Three Source Photometric stereo:
Step 1

Offline:
Using source directions & BRDF, construct reflectance map for each light source direction. \( R_1(p,q), R_2(p,q), R_3(p,q) \)

Online:
1. Acquire three images with known light source directions. \( E_1(x,y), E_2(x,y), E_3(x,y) \)
2. For each pixel location \((x,y)\), find \((p,q)\) as the intersection of the three curves
   \[ R_1(p,q)=E_1(x,y) \]
   \[ R_2(p,q)=E_2(x,y) \]
   \[ R_3(p,q)=E_3(x,y) \]
3. This is the surface normal at pixel \((x,y)\). Over image, the normal field is estimated.

Reflectance Map of Lambertian Surface

What does the intensity (Irradiance) of one pixel in one image tell us?
(For example, let’s say the intensity is 0.5)

Recovering the surface \( f(x,y) \)

Many methods: Simplest approach
1. From estimate \( \mathbf{n} = (n_x, n_y, n_z) \), \( p = n_x/n_z \), \( q = n_y/n_z \)
2. Integrate \( p = df/dx \) along a row \((x,0)\) to get \( f(x,0) \)
3. Then integrate \( q = df/dy \) along each column starting with value of the first row
Lambertian Surface

At image location \((u,v)\), the intensity of a pixel \(x(u,v)\) is:

\[ e(u,v) = [a(u,v) \hat{n}(u,v)] \cdot [s_0 s] = b(u,v) \cdot s \]

where

- \(a(u,v)\) is the albedo of the surface projecting to \((u,v)\).
- \(n(u,v)\) is the direction of the surface normal.
- \(s_0\) is the light source intensity.
- \(s\) is the direction to the light source.

Lambertian Photometric stereo

- If the light sources \(s_1, s_2,\) and \(s_3\) are known, then we can recover \(b\) at each pixel from as few as three images. (Photometric Stereo: Silver 80, Woodham81).

\[ [e_1 e_2 e_3] = b^T [s_1 s_2 s_3] \]

i.e., we measure \(e_1, e_2,\) and \(e_3\) and we know \(s_1, s_2,\) and \(s_3\). We can then solve for \(b\) by solving a linear system.

\[ b^T = [e_1 e_2 e_3 s_1 s_2 s_3] \]

- Surface normal is: \(n = b/|b|\), albedo is: \(|b|\).

What is the set of images of an object under all possible lighting conditions?

In answering this question, we’ll arrive at a photometric stereo method for reconstructing surface shape w/ unknown lighting.

The Space of Images

- Consider an \(n\)-pixel image to be a point in an \(n\)-dimensional space, \(x \in \mathbb{R}^n\).
- Each pixel value is a coordinate of \(x\).
- Many results will apply to linear transformations of image space (e.g., filtered images).
- Other image representations (e.g., Cayley-Klein spaces, See Koenderink’s “pixel f#@king paper”) short.

Assumptions

For discussion, we assume:

- Lambertian reflectance functions.
- Objects have convex shape.
- Light sources at infinity.
- Orthographic projection.

- Note: many of these can be relaxed....
Model for Image Formation

Lambertian Assumption with shadowing:
\[ x = \max(Bs, 0) \]
where
- \( x \) is an \( n \)-pixel image vector
- \( B \) is a matrix whose rows are unit normals scaled by the albedos
- \( s \in \mathbb{R}^3 \) is vector of the light source direction scaled by intensity

3-D Linear subspace

The set of images of a Lambertian surface with no shadowing is a subset of 3-D linear subspace.

\[ \mathbb{L} = \{ x \mid x = Bs, \forall s \in \mathbb{R}^3 \} \]
where \( B \) is a \( n \) by 3 matrix whose rows are product of the surface normal and Lambertian albedo

How do you construct subspace?

With more than three images, perform least squares estimation of \( B \) using Singular Value Decomposition (SVD)

Rendering Images: \( \sum_i \max(Bs_i, 0) \)

Still Life

Original Images

Basis Images

3-D Linear subspace

The set of images of a Lambertian surface with no shadowing is a subset of 3-D linear subspace.

\[ \mathbb{L} = \{ x \mid x = Bs, \forall s \in \mathbb{R}^3 \} \]
where \( B \) is a \( n \) by 3 matrix whose rows are product of the surface normal and Lambertian albedo
Let \( L_i \) denote the intersection of \( L \) with an orthant \( i \) of \( \mathbb{R}^n \).

Let \( P_i(L_i) \) be the projection of \( L_i \) onto a "wall" of the positive orthant given by \( \max(x, 0) \).

Then, the set of images of an object produced by a single light source is:

\[
U = \bigcup_{i=0}^{M} P_i(L_i)
\]

With two lights on, resulting image along line segment between single source images:

- Superposition of images, non-negative lighting
- For all numbers of sources, and strengths, rest is convex hull of \( U \).

The image \( x \) produced by multiple light sources is

\[
x = \sum \max(B s_i, 0)
\]

where

- \( x \) is an \( n \)-pixel image vector.
- \( B \) is a matrix whose rows are unit normals scaled by the albedo.
- \( s_i \) is the direction and strength of the light source \( i \).

The set of images of any object in fixed posed, but under all lighting conditions, is a convex cone in the image space.

(Belhumeur and Kriegman, IJCV, 98)

Can the cones of two different objects intersect?

Can two different objects have the same cone?

How big is the cone?

How can cone be used for recognition?

Can two objects of differing shapes produce the same illumination cone?
**Do Ambiguities Exist? Yes**

- Cone is determined by linear subspace $L$
- The columns of $B$ span $L$
- For any $A \in \text{GL}(3)$, $B^* = BA$ also spans $L$.
- For any image of $B$ produced with light source $S$, the same image can be produced by lighting $B^*$ with $S^* = A^{-1}S$ because $X = B^*S^* = BAA^{-1}S = BS$
- When we estimate $B$ using SVD, the rows are NOT generally normal to albedo.

**Surface Integrability**

In general, $B^*$ does not have a corresponding surface.

Linear transformations of the surface normals in general do not produce an integrable normal field.

$B \xrightarrow{A} B^*$

**GBR Transformation**

Only Generalized Bas-Relief transformations satisfy the integrability constraint:

$$A = G^T = \begin{bmatrix} \lambda & 0 & -\mu \\ 0 & \lambda & -\nu \\ 0 & 0 & 1 \end{bmatrix}$$

**Generalized Bas-Relief Transformations**

Objects differing by a GBR have the same illumination cone.

Without knowledge of light source location, one can only recover surfaces up to GBR transformations.

**Uncalibrated photometric stereo**

1. Take $n$ images as input, perform SVD to compute $B^*$.
2. Find some $A$ such that $B^*A$ is close to integrable.
3. Integrate resulting gradient field to obtain height function $f^*(x,y)$.

Comments:
- $f^*(x,y)$ differs from $f(x,y)$ by a GBR.
- Can use specularities to resolve GBR for non-Lambertian surface.
GBR Preserves Shadows

Given a surface \( f \) and a GBR transformed surface \( f' \) then for every light source \( s \) which illuminates \( f \), there exists a light source \( s' \) which illuminates \( f' \) such that the attached and cast shadows are identical.

GBR is the only transform that preserves shadows.

[Criegman, Belhumeur 2001]

Codex Urbinas

As far as light and shade are concerned low relief fails both as sculpture and as painting, because the shadows correspond to the low nature of the relief, as for example in the shadows of foreshortened objects, which will not exhibit the depth of those in painting or in sculpture in the round.

Leonardo da Vinci
Treatise on Painting (Kemp)

The Illumination Cone

Thm: The span of the extreme rays of the illumination cone is equal to the number of distinct surface normals – i.e., as high as \( n \).
The number of extreme rays of the cone is \( n(n-1)+2 \)
(Belhumeur and Kriegman, IJCV '98)

Recent results

- Illumination cone is well capture by nine dimensions for a convex Lambertian surface.
  - Spherical Harmonic representation of lighting & BRDF.

Observation: The illumination cone is flat with most of its volume concentrated near a low-dimensional linear subspace.

<table>
<thead>
<tr>
<th>Source</th>
<th>Ball</th>
<th>Face</th>
<th>Phone</th>
<th>Parrot</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>48.2</td>
<td>53.7</td>
<td>67.9</td>
<td>42.8</td>
</tr>
<tr>
<td>#3</td>
<td>94.4</td>
<td>90.2</td>
<td>88.2</td>
<td>76.3</td>
</tr>
<tr>
<td>#5</td>
<td>97.9</td>
<td>93.5</td>
<td>94.1</td>
<td>84.7</td>
</tr>
<tr>
<td>#7</td>
<td>99.1</td>
<td>95.3</td>
<td>76.3</td>
<td>88.5</td>
</tr>
<tr>
<td>#9</td>
<td>99.5</td>
<td>96.3</td>
<td>97.2</td>
<td>90.7</td>
</tr>
</tbody>
</table>

Dimension:

[ Epstein, Hallinan, Yule 95]
Consider two objects with the same geometry, but different albedo patterns. Their 3-D linear subspaces can be expressed as:
\[ B_1 = R_1 N \]
\[ B_2 = R_2 N \]
where \( R_i \) is an \( n \) by \( n \) diagonal matrix of albedos (i.e. positive numbers).

Proposition: If \( e_1 \) is the illumination cone for an object defined by \( B_1 = R_1 N \) and \( e_2 \) is the cone for an object defined by \( B_2 = R_2 N \), then:
\[ e_1 = \{ x : x \in e_2 \} \]
\[ e_2 = \{ R_1^{-1} x : x \in e_1 \} \]

Where’d the moguls go?

• When a convex Lambertian surface is illumination by perfectly diffuse lighting, the resulting image is directly proportional to the albedo.
• For a convex object, the \( n \)-dimensional vector of albedos (and image) is contained within the object’s cone.
• For two objects with the same albedo pattern but different shape, their cones intersect in the interior.
• Two objects differing by a generalized bas relief transformation have the same cone.

Color Image Formation

A color image of a Lambertian surface without shadowing can be expressed as:
\[ x_i = \int \rho_i(\lambda) (R(\lambda) N)(s(\lambda) \hat{s}) d\lambda \]
where:
• \( \lambda \) is the wavelength.
• \( \rho_i(\lambda) \) is pixel response of the \( i \)-th color channel.
• \( R(\lambda) \) is an \( n \) by \( n \) diagonal matrix of the spectral reflectance of the facets.
• \( N \) is the \( n \) by \( 3 \) matrix of surface normals.
• \( s(\lambda) \) is the spectrum of the light source.
• \( \hat{s} \) is the direction of the light source.

Narrow Band Cameras

For a single channel with \( \rho(\lambda) \) modeled as a delta function at \( \lambda_i \), the image is:
\[ x_i = \rho(\lambda_i) (R(\lambda_i) N)(s(\lambda_i) \hat{s}) \]
\[ = \rho_i(\lambda) (R N)(s(\lambda) \hat{s}) \]

For \( c \) channels, the image \( x \in \textbf{R}^n \) is
\[ x = \left[ \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_c \end{array} \right] \]

Color Image Formation

A color image can be expressed as:
\[ x_i = \int \rho_i(\lambda) (R(\lambda) N)(s(\lambda) \hat{s}) d\lambda \]
where:
• \( \lambda \) is the wavelength.
• \( \rho_i(\lambda) \) is pixel response of the \( i \)-th color channel.
• \( R(\lambda) \) is an \( n \) by \( n \) diagonal matrix of the spectral reflectance of the facets.
• \( N \) is the \( n \) by \( 3 \) matrix of surface normals.
• \( s(\lambda) \) is the spectrum of the light source.
• \( \hat{s} \) is the direction of the light source.
Light Sources with Identical Spectra

Consider \( c \) color channels (not necessarily narrow band) and a scene illuminated by sources with identical spectra.

\[
x_i = \left( \int \rho_i(\lambda) s(\lambda) R(\lambda) d\lambda \right) s
\]

\[= R_i N_s
\]

For \( c \) channels, the image \( x \in \mathbb{R}^{cn} \) is

\[
x = [R_1 \mid R_2 \mid \ldots \mid R_c] N_s
\]

Major Issues

• Is Euclidean metric best choice for recognition?
• Should recognition metric be posed in image space?
• All lighting conditions are not equally likely - Can and should probabilities be attached to lighting?
• Does it make sense to pose recognition in image space when pose and within class variation are considered?

Illumination Cones: Recognition Method

• The set of images without shadowing lies in a 3-D linear subspace of \( \mathbb{R}^{cn} \).
• The illumination cone spans \( m \) dimensions of \( \mathbb{R}^{cn} \).
• By restricting the spectra of the light source, the cone is much smaller (\( m \) dimensions vs. \( cm \) dimensions), and the camera does not have to be narrow band.

Generating the Illumination Cone

Original (Training) Images

Surface \( f(x,y) \)
(albedo textured mapped on surface).

3D linear subspace

\( N_s \)
(albedos)
(surface normals)

Cone - Attached

Cone - Cast

[Georghiades, Belhumeur, Kriegman 01]

Predicting Lighting Variation

Single Light Source
**Yale Face Database B**

- 64 Lighting Conditions
- 9 Poses
- => 576 Images per Person

**Face Recognition: Test Subsets**

- Subset 1: 0-12°
- Subset 2: 12-25°
- Subset 3: 25-50°
- Subset 4: 50-77°

Test images divided into 4 subsets depending on illumination.

**Face Recognition: Lighting & Pose**

1. Union of linear subspaces
   1. Sample the pose space, and for each pose construct illumination cone.
   2. Cone can be approximated by linear subspace (used 11-D)
2. For computational efficiency, project using PCA to 100-D

**Geodesic Dome Database - Frontal Pose**

- [Georghiades, Belhumeur, Kriegman 01]

**Illumination Variability Reveals Shape**

**Pose Variability in test set: Up to 24°**
Illumination & Image Set

- Lack of illumination invariants
  [Chen, Jacobs, Belhumeur 98]
- Set of images of Lambertian surface w/o shadowing is 3-D linear subspace
  [Moser 93], [Nayar, Murase 96], [Shashua 97]
- Empirical evidence that set of images of object is well-approximated by a low-dimensional linear subspace
  [Hallinan 94], [Epstein, Hallinan, Yuille 95]
- Illumination cones
  [Belhumeur, Kriegman 98]
- Spherical harmonics lighting & images
  [Basri, Jacobs 01], [Ramamoorthi, Hanrahan 01]
- Analytic PCA of image over lighting
  [Ramamoorthi 02]

Some subsequent work

1. "Face Recognition Under Variable Lighting using Harmonic Image Exemplars," Zhang, Samaras, CVPR03
2. "Clustering Appearances of Objects Under Varying Illumination Conditions," Ho, Lee, Lim, Kriegman, CVPR 03
3. "Low-Dimensional Representations of Shaded Surfaces under Varying Illumination," Nillius, Eklundh, CVPR03
4. "Using Specularities for Recognition," Osadchy, Jacobs, Ramamoorthi, ICCV 03