Lecture 17: Photometric Stereo

COMP 590/776: Computer Vision
Instructor: Soumyadip (Roni) Sengupta
TA: Mykhailo (Misha) Shvets

Course Website: Scan Me!
Can we determine shape from lighting?

- Are these spheres?
  - Or just flat discs painted with varying color (albedo)?
  - There is ambiguity between *shading* and *reflectance*
  - But still, as humans we can understand the shapes of these objects
What we know: Stereo

Key Idea: use camera motion to compute shape
Next: Photometric Stereo

Key Idea: use pixel brightness to understand shape
Photometric Stereo

What results can you get?

Input (1 of 12)  Normals (RGB colormap)  Normals (vectors)  Shaded 3D rendering  Textured 3D rendering
Today’s class

• Measuring Light (recap)
• Image formation with shape, reflectance, and illumination
• Shape from Shading
• Photometric Stereo
• Uncalibrated Photometric Stereo
• Generalized Bas-Relief Ambiguity
• Photometric Stereo in ‘deep learning era’.
Today’s class

• Measuring Light (recap)
  • Image formation with shape, reflectance, and illumination
  • Shape from Shading
  • Photometric Stereo
  • Uncalibrated Photometric Stereo
  • Generalized Bas-Relief Ambiguity
  • Photometric Stereo in ‘deep learning era’.
Radiometry

• What determines the brightness of a pixel?
Radiometry

• What determines the brightness of a pixel?
Radiometry

- What determines the brightness of a pixel?
What is light?

Electromagnetic radiation (EMR) moving along rays in space
- \( R(\lambda) \) is EMR, measured in units of power (watts)
  - \( \lambda \) is wavelength

Light field
- We can describe all of the light in the scene by specifying the radiation (or “radiance” along all light rays) arriving at every point in space and from every direction

The **plenoptic function** describes all of this light: \[ R(X, Y, Z, \theta, \phi, \lambda, t) \]
Visible light

We “see” electromagnetic radiation in a range of wavelengths
Light transport

Illumination → Reflectance → Perception

[Diagram of pears and human eye]
Light sources

• Basic types
  • point source
  • directional source
    • a point source that is infinitely far away
  • area source
    • a union of point sources

• More generally
  • a light field can describe *any* distribution of light sources
  • Environment map

• What happens when light hits an object?
Today’s class

• Measuring Light (recap)
• Image formation with shape, reflectance, and illumination
• Shape from Shading
• Photometric Stereo
• Uncalibrated Photometric Stereo
• Generalized Bas-Relief Ambiguity
• Photometric Stereo in ‘deep learning era’.
Modeling Image Formation

We need to reason about:

• How light interacts with the scene
• How a pixel value is related to light energy in the world

Track a “ray” of light all the way from light source to the sensor
Directional Lighting

- Key property: all rays are parallel
- Equivalent to an infinitely distant point source
Lambertian Reflectance

Image intensity \( I \) is equal to the dot product of the surface normal \( N \) and the light direction \( L \):

\[
I = N \cdot L
\]

Image intensity \( I \) is proportional to \( \cos(\text{angle between } N \text{ and } L) \).
Materials - Three Forms

Ideal diffuse (Lambertian)

Ideal specular

Directional diffuse
Reflectance—Three Forms

Ideal diffuse (Lambertian)

Ideal specular

Directional diffuse
Ideal Diffuse Reflection

- Characteristic of multiple scattering materials
- An idealization but reasonable for matte surfaces
Lambertian Reflectance

1. Reflected energy is proportional to cosine of angle between L and N (incoming)

2. Measured intensity is viewpoint-independent (outgoing)

\[ I = N \cdot L \]
Final Lambertian image formation model

1. **Diffuse albedo**: what fraction of incoming light is reflected?
   - Introduce scale factor $k_d$

2. **Light intensity**: how much light is arriving?
   - Compensate with camera exposure (global scale factor)

3. **Camera response function**
   - Assume pixel value is linearly proportional to incoming energy (perform radiometric calibration if not)

\[ I = k_d \mathbf{N} \cdot \mathbf{L} \]
Albedo

Objects can have varying albedo and albedo varies with wavelength

Source: [https://en.wikipedia.org/wiki/Albedo](https://en.wikipedia.org/wiki/Albedo)

<table>
<thead>
<tr>
<th>Surface</th>
<th>Typical albedo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fresh asphalt</td>
<td>0.04[^4]</td>
</tr>
<tr>
<td>Open ocean</td>
<td>0.06[^5]</td>
</tr>
<tr>
<td>Worn asphalt</td>
<td>0.12[^4]</td>
</tr>
<tr>
<td>Conifer forest (Summer)</td>
<td>0.08,[^6] 0.09 to 0.15[^7]</td>
</tr>
<tr>
<td>Deciduous trees</td>
<td>0.15 to 0.18[^7]</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.17[^8]</td>
</tr>
<tr>
<td>Green grass</td>
<td>0.25[^8]</td>
</tr>
<tr>
<td>Desert sand</td>
<td>0.40[^9]</td>
</tr>
<tr>
<td>New concrete</td>
<td>0.55[^8]</td>
</tr>
<tr>
<td>Ocean ice</td>
<td>0.5–0.7[^8]</td>
</tr>
<tr>
<td>Fresh snow</td>
<td>0.80–0.90[^8]</td>
</tr>
</tbody>
</table>
Today’s class

- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.
Human Perception
Examples of the classic bump/dent stimuli used to test lighting assumptions when judging shape from shading, with shading orientations (a) 0° and (b) 180° from the vertical.

Our brain often perceives shape from shading.

Mostly, it makes many assumptions to do so.

For example:

Light is coming from above (sun).

Biased by occluding contours.
A Single Image: Shape from shading

Suppose (for now) \( k_d = 1 \)

\[
I = k_d N \cdot L = N \cdot L = \cos \theta_i
\]

You can directly measure angle between normal and light source

- Not quite enough information to compute surface shape
- But can be if you add some additional info, for example
  - assume a few of the normals are known (e.g., along silhouette)
  - constraints on neighboring normals—“integrability”
  - smoothness
- Hard to get it to work well in practice
  - plus, how many real objects have constant albedo?
  - But, deep learning can help
Deep Learning for Shape from Shading

InverseRenderNet: Learning single image inverse rendering

Ye Yu and William A. P. Smith
Department of Computer Science, University of York, UK
{yy1571, william.smith}@york.ac.uk

Figure 1: From a single image (col. 1), we estimate albedo and normal maps and illumination (col. 2-4); comparison multi-view stereo result from several hundred images (col. 5); re-rendering of our shape with frontal/estimated lighting (col. 6-7).
Application: Detecting composite photos

Fake photo

Real photo
Today’s class

• Measuring Light (recap)
• Image formation with shape, reflectance, and illumination
• Shape from Shading
• **Photometric Stereo**
  • Uncalibrated Photometric Stereo
  • Generalized Bas-Relief Ambiguity
  • Photometric Stereo in ‘deep learning era’.
Photometric stereo

Can write this as a matrix equation:

\[
\begin{bmatrix}
I_1 \\
I_2 \\
I_3
\end{bmatrix}
= k_d \begin{bmatrix}
L_1^T \\
L_2^T \\
L_3^T
\end{bmatrix} N
\]

\[
I_1 = k_d N \cdot L_1 \\
I_2 = k_d N \cdot L_2 \\
I_3 = k_d N \cdot L_3
\]
Solving the equations

\[
\begin{bmatrix}
I_1 \\
I_2 \\
I_3
\end{bmatrix}
= 
\begin{bmatrix}
L_1^T \\
L_2^T \\
L_3^T
\end{bmatrix}
k_d N
\]

\[
G = L^{-1} I
\]

\[
k_d = \|G\|
\]

\[
N = \frac{1}{k_d} G
\]

Solve one such linear system per pixel to solve for that pixel’s surface normal
More than three lights

Can get better results by using more than 3 lights

\[
\begin{bmatrix}
I_1 \\
\vdots \\
I_n
\end{bmatrix}_{nx3} =
\begin{bmatrix}
L_1 \\
\vdots \\
L_n
\end{bmatrix}_{nx3} k_d N_{3x1}
\]

Least squares solution:

\[
I = LG
\]

\[
L^T I = L^T L G
\]

\[
G = (L^T L)^{-1}(L^T I)
\]

Solve for N, k_d as before
Calibrating Lighting Directions

Trick: place a chrome sphere in the scene

- the location of the highlight tells you where the light source is
Example

Forsyth & Ponce, Sec. 5.4

Input views

Recovered albedo

Recovered normal field
Depth from normals

- Solving the linear system per-pixel gives us an estimated surface normal for each pixel

- How can we compute depth from normals?
  - Normals are like the “derivative” of the true depth
Depth from normals

Get a similar equation for $V_2$

- Each normal gives us two linear constraints on $z$
- compute $z$ values by solving a matrix equation

\[
V_1 = (x + 1, y, z_{x+1,y}) - (x, y, z_{xy})
= (1, 0, z_{x+1,y} - z_{xy})
\]

\[
0 = N \cdot V_1
= (n_x, n_y, n_z) \cdot (1, 0, z_{x+1,y} - z_{xy})
= n_x + n_z(z_{x+1,y} - z_{xy})
\]
Normal Integration

\[ \nabla z = [p, q]^\top \]

where:

\[
\begin{align*}
p &= -\frac{n_1}{n_3} \\
q &= -\frac{n_2}{n_3}
\end{align*}
\]

Linear Partial Differential Equations

Integrability Constraint:

\[ \partial_v p = \partial_u q \]

The order of taking 2\textsuperscript{nd} order partial derivative with \( u \) & \( v \) (or \( x \) & \( y \)) shouldn’t matter!

\[
z(u, v) = z(u_0, v_0) + \int_{(r,s) = (u_0,v_0)}^{(u,v)} [p(r, s) \, dr + q(r, s) \, ds]
\]

Read Normal Integration: A Survey (if interested)
Results

from Athos Georghiades
Results
Extension

• Photometric Stereo from Colored Lighting

Video Normals from Colored Lights
Gabriel J. Brostow, Carlos Hernández, George Vogiatzis, Björn Stenger, Roberto Cipolla

Fig. 2. Applying the original algorithm to a face with white makeup. Top: example input frames from video of an actor smiling and grimacing. Bottom: the resulting integrated surfaces.
Today’s class

• Measuring Light (recap)
• Image formation with shape, reflectance, and illumination
• Shape from Shading
• Photometric Stereo
• Uncalibrated Photometric Stereo
• Generalized Bas-Relief Ambiguity
• Photometric Stereo in ‘deep learning era’.
What if the light directions are unknown?

\[ a = \text{albedo.} \]

Previously \( k_d \) was used for albedo.

\[
\begin{align*}
I_1 &= a\hat{n}^\top \hat{l}_1 \\
I_2 &= a\hat{n}^\top \hat{l}_2 \\
&\vdots \\
I_N &= a\hat{n}^\top \hat{l}_N
\end{align*}
\]

define “pseudo-normal” \( \vec{b} \triangleq a\hat{n} \)

solve linear system for pseudo-normal

\[
\begin{bmatrix}
I_1 \\
I_2 \\
\vdots \\
I_N
\end{bmatrix}_{N \times 1} = \begin{bmatrix}
\hat{l}_1^\top \\
\hat{l}_2^\top \\
\vdots \\
\hat{l}_N^\top
\end{bmatrix}_{N \times 3} \begin{bmatrix}
\vec{b}
\end{bmatrix}_{3 \times 1}
What if the light directions are unknown?

\[ a = \text{albedo.} \]

Previously \( k_d \) was used for albedo.

\[
\begin{align*}
I_1 &= a \hat{n}^\top \ell_1 \\
I_2 &= a \hat{n}^\top \ell_2 \\
\vdots \\
I_N &= a \hat{n}^\top \ell_N
\end{align*}
\]

Define “pseudo-normal” \( \vec{b} \triangleq a \hat{n} \)

Solve linear system for pseudo-normal at each image pixel

\[
\begin{bmatrix}
I_1 \\
I_2 \\
\vdots \\
I_N
\end{bmatrix}_{N \times M} =
\begin{bmatrix}
\ell_1^\top \\
\ell_2^\top \\
\vdots \\
\ell_N^\top
\end{bmatrix}_{N \times 3} \begin{bmatrix}
B
\end{bmatrix}_{3 \times M}
\]

M: number of pixels
What if the light directions are unknown?

\( a = \text{albedo}. \)

Previously \( k_d \) was used for albedo.

\[
\begin{align*}
I_1 &= a \hat{n}^T \ell_1 \\
I_2 &= a \hat{n}^T \ell_2 \\
&\vdots \\
I_N &= a \hat{n}^T \ell_N
\end{align*}
\]

Define "pseudo-normal" \( \vec{b} \triangleq a \hat{n} \)

Solve linear system for pseudo-normal at each image pixel

\[
\begin{bmatrix}
I_1 \\
I_2 \\
\vdots \\
I_N
\end{bmatrix}_{N \times M} =
\begin{bmatrix}
\ell_1^T \\
\ell_2^T \\
\vdots \\
\ell_N^T
\end{bmatrix}_{N \times 3} \begin{bmatrix}
B
\end{bmatrix}_{3 \times M}
\]

How do we solve this system without knowing light matrix \( L \)?
Factorizing the measurement matrix

Measurements = Lights \times \text{Pseudonormals}

What are the dimensions?
Factorizing the measurement matrix

• Singular value decomposition:

\[ D = U W V^T \]

To reduce to rank 3, we just need to set all the singular values to 0 except for the first 3.

This decomposition minimizes \( |I - LB|^2 \)
Are the results unique?

We can insert any 3x3 matrix $Q$ in the decomposition and get the same images:

$$I = LB = (LQ^{-1})(QB)$$

Can we use any assumptions to remove some of these 9 degrees of freedom?
Today’s class

• Measuring Light (recap)
• Image formation with shape, reflectance, and illumination
• Shape from Shading
• Photometric Stereo
• Uncalibrated Photometric Stereo
• Generalized Bas-Relief Ambiguity
• Photometric Stereo in ‘deep learning era’.
Generalized Bas-Relief ambiguity

We can insert any 3x3 matrix $Q$ in the decomposition and get the same images:

$$I = L B = (L Q^{-1})(Q B)$$

Can we use any assumptions to remove some of these 9 degrees of freedom?

Generalized Bas-Relief ambiguity to rescue!

$G$ has 3 degrees of freedom.

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}$$

What does $G$ mean?

How do we obtain $G$? What constraints lead us to $G$?
Generalized Bas-Relief ambiguity

Artists have exploited GBR ambiguity in creating statues!

- On can flatten a surface and yet give an impression of full 3D to a viewer

“The Bas-Relief Ambiguity”, Peter N. Belhumeur, David J. Kriegman, Alan L. Yuille, IJCV 99
Generalized Bas-Relief ambiguity

\[ z = f(x, y) \]

\[ \mathbf{n}(x, y) = \begin{bmatrix} -f_x \\ -f_y \\ 1 \end{bmatrix} \]

\[ \bar{f}(x, y) = \lambda f(x, y) + \mu x + \nu y \]
Generalized Bas-Relief ambiguity

Note that if \( \mathbf{p} = (x, y, f(x, y)) \) and \( \bar{\mathbf{p}} = (x, y, \bar{f}(x, y)) \), then \( \bar{\mathbf{p}} = G\mathbf{p} \) where

\[
G = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\mu & \nu & \lambda
\end{bmatrix}.
\]

\[
\bar{\mathbf{n}} = G^{-T}\mathbf{n} \quad \quad G^{-1} = \frac{1}{\lambda} \begin{bmatrix}
\lambda & 0 & 0 \\
0 & \lambda & 0 \\
-\mu & -\nu & 1
\end{bmatrix}.
\]
We can insert any 3x3 matrix $Q$ in the decomposition and get the same images:

$$I = LB = (LQ^{-1})(QB)$$

Can we use any assumptions to remove some of these 9 degrees of freedom?

Generalized Bas-Relief ambiguity to rescue!

$G$ has 3 degrees of freedom.

$G$ indicates integrable surface:

The order of taking 2nd order partial derivative with $u$ & $v$ (or $x$& $y$) shouldn’t matter!
Enforcing integrability

What does the integrability constraint correspond to?

• Differentiation order should not matter:

\[
\frac{d}{dy} \frac{df(x, y)}{dx} = \frac{d}{dx} \frac{df(x, y)}{dy}
\]

\[
I = LB = (L Q^{-1})(Q B)
\]

If B is integrable, then:

• \(B' = G^{-T} \cdot B\) is also integrable for all \(G\) of the form \((\lambda \neq 0)\)

\[
G = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\mu & \nu & \lambda
\end{bmatrix}
\]
For now, ignore specular reflection
And Refraction...
And Interreflections...
And Subsurface Scattering...
What assumptions have we made for all this?

- Lambertian BRDF
- Directional lighting
- Distant Lighting
- Orthographic camera
- No interreflections or scattering
Limitations

Bigger problems
• doesn’t work for shiny things, semi-translucent things
• shadows, inter-reflections

Smaller problems
• camera and lights have to be distant
• calibration requirements
  • measure light source directions, intensities
  • camera response function

Newer work addresses some of these issues

Some pointers for further reading:
• Hertzmann & Seitz, “Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs.” IEEE Trans. PAMI 2005
Today’s class

• Measuring Light (recap)
• Image formation with shape, reflectance, and illumination
• Shape from Shading
• Photometric Stereo
• Uncalibrated Photometric Stereo
• Generalized Bas-Relief Ambiguity

• Photometric Stereo in ‘deep learning era’.
Photometric Stereo now ... in Deep Learning era!

- Exploiting High-quality CG rendering for training data
- Designing deep neural network architectures
- Designing loss functions

- GBR ambiguity is still a problem! -> Flattened objects reconstructed.
Using lighting as a cue for 3D reconstruction (Photometric Stereo)

“Shape & Material Capture at Home”, Lichy, Wu, Sengupta, Jacobs, CVPR 2021

“Real-Time Light-Weight Near-Field Photometric Stereo”, Lichy, Sengupta, Jacobs, CVPR 2022
Single iPhone Image with Built-In Flash
Photometric Stereo + SLAM for colon reconstruction in colonoscopy

Johnson and Adelson, 2009
Johnson and Adelson, 2009
Lights, camera, action
Figure 7: Comparison with the high-resolution result from the original retrographic sensor. (a) Rendering of the high-resolution $20 bill example from the original retrographic sensor with a close-up view. (b) Rendering of the captured geometry using our method.
Figure 9: Example geometry measured with the bench and portable configurations. Outer image: rendering under direct lighting. Inset: macro photograph of original sample. Scale shown in upper left. Color images are shown for context and are to similar, but not exact scale.
Sensing Surfaces with GelSight

https://www.youtube.com/watch?v=S7gXih4XS7A