What’s In A Patch, II: Visualizing generic surfaces

Benjamin S. Kunsberg
Department of Applied Mathematics
Brown University

Daniel Niels Holtmann-Rice
Department of Computer Science
Yale University

Steven W. Zucker
Depts. of Computer Science and Biomedical Engineering
Yale University

May 18, 2017

Abstract

We continue the development of a linear algebraic framework for the shape-from-shading problem, exploiting the manner in which tensors arise when scalar (e.g. image) and vector (e.g. surface normal) fields are differentiated multiple times. In this paper we apply that framework to develop Taylor expansions of the normal field and build a boot-strapping algorithm to find these polynomial surface solutions (under any light source) consistent with a given patch to arbitrary order. A generic constraint on the image derivatives restricts these solutions to a 2-D subspace, plus an unknown rotation matrix. The parameters for the subspace and rotation matrix encapsulate the ambiguity in the shading problem.
1 Introduction

We have been developing a linear-algebraic approach to the shape-from-shading problem [7], and this paper is a continuation of that effort. The key intuition underlying the research effort is that, if one were to ‘drill down’ in derivatives for the surface, then this should correspond to analogous derivatives for the image (Fig. 1). We earlier considered which image (derivatives) are most likely given certain shape (normal) derivatives, for similar levels of differentiation. We now combine the derivatives in the Taylor sense, to calculate a representation of the full space of possible surface patches that could correspond to a given image patch, up to some order of differentiation. The algorithmic formulation allows us to ‘bootstrap’ the next-order structure from the previously calculated structure, similar to the use of recursion in series solutions to ordinary differential equations. It is important because, as shown in Fig. 3, the differential order of the image and of the surface must somehow be coupled. For convenience, we work in ambient Euclidean space; for a related abstract treatment using covariant derivatives, see [9].

In general there are infinite surfaces that can correspond to a given image patch, given the ill-posedness of the problem. The contribution here is that we are able to ‘control’ it with linear algebraic machinery [7] plus a generic lighting constraint [5]. Importantly, the notion of control that emerges is algorithmic, and indicates that, while the generic assumptions can structure solutions, even when they are linked to a particular (differential) level ambiguities remain. In particular, for a smooth Lambertian image patch (with no additional information regarding the light source or boundary), we develop a characterization of the set of possible underlying surface patches. Unlike previous approaches, we do not restrict the underlying surface patch to be a polynomial of a fixed degree, but rather allow the surface patch to account for all intensity variation in the image patch. Working with Taylor approximations, we let the order be dependent on each image patch and introduce an algorithm that realizes the set of generic surface patches corresponding to that image patch. In effect this provides a visualization of the ambiguity in the problem well beyond bas-relief [2]; see Fig. 1.

An illustration of the family of solutions is shown in Fig. 2 for a Lambertian image patch of a cylinder. This is, of course, a very special object

\footnote{We exploit the tensor structure relevant to relate image derivatives to normal field derivatives. The companion paper [7] contains two Appendices with appropriate background material.}
Figure 1: By Taylor’s theorem, increased derivatives at a point provide additional information within a neighborhood. We apply this idea to the shape inference problem, by noting that derivatives of the image $D^1 I$ can be related to derivatives on the surface $D^j N$, although the map is not one-to-one. In this paper we characterize the ambiguity in possible surface reconstructions even with a generic lighting assumption. The results further illustrate the rich interactions between linear algebra and differential geometry.
in which the surface normal variation is restricted to the radial direction and the curvature forms are low rank; see discussion in [7]. But it is also an important object in shape-from-shading research. Algorithmically it has been invoked to motivate a mean-curvature prior [1] and used to estimate human ‘reflectance functions’ [13]. Importantly, even in this special case, there is enormous variation in the perceived shape [14, 11, 10] and it plays a key role in light-source identification algorithms (e.g., [12]). Nevertheless, even among these many possibilities, neither the the bas-relief nor the generic lighting parameters, are invoked, so the variation is truly impressive. These concepts are developed in the course of this paper; we end with the algorithm that was used to compute these surfaces.

2 Background

A background review on shape-from-shading algorithms is provided in [17] and in the companion paper [7]. Here we concentrate on the few papers that are explicitly based on a patch model. For this approach the idea is to solve for local patches individually and then “stitch” them together [3, 15, 8].

In [4], the authors formulate the problem as solving a (large) system of polynomials using modern homotopy solvers over a triangulation. This is feasible for small images, involves (up to) quartic interactions, and leads to exact recovery of all possible solutions. For general patches, one can model the associated pixel values with various degrees of underlying surface complexity, represented as a Taylor polynomial in either heights or normals. [15] assume the image patch derives from a second-order surface and, therefore seek a quadratic solution; in [8] the image patch is modeled from a third-order surface. Assuming a local surface patch is exactly modeled by a quadratic, there are in general only four solutions to the local image formation model, i.e., the coefficients of the quadratic and the light source. If the image patch is large enough (i.e., number of pixels) relative to the number of coefficients of the Taylor polynomial, then the local patch can be determined up to a four-fold ambiguity [15, 8]. Clearly over-fitting can be a problem if the image patch is taken to be too large; e.g., errors will arise in fitting a quadratic surface to an image patch that arose from a quartic surface. In general, fixing the underlying surface complexity while considering successively larger image patches creates overfitting, whereas fixing the constraints (image patch size) while increasing the degree of the Taylor polynomial leads to increasing am-
Figure 2: A set of possible solutions to a cylindrical image patch. The rows depict variation in the generic solution space; the columns correspond to the bas-relief ambiguity (e.g. changing the attitude of the tangent plane and light source). The richness of these possible solutions, given that we will show they all are ‘generic,’ is well beyond what is normally expected [13]. The Taylor approximating polynomials $I$ and $N$ have degree 5. Generic parameters $(c_1, c_2)$ chosen by $c_1 = c_2$ varying linearly from 0 to 1.
Figure 3: A cartoon illustrating the coupling between image patch size and surface order. (a) As the patch size increases (here illustrated with additional image intensity normals), different intensities at different positions indicate different surface normals; as this brightness variation increases, very low-order surfaces become unlikely. (Brightness is the projection of the surface normal onto a (possibly unknown) light source vector.) (b) For a given (in this case, constant) image patch, different surfaces are possible if sampled appropriately.
biguity. See Figure 3. The variability in possible solutions is also biologically relevant [16], where researchers have attempted to identify the response to an image for a given surface. The variability that we describe must be kept in mind when attempting to assess the uniqueness (or lack thereof) in such neural responses.

3 Conditions on Generic Normal Fields Associated with a Lambertian Image Patch

Branching off from the companion paper [7], we now want to determine the set of possible underlying generic surface patches corresponding to a given image patch. We assume orthogonal projection, so the surface $S$ can be thought of as height field over the image plane. For more background on the setup and relevant notation, see [7].

The image patch $I(x, y)$ centered at the origin is modeled by a Taylor approximation $\bar{I}(x, y)$:

$$
\bar{I}(x, y) = I(0, 0) + x \mathcal{D}_p^1 I(e_1) + y \mathcal{D}_p^1 I(e_2) + \\
+ \frac{1}{2} \left( x^2 \mathcal{D}_p^2 I(e_1, e_1) + xy \mathcal{D}_p^2 I(e_1, e_2) + y^2 \mathcal{D}_p^2 I(e_2, e_2) \right) \\
+ \text{third and higher order terms}
$$

The $e_i$ represent the standard basis vectors. As in [7], we use the notation $\mathcal{D}_p^j I(\cdot, \cdot, \ldots, \cdot)$ to represent the $j^{\text{th}}$ derivative of the function $I(x, y)$ at the point $p$. It is a multilinear $j-$form that requires $j$ vector inputs to return a scalar value in $\mathbb{R}$ (See Appendix A in [7]). Thus, $\mathcal{D}_p^j I : \mathbb{R}^{2^j} \to \mathbb{R}$ for each $j$. Our goal is to understand the under-determined map from $\bar{I}(x, y)$ to a Taylor approximation of the surface normal field, $\bar{N}(x, y)$. $\bar{N}(x, y)$ is a multivariate polynomial with coefficients $\mathcal{D}_p^j N(\cdot, \cdot, \ldots, \cdot) : \mathbb{R}^{2^j} \to \mathbb{R}^3$.

Relating the two Taylor approximations can be understood by relating the coefficients. Thus, we seek an algorithm that takes the known values $\{\mathcal{D}_p^j I\}_{j=1}^n$ as inputs and and outputs the $\{\mathcal{D}_p^j N\}_{j=1}^n$. Building on the earlier results, this one to many map carries the ambiguity when going from the Taylor approximation of the image to the Taylor approximation of the surface. For notational simplicity, we will drop the $p$ subscript for the rest of the analysis.
A subtlety arises because the normal vector is unit length. First, represent the surface as a height function over the image plane: \( F(x,y) = \{x,y,f(x,y)\} \). Using subscripts to denote partial differentiation, the associated normal field is

\[
N(x,y) = \frac{1}{\sqrt{1 + f_x^2 + f_y^2}} \{-f_x, -f_y, 1\}
\] (2)

### 3.1 Image intensity gives a projection of the normal field

A Lambertian image intensity is given by \( I(x,y) = \alpha L \cdot N(x,y) \). Assuming constant albedo, we set \( \alpha = 1 \). We apply derivative operators \( j \) times to both sides to obtain:

\[
\mathcal{D}^j I(\cdot,\cdot,\ldots,\cdot) = L^T \mathcal{D}^j N(\cdot,\cdot,\ldots,\cdot)
\] (3)

To appreciate \( \mathcal{D}^j N \), think of e.g., \( \mathcal{D}^2 N \): this requires two \( \mathbb{R}^2 \) vectors \( \{v_1,v_2\} \) as input and outputs a vector in \( \mathbb{R}^3 \). This output is the “change in the change in the normal field” as we differentiate first in the \( v_1 \) direction and then in the \( v_2 \) direction. Note that, due to the symmetry of the derivative, the order of the inputs does not matter.

The square root term in the denominator creates difficulties in relating the \( \{\mathcal{D}^j I\}_{j=1}^n \) to the parameters \( \{f_x, f_y, f_{xx}, \ldots\} \) (or other surface parameters). According to the above equation (3), we see that the relationships between \( \{\mathcal{D}^j I\}_{j=1}^n \) and \( \{\mathcal{D}^j N\}_{j=1}^n \) is a projection along the (unknown) light source vector \( L \). We will need two more linearly independent projections in order to uniquely define the remainder of \( \{\mathcal{D}^j N\}_{j=1}^n \) and thus recover the normal field \( \bar{N}(x,y) \). As we now show, one of these additional projections will be set by the unit length condition of \( \bar{N}(x,y) \). The final projection can be freely set and represents the ambiguity in the problem.

### 3.2 Normalization constraints yield another projection of the normal field

The normalization constraint can be expanded as a system of linear constraints in the Taylor series. By enforcing this linear system of constraints, we can ensure an approximately unit length normal field (up to error \( O(x^{n+1}) \)) in the following manner.
The normalization constraint is:

\[ 1 = \langle \mathbf{N}(x, y), \mathbf{N}(x, y) \rangle \tag{4} \]

Here, we write \( \langle \cdot, \cdot \rangle \) as the standard dot product in \( \mathbb{R}^3 \) and we write \( \mathbf{N}_0 = \mathbf{N}(p) \) as the normal vector at the center of our patch. The above equation can be differentiated in an arbitrary image vector direction \( \mathbf{u} \) and evaluated at \( p \):

\[ 0 = \langle \mathcal{D}_u \mathbf{N}, \mathbf{N}_0 \rangle \tag{5} \]

Differentiate in another direction \( \mathbf{v} \):

\[ 0 = \langle \mathcal{D}_{vu} \mathbf{N}, \mathbf{N}_0 \rangle + \langle \mathcal{D}_v \mathbf{N}, \mathcal{D}_u \mathbf{N} \rangle \tag{6} \]

We do not choose \( \{ \mathbf{u}, \mathbf{v} \} \) before we differentiate; we could keep these directions unknown and general. That is, we consider the \( \mathcal{D}_j \mathbf{N} \) as a \((1, j)\) tensor – a linear machine seeking \( j \) vectors and outputting a vector in \( \mathbb{R}^3 \). We create \( \langle \mathcal{D}_j \mathbf{N}, \mathcal{D}_k \mathbf{N} \rangle \) as a new \((0, j + k)\) tensor in the following way:

1. Construct \( \mathcal{D}_j \mathbf{N} \otimes \mathcal{D}_k \mathbf{N} \) as a \((2, j + k)\) tensor. The two contravariant parts correspond to the \( \mathbb{R}^3 \) vectors \( \mathcal{D}_j \mathbf{N}, \mathcal{D}_k \mathbf{N} \) once \( j + k \) inputs have been chosen.

2. Lower an index associated with the unique contravariant part (in \( \mathbb{R}^3 \)) of \( \mathcal{D}_k \mathbf{N} \) to get a \((1, j + k + 1)\) tensor.

3. Contract the two indices associated with the \( \mathbb{R}^3 \) parts (we now have one covariant and one contravariant) to perform the dot product.

Thus,

\[ \langle \mathcal{D}_j \mathbf{N}, \mathcal{D}_k \mathbf{N} \rangle = C(\flat(\mathcal{D}_j \mathbf{N} \otimes \mathcal{D}_k \mathbf{N})) \tag{7} \]

where \( C \) is the contraction operator and \( \flat \) lowers the appropriate index.

Examining the Equations 5, 6, a pattern emerges: if \( \{ \mathcal{D}_j \mathbf{N} \}_{j<m} \) were known, then we could calculate \( \langle \mathcal{D}_m \mathbf{N}, \mathbf{N}_0 \rangle \). This key point will allow us to solve for \( \langle \mathcal{D}_m \mathbf{N}, \mathbf{N}_0 \rangle \) for each \( m \) inductively and thereby gain knowledge of the projection of the \( \mathcal{D}_j \mathbf{N} \) coefficients onto the central normal \( \mathbf{N}_0 \). Continuing to take derivatives and rearranging, we get the following proposition.
Proposition. The constraint $\langle \mathbf{N}, \mathbf{N} \rangle = 1$ can be Taylor approximated up to order $k$ by enforcing series of linear constraints,

$$\langle \mathcal{D}^j \mathbf{N}, \mathbf{N}_0 \rangle = - \sum_{1 \leq a \leq \left\lfloor \frac{j}{2} \right\rfloor} \sum_{\pi_a \in \Pi} \langle \mathcal{D}^a \mathbf{N}(\pi_a(V)), \mathcal{D}^{j-a} \mathbf{N}(\pi_a(V)^C) \rangle$$

for every $j \leq k$.

Here $\Pi$ is the set of combinations of $a$ objects chosen from $j$ objects. These combinations arise from the application of the product rule multiple times. We let $V$ represent the space $\mathbb{R}^{2j}$ of $j$ 2D image vector inputs into the tensor $\mathcal{D}^j \mathbf{N}$ and then $\pi_a(V)$ represents a subset of $a$ inputs out of the $j$ possible ones. Thus, $\pi_a(V)^C$ represents the remaining $j-a$ inputs. Note that since the order of the inputs doesn’t matter (but whether they get fed to $\mathcal{D}^a \mathbf{N}$ or $\mathcal{D}^{j-a} \mathbf{N}$ does), we use combinations.

In conclusion, if $\{\mathcal{D}^j \mathbf{N}\}_{j<m}$ were known, we could acquire the projections of $\mathcal{D}^m \mathbf{N}$ onto the vector $\mathbf{N}_0$. As $\mathcal{D}^m \mathbf{N}$ is fully defined when its projection onto three linearly independent vectors is known, it remains to search for one more projection. Unfortunately, there is no other information in the shape from shading problem that allows us to directly set a third projection of the $\mathcal{D}^j \mathbf{N}$ coefficients. This inherent ambiguity in the problem is due to the normal field being a higher dimensional entity (taking values on $\mathbb{S}^2$) than the intensity function (taking values on $\mathbb{R}$).

### 3.3 Using generic lighting to obtain a third projection

We now introduce a device that will allow us to calculate a third projection of $\mathcal{D}^j \mathbf{N}$. Let $G(x,y) : \Omega \rightarrow \mathbb{R}$ be any smooth function and let $b$ be any direction in $\mathbb{R}^3$ not in the span of $\{L, \mathbf{N}_0\}$. Suppose we choose to set $b^T \mathcal{D}^j \mathbf{N} = \mathcal{D}^j G, \forall j \in \{1, \ldots, n\}$. Provided we ensure $L^T \mathcal{D}^j \mathbf{N} = \mathcal{D}^j I, \forall j \in \{1, \ldots, n\}$ and equation 8 holds for each $j$, we will construct a Taylor series $\tilde{N}(x,y)$ that is approximately unit length and approximately matches the image. (For a discussion of these Taylor remainder errors, please see the Appendix.) Integrating this normal field would provide a surface patch matching the original image patch for any $G$; however, some choices of $G$ may be better (i.e. more robust and probable) than others.

A natural choice for $G$ comes from the generic framework and notation...
developed in [5, 6], which we now develop for our case. Define

$$\beta = [D^0N \quad D^1N \quad \ldots \quad D^nN]$$

(9)

$$Y = [D^0I \quad D^1I \quad \ldots \quad D^nI]$$

(10)

Here, we have unfolded (see Appendix in [7]) the various tensors $D^jN$ into $3 \times 2^j$ matrices and then appended them together. Let $m = \sum_{i=1}^{n} 2^i$. Then, $\beta$ is a linear map from $\mathbb{R}^m$ to $\mathbb{R}^3$ and $Y$ is a linear map from $\mathbb{R}^m$ to $\mathbb{R}$.

Now, consider the following rendering function:

$$f(L, \beta) = Y$$

(11)

$$= L^T \beta$$

(12)

Following [5, 6], we call $L$ the generic variable, $\beta$ the scene parameters, and $Y$ the observations. $L$ is an unknown vector in $\mathbb{R}^3$. Following [5, 6], we assume a Gaussian noise model on the observations $Y$:

$$Y = \hat{Y} + T$$

(13)

where $\hat{Y}$ is the ideal rendered observation and $T \sim N(0, \Sigma)$ with $\Sigma = \text{diag}(\sigma^2)$ for some $\sigma \in \mathbb{R}^+$. For the noise model, we have

$$P(Y|\beta, L) = \frac{1}{(\sqrt{2\pi}\sigma^2)^m} e^{-\frac{||Y-f(L, \beta)||^2}{2\sigma^2}}$$

(14)

Applying Bayes’ theorem and integrating over the generic variable $L$ yields the posterior distribution:

$$P(\beta|Y) = k \exp\left(-\frac{||Y-f(L_0, \beta)||^2}{2\sigma^2}\right) \frac{1}{\sqrt{\det(A)}}$$

(15)

$$= k \begin{cases} \text{fidelity} & \text{(prior probability)} & \text{(genericity)} \end{cases}$$

(16)

where $P_\beta(\beta), P_L(L_0)$ are prior distributions on the surface and light source parameters, $L_0$ is the light source that can best account for the observations given a chosen $\beta$ and $A$ is a matrix with the following elements:

$$A_{ij} = f_i' \cdot f_j' - (Y - f(L_0, \beta)) \cdot f_{ij}''$$

(17)
with

\[ f'_i = \left. \frac{\partial f(L, \beta)}{\partial l_i} \right|_{L=L_0} \quad (18) \]

\[ f''_{ij} = \left. \frac{\partial^2 f(L, \beta)}{\partial l_i \partial l_j} \right|_{L=L_0} \quad (19) \]

We now seek the solutions that maximize the posterior probability \( P(\beta | Y) \).

From equation (15), we maximize by choosing \( \beta \) and \( L_0 \) so that \( ||Y - f(L_0, \beta)|| = 0 \) while at the same time setting \( \det(A) = 0 \).

Where, \( ||Y - f(L_0, \beta)|| = 0 \), the condition that \( \det(A) = 0 \) is equivalent to the constraint that \( \beta \) is a low rank \( 3 \times m \) matrix. Under this condition, \( \beta \) is determined (up to two constants \( c_1, c_2 \)) by its projection onto two linearly independent vectors. (We ignore the rank 1 case, as it’s infinitesimally unlikely compared to the rank 2 case.) \( \beta \)'s projection onto two linearly independent vectors can already be obtained, as the components of \( \beta \) are each \( D_j N \). Thus, if we restrict \( \bar{N}(x, y) \) to the generic solutions, we can solve for it unambiguously up to the unknowns \( \{c_1, c_2, L_0, N_0\} \). We now show this.

### 3.4 Representing unknown lighting and tangent plane orientation via change of basis

As the projections of \( D_j N \) are onto the vectors \( \{L, N_0\} \), we will work in a basis defined by those vectors. Define \( l_t \) to be the unit length projection of \( L_0 \) onto the tangent plane perpendicular to \( N_0 \). That is, \( l_t = \frac{1}{\sqrt{1 - (L - (L \cdot N) N)}} \). Let \( b = N_0 \times l_t \). Then, define \( P \in SO_3(\mathbb{R}) \):

\[
P = \begin{pmatrix} N_0 & l_t & b \end{pmatrix}^T \quad (20)
\]

Note that \( P \) is an unknown orthogonal matrix, since we don’t know either the normal or the direction of the light source. However, rather than computing the Taylor surface \( \bar{N}(x, y) \) in the standard \( \mathbb{R}^3 \) basis, we will instead compute the modified Taylor surface \( P^T \bar{N}(x, y) \). This is merely considering the output of \( \bar{N}(x, y) \) in a different frame. In this fashion, we solve for a family of surfaces that will all match the Taylor image polynomial. To obtain a single member of that family, we choose an element \( Q \in SO_3(\mathbb{R}) \) and
Figure 4: A factorization to separate the unknown portion of the surface from its geometry. We precompose $P^T$ with the unfolded $D^j N$ tensors of size $3 \times 2^\otimes j$. This is equivalent to orthogonally changing the basis of the output space.

multiply to get $Q(P^T \tilde{N}(x, y))$. This is equivalent to choosing a normal and light source for the scene.

Thus our new goal is to solve for $P^T \tilde{N}(x, y)$ by solving for the coefficients of the Taylor series $P^T D^j N$, $1 \leq j \leq n$; see Fig. 4. We do this in an inductive manner, as we will need the $\{P^T D^j N\}_{j<k}$ in order to solve for $P^T D^k N$.

4 Algorithm

In this section, we will put the pieces described in the above sections together in order to create an inductive algorithm that can solve for all generic surfaces corresponding to a single image patch. To do this efficiently, we use an unfolding of the tensors $D^j I$ and $D^j N$.

We recall that $D^j N$ is a $(1, j)$ tensor, that is, a multilinear map from $\mathbb{R}^{2j}$ to $\mathbb{R}^3$. It has a matrix representation—a rank 1 unfolding—whose dimensions are $3 \times 2^j$. To calculate the action of this tensor on our inputs $\{v_1, v_2, \ldots, v_j\}, v_i \in \mathbb{R}^2$, we apply its matrix representation to the Kronecker product of the inputs $w = v_1 \otimes v_2 \ldots \otimes v_j$. We seek these matrix representations. Let $r_i^j$ stand for row $i$ of $P^T D^j N$. 
Figure 5: Similar to Figure 2, we show many solution surfaces to a given image. Across columns, we change the rotation matrix $P$, which amounts to changing the light source and central normal (e.g. bas-relief ambiguity). The Taylor approximating polynomials $\bar{I}$ and $\bar{N}$ have degree 5. Across rows, we change the values of the generic parameters $c_1, c_2$, chosen by $c_1 = c_2$ varying linearly from $-1$ to $1$. However, all solutions are considered ‘generic” according to Freeman’s definition [5].
Figure 6: Similar to Figure 2, we show many solution surfaces to a given image. Across columns, we change the rotation matrix $P$, which amounts to changing the light source and central normal (e.g. bas-relief ambiguity). The Taylor approximating polynomials $\bar{I}$ and $\bar{N}$ have degree 5. Across rows, we change the values of the generic parameters $c_1, c_2$, chosen by $c_1 = c_2$ varying linearly from $-1$ to $1$. However, all solutions are considered ‘generic” according to Freeman’s definition [5].
\[
P^{T}\mathcal{D}^{j}N = \begin{pmatrix}
    r_1^j \\
    r_2^j \\
    r_3^j
\end{pmatrix}
\]  \hspace{1cm} (21)

Suppose \( \{P^{T}\mathcal{D}^{j}N\}_{j<k} \) were known and the linear combination constants \( \{c_1, c_2\} \) were chosen. We will use an inductive algorithm to define the matrices \( P^{T}\mathcal{D}^{j}N \) consecutively in \( j \).

By equation 8, we can calculate \( r_{j+1}^1 \) by noting that an orthogonal transformation does not change inner products. Thus, the RHS of equation 8 can be calculated and \( r_{j+1}^1 \) can be set equal to it. Next, we define \( r_2^j = \mathbf{l}_t^T\mathcal{D}^jN \) in the following manner:

\[
r_2^j = \mathbf{l}_t^T\mathcal{D}^jN = \frac{1}{\sqrt{1-I^2}}(L - (L \cdot N)N)^T\mathcal{D}^jN
\]  \hspace{1cm} (22)

\[
= \frac{1}{\sqrt{1-I^2}}(L^T\mathcal{D}^jN - IN^T\mathcal{D}^jN)
\]  \hspace{1cm} (23)

\[
= \frac{1}{\sqrt{1-I^2}}(\mathcal{D}^jI - IR_1^j)
\]  \hspace{1cm} (24)

Thus, given \( r_1^j \) from the normalization constraints and \( \mathcal{D}^jI \) from the image information, we can find the next row \( r_2^j \) uniquely. Note that we are assuming that \( \mathbf{l}_t \) exists, which it will at every regular point. It remains to define \( r_3^j \) as the linear combination of the previous two rows \( r_3^j = c_1r_1^j + c_2r_2^j \).

Now, we have defined \( P^{T}\mathcal{D}^{j}N \) by its rows \( \{r_1^j, r_2^j, r_3^j\} \) and we continue on to \( P^{T}\mathcal{D}^{j+1}N \).

It remains to define the base case: \( P^{T}\mathcal{D}^{1}N \). From equation 5, we know its first row \( r_1^1 \) must be the 0 vector. From equation 25, we find that \( r_2^1 \) is just the weighted brightness gradient \( r_2^1 = \frac{\mathcal{D}^1I}{\sqrt{1-I^2}} \). Finally, due to the generic assumption, we set \( r_3^1 = c_2r_2^1 \).

**Proposition.** Following the algorithm described above (and summarized in Algorithm 1) yields coefficients \( \{\mathcal{D}^jN\}_1^m \) defining a multivariate Taylor polynomial \( \tilde{N}(x, y) \) that is generic according to \( \mathcal{G} \), matches exactly the image Taylor approximation \( \tilde{I}(x, y) \), and is unit length (up to error \( O(h^m) \)).

See the Appendix for precise details regarding the unit length error.
Algorithm 1 $\mathcal{D}^j\mathbf{N}$ Induction

1: Input: $\{\mathcal{D}^m\mathbf{I}\}_{m=0}^k$, $c_1, c_2$
2: For $0 \leq j \leq k - 1$
3: If $j = 0$
4: \[ r_1^j = \{0, 0\} \]
5: Else
6: \[ r_{j+1}^1 \leftarrow \text{calculated from } \{\mathbf{P}^T \mathcal{D}^k \mathbf{N}\}_{k<j} \text{ via Eqn (8)} \]
7: \[ r_{j+1}^2 \leftarrow \text{calculated from } \{r_{j+1}^1, \mathcal{D}^{j+1} \mathbf{I}\} \text{ via Eqn (25)} \]
8: \[ r_{j+1}^3 \leftarrow c_1 r_{j+1}^1 + c_2 r_{j+1}^2 \text{ using generic constants } c_1, c_2 \]
9: \[ \mathbf{P}^T \mathcal{D}^{j+1} \mathbf{N} \leftarrow \{r_{j+1}^1, r_{j+1}^2, r_{j+1}^3\} \]
10: End

5 Discussion

There have been many attempts to relate local image derivatives to local surface derivatives in Lambertian shading, but due to the $\sqrt{1 + f^2_x + f^2_y}$ in the denominator, the derivatives become more and more complex and the analysis soon becomes intractable. There have also been approaches towards representing the surface in a different way (principal directions basis, covariant derivatives, stereographic projections) but all tend to gain complexity as more derivatives are considered. In this work, we have described a method of representation that does not increase in analytic complexity as more derivatives are considered – this allows us to calculate all generic Taylor expansions (of any order) of a surface for a given image.

Essentially, we project each set of Taylor coefficients $\mathcal{D}^j\mathbf{N}$ onto an unknown plane (the “visible attribute plane”) defined by some unknown rotation matrix $\mathbf{P}$. This is a plane in $\mathbb{R}^3$ that is spanned by the light source $\mathbf{l}_t$ and the central normal $\mathbf{N}_0$. We know the projections as the $\langle \mathcal{D}^j\mathbf{N}, \mathbf{N}_0 \rangle$ is determined by the unit normal constraint. The $\langle \mathcal{D}^j\mathbf{N}, \mathbf{l}_t \rangle$ is a function of the image and $\langle \mathcal{D}^j\mathbf{N}, \mathbf{N}_0 \rangle$. We do not know this visible attribute plane but we can calculate these projections onto it from the image. Thus, any two normal fields with the same Taylor coefficients $\mathcal{D}^j\mathbf{N}$ will have equivalent projections onto this plane at each differential level $j$. The ambiguity stems from the fact that we cannot know the heights of these tensors above the plane (the projection of $\mathcal{D}^j\mathbf{N}$ along $\mathbf{b}$).

Two final remarks: First, it is possible to make the problem more well posed (less ambiguous) if information from the boundaries is included. Sec-
Figure 7: Illustration of method for generating the equivalence class of Taylor polynomial surfaces with equivalent images. Given a surface, we calculate its $\{D^j N\}_{j=1}^m$ tensors and project them onto the plane spanned by $\{N_0, l_i\}$ (the visible attribute plane). We then choose arbitrary (or generic) heights above each plane and generate a new set of $\{D^j N\}_{j=1}^m$, which corresponds to another surface with the same image Taylor approximation of order $m$. 

\[ D^1 N \quad D^2 N \quad D^3 N \]

\[ \downarrow \quad \downarrow \]

\[ \{ \text{Visible Attribute Plane} \} \quad \{ \text{Possible Reconstruction} \} \]

\[ \equiv \]

\[ \{ \text{Building Equivalence Classes of Image Patches} \} \]
ond, our algorithm can be modified to generate all Taylor polynomial normal fields \( \vec{N}(x, y) \) (instead of only the generic ones) for a given image patch by modifying the above algorithm: simply select \( r_3^j \) arbitrarily for each \( j \). That is, rather than selecting \( r_3^j \) as a linear combination of \( r_1^j \) and \( r_2^j \), \( r_3^j \) can be set equal to any \( v \in \mathbb{R}^2 \) (satisfying the appropriate symmetry according to differentiation, e.g. \( N_{xy} = N_{yx} \)). Thus, we see that the generic assumption significantly reduces the ambiguity of our Taylor expansion: we either must choose 2 generic constants or \( \sum_{j=1}^{m} j + 1 \) other values.

In summary, image patches are related to surface patches by some rendering function. Lambertian models with orthographic projection are the simplest such functions, and amount to projection of surface normals against a light source vector. This suggests a linear algebraic approach. We here take this one step further, for the inverse problem, by asking which surface patches could be consistent with a given image patch up to some number of derivatives. This allows identification of structure at different levels, although these ‘Taylor tensors’ can only be determined up to unknown generic constants \( c_1, c_2 \) and unknown pose \( P \in SO_3 \). Nevertheless, it leads to an algorithm for calculating every possible consistent normal field, even with a generic lighting assumption. Perhaps not surprisingly, such visualizations go well beyond the bas-relief ambiguity, even with generic constraints.
References

[1] Jonathan T Barron and Jitendra Malik. Shape, albedo, and illumination from a single image of an unknown object. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 334–341. IEEE, 2012.

[2] Peter N Belhumeur, David J Kriegman, and Alan L Yuille. The Bas-Relief Ambiguity. *International Journal Of Computer Vision*, 35(1):33–44, 1999.

[3] Pierre Breton, Lee A. Iverson, Michael S. Langer, and Steven W. Zucker. Shading flows and scenel bundles: A new approach to shape from shading. In *Proceedings of the Second European Conference on Computer Vision*, ECCV ’92, pages 135–150, London, UK, UK, 1992. Springer-Verlag.

[4] A. Ecker and A. D. Jepson. Polynomial shape from shading. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 145–152, June 2010.

[5] William T Freeman. The generic viewpoint assumption in a framework for visual perception. *Nature*, 368(6471):542–545, April 1994.

[6] William T Freeman. Exploiting the generic viewpoint assumption. *International Journal Of Computer Vision*, 20(3):243–261, 1996.

[7] Daniel Holtmann-Rice, Benjamin Kunsberg, and Steven W. Zucker. What’s In A Patch, I: Tensors, Differential Geometry and Statistical Shading Analysis. *ArXiv e-prints*, 2017.

[8] Benjamin Kunsberg and Steven W Zucker. How Shading Constrains Surface Patches without Knowledge of Light Sources. *SIAM Journal on Imaging Sciences*, 7(2):641–668, April 2014.

[9] Serge Lang. *Differential and Riemannian Manifolds*. Springer, New York, second edition, 1995.

[10] Pascal Mamassian and Daniel Kersten. Illumination, shading and the perception of local orientation. *Vision Research*, 36(15):2351 – 2367, 1996.
A Discussion of Taylor remainder errors

This paper documents an algorithm taking Taylor approximations to the image $\hat{I}$ to Taylor approximations to the normal field $\hat{N}$. Here we discuss the potential errors from using the Taylor approximations rather than the true values.

Recall the multivariate Taylor remainder formula with multi-index notation:

$$R_{a,k}(\mathbf{h}) = \sum_{|\alpha|=k+1} \partial^{\alpha} f(a + c\mathbf{h}) \frac{\mathbf{h}^\alpha}{\alpha!} \quad \text{for some } c \in (0, 1) \quad (26)$$
where \( a \) is the point of expansion, \( h \) is a vector in \( \mathbb{R}^2 \), \( k \) is the order of the Taylor expansion, and \( \alpha \) is a multi-index. We can use this equation to calculate errors for the two Taylor approximations used in the paper. We apply it to an image patch \( \Omega \subset \mathbb{R}^2 \), where we normalize so that \( \Omega = \{ a + h \mid ||h|| \leq 1 \} \). The error \( \delta = I - \bar{I} \) will be bounded proportional to the largest value of the \((k+1)\)th derivative of the image \( I \) in the image patch centered at \( a \):

\[
\delta a, k(h) = \sum_{|\alpha| = k+1} \partial^\alpha \delta(a + ch) \frac{h^\alpha}{\alpha!} \quad \text{for some } c \in (0, 1) \tag{27}
\]

\[
\leq \max_{v \in S^1, c \in (0, 1)} 2^k \left\| D^j I_{a+ch} \left( v \otimes 2^j \right) \right\|^2 \tag{28}
\]

where we have bounded each term in the sum by the largest value and bounded \( h^\alpha \) by 1.

Similarly, from Section 3.2, we enforce the unit length condition for the Taylor approximation \( \bar{N} \) via a sequence of linear constraints on the derivatives \( D^j \bar{N} \). This will result in an error corresponding to a deviation from unit length of the Taylor approximation \( \bar{N} \). To analyze this error, let \( \epsilon = \langle \bar{N}, \bar{N} \rangle - 1 \) and apply the above remainder formula. We see that, for a fixed Taylor order \( k \), the error \( \epsilon \) is bounded:

\[
\epsilon a, k(h) = \sum_{|\alpha| = k+1} \partial^\alpha \epsilon(a + ch) \frac{h^\alpha}{\alpha!} \quad \text{for some } c \in (0, 1) \tag{29}
\]

\[
\leq \max_{v \in S^1, c \in (0, 1)} 2^k \left\| D^j \bar{N}_{a+ch} \left( v \otimes 2^j \right) \right\|^2 \tag{30}
\]

Note: Although the error \( \delta = I - \bar{I} \) can be calculated before applying the algorithm, the error \( \epsilon = \langle \bar{N}, \bar{N} \rangle - 1 \) can only be calculated exactly from ground truth as it requires \( \bar{N} \). One must (using the described algorithm) first solve for \( \bar{N} \) and then verify that it is nearly norm 1. For the examples shown in the paper, errors \( \epsilon(h), \delta(h) \) were about 1%.