J}\mathbf{J}^T$. We can generalize equation (4) to:

$$\begin{aligned} \mathbf{b}(\mathbf{x}) &= \mathbf{P}_3 \mathbf{e}(\mathbf{x}) + \mathbf{p}_4 e_4(\mathbf{x}), \\ \tilde{a}(\mathbf{x}) &= \mathbf{v} \cdot \mathbf{e}(\mathbf{x}) + v_4 e_4(\mathbf{x}), \\ \mathbf{s}(\mu) &= \mathbf{Q}_3 \mathbf{f}(\mu) + \mathbf{q}_4 f_4(\mu), \\ 1 &= \mathbf{w} \cdot \mathbf{f}(\mu) + w_4 f_4(\mu). \end{aligned} \tag{7}$$

As before, there is a linear ambiguity. The difference is that it is now a four by four linear transformation instead of a three by three. It will turn out, however, that the integrability constraint in combination with assuming constant magnitude of the light source will be sufficient to remove this ambiguity. Note that the last equation of (7) already gives conditions on $\mathbf{w}$ and w_4.

4 Using Surface Integrability

From the analysis in [13] we conclude that symmetry assumptions on the dataset place powerful constraints on the linear transformation, but are not generally applicable.

However, there is another constraint which can always be applied. This is the *surface integrability equations*, see various chapters such as Frankot and Chellappa in [6]. These integrability equations are used to ensure that the set of surface normals forms a consistent surface. As discussed in [2],[1] these constraints are powerful enough, theoretically, to determine the surface and albedo up to a generalized bas relief transformation (GBR). In this section we introduce a method for imposing integrability and demonstrate that it will indeed work in practice.

The integrability constraints are usually expressed in terms of the surface normals but, as shown in [2], they can be generalized to apply to the $\mathbf{b}(\mathbf{x})$ vectors. The constraints can be expressed in differential form:

$$\frac{\partial}{\partial x}\Big(\frac{b_2(\mathbf{x})}{b_3(\mathbf{x})}\Big) = \frac{\partial}{\partial y}\Big(\frac{b_1(\mathbf{x})}{b_3(\mathbf{x})}\Big) \tag{8}$$

Expanding this out we get

$$b_3\frac{\partial b_2}{\partial x} - b_2\frac{\partial b_3}{\partial x} = b_3\frac{\partial b_1}{\partial y} - b_1\frac{\partial b_3}{\partial y} \tag{9}$$

For simplicity of mathematics, we now define $\mathbf{P}$ to be a 3×4 matrix equal to $\big(\mathbf{P}_3, \mathbf{p_4}\big)$, where $\mathbf{P}_3$ and $\mathbf{p}_4$ are defined in equation (7). The three rows of $\mathbf{P}$ are three four-vectors which we denote by $\check{\mathbf{p}}_1$, $\check{\mathbf{p}}_2$, $\check{\mathbf{p}}_3$.

Now we substitute the following values for $\mathbf{b}(\mathbf{x})$, $b_i(\mathbf{x}) = \sum_{\tau=1}^{4} \mathbf{P}_{i\tau} e_\tau(\mathbf{x}), \quad i = 1, 2, 3.$

$$\sum_{\mu<\nu}\{P_{3\mu}P_{2\nu} - P_{2\mu}P_{3\nu}\}\ \{e_\mu \tfrac{\partial e_\nu}{\partial x} - e_\nu \tfrac{\partial e_\mu}{\partial x}\} = \sum_{\mu<\nu}\{P_{3\mu}P_{1\nu} - P_{1\mu}P_{3\nu}\}\ \{e_\mu \tfrac{\partial e_\nu}{\partial y} - e_\nu \tfrac{\partial e_\mu}{\partial y}\} \tag{10}$$

This gives us $|\Omega|$ linear equations for twelve unknowns. These equations are therefore over constrained but they can be solved by least squares to determine the $P_{3\mu}P_{2\nu} - P_{2\mu}P_{3\nu}$ and $P_{3\mu}P_{1\nu} - P_{1\mu}P_{3\nu}$ up to a constant scaling factor. These correspond to the cross products in four dimensions $\check{\mathbf{p}}_1 \times \check{\mathbf{p}}_3$, and $\check{\mathbf{p}}_2 \times \check{\mathbf{p}}_3$, . By inspection, the only transformation which preserves these cross products is:

$$\begin{aligned} \check{\mathbf{p}}_1 &\rightarrow \lambda\check{\mathbf{p}}_1 + \alpha\check{\mathbf{p}}_3 \\ \check{\mathbf{p}}_2 &\rightarrow \lambda\check{\mathbf{p}}_2 + \beta\check{\mathbf{p}}_3 \\ \check{\mathbf{p}}_3 &\rightarrow \frac{1}{\lambda}\check{\mathbf{p}}_3 \end{aligned} \tag{11}$$

which corresponds to the GBR [2],[1].

This means, consistent with the generalized bas relief ambiguity, that knowing these cross products will only allow us to solve for the $\mathbf{P}$ up to a generalized bas relief transformation. We now describe an explicit procedure to solve for the $\mathbf{P}$ in terms of the cross products.

First, observe that the cross product terms include some of the co-factors of the three by three submatrix $\mathbf{P}_3$. Recall that this matrix is related to co-factors:

$$k\begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} = \begin{pmatrix} \Delta_{11} & \Delta_{21} & \Delta_{31} \\ \Delta_{12} & \Delta_{22} & \Delta_{32} \\ \Delta_{13} & \Delta_{23} & \Delta_{33} \end{pmatrix}^{-1} \tag{12}$$

where the co-factors are given by $\Delta_{11} = P_{22}P_{33} - P_{23}P_{32}$, etc. and k is a normalization constant.

In fact, the cross products determine the co-factors Δ_{11},Δ_{12},Δ_{13},Δ_{21},Δ_{22}, Δ_{23}. The remaining three co-factors $\Delta_{31}, \Delta_{32}, \Delta_{33}$ are unknown. These correspond to the parameters λ, α, β of the generalized bas relief transformation. Specific choices of them will correspond to specific transformations. We therefore select values for $\Delta_{31}, \Delta_{32}, \Delta_{33}$ which will later be modified as we solve for the generalized bas relief ambiguity.

We can now solve equation (12) to determine $\mathbf{P}_3$ up to GBR. To determine the remaining values of $\mathbf{P}$, the $\mathbf{p_4}$, we use the remaining cross product terms and least squares.

The results show that we can reconstruct the $\mathbf{b}$ up to a GBR. Figure (2) shows the result of the reconstruction in (a), but also shows the deformations that might arrise due to a GBR in (b).

5 The Full SVD solution

The previous section has shown how we can use integrability to solve for $\mathbf{P}$, and hence $\mathbf{b}$, up to a GBR. This means that the full solution is given by $\mathbf{GP}_3^* \; \mathbf{Gp}_4^*$ where $\mathbf{G}$ is an arbitrary GBR and $\mathbf{P}_3^*$ and $\mathbf{p}_4^*$ are the output of our algorithm imposing integrability.

In this section, we show that additional assumptions can be used to determine the full solution. There are several possible choices. The one we describe here assumes that the magnitude of the light source is constant.

First, observe that we can directly solve for $\mathbf{w}$ and w_4 using the last equation of (7) and applying least squares. This gives:

$$w_4 = \sum_{\mu} f_4(\mu), \quad \mathbf{w} = \sum_{\mu} \mathbf{f}(\mu). \tag{13}$$

In addition, we can use the assumption that the light sources have constant magnitude which, without loss of generality, can be set equal to 1. Using equation (7) we find that we can express the magnitude squared of the $\mathbf{s}(\mu)\cdot\mathbf{s}(\mu)$ in terms of unknown quantities such as $\mathbf{Q}_3^T\mathbf{Q}_3$, $\mathbf{Q}_3^T\mathbf{q}_4$, and $\mathbf{q}_4^T\mathbf{q}_4$ and known quantities such as the eigenvectors. This, extending our analysis of Hayakawa, allows us to determine $\mathbf{Q}_3$ and $\mathbf{q}_4$ up to an arbitrary rotation matrix $\mathbf{R}$ (using a least squares cost function solved by a mixture of SVD and steepest descent).

By now, we have solved for the $\mathbf{P}$ up to a GBR $\mathbf{G}$, the $\mathbf{w}$ and w_4, and $\mathbf{Q}_3$ and $\mathbf{q}_4$ up to a rotation $\mathbf{R}$. We have, as yet, no knowledge of $\mathbf{v}$ and v_4.

But we still have the constraint that $\mathbf{P}^T\mathbf{Q} = \mathbf{D}$. We see that $\mathbf{G}$ and $\mathbf{R}$ only appear in this equation in the form $\mathbf{M} \equiv \mathbf{G}^T\mathbf{R}$. Indeed the equations $\mathbf{P}^T\mathbf{Q} = \mathbf{D}$ reduce to linear equations for $\mathbf{M}, \mathbf{v}, v_4$. They can therefore be solved by least squares.

It now remains to determine $\mathbf{G}$ and $\mathbf{R}$ from $\mathbf{M}$. Recall that $\mathbf{M} = \mathbf{G}^T\mathbf{R}$, where $\mathbf{G}$ is a GBR and $\mathbf{R}$ is a rotation matrix. We therefore have that $\mathbf{M}\mathbf{M}^T = \mathbf{G}\mathbf{G}^T$ and so we can determine $\mathbf{G}\mathbf{G}^T$. From the form of a GBR it can be shown that $\mathbf{G}$ can be determined uniquely from $\mathbf{G}\mathbf{G}^T$ apart from a square root ambiguity (corresponding to the well-known concave/convex ambiguity). Now that $\mathbf{G}$ is known we can solve $\mathbf{M} = \mathbf{G}^T\mathbf{R}$ by least squares while imposing the condition that $\mathbf{R}$ is a rotation matrix. Figures (3 shows the results on the face.

6 Locating and Rejecting Shadows

So far, we have assumed that there are no shadows, or specularities, in the image. But it is clear from our dataset that this is a poor assumption. The least squares techniques we impose have given us some protection against the effect of shadows, but inevitably biases have been introduced.

In this section, we show that we can modify our method and eliminate shadows by an iterative process starting with the results given by the SVD method. Our strategy is to treat the shadows as outliers which can be removed by techniques from Robust Statistics [7]. We introduce a binary indicator variable $\mathbf{V}(\mathbf{x},\mu)$ which can be used to indicate whether a point $\mathbf{x}$ is in shadow when the illuminant is $\mathbf{s}(\mu)$. This variable $\mathbf{V}(\mathbf{x},\mu)$ must be estimated. To do so, we can use our current estimates of $\mathbf{b}(\mathbf{x})$ and $\mathbf{s}(\mu)$ to determine whether $\mathbf{x}$ is likely to be in shadow from light source $\mathbf{s}(\mu)$. This will be determined by setting $\mathbf{V}(\mathbf{x},\mu) = 0$ if $\mathbf{n}(\mathbf{x})\cdot\mathbf{s}(\mu) \leq T$, where T is a threshold. We then re-estimate $\mathbf{b}(\mathbf{x})$ and $\mathbf{s}(\mu)$ and repeat.

More precisely, we define a modified energy function:

$$E[\mathbf{V}(\mathbf{x},\mu),\mathbf{b}(\mathbf{x}),\tilde{a}(\mathbf{x}),\mathbf{s}(\mu)] = \sum_{\mathbf{x},\mu}\{\mathbf{I}(\mathbf{x},\mu)-\mathbf{V}(\mathbf{x},\mu)\mathbf{b}(\mathbf{x})\cdot\mathbf{s}(\mu)-\tilde{a}(\mathbf{x})\}^2$$
$$+c_1\sum_x\{(b_3\frac{\partial b_1}{\partial y}-b1\frac{\partial b_3}{\partial y})-(b_3\frac{\partial b_2}{\partial x}-b_2\frac{\partial b_3}{\partial x})\}^2$$
$$c_2\sum_\mu\{1-\mathbf{s}(\mu)\cdot\mathbf{s}(\mu)\}^2 \quad (14)$$

where c_1 and c_2 are constants.

We set

$$\mathbf{V}(\mathbf{x},\mu) = \begin{cases} 0 & \text{if } \mathbf{b}(\mathbf{x})\cdot\mathbf{s}(\mu)\le T \\ 1 & \text{if } \mathbf{b}(\mathbf{x})\cdot\mathbf{s}(\mu) > T \end{cases} \quad (15)$$

Then we minimize with respect to the variables $\mathbf{b}(\mathbf{x})$, $\mathbf{s}(\mu)$, and $\tilde{a}(\mathbf{x})$.

The energy is quadratic in $\tilde{a}(\mathbf{x})$ and can be minimize directly,

$$\frac{\partial E}{\partial\tilde{a}(\mathbf{x})} = -2\sum_\mu\{\mathbf{I}(\mathbf{x},\mu)-\mathbf{V}(\mathbf{x},\mu)\mathbf{b}(\mathbf{x})\cdot\mathbf{s}(\mu)-\tilde{a}(\mathbf{x})\}$$
$$\tilde{a}(\mathbf{x})^* = \frac{1}{N}\sum_\mu\{\mathbf{I}(\mathbf{x},\mu)-\mathbf{V}(\mathbf{x},\mu)\mathbf{b}(\mathbf{x})\cdot\mathbf{s}(\mu)\} \quad (16)$$

Where N is the number of images.

To minimize with respect to $\mathbf{s}(\mu)$ requires steepest descent.

$$\frac{\partial E}{\partial\mathbf{s}(\mu)} = \sum_{\mathbf{x}}\{\mathbf{I}(\mathbf{x},\mu)-\mathbf{V}(\mathbf{x},\mu)\mathbf{b}(\mathbf{x})\cdot\mathbf{s}(\mu)-\tilde{a}(\mathbf{x})\}(-2\mathbf{V}(\mathbf{x},\mu)\mathbf{b}(\mathbf{x}))$$
$$-4c_2\{1-\mathbf{s}(\mu)\cdot\mathbf{s}(\mu)\}\mathbf{s}(\mu) \quad (17)$$

For $\mathbf{b}(\mathbf{x})$ we also need steepest descent. Because of the derivative in $\mathbf{b}(\mathbf{x})$ we need to descretize the integrability terms.

$$\frac{\partial E}{\partial\mathbf{b}(\mathbf{x})} = \sum_\mu\{\mathbf{I}(\mathbf{x},\mu)-\mathbf{V}(\mathbf{x},\mu)\mathbf{b}(\mathbf{x})\cdot\mathbf{s}(\mu)-\tilde{a}(\mathbf{x})\}(-2\mathbf{V}(\mathbf{x},\mu)\mathbf{s}(\mu))$$
$$+\ \text{integrability terms} \quad (18)$$

The integrability energy terms are

$$(b_3\frac{\partial b_1}{\partial y}-b_1\frac{\partial b_3}{\partial y})\to\{b^3_{i,j}(b^1_{i,j+1}-b^1_{i,j})-b^1_{i,j}(b^3_{i,j+1}-b^3_{i,j})\}$$
$$(b_3\frac{\partial b_2}{\partial x}-b_2\frac{\partial b_3}{\partial x})\to\{b^3_{i,j}(b^2_{i+1,j}-b^2_{i,j})-b^2_{i,j}(b^3_{i+1,j}-b^3_{i,j})\}$$

The derivatives are given by:

$$\frac{\partial I_c}{\partial b^1_{i,j}} = -2c_1\sum_{i,j}\{b^3_{i,j}(b^1_{i,j+1}-b^1_{i,j})-b^1_{i,j}(b^3_{i,j+1}-b^3_{i,j})$$
$$-(b^3_{i,j}(b^2_{i+1,j}-b^2_{i,j})-b^2_{i,j}(b^3_{i+1,j}-b^3_{i,j}))\}$$
$$(b^3_{i,j+1}-b^3_{i,j}) \quad (19)$$

and similarly for the derivatives with respect to $b^2_{i,j}$ and $b^3_{i,j}$.

We repeat this process several times and obtain the shadows, see figure (4) and the albedo, ambient, and light source direction, see figure (5).

7 Conclusion

We have shown that the SVD approach can be generalized to include ambient background illumination. Furthermore, we have demonstrated the power of the integrability constraint and shown that it can reduce the amount of assumptions needed to solve for shape and albedo and light source directions. Finally, we have demonstrated that a robust version of SVD can be used to deal with shadows and specularities.

Our current work involves extending this algorithm to deal with more complex reflectance functions. In particular, by using bidirectional reflectance distribution functions (BRDF's) to model non-Lambertian components of the reflection.

Acknowledgements

It is a pleasure to acknowledge conversations with Peter Belhumeur, Russell Epstein, and David Kriegman. Support for this work was provided by the NSF Grant IRI 92-23676 and ARPA/ONR Contract N00014-95-1-1022.

References

[1] P. Belhumeur, D. Kriegman, and A.L. Yuille. "The Bas-Relief Ambiguity". Submitted to IEEE Conf. on Computer Vision and Pattern Recognition. 1997.

[2] R. Epstein, A.L. Yuille, and P.N. Belhumeur. "Learning object representations from lighting variations".In **Object Representation in Computer Vision II**. Eds. J. Ponce, A. Zisserman, and M. Hebert. Springer Lecture Notes in Computer Science 1144. 1996.

[3] P.W. Hallinan. "A low-dimensional lighting representation of human faces for arbitrary lighting conditions". In. *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pp 995-999. 1994.

[4] P.W. Hallinan. **A Deformable Model for Face Recognition under Arbitrary Lighting Conditions.** PhD Thesis. Division of Applied Sciences. Harvard University. 1995.

[5] K. Hayakawa. "Photometric Stereo under a light source with arbitrary motion". *Journal of the Optical Society of America* A, 11(11). 1994.

[6] B.K.P. Horn and M. J. Brooks, Eds. **Shape from Shading**. Cambridge MA, MIT Press, 1989.

[7] P.J. Huber. **Robust Statistics**. John Wiley and Sons. New York. 1980.

[8] Y. Iwahori, R.J. Woodham and A Bagheri "Principal Components Analysis and Neural Network Implementation of Photometric Stereo". In *Proceedings of the IEEE Workshop on Physics-Based Modeling in Computer Vision.* pp117-125 (1995).

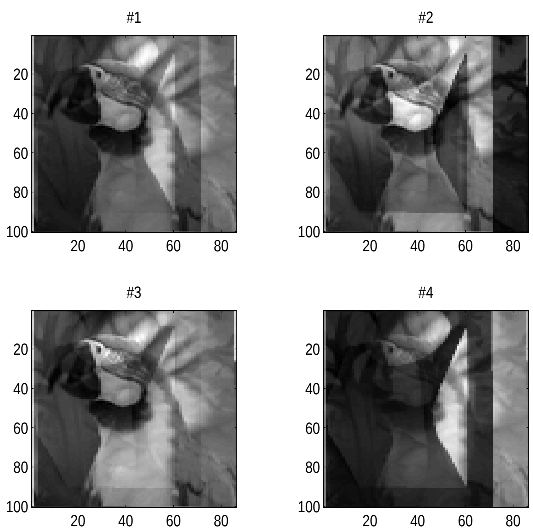
[9] H. Murase and S. Nayar. "Visual learning and recognition of 3-D objects from appearance". *Int. Journal of Computer Vision.* 14. pp 5-24. 1995.

[10] A. Shashua. **Geometry and Photometry in 3D Visual Recognition**. PhD Thesis. MIT. 1992.

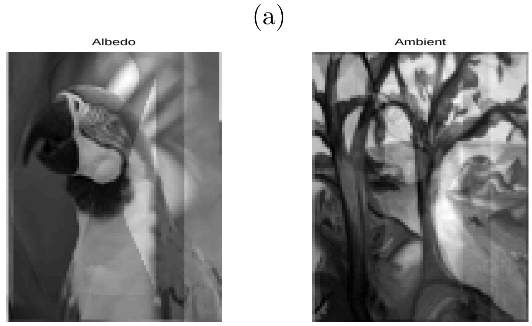
[11] W. Silver. *Determining Shape and Reflectance Using Multiple Images.* PhD Thesis. MIT, Cambridge, MA. 1980.

[12] R. Woodham. "Analyzing Images of Curved Surfaces". *Artificial Intelligence*, 17, pp 117-140. 1981.

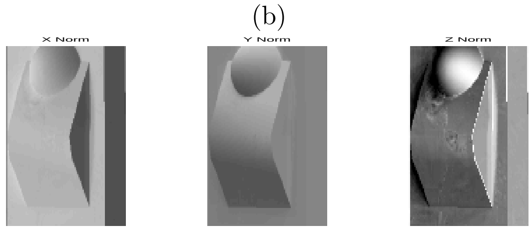
[13] A.L. Yuille. "Mathematical Results on SVD for Lighting and Warps". Technical Report. Smith-Kettlewell Eye Research Institute. San Francisco. California 94115.



(a)

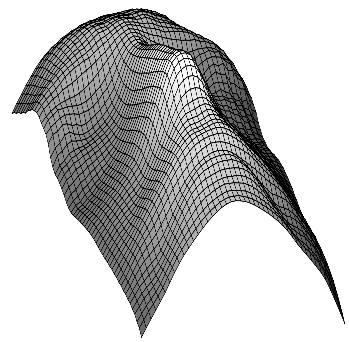


(b)

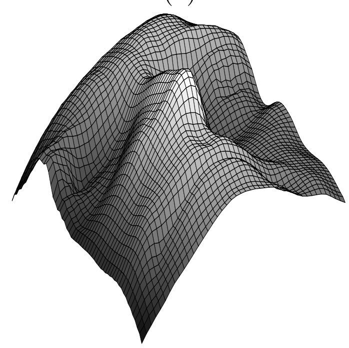


(c)

Figure 1: (a) shows four images of an object with the albedo of a parrot and ambient lighhting of a tree. (b) shows that our approach manages to dtermine the albedo (left) and the ambient illumination (right). Moreover, the x,y and z components of the normals of the object are also found correctly. This nice separation between albedo and ambient will start degrading if there are many shadows in the images.

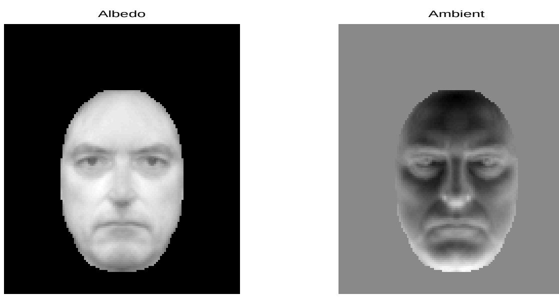


(a)

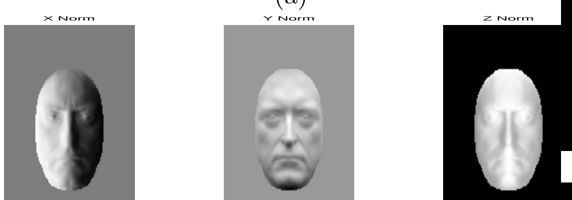


(b)

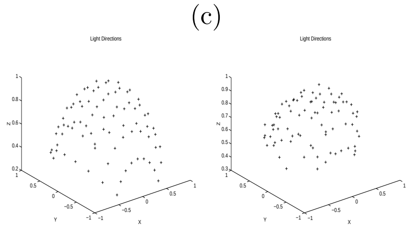
Figure 2: Figure (a) shows the face reconstructed up to an unknown GBR using integrability. Observe that the reconstruction is accurate. In figure (b) we introduced a GBR by hand to demonstrate the type of deformations which might arise.



(a)

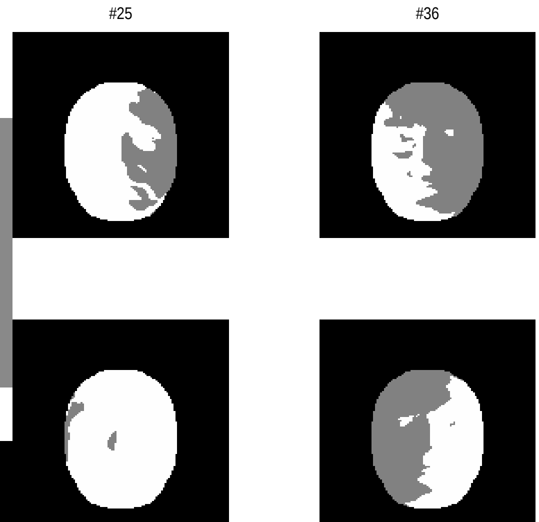


(c)

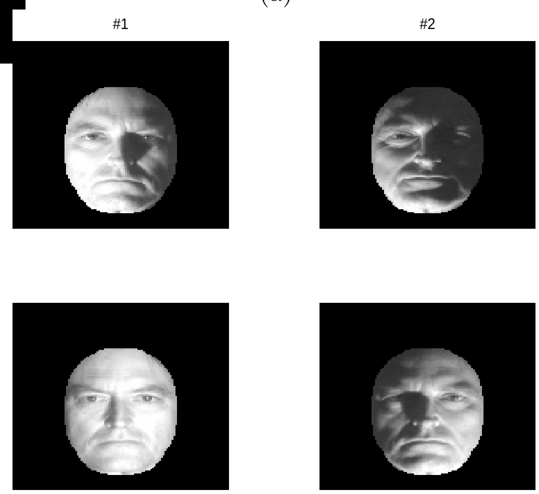


(b)

Figure 3: The full SVD on the face. (a) shows the estimated albedo and ambient term. (b) shows the true lighting positions (left) and the estimated light source positions (right). (c) shows the x,y and z components of the surface normals.

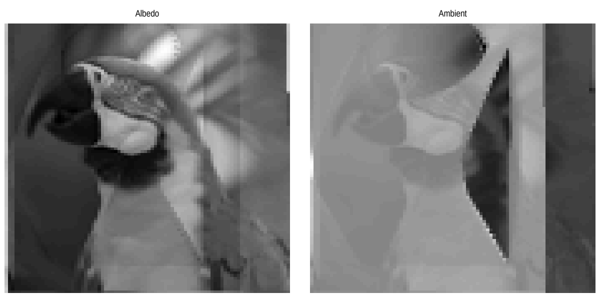


(a)

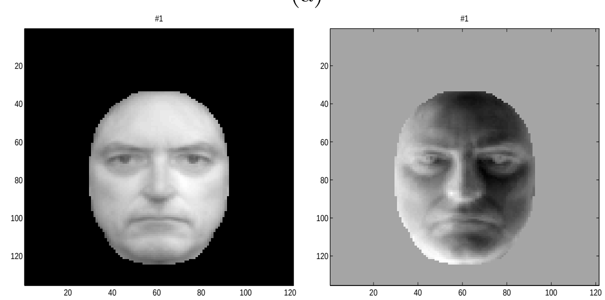


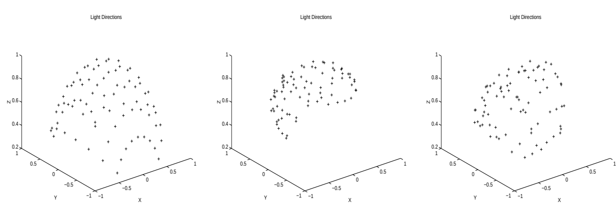
(b)

Figure 4: The top four images show the shadows extracted by our approach for the corresponding four input images at the bottom.



(a)





(b)

(b)

Figure 5: The albedo and ambient of the rarrot image and the face image are obtained using our shadow rejection technique. They are more accurate that those shown in figure 3. The bottom row shows the true lighting (left), the estimated lighting for the parrot (center) and the estimated lighting for the face (right). Again these are more accurate than in figures 3.