# Lambertain Models of Reflectance

- The Lambertian reflectance model is the simplest way to model how images are generated from three dimensional objects illuminated by various light sources. It does not model complex lighting effects, like specularity or inter-reflection. But it is a simple model which is a good approximation much of the time and is easy to analyze. Other models, e.g., BRDF models, are much harder to analyze.
- The Lambertian lighting model can be applied to shape from shading and photometric stereo.
- There are more complex models for generating realistic images, known as radiosity models (give a link to a paper). But these are hard to invert. From the computer vision perspective, complex lighting situations are probably best dealt with by learning methods.

# Linear Lambertian

- The linear Lambertian model is

$$I(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}, \tag{15}$$

  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., If $I_1(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}_1$ and $I_2(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}_2$ are the images of an object with albedo $a(\vec{x})$ and shape $\vec{n}(\vec{x})$ under lighting $\vec{s}_1$ and $\vec{s}_2$ respectively, then the image of the object when illuminated by both light sources is $I(\vec{x}) = a(\vec{x})\vec{n}(\vec{x}) \cdot (\vec{s}_1 + \vec{s}_2) = I_1(\vec{x}) + I_2(\vec{x})$.
- Most importantly, the image $I(\vec{x})$ does not depend directly on the viewpoint. resolve this ambiguity 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.

# Perceptual Ambiguity and Attached Shadows

- 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 resolve this ambiguity by tending to perceive objects to be convex. Surprisingly, an inverted face mask appears to be a real face. This remains true even of additional cues are present, e.g., 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 (where the light source is occluded). We can deal with attached shadows by modifying the equation to be:

$$I(\vec{x}) = \sum_{\mu} \min\{a(\vec{x})\vec{n}(\vec{x}) \cdot \vec{s}_{\mu}, 0\}. \tag{16}$$

- This is no longer linear in the light source directions. This makes analyzing the model much more difficult but still possible, see later in the lecture.
- To understand the importance of cast and attached shadows consider 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 Lambertian model will predict the same intensity for each (if cast shadows are ignored). But the intensity at the top of the mountain will be much brighter than the point at the bottom of the mineshaft because the bottom of the mineshaft will be occluded from most of the light sources –

# Perceptual Ambiguity and Attached Shadows

- 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.

# 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 (partly due to cast shadows but also atmospheric effects). In general, objects closer to the viewer tend to look brighter than objects further away.

# 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.. 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 $\sum_{\vec{x}'} K(\vec{x}, \vec{x}') e(\vec{x}') = \lambda e(\vec{x})$.
- This analysis shows that the first five eigenvectors typically contain 90 percent of the energy (sum of all the eigenvalues). 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.

# Experimental Analysis of Lambertian Models (2)

- The eigenvalues correspond (roughly) to images of the object lit from different directions. We can project the original image onto the eigenbases to determines its best reconstruction and see which parts of the image are not described by the model.
- Theoretical work – independently done in the computer vision community by Basri and Jacobs and in the computer graphics community by Ramamoothi and Hanrahan – showed that for convex objects nine eigenvectors captured much of the intensity variations and, if the lighting was restricted to come from the frontal hemisphere then only five were needed (in agreement with the experimental results).
- These theoretical studies were done by using spherical harmonics (fourier theory on a sphere) and showing that only a limited number were needed. (See later in this lecture for more details).

# The Lambertian Model and Singular Value Decomposition

- 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 can recover the albedo by $a(\vec{x}) = |\vec{b}(\vec{x})|$ and surface normal by $\vec{n}(\vec{x}) = \frac{\vec{b}(\vec{x})}{a(\vec{x})}$).
- We can formulate this by a Gibbs distribution $P(\vec{b}|I) = (1/Z)\exp\{-E[\vec{b}]\}$, where:

$$E[\vec{b}] = \sum_{\mu=1}^{N} \sum_{\vec{x}} (I^{\mu}(\vec{x}) - \vec{b}(\vec{x}) \cdot \vec{s}^{\mu})^2. \quad (17)$$

- This can be solved (up to an ambiguity described later) by linear algebra. We 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:

$$E[B, S] = \sum_{up} \{J_{\mu p} - \sum_{i=1}^{3} B_{pi} S_{i\mu}\}^2. \quad (18)$$

# The Lambertian Model and Singular Value Decomposition (2)

- The solution is obtained by decomposing $J$ using SVD:

$$J = UDV^T, \ \text{s.t.}\ UU^T = I, \ \ VV^T = I, \ \ D \text{ diagonal.} \tag{19}$$

- 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:

$$B_{pi} = \sum_{k=1}^{3} P_{ik} e_k(p), \ \ S_{i\mu} = \sum_{k=1}^{3} Q_{ik} f_k(\mu), \ \ \sum_{k=1}^{3} P_{ki} Q_{kj} = D_{ii}\delta_{ij} \ i = 1, 2, 3. \tag{20}$$

# The Lambertian Model and Singular Value Decomposition (3)

- 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 (20) 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.

# 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:

$$I(\vec{x}) = \sum_{\mu} \max\{\vec{b}(\vec{x}) \cdot \vec{s}_{\mu}, 0\}. \quad (21)$$

- 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. (This is similar to the ambiguity that arises when we apply SVD to solve structure from rigid motion).

# The Integrability Constraint

- 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.
- In differential form, this corresponds to the integrability condition:

$$\frac{\partial}{\partial x}\frac{n_2}{n_3} = \frac{\partial}{\partial y}\frac{n_1}{n_3}. \tag{22}$$

- 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 $z_{xy} = z_{yx}$.

# Ambiguity and the Integrability Constraint

- 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:

$$\begin{aligned} b_1(\mathbf{x}) &\mapsto \lambda b_1(\mathbf{x}) + \alpha b_3(\mathbf{x}) \\ b_2(\mathbf{x}) &\mapsto \lambda b_2(\mathbf{x}) + \beta b_3(\mathbf{x}) \\ b_3(\mathbf{x}) &\mapsto \tau b_3(\mathbf{x}) \end{aligned} \tag{23}$$

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

$$z(x, y) \mapsto \lambda z(x, y) + \mu x + \nu y. \tag{24}$$

- 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. It can be shown that cast shadows are also invariant to this transformation. This can also be extended to perspective projection.

# 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).

# 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).

# The Birectional Reflectance Distribution Function: Beyond Lambertian

- The BRDF is defined by the function $f(\theta_{in}, \phi_{in}; \theta_{out}, \phi_{out})$ which specifies how much light will be reflected in a direction $(\theta_{out}, \phi_{out})$ for a unit of incoming light in a direction $(\theta_{out}, \phi_{out})$.
- This can be expanded in a series of spherical harmonics, It can be shown that the Lambertian is the first term in the expansion. The higher order terms allow for more realistic reflectances and are used in the computer graphics community. But the Lambertian term is the only one where the lighting is independent of the viewpoint, which makes it much easier to analyze and to design algorithms which can exploit it.

# Harmonic Analysis: Lambertian Reflectance is Smooth

- We now describe the work by Basri and Jacobs – and by Ramamoothi and Hahrahan. This gives a way to deal with attached shadows for certain class of objects (those with convex geometry).
- The reflectance can be obtained by convolution in harmonic (spherical) space. $R(v) = \int_{S^2} k(u,v) I(u) du$, where $I(u)$ is the input light at spherical angle $u$, $k(u,v)$ is the amount reflected in direction $v$ per unit input from $u$, and $R(v)$ is the total output light at direction $v$.
- Spherical harmonics are orthogonal bases for functions defined on the sphere (similar to sinusoids on the plane). There is an equivalent convolution theorem (Funk-Hecke).
- Mathematically $Y_{nm}(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_{nm}(\cos\theta) \exp\{im\phi\}$.

  $P_{nm}(z) = \frac{(1-z^2)^{m/2}}{2^n n!} \frac{d^{n+m}}{dz^{n+m}} (z^2 - 1)^n$.
- The $n^{th}$ order harmonics have $2n+1$ components. Rotation corresponds to phase shift (same $n$, different $m$). In space coordinates these are polynomials of degree $n$.

# Harmonic Approximations

- We can express the lighting in terms of harmonics $l(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} l_{nm} Y_{nm}(\theta, \phi)$.
- The amount of lighting reflected is expressed (and approximated) by $r(\theta, \phi) = k * l \approx \sum_{n=0}^{2} \sum_{m=-n}^{n} k_n l_{nm} Y_{nm}(\theta, \phi)$.
- Keeping the first few terms – summing to $n = 2$ instead of to $\infty$ gives approximation accuracy 99.2%. This can be shown because the lambertian has the transformation kernel $k(\theta) = \max(\cos \theta, 0)$ which can be expressed as $\sum_{n=0}^{\infty} k_n Y_{n0}$ but the terms with $n = 0, 1, 2$ carry 99.2%. This has 9 basis images. If we use $N = 0, 1$ we have four basis images (ambient plus point source light) with accuracy 87.5%. (In practise due to self-occlusions almost 98% can be obtained with only 6 basis images (Ramamoorthi).
- Harmonic representations handle convex, lambertian objects with multiple light sources (including attached shadows). But do not model cast shadows and inter-reflections. They also do not represent specular objects.

# Applications of the harmonic analysis

- There have been applications to face recognition in unconstrained situations, to photometric stereo, and to 3D reconstruction.
- The basis functions can be found to be: (i) $n = 0$, $\rho(\vec{x})$, $n = 1$, $\rho(\vec{x})n_z(\vec{x})$, $\rho(\vec{x})n_x(\vec{x})$, $\rho(\vec{x})n_y(\vec{x})$, (iii) $n = 2$, $\rho(\vec{x})(2n_x^2(\vec{x}) - 1)$, $\rho(\vec{x})(n_x^2(\vec{x}) - n_y^2(\vec{x}))$, $\rho(\vec{x})n_x(\vec{x})n_y(\vec{x})$, $\rho(\vec{x})n_x(\vec{x})n_z(\vec{x})$, $\rho(\vec{x})n_y(\vec{x})n_z(\vec{x})$. Here $\rho(\vec{x})$ is the albedo, $\vec{n}(\vec{x}) = (n_x(\vec{x}), n_y(\vec{x}), n_z(\vec{x}))$ is the surface normal.
- Photometric stereo can be done as before. $M = LS$, where $M$ is the set of images, $L$ is the corresponding light sources and $S$ is the nine basis vectors (that represent the shape $\vec{n}(\vec{x})$ and the albedo $\rho(\vec{x})$. SVD recovers $L$ and $S$ up to an $r \times r$ ambiguity.
- Reconstruction with a prior (Kemelmacher and Basir). Reconstruction is impossible given a single image. But can be done with a prior over the depth $Z$ and the albedo $\rho$.