

# Lecture 12

Gidon Rosalki

2026-01-11

## 1 Vanishing points

### 1.1 Measuring height

Given an object in the scene, where we know its height, we may measure the height of the camera based off the point at which the object of known height meets the horizon line, and that is the height of the camera. We define *parallel lines* to be lines that meet at the horizon. A line on the ground is any line between two points, so its parallel line will be any line that meets it at the horizon, and goes through another point above the horizon line.

If a camera is looking up or down, it matters not, the horizon line will still be at the same point in the image, for a given height of the camera.

### 1.2 Perspective

In 3 point perspective, the vertical lines will meet at a vanishing point below the image if we are looking at it from above, and above the image if we are looking at it from below. The perpendicular horizontal lines will meet at the opposite sides of the image at their respective vanishing points. All the vertical lines will correspond to the *same* vertical vanishing point. There was a time before perspective in art, where there is no, or little feeling of depth, and lines do not meet at vanishing points, they are perfectly parallel. Original images in generative AI also had lots of vanishing point issues, though these problems have mostly been resolved these days. By Van Gogh artists had started working with perspective, and vanishing points correctly.

One may make use of this to change the apparent size of objects in images / drawings. For example, two things may be the same actual size, but if one is much closer to the vanishing points then it will appear much larger.

### 1.3 Visual cues

Based off things like occlusion, shading, texture, focus, highlights, shadows, and more, we can get cues about the location of different objects in an image.

For **depth cues**, consider occlusion. The occluding object is closer to the camera than the occluded object. Additionally, size, and the location of the bottom of the object can give further cues as to the location of the object. When one object occludes the other, we will note that there are 2 edges that continue in different directions, one occluded, and one continuous. We call these points *T junctions* (note, they are not necessarily perpendicular), and they also give us clues as to occlusion and depth.

Haze in images can also give us clues on the depth of areas. Hazy parts are almost always further away than less hazy parts of an image. This will be discussed further in the tutorial.

## 2 Photometric stereo

To do this, we firstly need to find the normal of a surface. The equation of a plane is given as

$$Ax + By + Cz + D = 0$$

So, the surface normal is

$$N = \left( \frac{A}{C}, \frac{B}{C}, 1 \right) = (p, q, 1)$$

The normal of this vector is simply

$$n = \frac{N}{|N|} = \frac{(p, q, 1)}{\sqrt{p^2 + q^2 + 1}}$$

The light source vector (as in, the vector from which light arrives in the scene), is as follows:

$$S = \left( \frac{X}{Z}, \frac{Y}{Z}, 1 \right) \tag{1}$$

$$s = \frac{S}{|S|} = \frac{(p_S, q_S, 1)}{\sqrt{p_S^2 + q_S^2 + 1}} \tag{2}$$

So, the cos of the angle between them is given by

$$\cos \theta_i = n \cdot s = \frac{(p p_S + q q_S + 1)}{\sqrt{p^2 + q^2 + 1} \sqrt{p_S^2 + q_S^2 + 1}}$$

The plane at  $z = 1$  is called the gradient space (pq plane), and every point in it corresponds to a particular surface orientation.

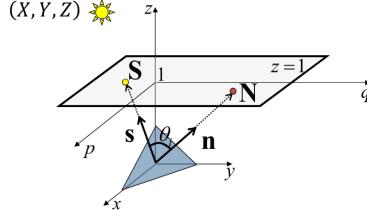


Figure 1: Gradient space

The **reflectance map** relates image irradiance  $I(x, y)$  to surface orientation  $(p, q)$  for a given light source direction, and surface reflectance. The Lambertian reflectance model, for matte / diffuse targets takes

$$k = \text{Light source brightness} \quad (3)$$

$$\rho = \text{Surface albedo (reflectance)} \quad (4)$$

$$c = \text{Constant (optical system)} \quad (5)$$

$$(6)$$

So the image irradiance is given as

$$I = \rho k c \cos \theta_i = \rho k c n \cdot s$$

If we let  $\rho k c = 1$  then  $I = \cos \theta_i = n \cdot s$ . We may state that these values are approximately equal, where  $\theta$  is the angle between the light source, and the surface normal.  $I$  does not depend on the viewing direction  $V$ .

There are 3 main simple reflectance models. Diffuse models assume that an inbound light ray is sent in every direction. Glossy models send the light in mostly the same direction, and in mirror, then the angle of incidence is equal to the angle of reflection.

From this calculated  $I$ , we can create the *reflectance map*  $R(p, q)$ , with  $p$  on the  $x$  axis, where  $R(p, q)$  is at a maximum when  $(p, q) = (p_S, q_S)$ .

Now, to create photometric stereo, we can take multiple images, each from a different light source direction. The points where the lines cross is a shared point in the image, from multiple different perspectives.

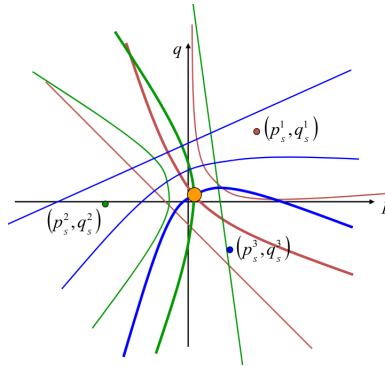


Figure 2:

Note that this means that photometric stereo does not work outside, since we cannot trivially move the position of the sun, and there are often many light sources outside, thanks to reflections. This method only works if we can control the light source, and ensure that there is only one, that we can move.

To perform Lambertian photometric stereo, then  $I = \rho k c \cos \theta_i = \rho n \cdot s$ , we have 3 image radiances:

$$I_1 = \rho n \cdot s_1 \quad (7)$$

$$I_2 = \rho n \cdot s_2 \quad (8)$$

$$I_3 = \rho n \cdot s_3 \quad (9)$$

$$(10)$$

We can write this as the matrix equation

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} s_1^T \\ s_2^T \\ s_3^T \end{bmatrix} \rho n \quad (11)$$

Which we can now solve.  $I$  is a 3 by 1 matrix,  $S$  a 3 by 3, and  $\tilde{n} = \rho n$  is a 3 by 1.

$$\tilde{n} = S^{-1}I \quad (12)$$

$$\rho = |\tilde{n}| \quad (13)$$

$$n = \frac{\tilde{n}}{|\tilde{n}|} = \frac{\tilde{n}}{\rho} \quad (14)$$

We can compute the light source direction by placing a chrome sphere in the scene, and the location of the highlight is the direction of the light source.

## 2.1 Normal map vs depth map

We have taken every point on the surface, and found its normal as a point on a plane that is above. We want to recreate the original shape from this plane. We can determine the depth map from the normal map by integrating over gradients  $p, q$  across the image. A normal map that produces a unique depth map, that is independent of the image plane direction that is used to sum over the gradients is called *integrable*. Integrability is enforced when the following condition holds:

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$$

The Penrose staircase violates integrability.

By taking many images under varying illuminations, we can reconstruct the x gradient, and y gradient images. From here, via 2D integration, we can recover the reconstructed shape. This is done by estimating the derivatives of the surface

$$\hat{p}(x, y) = \frac{(z(x+1, y) - z(x-1, y))}{2} \quad (15)$$

$$\hat{q}(x, y) = \frac{(z(x, y+1) - z(x, y-1))}{2} \quad (16)$$

Represented as a matrix multiplication. From here, we find  $z(x, y)$  that minimises

$$E = \sum_{x, y} \left( (\hat{p}(x, y) - p(x, y))^2 + (\hat{q}(x, y) - q(x, y))^2 \right)$$

Using the pseudo inverse, since there are 1 million unknowns for a 1K by 1K image.

There are some serious limitations of photometric stereo. It does not work for shiny / colourful / semi translucent surfaces, and has problems with shadows, and inter reflections (ie, non convex objects). Additionally, the camera, and lights have to be distant, and there are calibration requirements, of measuring the light source directions, and intensities, along with the camera response function.

## 2.2 Shape from shading

Here we are asking, given a **single** image of an object, with known surface reflectance, taken under a known light source, can we recover the shape of the object? This is equivalent to saying, given  $R(p, q)$ ,  $(p_s, q_s)$ , and surface reflectance, can we determine  $(p, q)$  *uniquely* for each image point? In short, no. There is no unique solution given this.

We may add *smoothness constraints* to help here, where we assume that the surface orientation  $(p, q)$  is similar at neighbouring points. So we want to minimise

$$\int \int (p_x^2 + p_y^2 + q_x^2 + q_y^2) + \lambda (I(x, y) - R(p, q))^2 dx dy$$

The first term is the smoothness term, and the second the data term. We want to minimise the sum of square derivatives of  $p$  and  $q$ , and so add a penalty for a rapid change in  $p$  and  $q$ . Note that this involves the first, and second derivatives of  $z$ .

We should also note that if we can see the edges of a shape, then we know what its normals are, and can resolve the issue that way.