Lecture 10: The Lambertian Reflectance Model

A.L. Yuille

February 18, 2012

1 Introduction

This lecture describes the Lambertian model – this is the most basic model for generating images. It is a simple model and is often a good first order approximation. Other models are significantly more complicated and difficult to deal with. Risking an oversimplification – either use a Lambertian model or a statistical model or give up (for computer vision – other models are very useful for computer graphics).

2 The Lambertian Model

The linear Lambertian model is

$$I(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}, \tag{1}$$

where $I(\vec{x})$ is the image, $a(\vec{x})$ is the albedo, $\vec{n}(\vec{x})$ is the surface normal, $\vec{s}$ is the light source ($|\vec{s}|$ is the magnitude – the illumination strength – and $\vec{s}/|\vec{s}|$ is the light source direction.

This model is sometimes called the cosine rule because the image depends on the cosine between the surface normal and the light source direction.

The model is linear in the light source – i.e., $I(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}_1 + I(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}_2 = I(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot (\vec{s}_1 + \vec{s}_2)$.

Importantly, the image $I(\vec{x})$ does not depend directly on the viewpoint (changing the viewpoint corresponds to spatial warping of the intensity $I(\vec{x}) \mapsto I(\vec{\phi}(\vec{x}))$ where $\vec{\phi}(\vec{x})$ is a spatial warping function). This is unlike a specular object, like a mirror, where the image is a function of the viewpoint, the light source direction and the orientation of the mirror.

There is a well-known perceptual ambiguity – it is impossible to distinguish between a convex object lit from above and a concave object lit from below. Humans have a tendency to perceive objects to be convex. Dramatically, an inverted face mask appears to be a real face. This is despite having additional depth cues such as binocular stereo or structure from motion. We will discuss this and other ambiguities later in this lecture.

This linear model must be modified to account for shadows. There are two types of shadows: (i) attached shadows, and (ii) cast shadows. We can deal with attached shadows by modifying the equation to be:

$$I(\vec{x}) = \sum_{\mu} \min\{a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}_{\mu}, 0\}. \tag{2}$$

This is no longer linear in the light source directions. This makes analyzing the model much more difficult. Cast shadows can be modeled also – see figure (5).

Here is an example which illustrates the importance of cast and attached shadows. Consider the a point at the top of a mountain and a point at the bottom of a mineshaft. Both points have the same surface normal and albedo, so a Lambertain model will predict the same intensity for each (ignoring shadows). But typically the intensity at the top of the mountain will be much brighter than the point at the bottom of the mineshaft. The reason is that the point at the bottom of the mineshaft will be occluded from most of the light sources – i.e. only a light source that points directly down the mine will illuminate it. By contrast, the

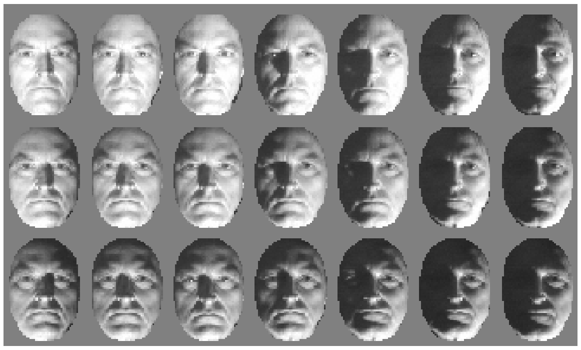


Figure 1: Examples of faces under different lighting conditions.

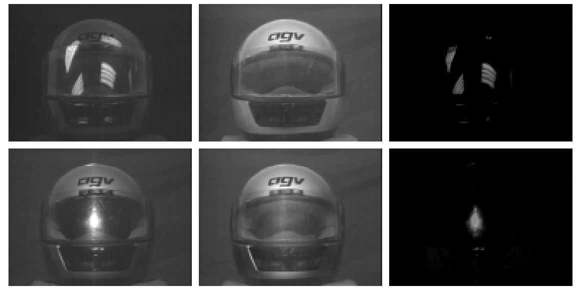


Figure 6: Top left is helmet under ambient lighting conditions. Top center is reconstruction from projection onto the first five eigenvectors. Note the specularities have been eliminated. Top right is the difference between original image and reconstruction. (Differences less than zero have been set to zero because the specularities are, by definition, positive.) The specularities are easily isolated. Bottom row shows the same effect for an image from the dense data set.

Figure 1: Exampleimages: Faces and Helmet (specular)

point on the top of the mountain will be illuminated from many directions and hence will usually be much brighter.

2.1 Other lighting models.

There are a range of other lighting models that have been used. These are particularly good for computer graphics where the goal is to generate an image knowing the positions and shapes of the objects and the positions of the light sources. Computer vision addresses the harder inverse problem of determining the objects and light sources from the image. Lambertian models are comparatively easier to invert – mainly because they are linear (if we ignore shadows) and the lighting is independent of the viewpoint direction. It is far harder to do this for the non-Lambertian models. Examples: (I) Radiosity Models. (II) Bidirectional Lighting Functions (BRDF's). (III) Specularity models.

2.2 Shape from Shading:

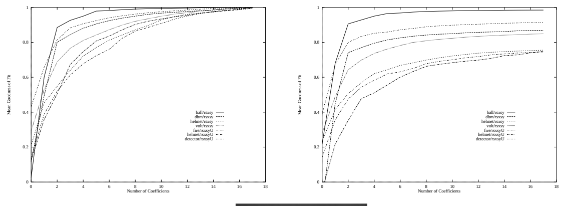
Humans have the ability to estimate the shape of an object from the image intensity – or shading pattern. There has been a lot of work on estimating shape from shading. But this is a severely ill-posed problem. Current methods make strong assumptions – e.g. constant albedo, known light source. These assumptions are typically not valid in many real world images. Classic shape from shading models assume that the light source directions are known and typically assume that the surface is locally smooth (Horn). More sophisticated methods were developed by Oliensis.

Statistical shape from shading (ref: Potetz and Lee) proceeds by learning the statistical relations between image intensity and depth/shape. This is learnt from a dataset obtained by using a laser-range finder and a camera. Their results suggest that Lambertian models rarely apply in natural images.

3 Experimental Analysis of Lambertian Models

The linear Lambertian model (i.e. ignoring shadows) implies that the image of an object lies in a three-dimensional space (Sha-ashua). This implies that the image can be modeled as $I(\vec{x}) = \sum_{i=1}^{3} \alpha_i e_i(\vec{x})$. Such a model can help for recognizing objects and for tracking an object when the lighting varies.

The linear model can be investigated empirically by taking photographs of an object from different lighting conditions, see figure (1). To get a set of images $\{I^\mu(\vec{x})\}$. We can then compute the correlation matrix $K(\vec{x}, \vec{x}') = \frac{1}{N} \sum_{\mu=1}^{N} I^\mu(\vec{x}) I^\mu(\vec{x}')$. Then calculate the eigenvectors and eigenvalues $\sum_{\vec{x}'} K(\vec{x}, \vec{x}') e(\vec{x}') = \lambda e(\vec{x})$.



Eigen-vector	ball	parrot	phone	face	helmet	helmet	helmet	fire ext	function generator	infrared detector
	sparse (right)	dense (both)	dense (right)	sparse (right)	sparse (right)	dense (right)	dense (both)	dense (right)	dense (right)	dense (right)
#1	0.482	0.428	0.679	0.537	0.320	0.340	0.388	0.536	0.806	0.624
#2	0.844	0.697	0.832	0.752	0.569	0.540	0.474	0.687	0.879	0.805
#3	0.944	0.763	0.882	0.902	0.651	0.628	0.581	0.765	0.922	0.885
#4	0.965	0.815	0.920	0.921	0.728	0.746	0.655	0.816	0.936	0.915
#5	0.979	0.847	0.941	0.935	0.798	0.772	0.722	0.852	0.948	0.927
#6	0.989	0.872	0.952	0.945	0.845	0.794	0.750	0.882	0.956	0.939
#7	0.991	0.885	0.963	0.953	0.881	0.816	0.795	0.901	0.962	0.948
#8	0.993	0.897	0.968	0.958	0.905	0.833	0.811	0.913	0.968	0.954
#9	0.995	0.907	0.972	0.963	0.924	0.848	0.824	0.925	0.972	0.960
#10	0.996	0.917	0.975	0.966	0.943	0.861	0.837	0.933	0.975	0.965

Figure 2: Eigengraph and Eigentable

Figure 3: Left Five panels: First five Eigenfaces. Right four panels from left to right: (i) original image, (ii) reconstruction using first three eigenvectors, (iii) reconstruction using first five eigenvectors, (iv) residual – the difference between the first and third panel. Observe that the errors are in "variable" parts of the image – e.g., strands of hair, shadows near the eyes and under the nose.

This analysis shows that the first five eigenvectors typically contain 90 percent of the energy (sum of all the eigenvalues), see figure (2). This plots $\sum_{i=1}^{n} \lambda_i / (\sum_{i=1}^{N} \lambda_i)$ as a function of n where N is the total number of eigenvalues.

The eigenvalues correspond (roughly) to images of the object lit from different directions, see figure (3). We can project the original image onto the eigenbasis to determines its best reconstruction and see which parts of the image are not described by the model, see figure (3).

Analysis by Basri and Jacobs and by Hanrahan and Ramamoothi gives a better fit to the data and explains why five components are needed for convex objects illuminated from light sources in the front.

4 Singular Value Decomposition (SVD)

To test the Lambertian model more for realistic objects we can try to estimate the albedo and shape from a set of images of the same object taken under different lighting conditions. We first make a change of notation and set $\vec{b}(\vec{x}) = a(\vec{x})\vec{n}(\vec{x})$ (we recover $a(\vec{x}) = |\vec{b}(\vec{x})|$ and $\vec{n}(\vec{x}) = \frac{\vec{b}(\vec{x})}{a(\vec{x})}$).

We can formulate this by a Gibbs distribution $P(\vec{b}|I) = (1/Z)\exp\{-E[\vec{b}]\}$, where:

$$E[\vec{b}] = \sum_{\mu=1}^{N} \sum_{\vec{x}} (I^{\mu}(\vec{x}) - \vec{b}(\vec{x}) \cdot \vec{s}^{\mu})^2. \tag{3}$$

This can be solved (up to an ambiguity described later) by linear algebra. Change notation so that $\mathbf{J} = \{J_{p\mu}\}$ is an $M \times N$ matrix (replacing $\vec{x}$ by p). Similarly replace $\vec{b}(\vec{x})$ by a matrix $\mathbf{B} = \{B_{pi}\}$ and $\mathbf{S} = \{S_{i\mu}\}$ ($i = 1, 2, 3$ labels the component of the vectors $\vec{b}$ and $\vec{s}$). Then the goal is to minimize the energy function:

$$E[B, S] = \sum_{up} \{J_{\mu p} - \sum_{i=1}^{3} B_{pi} S_{i\mu}\}^2. \tag{4}$$

The solution is obtained by decomposing J using SVD:

$$J = UDV^T, \quad \text{s.t. } UU^T = I, \quad VV^T = I, \quad D \text{ diagonal}. \tag{5}$$

Note that $JJ^T = UDV^TVDU^T = UD^2U^T$ and $J^TJ = VD^2V^T$. It follows that the columns of U and V are the eigenvectors of JJ^T and J^TJ respectively with eigenvalues being the diagonal elements of D^2 (D is diagonal so D^2 is also).

Let $e_k(p) : k = 1, 2, 3$ and $f_k(\mu) : k = 1, 2, 3$ be the first three columns of U and V respectively. Then the solutions are given by:

$$B_{pi} = \sum_{k=1}^{3} P_{ik} e_k(p), \quad S_{i\mu} = \sum_{k=1}^{3} Q_{ik} f_k(\mu), \quad \sum_{k=1}^{3} P_{ki} Q_{kj} = D_{ii}\delta_{ij} \ i = 1, 2, 3. \tag{6}$$

This result depends only on the first three elements of the diagonal matrix D. If we substitute this result back into the energy we obtain the result $\sum_{i=4}^{N} D_{ii}^2$. Hence this sum helps validate whether the images really do lie on a three-dimensional bilinear space. If they do, then $\sum_{i=4}^{N} D_{ii}^2 = 0$.

The solution given by equation (6) only estimates the variables – lighting, albedo and surface normal up to an ambiguity. This is because the matrices P and Q are only specified up to an arbitrary invertible matrix A – we can send $P \mapsto A^T P$ and $Q \mapsto A^{-1} Q$ and $P^T Q \mapsto P^T A A^{-1} Q = P^T Q$.

This ambiguity has nothing to do with the SVD approach. Instead it is inherent in the lambertian lighting model, even if we model shadows, as we will now describe.

5 Ambiguity in Lambertian Models

We now describe an invariant inherent in the Lambertian lighting model. This invariance occurs even if we include both cast and attached shadows. It relates, as we will discuss, to the ambiguity of viewing objects from different viewpoints.

Formulate the imaging equation as:

$$I(\vec{x}) = \sum_{\mu} \max\{\vec{b}(\vec{x}) \cdot \vec{s}_\mu, 0\}. \tag{7}$$

Observe that $\vec{b} \cdot \vec{s}_\mu$ is invariant to the transformation $\vec{b} \mapsto A^T \vec{b}$, $\vec{s}_\mu \mapsto A^{-1} \vec{s}_\mu$ (where A is any invertible matrix). This suggests that we can only estimate $\vec{b}(\vec{x})$ up to the group of transformations A which are invertible. But there is a constraint which reduces the ambiguity because the surface normal $\vec{n} = (n_1, n_2, n_3)$ must satisfy the *surface consistency constraint*. Any surface must obey the – apparently trivial – constraint that if you travel in a closed loop on the surface you end up with the same height that you started with (FIGURE!!). In differential form, this corresponds to the integrability condition:

$$\frac{\partial}{\partial x} \frac{n_2}{n_3} = \frac{\partial}{\partial y} \frac{n_1}{n_3}. \tag{8}$$

This can be obtained by the following argument. Represent the surface as $(x, y, z(x, y))$. By taking derivatives wrt x and y, we see that the vectors $(1, 0, z_x)$ and $(0, 1, z_y)$ must be tangent to the surface (z_x, z_y are the derivatives wrt x and y respectively). Hence the surface normal $\vec{n}$ can be obtained as the (normalized) cross product: $\vec{n} = \frac{1}{(1+z_x^2+z_y^2)^{1/2}}(-z_x, -z_y, 1)$. It follows that $n_1/n_3 = -z_x$ and $n_2/n_3 = -z_y$ and the integrability condition follows from the identity $z_{xy} = z_{yx}$.

The integrability condition is not consistent with the subgroup of invertible linear transformations. But it is consistent with a subgroup of three-dimensional linear transformations given by:

$$\begin{aligned} b_1(\mathbf{x}) &\mapsto \lambda b_1(\mathbf{x}) + \alpha b_3(\mathbf{x}) \\ b_2(\mathbf{x}) &\mapsto \lambda b_2(\mathbf{x}) + \beta b_3(\mathbf{x}) \\ b_3(\mathbf{x}) &\mapsto \tau b_3(\mathbf{x}) \end{aligned} \tag{9}$$

To get better understanding, it can be shown that this corresponds to a transformation on the surface of form:

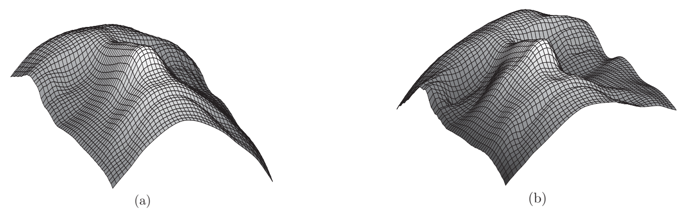


Figure 4: Left Panel: the correct reconstruction of a face. Right panel: a GBR transformation of a face.

$$z(x, y) \mapsto \lambda z(x, y) + \mu x + \nu y. \quad (10)$$

Here the first term $\lambda z(x, y)$ is a bas relief transformation and is known to sculptures (to save material by making sculptures with small changes in $z(x, y)$) and to Koenderink (see later). The other two terms correspond to adding a plane. The whole effect is called the Generalized Bas Relief ambiguity, see figure (4). It can be shown that cast shadows are also invariant to this transformation, see figure (5. (DETAILS).

This analysis has been extended to perspective projection (Belhumeur and Kriegman).

6 Resolving the Ambiguity

Shape from shading methods typically assume that the albedo is constant and the light source is known. This is often sufficient to yield a unique solution (see Oliensis).

Other ways make strong assumptions about the geometry of the object being viewed. For example, Atick developed a shape from shading method for faces which made use of a prior on the surface shade of faces – eigenheads, doing principal component analysis on three-dimensional depth maps of faces.

Other ways to resolve these ambiguities are to make assumptions about the light source directions or the albedo properties (for example, that albedo is piecewise smooth).

It is known that humans do not always estimate shape from shading correctly. Make-up can alter the albedo of a face and makes its shape appear different – for example, by enhancing cheek bones. More rigourous studies have been performed (Bulthoff, Koenderink, others) which show that there are biases in the perception of shape. In particular, Koenderink shows biases similar to the bas relief ambiguity (which relates to his theoretical studies on shape ambiguity).

7 Geometric and Lighting Invariances

First consider geometry alone.

Suppose we have a set of points $\{\vec{r}_i\}$ in three-dimensional space. If we transform then by an affine transformation $\vec{r}_i \mapsto A\vec{r}_i + \vec{a}$ (where A is an invertible matrix and $\vec{a}$ is a vector). Then their orthographic projection onto any two-dimension image plane is transformed by a two-dimensional affine transformation.

An affine transformation is a good approximation to the nonlinear perspective transformation (sometimes called weak perspective). This is only valid for a limited range of the parameters of the affine transform, but we ignore this for now.

Now consider the geometry and the lighting. This leads to the KGBR tranformation where the shape of the object is transformed by an affine transformation and the albedo of the object is changed similarly.

It can then be shown that for two objects related by a KGBR we can always find corresponding viewpoints and lighting so that the objects look identical. (Where orthographic projections related by two-dimensional affine transformations are considered to be identical).

The GBR transformation is obtained as a special case of KGBR. Where we make the additional requirements that the two object must appear to be the same for the same viewpoint (but different lighting for each object).

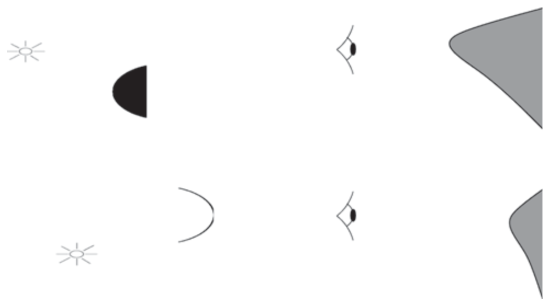


Fig. 1. Left panel: convex versus concave ambiguity. A convex object lit from above looks like a convex object lit from below. Right panel: the bas-relief ambiguity. The perception of shape is relatively insensitive to a linear scaling in the viewing direction.

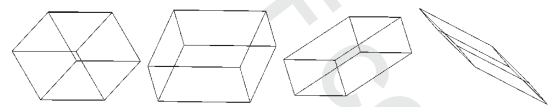


Fig. 2. Cube viewed from direction (0.51, 0.63, 0.58) (far-left panel) and the same cube undergoing affine transformations (remaining panels) seen from the same viewpoint.

Figure 5: KGBR1

Fig. 3. If the lighting conditions are unknown, then it is impossible to distinguish between two objects related by a GBR transform.[5] For any image of the first object, under one illumination condition, we can always find a corresponding illumination condition that makes the second object appear identical (i.e., we can generate an identical image). We show two objects under three different, but corresponding, lighting conditions.

the object.

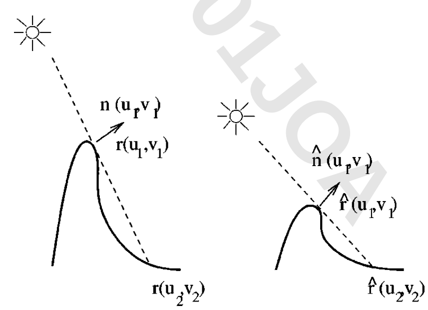


Fig. 5. The cast shadow boundaries, and hence the cast shadows, are preserved by the KGBR. Similar results were shown for the GBR.[5]

Figure 6: KGBR2

SVD can also be applied to estimating the three-dimensional shape (Kontsevich and Kontsevich, Tomasi and Kanade).