Photometric Stereo

Shree K. Nayar

Monograph: FPCV-3-2

Module: Reconstruction I

Series: First Principles of Computer Vision

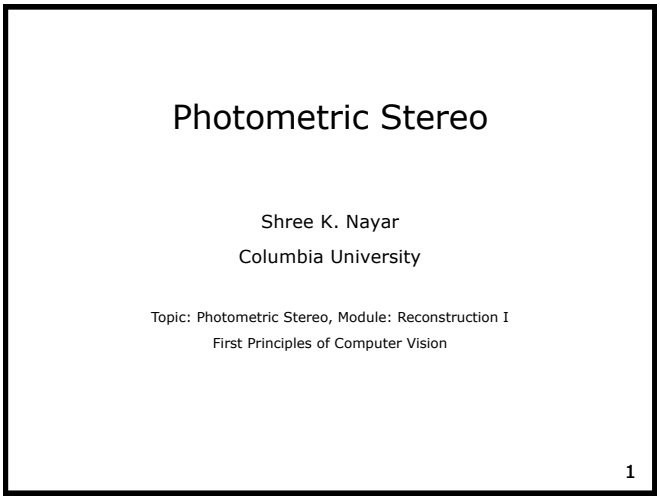
Computer Science, Columbia University

February, 2025

FPCV Channel

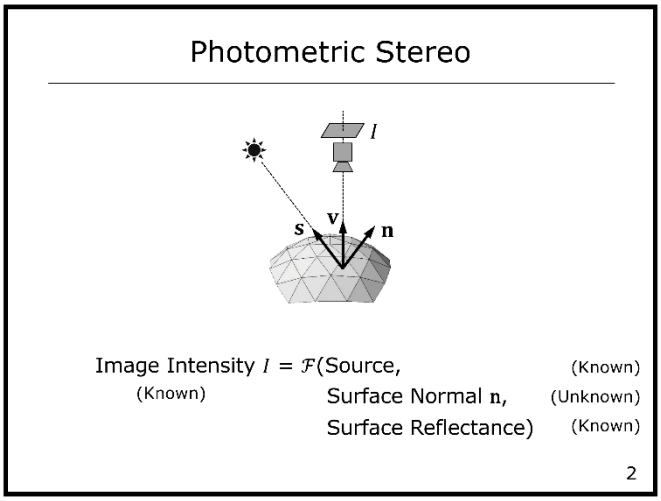
FPCV Website

In the previous lecture, we discussed various radiometric concepts. We know how to describe the brightness of a light source by its radiant intensity, the illumination of a surface by its irradiance, and the brightness of a surface by its radiance. We also described the bidirectional reflectance distribution function (BRDF), which describes the ability of a surface to take light from one direction and reflect it in another direction. Finally, we presented a few reflectance models that are commonly found in practice.



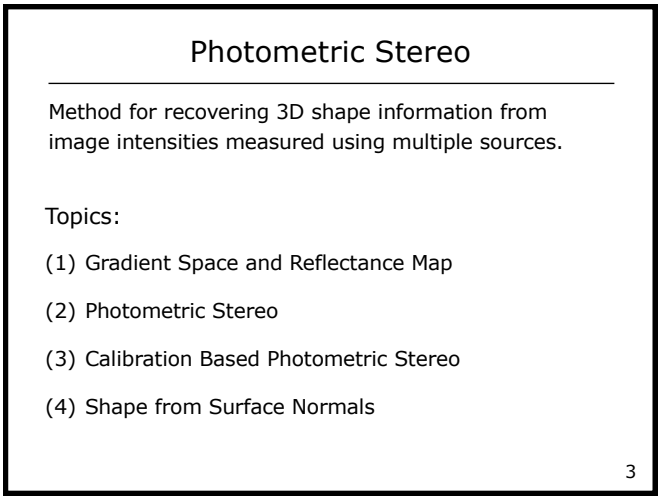
Using these concepts as our foundation, we will now develop our first method for recovering 3D shape information from image intensity values. This method is called photometric stereo and is widely used in several application domains.

Consider the imaging setup shown here. It consists of a camera with a viewing direction v, which is aligned with the *z*-axis of the coordinate frame. Let us assume that the object of interest is illuminated by a single point source in the direction $\mathbf{s}$, where $\mathbf{s}$ is a unit vector. Let us focus on a single point on the object with a surface normal vector $\mathbf{n}$, which is also a unit vector. This point produces the image intensity I in the image. What can be said about the normal vector $\mathbf{n}$ from the intensity I?



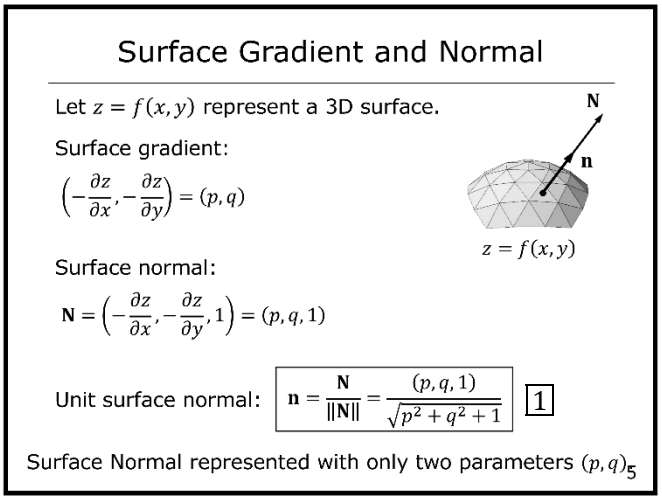
As shown here, I is affected by several factors—the brightness and direction of the source, the surface normal, and the BRDF of the surface. To make the problem more tractable, we will assume that the parameters of the source are known and that the source is distance from the object. Therefore, the source direction $\mathbf{s}$ can be assumed to be the same for all points on the object. This known source assumption can be satisfied in structured settings, such as factory automation, where illumination can be controlled. Furthermore, we will assume that the BRDF of the surface is known, as is often the case in applications such as factory automation. Even with these assumptions in place, we note that unit normal vector $\mathbf{n}$ has two parameters, while we have a single measurement I.

Photometric stereo is a method of estimating the surface normal at each point on an object's surface, when the illumination can be controlled and reflectance properties of the surface are known. First, we will discuss the gradient space, which is a convenient representation of surface normal. Using the gradient space, we will develop the concept of a reflectance map, which gives us the mapping between surface normal and image intensity, when the reflectance properties (BRDF) of the surface and the parameters of the light source are known.

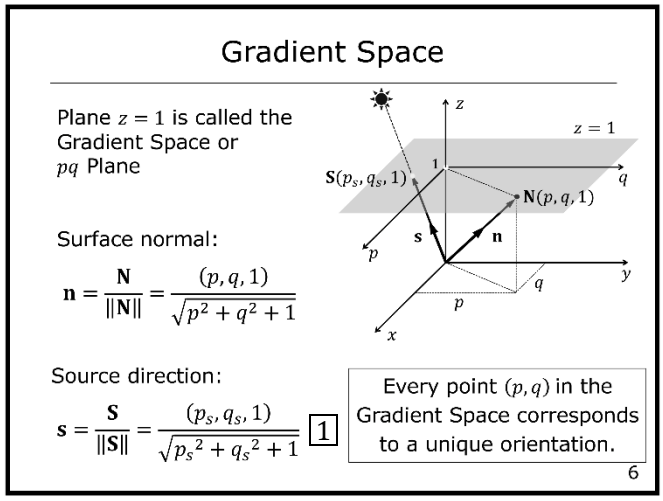


Given an image intensity measured using a single light source, there are an infinite number of surface normals that could have produced that intensity. In order to resolve this ambiguity, photometric stereo uses multiple light sources. We compute a reflectance map for each of the light sources. Using these maps, we can uniquely determine the surface normal at each point on the objects' surface from the set of intensities produced by the multiple light sources.

The above approach of using reflectance maps assumes that we can write an expression for the reflectance model of the surface. This is possible when we know *a-priori* that the material we are dealing with has a BRDF with a simple form, such as the Lambertian model. However, in practice, there are many scenarios where such an assumption cannot be made. For such cases, we will develop a calibration-based (data driven) approach to photometric stereo, where the mapping between image intensities and surface normal is precomputed using an object of known shape with a BRDF that is identical to that of the object whose shape we wish to estimate. Irrespective of the approach used, calibration or not, photometric stereo gives us a normal vector at each point on a surface. We conclude with a method for reconstructing the 3D shape of a surface from its measured surface normals.



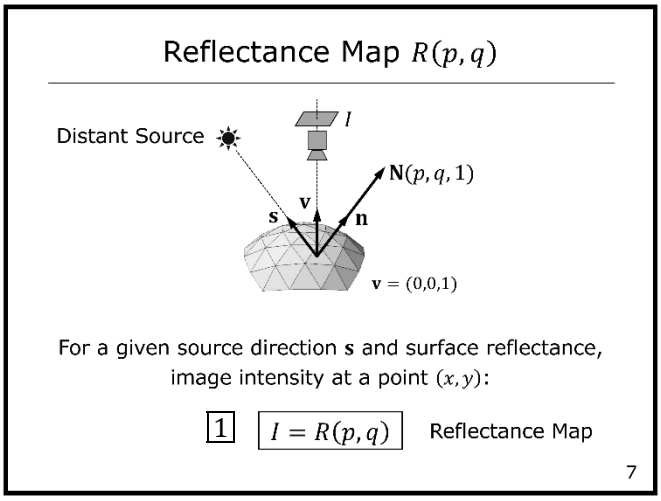
We begin with a simple way to represent surface orientation called the gradient space. Consider the surface shown here, whose depth z is a function of the spatial coordinates (x, y). The surface gradient can be represented as (p, q), where p is the negative of the derivative of z with respect to x, and q is the negative of the derivative of z with respect to y. Using this notation, we can represent the surface normal $\mathbf{N}$ at any point on a surface as $(p, q, 1)$. The unit surface normal $\mathbf{n}$ is the surface normal vector $\mathbf{N}$ divided by its magnitude, $\|\mathbf{N}\|$**.** Note that both $\mathbf{N}$ and $\mathbf{n}$ can be expressed using just the two parameters (p, q) of the gradient [1].



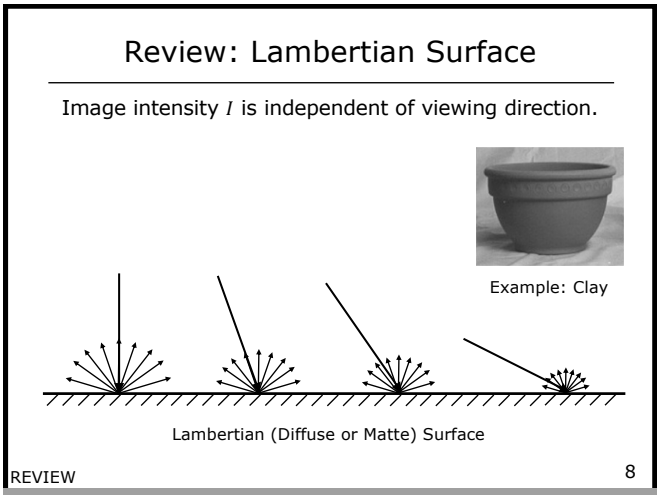
Here, we provide a geometric interpretation for the relationship between the normal vector and its corresponding gradient (p, q). We have a coordinate frame, represented by x, y, and z axes. We can erect a plane parallel to the x-y plane at $z = 1$, and draw the p and q axes parallel to the x and y axes, respectively. This plane defines the $p - q$ space, also called the gradient space.

Consider the unit surface normal $\mathbf{n}$. When it is extended to intersect the $p - q$ plane, we obtain the coordinates $(p, q, 1)$ of the normal vector $\mathbf{N}$. This holds true for any unit normal that lies in the upper hemisphere (i.e., $z \geq 0$). Conversely, each point on the $p - q$ plane corresponds to a unique surface normal. Note that the (p, q) gradient space can be used to express any direction; for instance, the direction of a light source. Given a source direction vector $\mathbf{S} = (p_s, q_s, 1)$, the corresponding unit vector $\mathbf{s}$ is obtained dividing $\mathbf{S}$ by its magnitude [1].

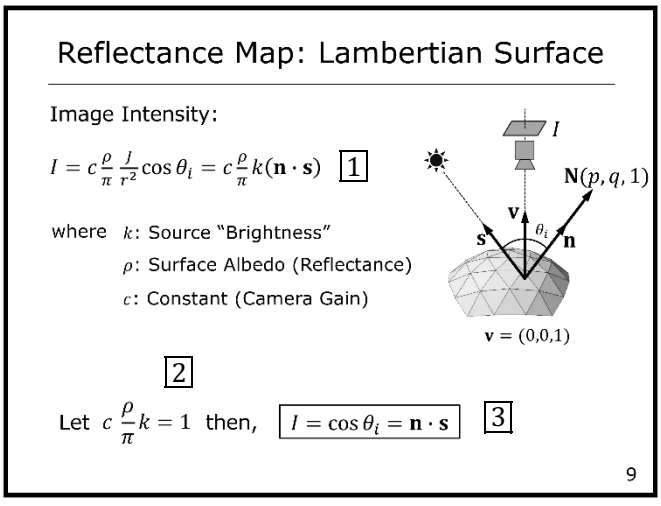
In photometric stereo, as stated earlier, we assume that the location of the light source and the BRDF of the object of interest are known. Both of these pieces of information are used to pre-compute a reflectance map $R(p,q)$, which related the image intensity I measured at a point to its surface normal (p,q) [1]. In other words, given a surface normal, the reflectance map tells us the image intensity it would produce. Note that we would like to solve the inverse problem—given an image intensity, find the surface normal.



Assume the BRDF is Lambertian. We know that the radiance (brightness) of a Lambertian surface is independent of the direction from which it is viewed; the intensity depends only on the irradiance of the surface, which is determined by the angle of the light source with respect to the normal. As this incidence angle increases, the irradiance of the surface decreases, and consequently, the radiance also decreases. For any given incidence angle, the radiance is the same in all directions. The clay pot shown here is an example of a Lambertian-like object.

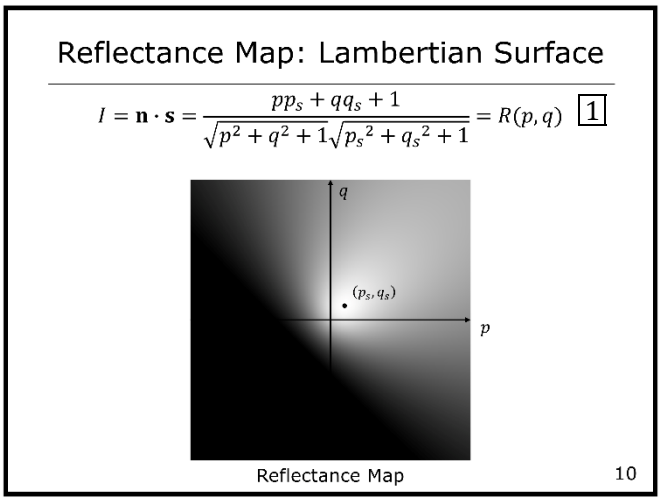


Now, let us derive the reflectance map for a Lambertian surface. The intensity measured on a Lambertian surface was derived in the previous lecture on radiometry and reflectance [1]. Here, θ_i is the angle between the surface normal and the source direction, J is the radiant intensity of the source, r is the distance of the source from the surface point, and ρ is the albedo, or reflectivity, of the Lambertian surface. An albedo equal to 0 corresponds to a black surface, whereas an albedo of 1 corresponds to a white surface, which reflects all the light that falls on it.

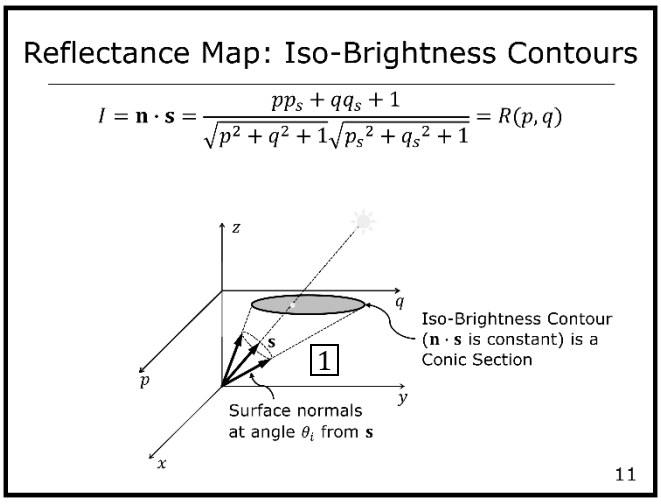


The constant c is the gain of the camera, which accounts for the diameter of the lens, as well as the exposure time of each camera pixel. Note that c is known in any imaging system. Since we assume that the light source is known, J and r are also known, and can be expressed as a single k constant . Since the BRDF of the surface is known, ρ is also known. Therefore, without loss of generality, we can set c times ρ times k divided by π equal to 1 [2]. Hence, we get a simple expression where I is equal to the cosine of the angle of incidence θ_i, which in turn is equal to $\mathbf{n} \cdot \mathbf{s}$ [3].

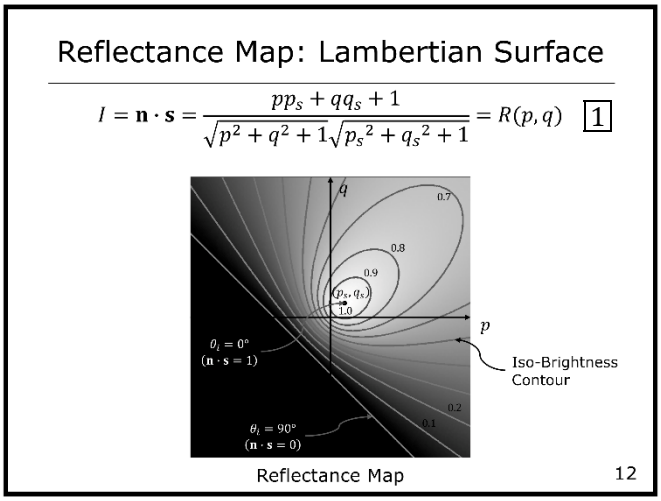
The term $\mathbf{n} \cdot \mathbf{s}$ can be expressed in terms of the normal vector $(p, q, 1)$ and the source vector $(p_s, q_s, 1)$. The result is the reflectance map $R(p, q)$ of a Lambertian surface lit by a known point source. For any given normal (p, q), R reveals the image intensity it would produce. Below, we see a plot of $R(p, q)$, where the brightness at each (p, q) is proportional to its image intensity. Note that the brightest point (black dot) corresponds to a surface normal pointing in the direction of the source, (p_s, q_s). It produce the highest brightness of 1 since, in this case, the normal and source vectors are aligned ($\mathbf{n} \cdot \mathbf{s}$ = 1).



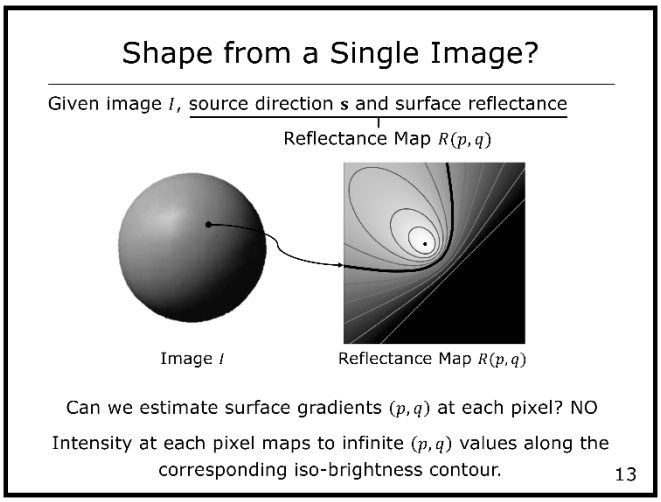
We can see that the above reflectance map is smooth. Let us explore its structure by understanding what a contour of constant brightness in the map represents. Here is a visualization of the (p, q) gradient space that lies at $z = 1$. The unit vector $\mathbf{s}$ points in the direction of the source. The intersection of $\mathbf{s}$ with the gradient space is (p_s, q_s). In the case of a Lambertian surface, $cos\theta_i$ is the same for all surface normals that subtend the same angle with respect to the source direction $\mathbf{s}$. This set of normals lie on a cone [1]. The intersection of this cone with the (p, q) plane is a conic section. Depending on the apex angle of the cone and its tilt, this intersection could be a circle, an ellipse, a parabola, or a hyperbola. Along such an intersection the reflectance map value is constant. We refer to this as an iso-brightness contour.



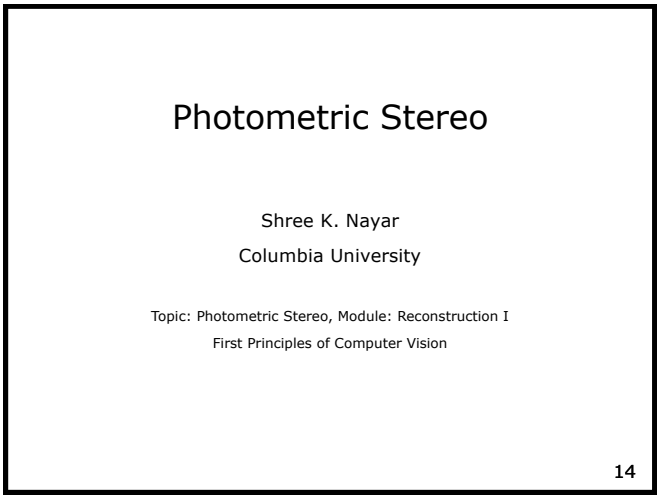
Here is the reflectance map with a few iso-brightness contours overlaid on it. The brightest point in the reflectance map corresponds to the surface normal that points in the direction of the source ($cos\theta_i$= 1). As we move away from the source direction, we get iso-brightness contours that drop in brightness. To determine the (p, q) values for which the brightness equals 0, we can substitute $I = 0$ in equation $\boxed{1}$. The result is the straight line: $pp_s + qq_s + 1 = 0$.



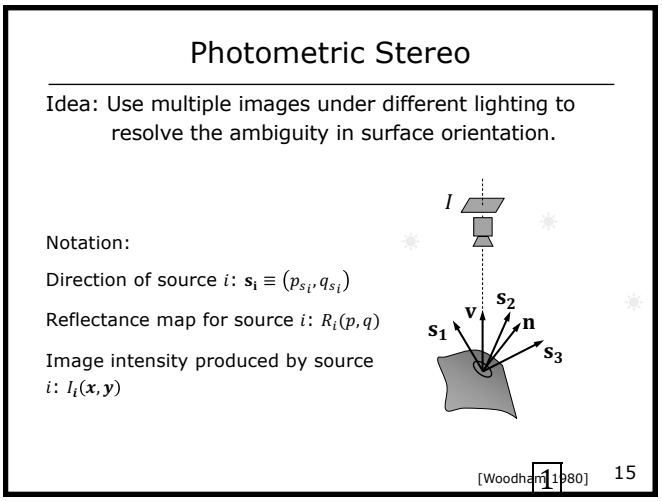
Can we recover the shape of an object from a single image? In other words, given an image intensity I and the reflectance map $R(p, q)$, can we find the surface normal (p, q)? The answer is no. A single intensity at a pixel corresponds to an iso-brightness contour in the reflectance map. Thus, there are an infinite number of (p, q) values that could have produced the same image intensity, resulting in an ambiguity. The idea behind photometric stereo is to use multiple light sources to resolve this ambiguity and compute a unique surface normal at each point on the surface.



Photometric stereo, developed by Robert Woodham, uses multiple images of the object taken from the same camera location under different lighting conditions. It makes use of the multiple intensity values measured at each pixel to resolve the ambiguity in surface normal. Photometric stereo is widely used in structured environments.

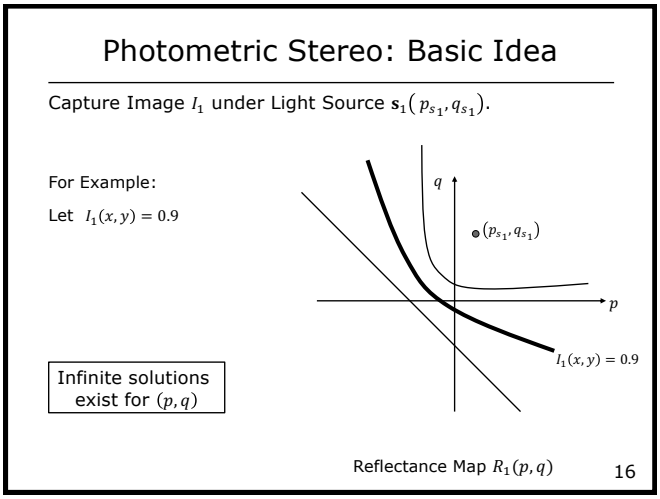


In our setup shown here [1], we consider a single surface point, bearing in mind that the same method is applied to every point on the surface, i.e., every pixel in the image. We will assume that the camera and sources are far away from the surface. This implies that the viewing direction of the camera is the same for all points on the surface, and the direction of each source is the same for all points on the surface.



The setup here consists of three light sources in the directions $\mathbf{s}_1$, $\mathbf{s}_2$ and $\mathbf{s}_3$, the point of interest has a normal $\mathbf{n}$, and the viewing direction is $\mathbf{v}$. For each of the light sources represented by (p_{s_i}, q_{s_i}), we assume that we know the BRDF of the surface and thus can compute a corresponding reflectance map $R_i(p,q)$. Three images are captured by activating the light sources in sequence. At each image point (x,y), we have a set of measured intensities, where $I_i(x,y)$ is the intensity produced by source i.

We want to uniquely determine the surface normal (p,q) using the multiple intensities we have measured. Assume that the reflectance map corresponding to $\mathbf{s}_1$ with direction (p_{s_1}, q_{s_1}) is $R_1(p,q)$. Now, let the brightness of a particular point measured due to source $\mathbf{s}_1$ be 0.9. This brightness value corresponds to an iso-brightness contour (bold curve), which implies that any of the (p,q) values on this contour could have produced the brightness value of 0.9. Thus, we have an infinite number of solutions.



We now use a second light source $\mathbf{s}_2$, which corresponds to a different reflectance map $R_2(p,q)$. For visualization purposes, we have overlaid both reflectance maps $R_1(p,q)$ and $R_2(p,q)$. For the same surface point, assume that the brightness value measured for the source $\mathbf{s}_2$ 0.7. This corresponds to an iso-brightness contour in $R_2(p,q)$, that intersects the iso-brightness contour in $R_1(p,q)$ at two points. Now, we have gone from infinite solutions using only one light source to two solutions.

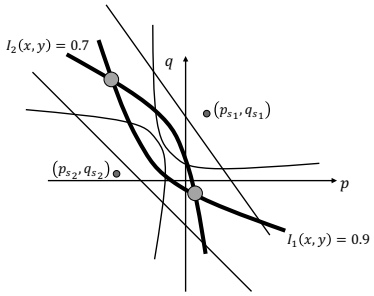
Photometric Stereo: Basic Idea

Capture Image I_2 under Light Source $\mathbf{s}_2(p_{s_2}, q_{s_2})$.

For Example:

Let $I_1(x,y) = 0.9$

$I_2(x,y) = 0.7$



Two solutions exist for (p,q)

Reflectance Maps $R_1(p,q)$, $R_2(p,q)$ 17

Next, we use our third light source in the direction $\mathbf{s}_3$, the reflectance map for which is $R_3(p,q)$. The brightness value measured this time is 0.4, which corresponds to an iso-brightness contour that resolves the ambiguity between the two solutions, to produce a unique solution for the normal of the surface point.

Photometric Stereo: Basic Idea

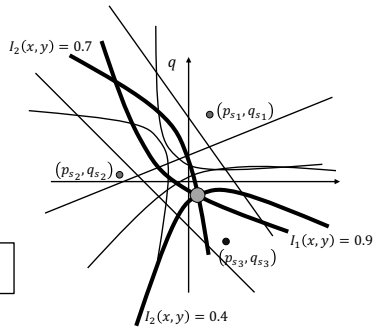
Capture Image I_3 under light source $\mathbf{s}_3(p_{s_3}, q_{s_3})$.

For Example:

Let $I_1(x,y) = 0.9$

$I_2(x,y) = 0.7$

$I_3(x,y) = 0.4$



Unique solution for surface orientation: (p', q')

Reflectance Maps $R_1(p,q)$, $R_2(p,q)$, $R_3(p,q)$ 18

Let us summarize the basic idea behind photometric stereo. We acquire K images with K known light sources. Using the known directions of the sources and the BRDF of the surface, we construct a reflectance map for each of the K sources. Finally, for each image pixel (x,y), we find the normal (p,q) of the corresponding surface point as the intersection of K reflectance map curves.

Photometric Stereo: Basic Idea

Step 1: Acquire K images with K known light sources.

Step 2: Using known source direction and BRDF, construct reflectance map for each source direction.

Step 3: For each pixel location (x,y), find (p,q) as the intersection of K curves. This (p,q) gives the surface normal at pixel (x,y).

Smallest K needed depends on the material properties.

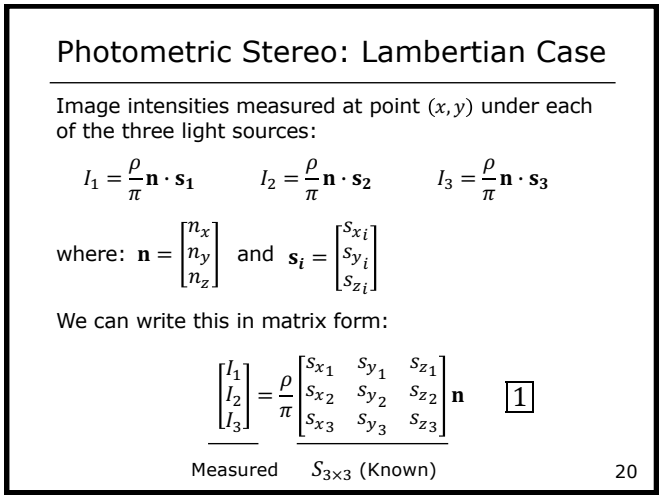
Example: $K = 3$ for Lambertian Surfaces.

19

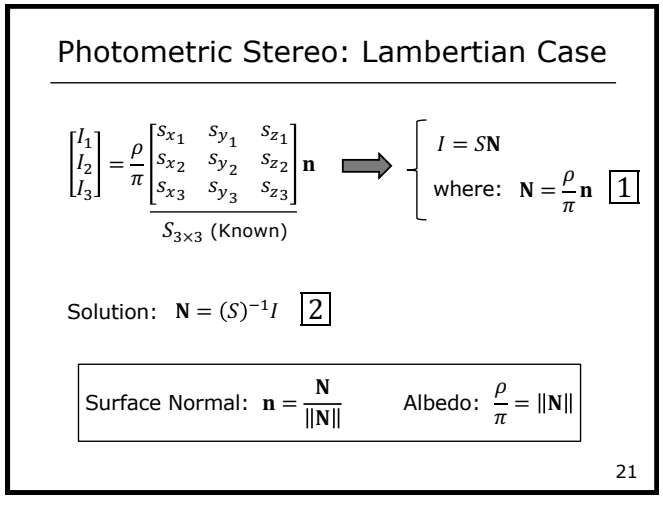
The smallest K needed depends on the complexity of the BRDF of the surface. For a Lambertian surface, three light sources are sufficient to compute the normal at each point. If the BRDF is more

complex and has both specular and diffuse components, more than three light sources may be needed. An object that has an anisotropic BRDF may require an even larger number of sources. The worst case, however, is when the object is a perfect mirror. In this case, for any light source, there is only one viewing direction for which a surface point produces a non-zero image intensity. Thus, we would need an infinite number of light sources to measure all the surface normals on a perfect mirror.

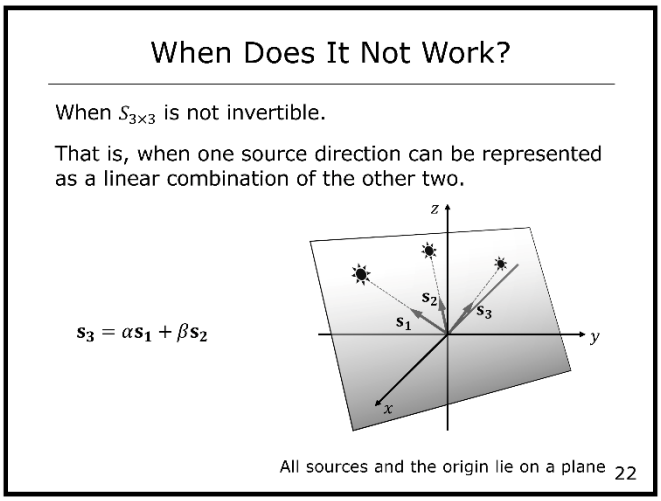
Let us take a closer look at the Lambertian case. In our analysis here, we will solve for surface normal using explicit expressions for the measured image intensities, instead of using the gradient space and reflectance maps. While we previously normalized our measured intensities to eliminate the brightness of each source, the camera gain, and the albedo, in this case, we reintroduce the albedo of the surface as an unknown in our expression for image intensity. The image intensity I_1 due to source $\mathbf{s}_1$ is therefore the albedo ρ divided by π, times $\mathbf{n} \cdot \mathbf{s}_1$. Similarly, we have equations for image intensities I_2 and I_3 corresponding to sources $\mathbf{s}_2$ and $\mathbf{s}_3$, respectively. We can represent the three image intensity expressions together in matrix form [1]. In this equation, the 3x3 source matrix S and the vector of image intensity values are known. The albedo and surface normal vector are unknown.



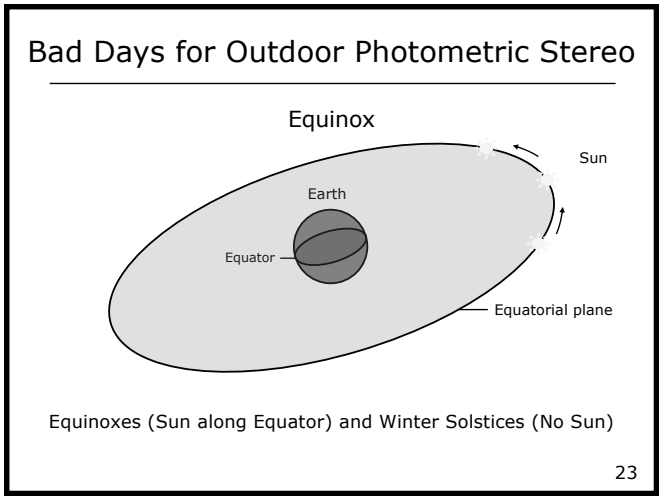
The matrix form can be represented as a simple expression where the vector $\boldsymbol{I}$ of measured intensities is equal to the source matrix S multiplied by the normal vector $\mathbf{N}$. Note that $\mathbf{N}$ is the unit normal vector $\mathbf{n}$ scaled by the ρ / π [1]. We solve this expression for $\mathbf{N}$, by multiplying both sides by the inverse of S [2]. From $\mathbf{N}$, we can compute the unit normal vector by dividing $\mathbf{N}$ by its magnitude. In addition, the magnitude of $\mathbf{N}$ corresponds to the albedo ρ divided by π. Thus, using just three light sources, we are able to find the normal and albedo at each point on a Lambertian object. This means that the normals of a Lambertian object can be estimated, even when it has an unknown texture.



This approach will only work when the source matrix S is invertible. Note that S is not invertible when the light sources lie on a plane that includes the surface point, as shown here. In this case, any one of the sources can be expressed as a linear combination of the other two, making the matrix S singular.



Photometric stereo can be applied outdoors as well by using the sun as a light source with varying direction. We can simply point a camera at an object of interest (say, a building) and take images at different times of the day. It turns out, however, that there are times of the year when this approach will not work because the source matrix S ends up not being invertible.



One example of a bad time for outdoor photometric stereo is an equinox. During an equinox, the sun moves along a trajectory that lies on a plane that passes through the equator of the Earth. In this case, irrespective of where any object lies on the surface of the Earth, it also lies on the plane on which all the sun positions lie, called the equatorial plane. In this case, any sun direction can be expressed as a linear combination of any two of its other directions, resulting in a singular matrix S.

Another bad time for photometric stereo is the winter solstice, during which time parts of the Earth do not get any sunlight. The parts that do get some sunlight, receive weak rays that subtend an angle of less than fifteen degrees above the horizon.

In our above discussion, we showed that a minimum of three source directions are needed to compute normals for a Lambertian surface. In practice, however, to get reliable normal estimates, we want to use as many light sources as the application allows us to. In this case, $\boldsymbol{I}$ is a Kx1 vector and S is a Kx3 matrix, where $K > 3$. Although S is no longer invertible, we can solve for $\mathbf{N}$ with the least squares method, which uses a pseudo-inverse. This is done by pre-multiplying both sides of $\boldsymbol{I} = S\mathbf{N}$ by S^T. Since $S^T S$ is invertible, we can solve for $\mathbf{N}$ [1]. As before, we can find the unit normal $\mathbf{n}$ by dividing $\mathbf{N}$ by its magnitude, and the magnitude itself is equal to the albedo ρ divided by π. This method can be used for any number K of light sources larger than three.

More Sources than Minimum Needed

Better results by using more ($K > 3$) light sources

$$\underbrace{\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_K \end{bmatrix}}_{I_{K\times 1}} = \frac{\rho}{\pi} \underbrace{\begin{bmatrix} s_{x_1} & s_{y_1} & s_{z_1} \\ s_{x_2} & s_{y_2} & s_{z_2} \\ \vdots & \vdots & \vdots \\ s_{x_K} & s_{y_K} & s_{z_K} \end{bmatrix}}_{S_{K\times 3}} \mathbf{n} \Rightarrow \begin{cases} I = S\mathbf{N} \\ \text{where: } \mathbf{N} = \frac{\rho}{\pi}\mathbf{n} \end{cases}$$

$S_{K\times 3}$ is not a square matrix and hence not invertible.

Solution: Use Least Squares Estimation

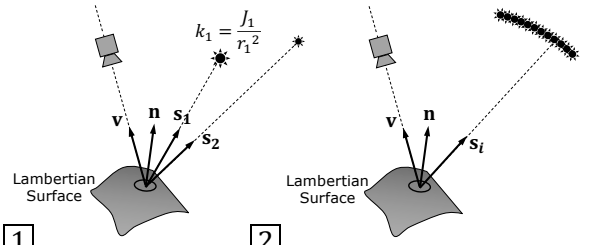
$$S^T I = \underbrace{S^T S}_{3\times 3}\mathbf{N}$$

$$\mathbf{N} = \underbrace{(S^T S)^{-1} S^T}_{\text{Pseudo Inverse}} I \quad [1]$$

24

Consider the image of a Lambertian surface lit simultaneously by multiple light source. Assume that the multiple sources are visible to all points on the surface. It turns out that the multiple sources can be replaced by a single point source to obtain an image of the surface that is identical to that obtained with the multiple sources. We refer to this point source as the "effective" source.

"Effective" Point Source for Multiple Sources



[1] $I_1 = \frac{\rho}{\pi} k_1 \mathbf{n} \cdot \mathbf{s_1}$ [2] $I_2 = \frac{\rho}{\pi} k_2 \mathbf{n} \cdot \mathbf{s_2}$ $I = \frac{\rho}{\pi} \mathbf{n} \cdot \sum_i k_i \mathbf{s_i}$

[3] $I = \frac{\rho}{\pi} \mathbf{n} \cdot (k_1 \mathbf{s_1} + k_2 \mathbf{s_2}) = \frac{\rho}{\pi} \hat{k}\, \mathbf{n} \cdot \hat{\mathbf{s}}$ $I = \frac{\rho}{\pi} \hat{k}\, \mathbf{n} \cdot \hat{\mathbf{s}}$ [4]

$(\hat{k}, \hat{\mathbf{s}})$: Effective Source is in direction of "Centroid" of Sources

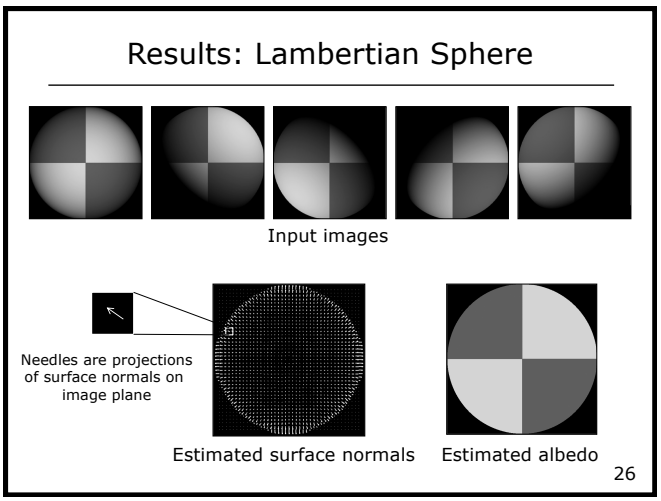
25

To illustrate this, consider a point on a Lambertian surface with unit normal $\mathbf{n}$ and viewing direction $\mathbf{v}$. We turn on a point light source in the direction $\mathbf{s}_1$ with brightness k_1. From our earlier discussion, the brightness of the source can be represented by its radiant intensity J_1 divided by the square of its distance r_1 from the surface. Thus, the image intensity of the point on a Lambertian surface due to this first source can be written as equation [1].

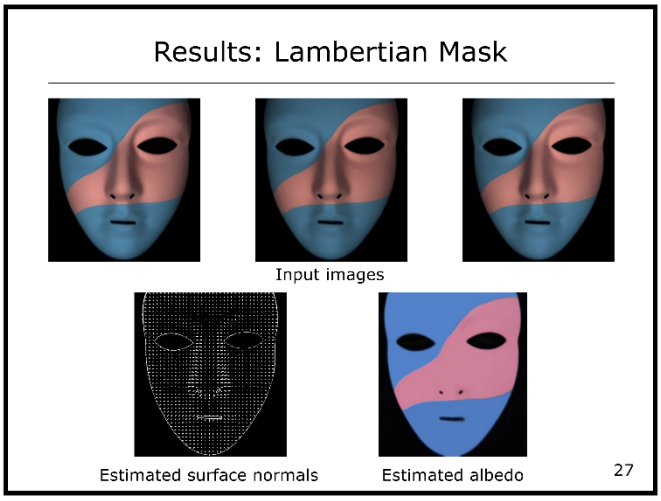
Now, with the above source on, we turn on a second point source with brightness k_2 and direction $\mathbf{s}_2$. The intensity I_2 at the surface point due to just the second source can be written as equation [2]. The total brightness of the scene point is therefore the sum of the brightnesses due to the two sources. The effective light source is a vector computed as the sum of $k_1\mathbf{s}_1$ and $k_2\mathbf{s}_2$, as in equation [3]. $\hat{k}$ and $\hat{\mathbf{s}}$ are the magnitude and direction of the effective light source, respectively. Therefore, for a Lambertian object viewed by a camera, if we turn on two point sources at the same time, as long as the two sources are visible to all points on the object, we can replace the two sources with a single point source to obtain

an image that is identical to the one produced by the two sources. By induction, this result holds true for any number of simultaneously activated point sources. Therefore, the above result applies to extended (area) sources as well. We can view the extended source, shown on the right, and a collection of an infinite number of point sources, and the effective point source is given equation 4.

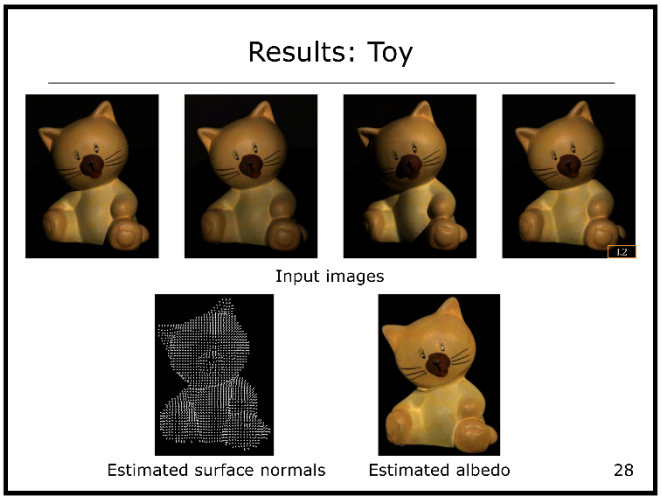
Here are the results of applying photometric stereo to a textured Lambertian sphere using five light sources. Note that, in any real setting, a surface point can be occluded from one or more sources by other points on the object. By using more than the minimum required number of three sources, we ensure that at least three non-zero brightness values are measured at each surface point, and use these values to compute its normal. In the surface normal map shown in the second row, the length of the vector (needle) represents the tilt of the normal with respect to the viewing direction, and the direction in which the needle points represents the second angle of the normal. The computed albedo map reveals that the sphere has four segments, each with uniform albedo.



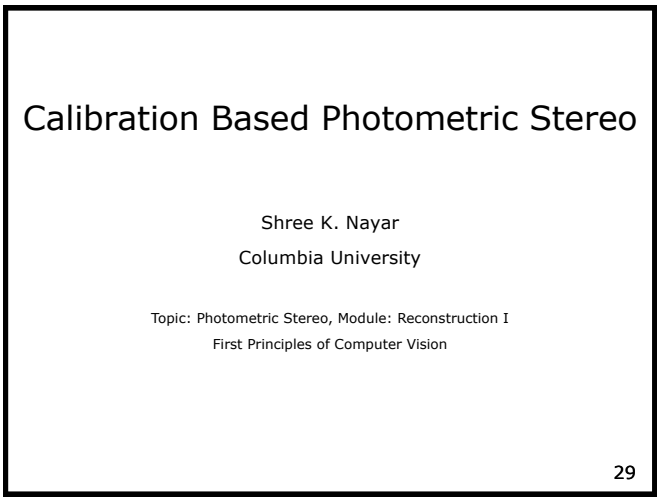
Here are results of photometric stereo applied to a painted mask. Since the mask is fairly shallow, three light sources were sufficient in this case. Since color images were captured in this case, we applied photometric stereo to each color channel separately, and the normal at each point was estimated as an average of the three computed normals. The albedo in this case is a color vector. Note that computed normal map turns out to be quite accurate. Even though there exists shading and brightness variations in each of the three captured images, the color albedo map is constant over each patch that was painted with the same paint.



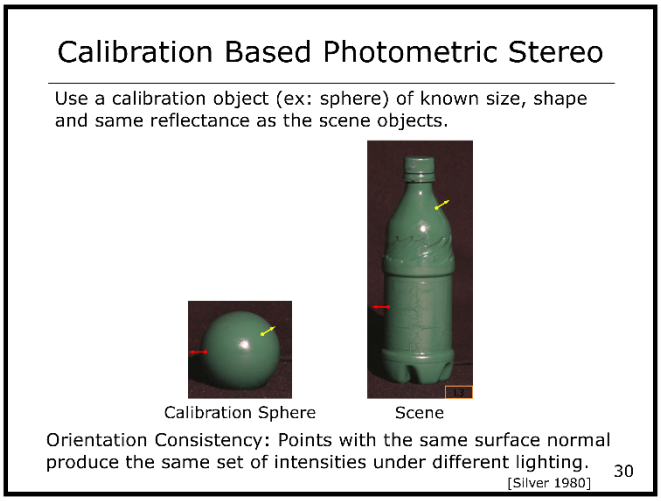
Here are results for a toy, computed using four light sources. The resultant normal map is reasonably accurate, except around the creases and edges. In fact, the albedo map also has higher than expected values in these regions and the forehead of the toy. This is likely because we have assumed the BRDF to be Lambertian, when it is non-Lambertian with a slight specular component. This example illustrates that there are many objects in the real-world where a simple BRDF model is inadequate and would cause errors in the computed normals and albedo values.



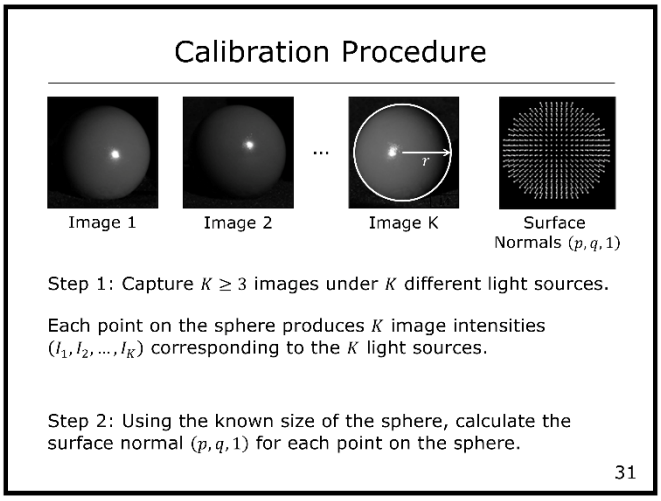
Even when we have access to the material of an object, we cannot always write a simple expression for its BRDF. However, since we have access to the material, we can fabricate an object of known shape with the same material properties and use it as a calibration object to apply photometric stereo to any other object made of the same material. This data-driven method is referred to as calibration-based photometric stereo. It is often used in the case of plastics or painted manufactured objects.



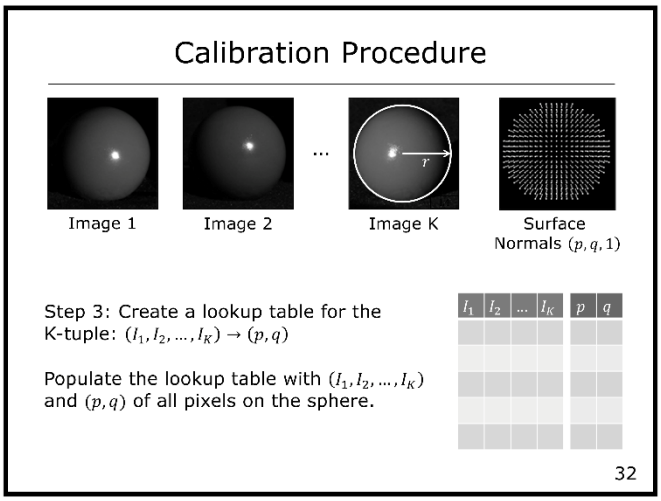
On the right is a bottle that has been spray-painted. We do not know the BRDF of the paint, but we can create the calibration sphere on the left and coat it with the same spray-paint. Now consider both these objects illuminated by a single source. We know that a surface normal on the bottle will produce a brightness that is the same as that produced by the point on the sphere with the same normal. This holds true for any lighting (source). We refer to this property as orientation consistency, and it is the basis for calibration-based photometric stereo.



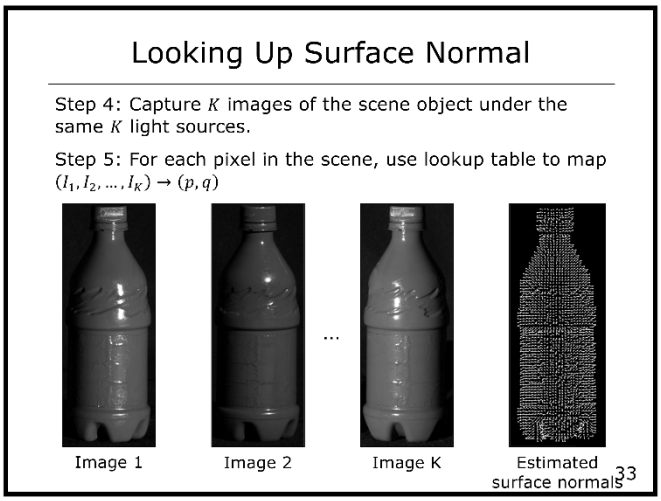
Here are the steps taken off-line (before applying photometric stereo to the objects of interest). First, we create a calibration object (a sphere) with the same BRDF as the object of interest. We take K images of the calibration object with the same K light sources that we are plan to use for the target object. At each point on the sphere, we now have K image intensities (I_1 through I_k) corresponding to the K light sources. Since the object is a sphere, by simply identifying the boundary of the sphere in an image, we can determine the normal (p, q) at each point on the scene.



Next, we create a lookup table, where the index of the table is the *K*-tuple of image intensities, I_1 through I_k, and the entry is the surface normal (p, q) that produced the image intensities. This lookup table is populated using all pixels (normals) on the calibration sphere.



Once we have the lookup table, the object of interest is illuminated with the same set of light K sources and K images are captured. For each point on the object, we use the *K*-tuple of image intensities to index the lookup table and read out the corresponding (p, q) value. The end result is a complete surface normal map, like the one shown on the right.



Photometric stereo gives us the normal at each point on a surface. We can recover the 3D shape of the surface by integrating the surface normals. Note that the estimated shape of the surface will be accurate only when the surface is continuous and does not include discontinuities. In the case of a discontinuity, although we know the normal on either side of it, we have no way of determining the amount by which the depth of the surface changed at a discontinuity. In developing our "shape from normal" method, we will therefore assume surface continuity.

Shape from Normals

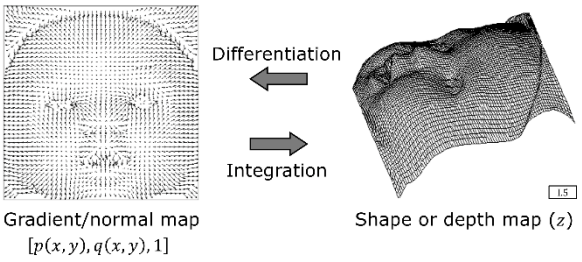
Shree K. Nayar
Columbia University

Topic: Photometric Stereo, Module: Reconstruction I
First Principles of Computer Vision

34

On the left here is a surface normal map measured using photometric stereo, where each needle represents a surface normal (p, q). Assume that the normal map corresponds to the surface $z(x, y)$ shown on the right, where z is the depth at position (x, y). From the definition of surface gradients, p is the negative of the derivative of z with respect to x, and q is the negative of the derivative of z with respect to y. Thus, we can compute the normal map from the depth map, by differentiation. However, we want to go from the normal map to the depth map, by integration.

Shape From Surface Normals



$$(p, q) = \left(-\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y}\right)$$

35

To recover shape, we begin by setting the depth of a reference point (x_0, y_0) on the surface to be equal to 0. Now, we can integrate the surface along any path starting from the reference point (x_0, y_0) and ending at some other point (x, y), by making use of the gradient values p, q that were estimated using photometric stereo. By integrating p in the x direction, and q in the y direction, we end up with a profile of the surface along that path. Shown here are two different paths over which integration is done. If the (p, q) values computed are accurate, and the surface is continuous, integration along both paths will result in the same depth z at (x, y).

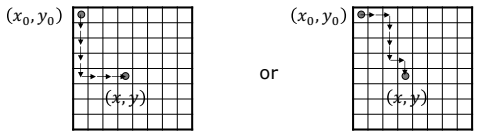
Shape From Surface Normals

Estimate surface by integrating surface gradient:

$$z(x, y) = z(x_0, y_0) + \int_{(x_0, y_0)}^{(x, y)} -(p\,dx + q\,dy)$$

where (x_0, y_0) is a reference point and $z(x_0, y_0) = 0$.

$z(x, y)$ obtained by integration along any path from (x_0, y_0).

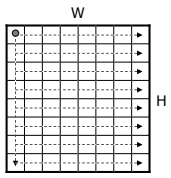


36

The above integration approach can be implemented as a naïve algorithm. First, we assign the depth value of 0 to the top-left corner pixel. Next, we integrate along the first column, using the q values, to compute the depths of all these pixels. These depths are then used as seed values to integrate along the horizontal direction, using the p values, to end up with a complete depth map.

Naïve Algorithm for Estimating Shape

1. Initialize reference depth
   $z(0,0) = 0$
2. Compute depth for first column
   for $y = 1$ to $(H-1)$
   $z(0,y) = z(0,y-1) - q(0,y)$
3. Compute depth for each row
   for $y = 0$ to $(H-1)$
   for $x = 1$ to $(W-1)$
   $z(x,y) = z(x-1,y) - p(x,y)$

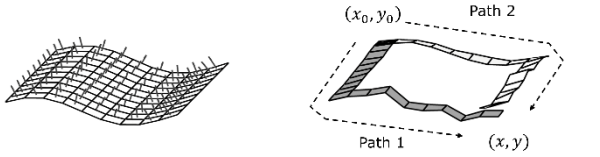


Computed Depth Map

37

The above naïve algorithm for estimating shape is simple. Unfortunately, it does not work well as our measured surface normals are not precise and will include errors, due to noise in our measured images. The errors could also be due to our assumed BRDF model not being precise. Irrespective of the source of the errors, the impact of the errors during integration is shown on the right. Here we have the results of integration between a reference point and an end point using two different paths. Due to the errors in the normals, we see that the two depth values computed for the point (x, y) differ. An ad-hoc way of dealing with noise in the normals is to choose many different paths to integrate the surface. This results in many depth values for each pixel. A simple average is then computed to obtain the depth at each pixel. The hope here is that averaging will reduce the effects of noise in the normals on the computed depth map.

Noise Sensitivity of Computed Shape



Actual surface shown with noisy estimates of surface gradient

Depth computed from noisy gradients depends on the integration path

One Solution: Compute depth maps using different paths. Then, find average of computed depth maps to reduce error.

38

It turns out that we can do much better by using a least squares approach. We know that the measured gradients (p, q) correspond to derivatives of a surface $z(x, y)$. Let us imagine for a moment that $z(x, y)$ has been computed. Then, we know that the negatives of its derivatives with respect to x and y should equal p and q, respectively. We therefore formulate an error measure D as the sum of the squares of the derivative errors, integrated over the image [1]. We then wish to find the surface $z(x, y)$ that minimizes D.

Estimating Shape Using Least Squares

Minimize the errors between measured surface gradients (p, q) and surface gradients of estimated surface $z(x, y)$.

Error Measure:

$$D = \iint_{Image} \left(\frac{\partial z}{\partial x} + p\right)^2 + \left(\frac{\partial z}{\partial y} + q\right)^2 dxdy \quad [1]$$

where $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are gradients of the estimated surface.

We need to find $z(x, y)$ that minimizes D

[Horn 1990] 39

The Frankot-Chellappa algorithm is an elegant algorithm that minimizes the error measure D in the Fourier domain. Let $Z(u, v)$, $P(u, v)$, and $Q(u, v)$ be the Fourier transforms of $z(x, y)$, $p(x, y)$, and $q(x, y)$, respectively. We can then represent z, p and q in terms of Z, P and Q using the inverse Fourier transform. These expressions for z, p and q are plugged into the equation for the error measure D in slide 39.

Frankot-Chellappa Algorithm

Minimize objective function D in Fourier Domain.

Let $Z(u, v)$, $P(u, v)$ and $Q(u, v)$ be the Fourier Transforms of $z(x, y)$, $p(x, y)$ and $q(x, y)$, respectively. Then:

$$z(x, y) = \iint_{-\infty}^{\infty} Z(u, v) e^{i2\pi(ux+vy)} dudv$$

$$p(x, y) = \iint_{-\infty}^{\infty} P(u, v) e^{i2\pi(ux+vy)} dudv$$

$$q(x, y) = \iint_{-\infty}^{\infty} Q(u, v) e^{i2\pi(ux+vy)} dudv$$

Substitute for $z(x, y)$, $p(x, y)$ and $q(x, y)$ in equation for D.

[Frankot 1988] 40

Since the error D now includes the Fourier transforms of z, p and q, finding the $z(x, y)$ that minimizes D is equivalent to finding the Fourier transform of z that minimizes D. Hence, we find the derivative of D with respect to the Fourier transform of z and set that equal to zero. This results in a very simple expression [1]. Since all the terms on the right hand side are known, we get the the Fourier transform of z, i.e., $\tilde{Z}$. Finally, we can find the inverse Fourier transform of $\tilde{Z}$ to get the surface $\tilde{z}$.

Frankot-Chellappa Algorithm

Find $Z(u, v)$ that minimizes D using $\frac{\partial D}{\partial z} = 0$.

Solution: $\tilde{Z}(u, v) = \frac{iuP(u, v) + ivQ(u, v)}{u^2 + v^2}$ [1]

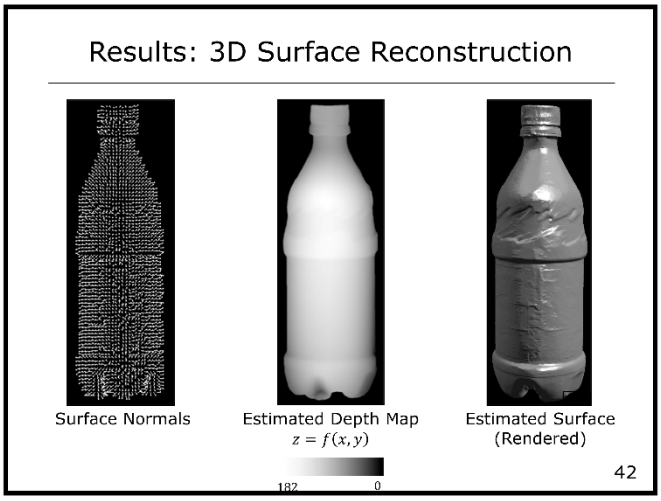
This is the Fourier Transform of the best fit surface.

Compute Inverse Fourier Transform to obtain $\tilde{z}(x, y)$.

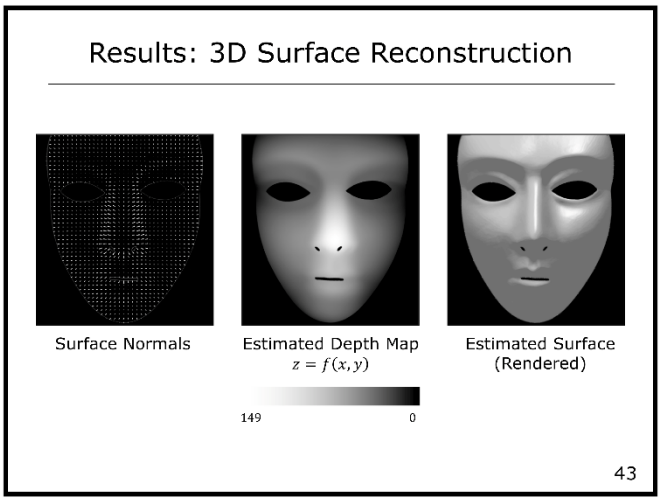
$$\tilde{Z}(u, v) \longrightarrow \boxed{\text{IFT}} \longrightarrow \tilde{z}(x, y)$$

[Frankot 1988] 41

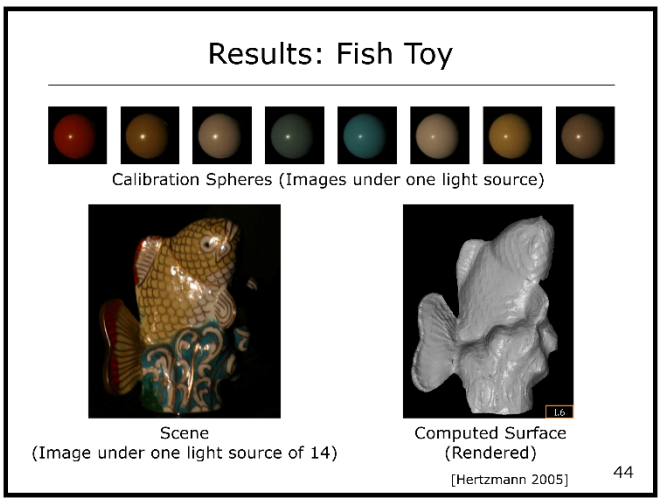
Here is an example of the Frankot-Chellappa algorithm applied to the computed normal map (left) of the bottle in slide 33 to get a depth map $z(x, y)$ (middle). The depth map is shown as an intensity map, where the brighter the point, the closer it is to the camera. Since it is difficult to figure out the quality of this depth map, computer graphics can be used to render an image (right) of the object using the computed depth map, a chosen lighting, and an assumed BRDF. This image reveals that details in the computed depth map.



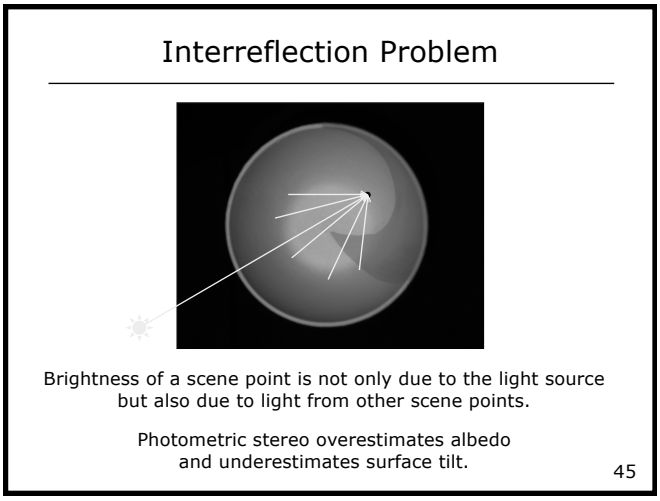
Here is another example where the Frankot-Chellappa algorithm is applied to the computed normal map of the mask in slide 27 to get a depth map (middle). Once again, computer graphics is used to render an image (right) of the depth map. We can see that the shape of the rendered surface is consistent with that of the mask in slide 27.



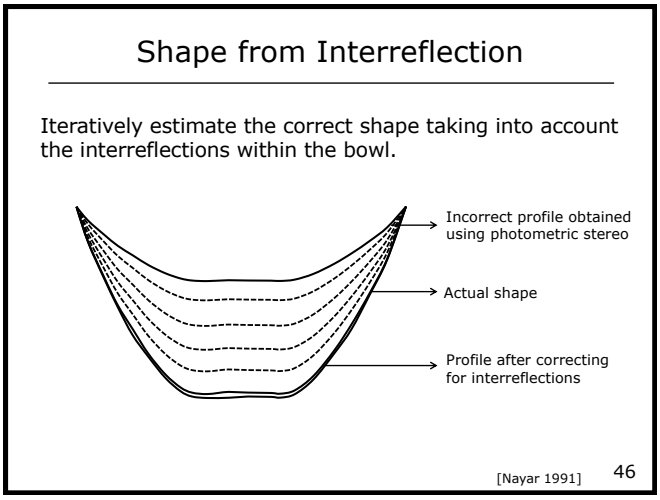
Here is an interesting example where the object is made up of many different materials (paints). In this case, a calibration object (sphere) is used for each one of the paints. In this example, 14 light sources where used to illuminate the object. An interesting question is how to determine which calibration object (lookup table) to use for each point on the object. A simple approach is to use the color of each point on the object to determine the calibration object for that point. Once the normal map is computed, the Frankot-Chellappa algorithm is used to get the depth map, a rendering of which is shown on the right.



Lastly, let us discuss the interreflection problem, which impacts the performance of photometric stereo. Here, we have a bowl illuminated by a single point light source. In photometric stereo, we assume that each surface point receives light only from the light source. However, when an object is concave, as in the case of this bowl, each point also receives light from all other object points that are visible to it. This effect is called interreflection. While developing photometric stereo, we ignored this problem as the light received by a point from other object points depends on the shape of the object, which is exactly what we wish to estimate. Due to interreflections, the albedo values of a concave object computed using photometric stereo tend to be overestimated, because each point reflects more light than it would if it had only received light from a source. In addition, interreflections cause the computed shape of a concave surface to be shallower than it actually is.



In order to deal with the effects of interreflection, we can make use of a simple iterative algorithm. Shown here is the cross-section of the actual shape of a Lambertian bowl. Upon applying photometric stereo, we end up with a shallow bowl and an inaccurate albedo map, due to interreflections. We can make use of this inaccurate shape and albedo map of the bowl to get an estimate of the interreflections received by each point on the bowl. We then remove these interreflection contributions from the images captured for each of the sources. We apply photometric stereo to these interreflection-compensated images to obtain a better estimate of the bowl's shape and albedo map. Note that since we used the initial (incorrect) shape and albedo map to estimate the interreflections, the new shape and albedo map are not the shape and albedo map of the actual bowl. Therefore, the above process is repeated until the shape and albedo map converge. We see here that the shape of the bowl computed using the above approach is very close to its actual shape.



The above result is for a Lambertian surface with a very simple geometry. In general, interreflections are very hard problem to deal with. That said, it is an important problem as virtually every 3D scene exhibits interreflections.

References: Textbooks

Textbooks:

Robot Vision (Chapter 10)
Horn, B. K. P., MIT Press

Computer Vision: A Modern Approach (Chapter 5)
Forsyth, D and Ponce, J., Prentice Hall

48

References: Papers

[Frankot 1988] R. Frankot and R. Chellappa. "A Method for Enforcing Integrability in Shape from Shading Algorithm." IEEE PAMI, 1988.

[Hertzmann 2005] A. Hertzmann, S. M. Seitz. "Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs." IEEE PAMI, 2005.

[Horn 1990] Horn, B.K.P. "Height and Gradient from Shading," IJCV, 1990.

[Nayar 1991] S.K. Nayar, K. Ikeuchi and T. Kanade. "Shape from Interreflections." IJCV, 1991.

[Silver 1980] W.M. Silver. "Determining Shape and Reflectance Using Multiple Images." Master's thesis, MIT, Cambridge, Mass., 1980.

[Woodham 1980] R. Woodham, "Photometric Method for Determining Surface Orientation from Multiple Images." Optical Engineering, 1980.

49

Image Credits

I.1 David Forsyth and Jean Ponce. Used with permission.
I.2 Dan Goldman. Used with permission.
I.3 Aaron Hertzmann. Used with permission.
I.4 Aaron Hertzmann. Used with permission.
I.5 Aaron Hertzmann. Used with permission.
I.6 Aaron Hertzmann. Used with permission.
I.7 Aaron Hertzmann. Used with permission.

50

Acknowledgements: Thanks to Joel Salzman, Tracy Cui, Kevin Chen and Nikhil Nanda for their help with editing this article.

References

[Horn 1986] B. K. P. Horn, Robot Vision, MIT Press, 1986.

[Forsyth 2003] D. Forsyth and J. Ponce, Computer Vision: A Modern Approach, Prentice Hall, 2003

[Frankot 1988] R. Frankot and R. Chellappa, A Method for Enforcing Integrability in Shape from Shading Algorithm, IEEE PAMI, 1988.

[Hertzmann 2005] A. Hertzmann and S. M. Seitz, Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs, IEEE PAMI, 2005.

[Horn 1990] Horn, B. K. P. "Height and Gradient from Shading," IJCV, 1990.

[Nayar 1991] S.K. Nayar, K. Ikeuchi and T. Kanade., Shape from Interreflections, IJCV, 1991.

[Silver 1980] W. M. Silver., Determining Shape and Reflectance Using Multiple Images, Master's thesis, MIT, Cambridge, Mass., 1980.

[Woodham 1980] R. Woodham, Photometric Method for Determining Surface Orientation from Multiple Images, Optical Engineering, 1980

[Nayar 2025C] Radiometry and Reflectance, Nayar, S. K., Monograph FPCV-3-1, First Principles of Computer Vision, Columbia University, New York, February 2025.

[Nayar 2025E] Shape from Shading, Nayar, S. K., Monograph FPCV-3-3, First Principles of Computer Vision, Columbia University, New York, March 2025.

[Nayar 2025G] Active Illumination Methods, Nayar, S. K., Monograph FPCV-3-5, First Principles of Computer Vision, Columbia University, New York, March 2025.