6.801/6.866: Machine Vision, Lecture 8

Professor Berthold Horn, Ryan Sander, Tadayuki Yoshitake
MIT Department of Electrical Engineering and Computer Science
Fall 2020

These lecture summaries are designed to be a review of the lecture. Though I do my best to include all main topics from the lecture, the lectures will have more elaborated explanations than these notes.

1 Lecture 8: Shape from Shading, Special Cases, Lunar Surface, Scanning Electron Microscope, Green's Theorem in Photometric Stereo

In this lecture, we will continue our discussion of photometry/radiometry, beginning with the concepts covered last lecture. We will then explore Hapke surfaces and how they compare to Lambertian surfaces, and show how they can be used to solve our shape from shading problem. Finally, we will introduce two new types of lenses and some of their applications: "thick" lenses and telecentric lenses.

1.1 Review of Photometric and Radiometric Concepts

Let us begin by reviewing a few key concepts from photometry and radiometry:

- **Photometry**: Photometry is the science of measuring visible radiation, light, in units that are weighted according to the sensitivity of the human eye. It is a quantitative science based on a statistical model of the human visual perception of light (eye sensitivity curve) under carefully controlled conditions [3].

- **Radiometry**: Radiometry is the science of measuring radiation energy in any portion of the electromagnetic spectrum. In practice, the term is usually limited to the measurement of ultraviolet (UV), visible (VIS), and infrared (IR) radiation using optical instruments [3].

- **Irradiance**: $E \triangleq \frac{\delta P}{\delta A}$ (W/m^2). This corresponds to light falling on a surface. When imaging an object, irradiance is converted to a grey level.

- **Intensity**: $I \triangleq \frac{\delta P}{\delta W}$ (W/ster). This quantity applied to a point source and is often directionally-dependent.

- **Radiance**: $L \triangleq \frac{\delta^2 P}{\delta A \delta \Omega}$ (W/m$^2 \times$ ster). This photometric quantity is a measure of how bright a surface appears in an image.

- **BRDF (Bidirectional Reflectance Distribution)**: $f(\theta_i, \theta_e, \phi_i, \phi_e) = \frac{\delta L(\theta_e, \phi_e)}{\delta E(\theta_i, \phi_i)}$. This function captures the fact that oftentimes, we are only interested in light hitting the camera, as opposed to the total amount of light emitted from an object. Last time, we had the following equation to relate image irradiance with object/surface radiance:

$$E = \frac{\pi}{4} L \left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

  Where the irradiance of the image E is on the lefthand side and the radiance of the object/scene L is on the right. The BRDF must also satisfy **Helmholtz reciprocity**, otherwise we would be violating the 2nd Law of Thermodynamics. Mathematically, recall that Helmholtz reciprocity is given by:

$$f(\theta_i, \theta_e, \phi_i, \phi_e) = f(\theta_e, \theta_i, \phi_e, \phi_i) \ \forall \ \theta_i, \theta_e, \phi_i, \phi_e$$

With this, we are now ready to discuss our first type of surfaces: ideal Lambertian surfaces.

1.2 Ideal Lambertian Surfaces

Ideal Lambertian surfaces are of interest because their sensed brightness (irradiance) only depends on the cosine of the incident angles (θ_i, ϕ_i). A few additional notes about these surfaces:

- Ideal Lambertian surfaces are equally bright from all directions, i.e.

$$f(\theta_i, \theta_e, \phi_i, \phi_e) = f(\theta_e, \theta_i, \phi_e, \phi_i) \ \forall \ \theta_i, \theta_e, \phi_i, \phi_e$$
$$\text{AND}$$

$$f(\theta_i, \theta_e, \phi_i, \phi_e) = K \in \mathbb{R} \text{ with respect to } \theta_e, \phi_e, \text{ i.e.} \frac{\partial f(\theta_i, \theta_e, \phi_i, \phi_e)}{\partial \theta_e} = \frac{\partial f(\theta_i, \theta_e, \phi_i, \phi_e)}{\partial \phi_e} = 0$$

- If the surface is ideal, the Lambertian surface reflects all incident light. We can use this to compute f. Suppose we take a small slice of the light reflected by a Lambertian surface, i.e. a hemisphere for all positive z. The area of this surface is given by: $\sin \theta_e \delta \theta_e \delta \phi_e$, which is the determinant of the coordinate transformation from euclidean to spherical coordinates. Then, we have that:

$$\int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} (fE\cos\theta_i \sin\theta_e d\theta_e)d\phi_e = E\cos\theta_i \ \textbf{(Integrate all reflected light)}$$

$$E\cos\theta_i \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} (f \sin\theta_e d\theta_e)d\phi_e = E\cos\theta_i \ \textbf{(Factor out } E\cos\theta_i\textbf{)}$$

$$\int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} (f \sin\theta_e d\theta_e)d\phi_e = 1 \ \textbf{(Cancel out on both sides)}$$

$$2\pi \int_{0}^{\frac{\pi}{2}} f \sin\theta_e \cos\theta_e d\theta_e = 1 \ \textbf{(No dependence on } \phi\textbf{)}$$

$$\pi f \int_{0}^{\frac{\pi}{2}} 2\sin\theta_e \cos\theta_e d\theta_e = 1 \ \textbf{(Rearranging terms)}$$

$$\pi f \int_{0}^{\frac{\pi}{2}} \sin(2\theta_e) d\theta_e = 1 \ \textbf{(Using identity } 2\sin\theta\cos\theta = \sin(2\theta)\textbf{)}$$

$$\pi f[-\frac{1}{2}\cos(2\theta_e)]_{\theta_e=0}^{\theta_e=\frac{\pi}{2}} = 1$$

$$\pi f = 1 \implies f = \frac{1}{\pi}$$

1.2.1 Foreshortening Effects in Lambertian Surfaces

Now let us discuss a key issue in radiation reflectance and machine vision - **foreshortening**. Foreshortening is the visual effect or optical illusion that causes an object or distance to appear shorter than it actually is because it is angled toward the viewer [1]. Let us show how this is relevant for computing f.

Suppose we have a point source radiating isotropically over the positive hemisphere in a 3-dimensional spherical coordinate system. Then, the solid angle spanned by this hemisphere is:

$$\Omega = \frac{A_{\text{surface}}}{r^2} = \frac{2\pi r^2}{r^2} = 2\pi \text{ Steradians}$$

If the point source is radiating isotropically in all directions, then $f = \frac{1}{2\pi}$. But we saw above that $f = \frac{1}{\pi}$ for an ideal Lambertian surface. Why do we have this discrepancy? Even if brightness is the same/radiation emitted is the same in all directions, this does not mean that the power radiated in all directions is hte same. This is due to **foreshortening**, since the angle between the object's surface normal and the camera/viewer changes.

1.2.2 Example: Distant Lambertian Point Source

Here, we will have that $L = \frac{1}{f}Es$ (where s is the location of the light source). Since we have to take foreshortening into account, the perceived area A of an object with world size A' is given by:

$$A = A' \cos\theta_i$$

Therefore, our expression for the the radiance (how bright the object appears in the image) is given by:

$$L = \frac{1}{\pi}Es = \frac{1}{\pi}E_0 \cos\theta_i$$

1.3 Hapke Surfaces

Hapke surfaces are another type of surface we study in this course. Formally, a Hapke surface is an object which has reflectance properties corresponding to **Hapke parameters**, which in turn are a set of parameters for an empirical model that is commonly used to describe the directional reflectance properties of the airless regolith surfaces of bodies in the solar system [2].

The BRDF of a Hapke surface is given by:

$$f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{1}{\sqrt{\cos\theta_e \cos\theta_i}}$$

We can see that the Hapke BRDF also satisfies Helmholtz Reciprocity:

$$f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{1}{\sqrt{\cos\theta_e \cos\theta_i}} = \frac{1}{\sqrt{\cos\theta_i \cos\theta_e}} = f(\theta_e, \phi_e, \theta_i, \phi_i)$$

Using the Hapke BRDF, we can use this to compute **illumination** (same as **radiance**, which is given by:

$$L = E_0 \cos\theta_i f(\theta_i, \phi_i, \theta_e, \phi_e)$$
$$= E_0 \cos\theta_i \frac{1}{\sqrt{\cos\theta_e \cos\theta_i}}$$
$$= E_0 \sqrt{\frac{\cos\theta_i}{\cos\theta_e}}$$

Where L is the radiance/illumination of the object in the image, $E_0 \cos\theta_i$ is the irradiance of the object, accounting for foreshortening effects, and $\frac{1}{\sqrt{\cos\theta_i \cos\theta_e}}$ is the BRDF of our Hapke surface.

1.3.1 Example Application: Imaging the Lunar Surface

What do the isophotes of the moon look like, supposing the moon can fall under some of the different types of surfaces we have discussed?

- **Lambertian**: We will see circles and ellipses of isophotes, depending on the angle made between the viewer and the the moon.

- **Hapke**: Because of the BRDF behavior, isophotes will run **longitudinally** along the moon in the case in which it is a Hapke surface. Recall isophotes are where the brightness of the object in the image, given by:

$$\sqrt{\frac{\cos\theta_i}{\cos\theta_e}} = \sqrt{\frac{\frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{s}}}{||\hat{\mathbf{n}}||_2 ||\hat{\mathbf{s}}||_2}}{\frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}}{||\hat{\mathbf{n}}||_2 ||\hat{\mathbf{v}}||_2}}} = c \implies \frac{\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}}{\hat{\mathbf{n}} \cdot \hat{\mathbf{v}}} = c^2 \text{ for some } c \in \mathbb{R}$$

Where $\hat{\mathbf{n}}$ is the surface normal vector of the object being imaged, $\hat{\mathbf{s}}$ is the unit vector describing the position of the light source, and $\hat{\mathbf{v}}$ is the unit vector describing the position of the vertical. We can derive a relationship between $\hat{\mathbf{n}}$ and $\hat{\mathbf{s}}$:

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{s}} = c\hat{\mathbf{n}} \cdot \hat{\mathbf{v}} \implies \hat{\mathbf{n}} \cdot (\hat{\mathbf{s}} - c\hat{\mathbf{v}}) = 0 \implies \hat{\mathbf{n}} \perp (\hat{\mathbf{s}} - c\hat{\mathbf{v}})$$

Next, we will use spherical coordinates to show how the isophotes of this surface will be longitudinal:

$$\hat{\mathbf{n}} = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix} \in \mathbb{R}^3$$

Since normal vectors $\hat{\mathbf{n}} \in \mathbb{R}^3$ only have two degrees of freedom (since $||\hat{\mathbf{n}}||_2 = 1$), we can fully parameterize these two degrees of freedom by the **azimuth** and **colatitude/polar** angles ϕ and θ, respectively. Then we can express our coordinate system given by the orthogonal basis vectors $\hat{\mathbf{n}}, \hat{\mathbf{s}}, \hat{\mathbf{v}}$ as:

$$\hat{\mathbf{n}} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{bmatrix}^T$$
$$\hat{\mathbf{s}} = \begin{bmatrix} \cos\theta_s & \sin\theta_s & 0 \end{bmatrix}^T$$
$$\hat{\mathbf{v}} = \begin{bmatrix} \cos\theta_v & \sin\theta_v & 0 \end{bmatrix}^T$$

We can use our orthogonal derivations from above to now show our longitudinal isophotes:

$$\sqrt{\frac{\cos\theta_i}{\cos\theta_e}} = \frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{s}}}{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}} = \frac{\sin\theta\cos\phi\cos\phi_s + \sin\theta\sin\phi\sin\phi_s}{\sin\theta\cos\phi\cos\phi_v + \sin\theta\sin\phi\sin\phi_v} = \frac{\cos(\phi-\phi_s)}{\cos(\phi-\phi_v)} = c \ \text{ for some } c \in \mathbb{R}$$

From the righthand side above, we can deduce that isophotes for Hapke surfaces correspond to points with the same azimuth having the same brightness, i.e. the isophotes of the imaged object (such as the moon) are along lines of constant longitude.

1.3.2   Surface Orientation and Reflectance Maps of Hapke Surfaces

Next, we will take another look at our "shape from shading" problem, but this time using the surface normals of Hapke surfaces and relating this back to reflectance maps and our (p, q) space. From our previous derivations, we already have that the BRDF of a Hapke surface is nothing more than the square root of the dot product of the surface normal with the light source vector divided by the dot product of the surface normal with the vertical vector, i.e.

$$\sqrt{\frac{\cos\theta_i}{\cos\theta_e}} = \sqrt{\frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{s}}}{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}}}$$

Next, if we rewrite our unit vectors as 3D vectors in (p, q) space, we have:

$$\hat{\mathbf{n}} = \frac{\begin{bmatrix} -p & -q & 1 \end{bmatrix}^T}{\sqrt{p^2 + q^2 + 1}}$$

$$\hat{\mathbf{s}} = \frac{\begin{bmatrix} -p_s & -q_s & 1 \end{bmatrix}^T}{\sqrt{p_s^2 + q_s^2 + 1}}$$

$$\hat{\mathbf{v}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$$

Then, substituting these definitions, we have:

$$\hat{\mathbf{n}}\cdot\hat{\mathbf{v}} = \frac{1}{\sqrt{1+p^2+q^2}}$$

$$\hat{\mathbf{n}}\cdot\hat{\mathbf{s}} = \frac{1+p_s p + q_s q}{\sqrt{1+p^2+q^2}\sqrt{1+p_s^2+q_s^2}}$$

$$\frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{s}}}{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}} = \frac{\frac{1+p_s p + q_s q}{\sqrt{1+p^2+q^2}\sqrt{1+p_s^2+q_s^2}}}{\frac{1}{\sqrt{1+p^2+q^2}}}$$

$$= \frac{1+p_s p + q_s q}{\sqrt{1+p_s^2+q_s^2}}$$

$$= \frac{1+p_s P + q_s q}{R_s}$$

(Where $R_s \triangleq \sqrt{1+p_s^2+q_s^2}$.) Note that p_s and q_s are constants, since they only reflect the spatial derivatives of the light source. This ratio of dot products is therefore linear in (p, q) space! Then, what do the isophotes look like in gradient space? Recall that isophotes take the form:

$$\sqrt{\frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{s}}}{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}}} = c \implies \frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{s}}}{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}} = \frac{1+p_s P + q_s q}{R_s} = c^2 \ \text{ for some } c \in \mathbb{R} \ .$$

Therefore, isophotes of Hapke surfaces are **linear** in spatial gradient space! This is very useful for building reflectance maps because, as we will see shortly, a simple change of coordinate system in (p, q) space tells us a lot about the surface orientation in one of the two directions.

For one thing, let us use our linear isophotes in gradient space to solve our photometric stereo problem, in this case with two measurements under different lighting conditions. Photometric stereo is substantially easier with Hapke surfaces than with Lambertian, because there is no ambiguity in where the solutions lie. Unlike Lambertian surfaces, because of the linearity in (p, q) space we are guaranteed by Bezout's Theorem to have only one unique solution.

1.3.3 Rotated (p', q') Coordinate Systems for Brightness Measurements

We can also derive/obtain a surface orientation in gradient space in a rotated coordinate system in gradient space, i.e. from $(p,q) \longrightarrow (p', q')$.

We can prove (see the synchronous lecture notes for this part of the course) that the transforming the points in (x, y) via a rotation matrix $\mathbf{R}$ is equivalent to rotating the gradient space (p, q) by the same matrix $\mathbf{R}$. I.e.

$$\mathbf{R} : (x, y) \implies (x', y'), \begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{R} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\mathbf{R} : (p, q) \implies (p', q'), \begin{bmatrix} p' \\ q' \end{bmatrix} = \mathbf{R} \begin{bmatrix} p \\ q \end{bmatrix}$$

Where $\mathbf{R}$ is given according to:

$$\mathbf{R} = \begin{bmatrix} \cos\alpha & sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \in SO(2)$$

(Where $SO(2)$ is the Special Orthogonal Group [4].) Note that $\mathbf{R}^{-1} = \mathbf{R}^T$ since $\mathbf{R}$ is orthogonal and symmetric.

By rotating our coordinate system from $(p, q) \longrightarrow (p', q')$, we are able to uniquely specify p', but not q', since our isophotes lie along a multitude/many q' values. Note that this rotation system is specified by having p' lying along the gradient of our isophotes.

Returning to our Lunar Surface application with Hapke surfaces, we can use this surface orientation estimation framework to take an iterative, incremental approach to get a profile of the lunar surface. This enables us to do **depth profile shaping** of the moon simply from brightness estimates! There are a few caveats to be mindful of for this methodology, however:

- We can only recover absolute depth values provided we are given initial conditions (this makes sense, since we are effectively solving a differential equation in estimating the depth z from the surface gradient $(p, q)^T$). Even without initial conditions, however, we can still recover the general shape of profiling, simply without the absolute depth.

- Additionally, we can get profiles for each cross-section of the moon/object we are imaging, but it is important to recognize that these profiles are effectively independent of one another, i.e. we do not recover any information about the relative surface orientation changes between/at the interfaces of different profiles we image. We can use **heuristic-based** approaches to get a **topographic mapping** estimate. We can then combine these stitched profiles together into a 3D surface.

1.4 "Thick" & Telecentric Lenses

We will now shift gears somewhat to revisit lenses, namely "thick" and telecentric lenses.

1.4.1 "Thick" Lenses

In practice, lenses are not thick, and are typically cascaded together to mitigate the specific aberrations and wavelength dependence of each individual lens. For thick lenses, we will leverage the following definitions:

- **Nodal/principal points**: The center points of the outermost (ending and beginning lenses).

- **Principal planes**: The distance from the object to the first imaging lens and the distance from the last imaging lens to the viewer or imaging plane. Abbreviated by b.

- **Focal length**: This definition is the same as before, but in this case, the focal length is the distance from the first imaging lens to the last imaging lens along the optical axis (i.e. orthogonal to the radial unit vectors of each lens in the "thick" lens).

From this, we have the formula:

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}$$

One "trick" we can do with this is achieve a long focal length f and a small field of view (FOV), i.e. as is done in a telephoto lens.

1.4.2 Telecentric Lenses

These types of lenses are frequently used in machine vision, such as for circuit board inspection or barcode reading.

Motivation for these types of lenses: When we have perspective projection, the degree of **magnification** depends on **distance**. How can we remove this unwanted distance-dependent magnification due to perspective projection? We can do so by effectively "moving" the Center of Projection (COP) to "$-\infty$". This requires using a lens with a lot of glass.

A few notes about this:

- To image with magnifying effects, the lens must be at least as wide in diameter as the width of the object.

- a A double telecentric lens can be created by moving the other "modal point" (the imaging plane/viewer) to "$+\infty$". In this case, our magnifying term given by $\cos^4 \alpha$ goes to zero, because:

$$\lim_{\text{COP}\to+\infty} \alpha = 0$$

- We can generalize this idea even more with "lenselets": These each concentrate light into a confined area, and can be used in an array fashion as we would with arrays of photodiodes. A useful term for these types of arrays is **fill factor**: The fill factor of an image sensor array is the ratio of a pixel's light sensitive area to its total area [5]. Lenselet arrays can be useful because it helps us to avoid **aliasing effects**. We need to make sure that if we are sampling an object discretely, that we do not have extremely high-frequency signals without some low pass filtering. This sort of low-pass filtering can be achieved with using large pixels and averaging (a crude form of lowpass filtering). However, this averaging scheme does not work well when light comes in at off-90 degree angles. We want light to come into the sensors at near-90 degrees (DSLRs find a way to get a around this).

- Recall that for an object space telecentric device, we no longer have a distance-based dependence. Effectively, we are taking **perspective projection** and making the **focal length** larger, resulting in **approximately orthographic projection**. Recall that for orthographic projection, projection becomes nearly independent in position, i.e. (x, y) in image space has a linear relationship with (X, Y) in the world. This means that we can measure sizes of objects independently of how far away they are!

  Starting with perspective projection:

  $$\frac{x}{f} = \frac{X}{Z}, \frac{y}{f} = \frac{Y}{Z}$$

  Having $|\Delta Z| << |Z|$, where ΔZ is the variation in Z of an object $\implies x = \frac{f}{Z_0}X, y = \frac{f}{Z_0}Y$

  Approximating $\frac{f}{Z_0} = 1 \implies x = X, y = Y$

1.5 References

1. Foreshortening, https://en.wikipedia.org/wiki/Perspective_(graphical)

2. Hapke Parameters, https://en.wikipedia.org/wiki/Hapke_parameters

3. Understanding Radiance (Brightness), Irradiance and Radiant Flux

4. Orthogonal Group, https://en.wikipedia.org/wiki/Orthogonal_group

5. https://en.wikipedia.org/wiki/Fill_factor_(image_sensor)