Illumination Cones and Uncalibrated Photometric Stereo

Computer Vision I
CSE252A
Lecture 9

Photometric stereo

- Single viewpoint, multiple images under different lighting.
  1. Arbitrary known BRDF, known lighting
  2. Lambertian BRDF, known lighting
  3. Lambertian BRDF, unknown lighting.

Announcement

- Reading on web page
- HW2 announced

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 (luminance) of one pixel in one image tell us?
(for example, let’s say the intensity is 0.5)

One viewpoint, two images, two light sources
Two super imposed reflectance maps

What does the intensity (luminance) 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 $n = (n_x, n_y, n_z)$, $p = n_x/n_z$, $q = n_y/n_z$
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:

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

where
- $a(u,v)$ is the albedo of the surface projecting to $(u,v)$.
- $\hat{n}(u,v)$ 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, Woodham 81).

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

- 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.

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

- Surface normal is: $n = b/|b|$, albedo is: $|b|$

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”)

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....
Lambertian Surface

At image location \((u,v)\), the intensity of a pixel \(x(u,v)\) is:
\[
x(u,v) = [a(u,v)n(u,v)]^T [s_0 s] = b(u,v)s^T
\]
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.

Model for Image Formation

Lambertian Assumption with shadowing:
\[
x = \max(B s, 0)
\]
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.

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

Rendering Images: \(\Sigma_i \max(B s_i, 0)\)

- 1 Light
- 2 Lights
- 3 Lights

How do you construct subspace?

- Any three images w/o shadows taken under different lighting span \(L\)
- Not orthogonal
- Orthogonalize with Gram-Schmidt
Matrix Decompositions

- Definition: The factorization of a matrix \( M \) into two or more matrices \( M_1, M_2, ..., M_r \), such that \( M = M_1 M_2 ... M_r \).
- Many decompositions exist...
  - QR Decomposition
  - LU Decomposition
  - LDU Decomposition
  - Etc.

SVD Properties

- In Matlab, \([u v] = svd(A)\), and you can verify that: \( A = u \Sigma v' \)
- \( r = \text{Rank}(A) \) = # of non-zero singular values.
- \( U, V \) give us orthonormal bases for the subspaces of \( A \):
  - 1st \( r \) columns of \( U \): Column space of \( A \)
  - Last \( n - r \) columns of \( U \): Left nullspace of \( A \)
  - 1st \( r \) columns of \( V \): Row space of \( A \)
  - Last \( n - r \) columns of \( V \): Nullspace of \( A \)
- For \( d \leq r \), the first \( d \) columns of \( U \) provide the best \( d \)-dimensional basis for columns of \( A \) in least squares sense.

Estimating \( B \) with SVD

1. Construct data matrix
   \[
   D = [ x_1 \ x_2 \ x_3 \ ... \ x_n ]
   \]
2. \([ U \ \Sigma \ V ] = svds(D)\)
   - If data had no noise, then rank(\( D \))=3, and the first three singular values (\( S \)) would be positive and rest would be zero.
   - Take first three column of \( U \) as \( B \).
Face Basis

Original Images

Basis Images

Movie with Attached Shadows

Single Light Source

Face Movie

3-D Linear subspace

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

\[ L = \{ x \mid x = Bs, \forall s \in \mathbb{R}^3 \} \]

where \( B \) is an \( n \) by 3 matrix whose rows are product of the surface normal and Lambertian albedo

\[ \text{N-dimensional Image Space} \]

Set of Images from a Single Light Source

\[ \text{Let } L_i \text{ denote the intersection of } L \text{ with an orthant } i \text{ of } \mathbb{R}^n. \]

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

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

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

Set of Images from Multiple Light Sources

\[ \text{With two lights on, resulting image along line segment between single source images: superposition of images, non-negative lighting} \]

\[ \text{For all numbers of sources, and strengths, rest is convex hull of } U. \]
The Illumination Cone

Theorem: 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)

Some natural ideas & questions

- 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?

Do Ambiguities Exist?

Can two objects of differing shapes produce the same illumination cone?

YES

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 * 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.

GBR Transformation

Only Generalized Bas-Relief transformations satisfy the integrability constraint:

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

$$f(x, y) \rightarrow \tilde{f}(x, y) = \lambda f(x, y) + \mu x + \nu y$$
**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.

**What about cast shadows for nonconvex objects?**

PP Reubens in Opticorum Libri Sex, 1613

**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.

[Krigman, Belhumeur 2001]

**Bas-Relief Sculpture**

**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)
Some natural ideas & questions

- 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?

Where’d the moguls go?

- When a convex Lambertian surface is illuminated 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.

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
(Bellman and Kriegman, ICV’98)

Shape of the Illumination Cone

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

<table>
<thead>
<tr>
<th>#</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>99.3</td>
<td>95.3</td>
<td>96.3</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>

[Epstein, Hallinan, Yuille 95]  
Dimension:
Subsequent results

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